# The Oncologist®

# The Relationship Between Eight GWAS-Identified Single-Nucleotide Polymorphisms and Primary Breast Cancer Outcomes

Soley Bayraktar,[a] Patricia A. Thompson,[b] Suk-Young Yoo,[c] Kim-anh Do,[c] Aysegul A. Sahin,[d] Banu K. Arun,[e] Melissa L. Bondy,[f] Abenaa M. Brewster[g]

[a]Department of Medical Oncology, Mercy Cancer Center, Ardmore, Oklahoma, USA; [b]Department of Cellular and Molecular Medicine, Arizona Cancer Center, University of Arizona, Tucson, Arizona, USA; Departments of [c]Biostatistics, [d]Pathology, [e]Breast Medical Oncology, and [g]Clinical Cancer Prevention, University of Texas MD Anderson Cancer Center, Houston, Texas, USA; [f]Baylor College of Medicine, Cancer Prevention and Population Sciences, Houston, Texas, USA
Disclosures of potential conflicts of interest may be found at the end of this article.

**Key Words.** Breast cancer • Prognosis • Single-nucleotide polymorphisms • TNRC9 • 17q23

| Learning Objectives | Describe the results of genome-wide association studies (GWAS) that have identified genetic variants associated with breast cancer risk. |
|---|---|
| | Discuss whether genetic risk variants identified through genome-wide association studies (GWAS) are also associated with breast cancer prognosis. |
| | Describe molecular mechanisms through which germline genetic variants may influence breast cancer survival. |

## Abstract

**Background.** Several single-nucleotide polymorphisms (SNPs) associated with breast cancer risk have been identified through genome-wide association studies (GWAS). We investigated whether eight risk SNPs identified in GWAS were associated with breast cancer disease-free survival (DFS) and overall survival (OS) rates.

**Patients and Methods.** A cohort of 739 white women with early-stage breast cancer was genotyped for eight GWAS-identified SNPs (rs2981582, rs1219648 [*FGFR2*], rs3803662, rs12443621, rs8051542 [*TOX3*], rs999737 [*RAD51L1*], rs6504950 [17q23], and rs4973768 [3p24]). Relationships between SNPs and breast cancer outcomes were evaluated using Cox proportional hazard regression models. The cumulative effects of SNPs on breast cancer outcomes were assessed by computing the number of at-risk genotypes.

**Results.** At a median follow-up of 121 months (range: 188–231 months) for survivors, 237 deaths (32%) and 186 breast cancer events (25%) were identified among the 739 patients. After adjusting for age, clinical stage, and treatment, rs12443621 (16q12; $p = $ .03) and rs6504950 (17q23; $p = $ .008) were prognostic for OS but not DFS. A higher risk for death was also found in the multivariable analysis of patients harboring three or four at-risk genotypes of the GWAS SNPs compared to patients carrying two or less at-risk genotypes (hazard ratio: 1.60, 95% confidence interval: 1.23–2.24; $p = $ .0008).

**Conclusion.** The study results suggest that previously identified breast cancer risk susceptibility loci, rs12443621 (16q12) and rs6504950 (17q23), may influence breast cancer prognosis or comorbid conditions associated with overall survival. The precise molecular mechanisms through which these risk SNPs, as well as others that were not included in the analysis, influence clinical outcomes remain to be determined. *The Oncologist* 2013;18:493–500

**Implications for Practice:** Prior genome-wide association studies (GWAS) have identified rare genetic variations in individual's DNA (SNPs) that increase the risk of breast cancer. However, the role that these GWAS discovered SNPs play in determining survival after a diagnosis of breast cancer is not clear. Here, we found that two GWAS identified SNPs were associated with overall but not breast cancer survival. Replication of these findings in different populations is needed to determine whether GWAS discovered SNPs can be used to develop prognostic and therapeutic approaches for patients with breast cancer.

## Introduction

Breast cancer is the most common female malignancy in the United States and is the second leading cause of death from cancer [1]. Despite the excellent 5-year overall survival rate of 83%–92% among women with stage I and II disease [2], the 10-

year risk of recurrence is estimated at 20%–40% [3]. Traditional prognostic factors, such as tumor size, grade, and lymph node metastasis status, are still the most important prognostic factors for long-term survival in breast cancer, but less is known about the effect of inherited genetic variation on recurrence risk and overall disease-free survival rates.

It has been estimated that there are 7 million common single-nucleotide polymorphisms (SNPs) in the human genome (with minor allele frequency [MAF] of >5%) [4]. Recently several genome-wide association studies (GWAS) [5–14] have identified multiple breast cancer susceptibility loci— of which the majority are SNPs—that contribute a small effect on breast cancer risk [15]. Some of the GWAS-identified SNPs are found to be at loci containing the genes *FGFR2* (rs2981582, rs1219648), *LSP1* (rs3817198, rs2771439), *MAP3K1* (rs726501, rs889312), *TOX3/TOX3* (rs3803662, rs8051542), *MRPS30* (rs10941679, rs7705343), *COX11* (rs10515083, rs16955329, rs2787487), *SLC4A7* (rs4973768), *TGFB1* (rs1982073), or *ESR1* (rs3020314), or at chromosomes 8p24 (rs13281615, rs283720) and 2q35 (rs13387042, rs10169372). Based on a recent meta-analysis and pooled analysis, which addressed the associations between 145 gene variants and breast cancer, rs2981582 represents the SNP in the *FGFR2* gene with the strongest association with breast cancer [16]. What remains unclear is whether these genetic variants continue to influence the biology of the disease after diagnosis and affect breast cancer survival. Although a number of studies have evaluated genetic polymorphisms in relation to prognosis [17–29], no study has examined the relative risks for breast cancer recurrence and death attributable to the recently discovered eight risk SNPs (rs2981582, rs1219648, rs3803662, rs12443621, rs8051542, rs999737, rs6504950, rs4973768).

The list of the GWAS-identified SNPs is growing exponentially. We selected eight risk SNPs that were identified based on a review of the literature [6, 10–11, 23, 30]. We conducted this retrospective analysis to elucidate a possible association between the eight GWAS-identified SNPs and clinical outcome of patients with early-stage breast cancer, adjusting for known clinical and tumor prognostic factors. We hypothesized that SNPs that are associated with increased breast cancer risk could be important genetic markers contributing to poorer breast cancer outcome.

## PATIENTS AND METHODS

The Early-Stage Breast Cancer Repository (ESBCR) is a retrospective cohort of 2,409 women diagnosed with American Joint Committee on Cancer pathologic stage I or II breast cancer and surgically treated (mastectomy or segmental mastectomy) at MD Anderson Cancer Center (MDACC) between 1985 and 2000. Criteria for eligibility and cohort details have been previously described [31]. Briefly, the ESBCR contains detailed information on patient (ethnicity, age, menopausal status) and tumor (clinical and pathological stage, estrogen receptor [ER] and progesterone receptor [PR] status, nuclear grade) characteristics, chemotherapy and endocrine treatment, radiation and surgery type (segmental mastectomy, mastectomy), and epidemiological risk factors, including family history [FH] of breast cancer in at least one first- or second-degree relative. Follow-up information was obtained by di-

**Table 1.** Patient demographic and disease characteristics

| Characteristic | n (%) |
|---|---|
| No. of patients | 739 |
| Age at diagnosis | |
| <50 yrs | 312 (42.2) |
| ≥50 yrs | 427 (57.8) |
| Family history of breast cancer | |
| No | 579 (78.3) |
| Yes | 150 (20.3) |
| Unknown | 10 (1.4) |
| Pathological stage | |
| Stage I | 211 (28.5) |
| Stage II | 528 (71.5) |
| ER and PR status | |
| Negative | 174 (23.5) |
| Positive | 528 (71.4) |
| Unknown | 20 (5.1) |
| Nuclear grade | |
| Grade I | 74 (10) |
| Grade II | 395 (53.5) |
| Grade III | 235 (31.8) |
| Unknown | 35 (4.7) |
| Treatment type | |
| Chemotherapy only | 243 (32.9) |
| Endocrine therapy only | 188 (25.4) |
| Chemotherapy and endocrine therapy | 147 (19.9) |
| None | 158 (21.4) |
| Unknown | 3 (0.4) |
| Year of diagnosis | |
| 1985–1990 | 206 (27.9) |
| 1991–1995 | 241 (32.6) |
| 1996–2002 | 292 (39.5) |

Abbreviations: ER, estrogen receptor, PR, progesterone receptor.

rect review of the medical records and linkage to the MDACC Tumor Registry, which mails annual follow-up letters to each registered patient who is known to be alive to determine their clinical status. The MDACC Tumor Registry checks the U.S. Social Security death index and the Texas Bureau of Vital Statistics for the status of patients who fail to respond to the letters.

We selected a subset of the ESBCR population with available formalin-fixed paraffin-embedded normal lymph nodes or blood samples available for DNA extraction for participation in the genotyping project. The subset of the ESBCR population was enriched to include all black (n = 196) and Hispanic patients (n = 208) and a random sample of white patients (n = 986). Because of the small percentage of patients from other ethnic groups, we included only white patients in the study analyses. A total of 247 white patients were excluded from study analysis because of DNA extraction or genotyping failure (n = 251) or insufficient clinical information (n = 6). Patients excluded from the analysis for sample failure were more likely to be diagnosed during 1985–1994 versus 1995–2000

**Table 2.** Univariate association of single-nucleotide polymorphisms with disease-free and overall survival times

| Single-nucleotide polymorphism | Disease-free survival | | | Overall survival | | |
|---|---|---|---|---|---|---|
| | No. of patients (events) | HR (95% CI) | *p* value | No. of patients (events) | HR (95% CI) | *p* value |
| **rs2981582[a] (FGFR2)** | | | | | | |
| GG | 210 (53) | 1.0 (reference) | | 210 (53) | 1.0 (reference) | |
| AG | 360 (98) | 1.1 (0.7–1.5) | .545 | 360 (98) | 1.2 (0.9–1.7) | .103 |
| AA | 153 (33) | 0.9 (0.5–1.3) | .645 | 153 (33) | 0.8 (0.5–1.2) | .306 |
| AG/AA | 513 (132) | 1.0 (0.7–1.4) | .772 | 513 (132) | 1.1 (0.8–1.5) | .372 |
| GG/AG | 570 (152) | 1.0 (reference) | | 570 (152) | 1.0 (reference) | |
| AA | 153 (33) | 1.0 (0.7–1.4) | .382 | 153 (33) | 0.6 (0.4–0.9) | .040 |
| **rs1219648[b] (FGFR2)** | | | | | | |
| AA | 205 (52) | 1.0 (reference) | | 205 (52) | 1.0 (reference) | |
| AG | 350 (95) | 1.0 (0.7–1.5) | .676 | 350 (95) | 1.1 (0.8–1.5) | .405 |
| GG | 154 (34) | 0.9 (0.5–1.3) | .647 | 154 (34) | 0.7 (0.5–1.1) | .216 |
| AG/GG | 504 (129) | 1.0 (0.7–1.4) | .886 | 504 (129) | 0.8 (0.7–1.3) | .860 |
| AA/AG | 555 (147) | 1.0 (reference) | | 555 (147) | 1.0 (reference) | |
| GG | 154 (34) | 0.8 (0.5–1.2) | .440 | 154 (34) | 0.7 (0.4–1.0) | .062 |
| **rs3803662[a] (TOX3)** | | | | | | |
| GG | 339 (85) | 1.0 (reference) | | 339 (85) | 1.0 (reference) | |
| AG | 312 (78) | 1.0 (0.7–1.4) | .768 | 312 (78) | 0.8 (0.6–1.0) | .191 |
| AA | 64 (19) | 1.2 (0.7–2.1) | .322 | 64 (19) | 0.8 (0.5–1.4) | .652 |
| AG/AA | 376 (97) | 1.0 (0.8–1.4) | .577 | 376 (97) | 0.8 (0.6–1.0) | .200 |
| GG/AG | 641 (163) | 1.0 (reference) | | 641 (163) | 1.0 (reference) | |
| AA | 64 (19) | 1.2 (0.7–2.0) | .343 | 64 (19) | 0.9 (0.6–1.5) | .922 |
| **rs12443621[b] (TOX3)** | | | | | | |
| AA | 172 (46) | 1.0 (reference) | | 172 (46) | 1.0 (reference) | |
| AG | 379 (89) | 0.8 (0.6–1.2) | .474 | 379 (89) | 0.7 (0.5–0.9) | .810 |
| GG | 162 (46) | 1.0 (0.7–1.6) | .680 | 162 (46) | 0.7 (0.5–1.0) | .120 |
| AG/GG | 541 (135) | 0.8 (0.6–1.1) | .309 | 541 (135) | 0.7 (0.5–0.9) | .022 |
| AA/AG | 551 (135) | 1.0 (reference) | | 551 (135) | 1.0 (reference) | |
| GG | 162 (46) | 1.0 (0.7–1.4) | .718 | 162 (46) | 0.9 (0.6–1.2) | .712 |
| **rs8051542[a] (TOX3)** | | | | | | |
| GG | 225 (53) | 1.0 (reference) | | 225 (53) | 1.0 (reference) | |
| AG | 341 (96) | 1.2 (0.8–1.7) | .197 | 341 (96) | 0.8 (0.6–1.1) | .228 |
| AA | 148 (32) | 0.9 (0.5–1.4) | .709 | 148 (32) | 0.7 (0.5–1.1) | .180 |
| AG/AA | 489 (128) | 1.1 (0.8–1.5) | .407 | 489 (128) | 0.8 (0.6–1.0) | .148 |
| GG/AG | 566 (149) | 1.0 (reference) | | 566 (149) | 1.0 (reference) | |
| AA | 148 (32) | 0.8 (0.5–1.1) | .257 | 148 (32) | 0.8 (0.6–1.2) | .385 |
| **rs999737[a] (RAD51L1)** | | | | | | |
| GG | 425 (109) | 1.0 (reference) | | 425 (109) | 1.0 (reference) | |
| AG | 266 (69) | 0.9 (0.7–1.2) | .761 | 266 (69) | 1.1 (0.8–1.4) | .383 |
| AA | 31 (7) | 0.8 (0.4–1.9) | .771 | 31 (7) | 1.3 (0.7–2.4) | .341 |
| AG/AA | 297 (76) | 0.9 (0.7–1.2) | .723 | 297 (76) | 1.1 (0.8–1.4) | .302 |
| GG/AG | 691 (178) | 1.0 (reference) | | 691 (178) | 1.0 (reference) | |
| AA | 31 (7) | 0.9 (0.4–1.9) | .796 | 31 (7) | 1.2 (0.7–2.3) | .418 |

(continued)

and were also more likely to have stage I versus stage II disease. The final study analysis consisted of 739 patients with breast cancer. The study was approved by the MD Anderson Cancer Center Institutional Review Board.

**SNP Selection and Genotyping**

We reviewed the literature and selected eight SNPs that were identified as novel breast cancer susceptibility loci by GWAS based on the available literature at the time of genotype anal-

CME

**Table 2.** (continued)

| Single-nucleotide polymorphism | Disease-free survival | | | Overall survival | | |
|---|---|---|---|---|---|---|
| | No. of patients (events) | HR (95% CI) | *p* value | No. of patients (events) | HR (95% CI) | *p* value |
| **rs6504950[a] (17q23)** | | | | | | |
| GG | 400 (100) | 1.0 (reference) | | 400 (100) | 1.0 (reference) | |
| AG | 265 (67) | 1.0 (0.7–1.3) | .916 | 265 (67) | 0.9 (0.6–1.1) | .476 |
| AA | 57 (16) | 1.2 (0.7–2.1) | .419 | 57 (16) | 1.7 (1.1–2.6) | .010 |
| AG/AA | 322 (83) | 1.0 (0.7–1.4) | .724 | 322 (83) | 1.0 (0.7–1.3) | .815 |
| GG/AG | 665 (167) | 1.0 (reference) | | 665 (167) | 1.0 (reference) | |
| AA | 57 (16) | 1.2 (0.7–2.0) | .421 | 57 (16) | 1.7 (1.1–2.6) | .004 |
| **rs4973768[b] (3p24)** | | | | | | |
| AA | 186 (42) | 1.0 (reference) | | 186 (42) | 1.0 (reference) | |
| AG | 356 (92) | 1.1 (0.7–1.6) | .514 | 356 (92) | 0.9 (0.7–1.3) | .909 |
| GG | 178 (50) | 1.2 (0.8–1.8) | .360 | 178 (50) | 0.8 (0.5–1.2) | .456 |
| AG/GG | 534 (142) | 0.8 (0.6–1.2) | .507 | 534 (142) | 1.1 (0.8–1.5) | .409 |
| AA/AG | 542 (134) | 1.0 (reference) | | 542 (134) | 1.0 (reference) | |
| GG | 178 (50) | 0.8 (0.6–1.2) | .407 | 178 (50) | 1.0 (0.7–1.4) | .704 |

[a]For each single-nucleotide polymorphism except rs1219648, rs12443621, rs4973768: AA, minor allele homozygote; AG, heterozygote; GG, major allele homozygote.

[b]For single-nucleotide polymorphism rs1219648, rs12443621, rs4973768: GG, minor allele homozygote; AG, heterozygote; AA, major allele homozygote.

Abbreviations: CI, confidence interval; HR, hazard ratio.

**Table 3.** Multivariate Cox proportional hazards model for overall survival in study cohort

| Variables | HR | 95% CI | *p* value |
|---|---|---|---|
| rs6504950 AA vs. AG+GG (reference) | 1.77 | 1.15–2.73 | .008 |
| rs12443621 AG+GG vs. AA (reference) | 0.72 | 0.53–0.78 | .035 |
| Pathological stage: Stage II vs. stage I | 2.08 | 1.45–2.96 | <.001 |
| Age at diagnosis: ≥50 yrs vs. <50 yrs | 1.33 | 0.95–1.85 | .09 |
| Treatment type | | | |
|   None vs. chemotherapy only | 0.50 | 0.34–0.75 | <.001 |
|   None vs. endocrine therapy only | 0.62 | 0.43–0.90 | .01 |
|   None vs. chemotherapy and endocrine therapy | 0.40 | 0.24–0.68 | <.001 |

Abbreviations: CI, confidence interval; HR hazard ratio.

ysis [6, 10–11, 23, 30, 32]. All selected SNPs were within genes or linkage disequilibrium (LD) blocks containing genes (10q26-rs2981582, rs1219648 [*FGFR2*]; 16q12-rs3803662, rs12443621, rs8051542 [*TOX3*]; 14q24-rs999737 [*RAD51L1*]; 17q23-rs6504950 [*COX11*]; 3p24-rs4973768 [*SLC4A7*]) based on the data from the International HapMap Project (www.hapmap.org, version 23). All SNPs met the following criteria: minor allele frequency (MAF) of ≥0.05, Illumina design score > 0.4, and $r^2 ≥ 0.8$ for binning.

Genotyping of the eight selected SNPs was performed on the Illumina GoldenGate platform (Illumina Inc., San Diego, CA, http://www.illumina.com) as part of a larger array of 1,536 SNPs of other candidate genes. Briefly, genomic DNA was extracted from peripheral blood using the QIAamp DNA Blood Mini kit (Qiagen, Valencia, CA, http://www.qiagen.com) according to the manufacturer's protocol [33]. Genomic DNA was extracted from normal formalin-fixed, paraffin-embedded (FFPE) tissue using the appropriate protocol in the PicoPure DNA extraction kit (Applied Biosystems, Foster City, CA, http://www.appliedbiosystems.com) [34]. The call rates

for the fixed FFPE samples and non-FFPE samples were 92% and 99%, respectively. Blinded duplicate samples (5%) were included in the array platform, and the duplication concordance was 100%. All genotyping information was analyzed and exported using the GenomeStudio software (Illumina).

**Statistical Analysis and Outcome Measures**

Disease-free survival (DFS) was defined as the elapsed time between the date of initial treatment until the first date of documented disease recurrence or death by breast cancer. Overall survival (OS) was calculated from the date of initial treatment until the date of death from any cause. Clinical variables included age, stage, ER/PR status, treatment (chemotherapy and/or endocrine therapy), and FH of breast cancer. For each SNP and clinical variable, including age, family history of breast cancer, stage, ER/PR status, nuclear grade, treatment type, year of diagnosis, and surgery type (mastectomy and lumpectomy), univariable Cox proportional hazard regression analysis was used to identify the association with DFS
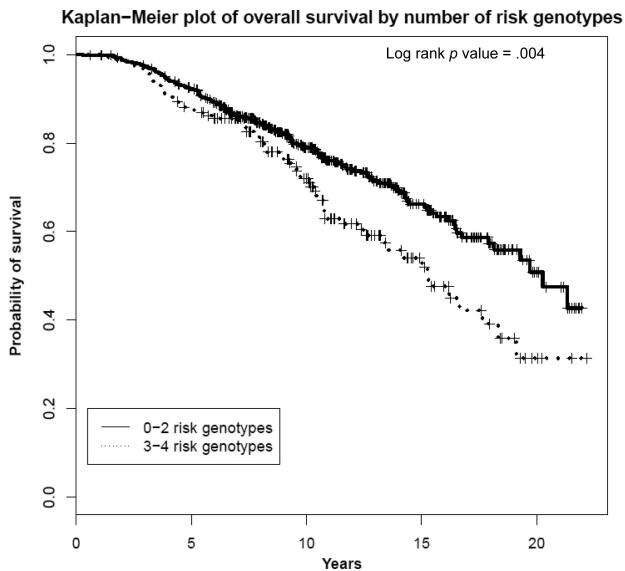
CME

**Figure 1.** Kaplan-Meier plot of overall survival by number of at-risk genotypes.

and OS. Because this was an exploratory study, no adjustments for multiple testing were performed [35].

The genetic and clinical variables that were possibly related to survival outcomes by univariable analysis ($p \leq .1$) were then included into a multivariable Cox proportional hazard regression model. In the process of fitting the multivariable Cox regression model, a backward selection was used with a $p$ value $\leq.05$ for the likelihood ratio test. Age at diagnosis was retained in the multivariable model, given its importance in determining survival. Interactions between the SNPs and hormone receptor (ER+ or PR+ versus ER− and PR−) and treatment (none, chemotherapy alone, endocrine therapy alone, chemotherapy, and endocrine therapy) statuses were assessed using the likelihood ratio test ($p < .05$), comparing the model including an interaction term with the reduced model without the interaction term. Furthermore, the cumulative effects of SNPs were assessed by computing the number of at-risk genotypes based on the results from the univariable analyses. The median survival time was estimated using Kaplan-Meier methods, and the log-rank analysis was used to compare survival between the at-risk genotype groups. Hazard ratios (HRs) and their 95% confidence intervals (CIs) were estimated for all associations between SNPs, clinical variables, and survival outcomes. Finally, we produced a recursive partitioning tree to evaluate higher order interactions between clinical variables and SNPs. Statistical analyses were carried out using SAS version 9.1 (SAS Institute, Cary, NC, http://www.sas.com) and R (http://www.r-project.org).

### RESULTS

Patient demographic and clinical characteristics are summarized in Table 1. At a median follow-up of 121 months (range: 188−231 months) for survivors, 237 deaths (32%) and 186 breast cancer events (25%) were identified among the 739 patients. The median OS was 217 months. The majority of patients were postmenopausal (57%) and had stage II disease (72%). Approximately 20% of patients did not receive any systemic treatment for their disease and 25% received endocrine therapy alone. There was

minimal missing data for the cohort, ranging from 5.1% for ER/PR status to 0.4% for systemic treatment.

### Univariable Associations Between SNPs and Breast Cancer Outcomes

In this cohort of women with a history of early-stage breast cancer, the genotype distributions of eight SNPs were in agreement with Hardy-Weinberg equilibrium. In univariable analysis, none of the SNPs were associated with DFS (Table 2). Significant associations were observed for OS for four SNPs (rs2981582, rs1219648, rs12443621, rs6504950). Homozygous carriers for the minor alleles for rs2981582 and rs1219648 had a decreased risk for death (HR AA vs. GG/AG: 0.6, 95% CI: 0.47−0.98, $p = .040$; HR GG vs. AA/AG: 0.70, 95% CI: 0.49−1.01). In addition, a decreased risk of death was observed among patients homozygous for the minor alleles for rs121443621 in the dominant model (HR GG vs. AA/AG: 0.70, 95% CI: 0.53−0.95, $p = .022$). In contrast, OS was worse for homozygous carriers of the minor A allele for rs6504950 both in the codominant model (HR AA vs. GG: 1.70, 95% CI: 1.13−2.61, $p = .01$) and the dominant model (HR AA vs. GG/AG: 1.70, 95% CI: 1.19−2.68, $p = .004$). The other investigated SNPs were not associated with OS.

### Multivariable Associations Between SNPs and Breast Cancer Outcomes

In multivariable models adjusted for age, ER/PR status, stage, and treatment type, we found two SNPs, rs12443621 and rs6504950, which were independent prognostic markers for OS. Compared with the GG/AG reference group, homozygous carriers of minor A allele for rs6504950 had a higher risk for death (HR: 1.77, 95% CI: 1.15−2.73, $p = .008$). Compared with the AA reference group, homozygous carriers for the AG/GG for rs12443621 had a decreased risk for death (HR: 0.72, 95% CI: 0.53−0.78, $p = .035$). In addition, stage II disease was associated with a worse OS (HR: 2.08, 95% CI: 1.45−2.96, $p < .001$) compared to patients with stage I disease, and patients who received chemotherapy and/or endocrine therapy had a better OS (Table 3) than patients who did not receive systemic treatment. We performed a stratified multivariable analysis to explore the relationship between prognostic SNPs and ER/PR tumor status. The increased risk of death associated with the SNP rs6504950 was seen predominantly among patients with ER- or PR-positive tumors ($p < .001$) compared to patients with ER-negative and PR-negative tumors ($p = .30$; data not shown). However, the test for interaction between SNPs rs12443621, rs6504950, and ER/PR status on overall survival was not statistically significant. There was no evidence of interaction by treatment status (data not shown).

We investigated the cumulative effect of carrying at-risk genotypes for the four SNPs significantly associated with OS in the univariable analysis. We categorized patients into two groups ($\leq$2 vs. 3−4 at-risk genotypes) and used patients carrying $\leq$2 at-risk genotypes as the reference group. Kaplan-Meier analysis revealed statistical significant differences in OS according to the number of at-risk genotypes (log-rank $p = .004$; Fig. 1). In multivariable analysis, after adjusting for significant clinical variables, a higher risk for death was found in the subcohort of patients harboring 3−4 at-risk genotypes compared to patients carrying $\leq$2 at-risk genotypes (HR: 1.60, 95% CI: 1.23−2.24, $p = .0008$; Table 4).

CME

**Table 4.** Multivariate Cox proportional hazards model for overall survival by number of at-risk genotypes in study cohort

| Variables | HR | 95% CI | *p* value |
|---|---|---|---|
| Single-nucleotide polymorphisms: 3–4 vs. ≤2 | 1.66 | 1.23–2.24 | .0008 |
| Age at diagnosis: ≥50 yrs vs. <50 yrs | 1.35 | 0.97–1.88 | .073 |
| Pathological stage: Stage II vs. stage I | 2.14 | 1.49–3.05 | <.001 |
| Treatment type | | | |
|    None vs. chemotherapy only | 0.50 | 0.34–0.75 | .007 |
|    None vs. endocrine therapy only | 0.62 | 0.43–0.91 | .01 |
|    None vs. chemotherapy and endocrine therapy | 0.40 | 0.24–0.68 | <.0005 |

Abbreviations: CI, confidence interval; HR hazard ratio.

## DISCUSSION

We evaluated the possible relation between eight GWAS-identified SNPs and breast cancer outcomes in a large cohort of women with early-stage breast cancer. Two GWAS risk SNPs, rs12443621(16q12) and rs6504950 (17q23), were associated with OS, independent of clinical factors. Furthermore, we observed that the risk SNPs (rs2981582, rs1219648, rs12443621, rs6504950) in univariable analysis appeared to act in a codominant (additive) fashion, with mortality risk increasing with the number of at-risk genotypes. To our knowledge, our study is the first to identify the association of these polymorphisms and OS and breast cancer DFS. Replication of our findings in larger study sets and validation is needed before these SNPs can be considered for prognostic risk stratification of patients with breast cancer.

Recent GWAS have identified several SNPs in novel independent loci as markers of breast cancer susceptibility in populations of diverse ethnicity [5, 36]. The rs12443621 prognostic SNP has been identified in an LD block containing the 59 end of the trinucleotide repeat containing 9 (*TOX3*) gene, located at chromosome 16q12 [5]. The function of the *TOX3* gene is unclear; however, it contains a putative high-mobility group box motif, suggesting that it may act as a transcription factor [8]. Of the three polymorphisms identified in the *TOX3* gene (rs3803662, rs12443621, and rs8051542) [37], SNP rs3803662 exhibited a stronger association with breast cancer risk in white individuals [5, 32, 38], whereas Chinese [36, 39] studies have showed no association between the rs12443621 or rs3803662 alleles and breast cancer risk [5, 32, 36, 38–39]. Therefore, the clinical relevance of SNP rs12443621 remains unknown. It is not clear why rs12443621 is associated with a decreased OS in this study population. It has been pointed out that genes do not generally act in a simple additive manner but through complex networks involving gene-gene and gene-environment interactions [40]; therefore, it is possible that a minor homozygous genotype or allele associated with increased breast cancer risk may have gene-gene or gene-environment interactions that yield an opposite effect on survival. The other prognostic SNP rs6504950 lies in a 300-kb LD block on 17q23 [11]. Cytochrome C assembly protein 11 (*COX11*), approximately 10 kb upstream of rs6504950 is the potential causative gene as suggested by the higher levels of *COX11* expression in lymphocytes in the HapMap samples (*p* = .000014) [41]. Of interest, both rs12443621A/G [8, 11, 36] and rs6504950 [42] showed a stronger association with risk of ER-positive than ER-negative breast tumors in

those studies. We found that the SNP rs6504950 was associated with a twofold increased risk of death predominately among patients with ER/PR-positive tumors, but the test for interaction was not statistically significant.

There have been an increasing number of epidemiological studies examining the effects of SNPs on cancer outcomes using a candidate gene approach [17–21, 43–44]. These studies have provided insights into the molecular mechanisms through which germline SNPs may influence breast cancer survival [19–20, 22, 24–26, 41, 43, 45–48]. To date, the functional role of the germline polymorphisms in the 16q12 (*TOX3*) and 17q23 loci is unknown. Smid et al. [49] identified a 69 panel of genes found in tumor DNA relevant to bone metastasis in breast cancer; among those, increased expression of *TOX3* was found to be a predictive factor. In prior studies, increased copy numbers of 17q23 chromosomal region have been associated with tumor progression and with poor prognosis in breast cancer [50–52]. Inaki et al. [53] found that 41% (28/69) of breast cancer tumors in their series showed evidence of gene amplification at 17q23 locus. Expression of recurrent fusion gene transcript (RPS6KB1-VMP1) created by a tandem duplication in the 17q23 locus showed a trend toward correlation (*p* = .06) with poor DFS. In our study, germline SNPs in the 16q12 and 17q23 loci did not influence breast cancer DFS but played a significant role in OS among a favorable prognostic group of older patients with breast cancer who did not receive chemotherapy.

Several limitations of this study should be acknowledged. First, the number of SNPs tested was limited to eight, which is not a comprehensive evaluation of the association between GWAS-identified risk SNPs and breast cancer prognosis. The study was exploratory and we did not adjust for multiple testing; therefore, some of our findings may be due to chance. In addition, the sample size was too small to definitively evaluate for breast cancer outcomes according to ER and PR status or systemic treatment received. In addition, the sample size was too small to definitively evaluate for breast cancer outcomes according to breast tumor subtypes. The Breast Cancer Association Consortium has taken the first step in investigating GWAS-identified susceptibility loci in relation to risk of developing specific breast tumor subtypes; however, information on outcomes is not yet available.

In addition, the ESBCR cohort represents a select group of white women with early-stage disease treated exclusively at a single institution, limiting generalizability of the findings to all patients with breast cancer. Azzatto et al. performed a GWAS

The Oncologist

of prognosis in patients with breast cancer enrolled in the Nurses' Health Cohort and found no genetic variants associated with breast cancer-specific death. However, that study had few breast cancer deaths ($n = 92$) and may have been underpowered in the discovery analysis to detect SNPs with modest effects [54]. Similarly, in our study we found no association with breast cancer DFS, but there was an association with OS. The reasons for this specific association are unclear but suggest that any additional biological effect of the genetic variants may be explained by influences on host comorbid conditions, lifestyle factors, or environmental factors that contribute to mortality.

The GWAS have identified several common SNPs that convey breast cancer risk, but it is unclear whether these genetic variants continue to influence the biology of the disease after diagnosis. In our study, we found a significant association between two GWAS-identified SNPs 16q12 (*TOX3*) and 17q23 loci and OS in a cohort of patients with early-stage breast cancer; however, they were not associated with breast cancer DFS. The reasons for the specific associations are unclear but suggest that any additional biological effect of the genetic variants may be explained by influences on host comorbid conditions, lifestyle factors, or environmental factors that contribute to mortality. These findings provide support for the hypothesis that SNPs in these loci may be biologically important and may be promising targets for diagnostic, prognostic, and therapeutic approaches. These results need confirmation in independent cohorts. Future studies that evaluate the prognostic relationship of these SNPs by breast cancer subtype and type of treatment received may provide additional information that could be used to individualize the care of patients with breast cancer.

## AUTHOR CONTRIBUTIONS

**Conception/Design:** Patricia A. Thompson, Abenaa M. Brewster
**Provision of study material or patients:** Patricia A. Thompson, Melissa L. Bondy, Abenaa M. Brewster
**Collection and/or assembly of data:** Patricia A. Thompson, Aysegul A. Sahin, Melissa L. Bondy, Abenaa M. Brewster
**Data analysis and interpretation:** Soley Bayraktar, Suk-Young Yoo, Kim-anh Do, Aysegul A. Sahin, Banu K. Arun, Melissa L. Bondy, Abenaa M. Brewster
**Manuscript writing:** Soley Bayraktar, Suk-Young Yoo, Kim-anh Do, Aysegul A. Sahin, Banu K. Arun, Melissa L. Bondy, Abenaa M. Brewster
**Final approval of manuscript:** Soley Bayraktar, Patricia A. Thompson, Suk-Young Yoo, Kim-anh Do, Aysegul A. Sahin, Banu K. Arun, Melissa L. Bondy, Abenaa M. Brewster

## REFERENCES

**1.** Jemal A, Siegel R, Ward E et al. Cancer statistics, 2009. CA Cancer J Clin 2009;59:225–249.

**2.** National Cancer Institute. SEER Cancer Statistics Review. Available at http://seer.cancer.gov/csr/1975_2009_pops09/index.html. Accessed February 27, 2013.

**3.** Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: An overview of the randomised trials. Lancet 2005;365:1687–1717.

**4.** Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet 2001;27:234–236.

**5.** Easton DF, Pooley KA, Dunning AM et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007;447:1087–1093.

**6.** Thomas G, Jacobs KB, Kraft P et al. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). Nat Genet 2009;41:579–584.

**7.** Han J, Kraft P, Nan H et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. PLoS Genet 2008;4:e1000074.

**8.** Stacey SN, Manolescu A, Sulem P et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 2007;39:865–869.

**9.** Gold B, Kirchhoff T, Stefanov S et al. Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. Proc Natl Acad Sci U S A 2008;105:4340–4345.

**10.** Zheng W, Long J, Gao YT et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 2009;41:324–328.

**11.** Ahmed S, Thomas G, Ghoussaini M et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 2009;41:585–590.

**12.** Stacey SN, Manolescu A, Sulem P et al. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 2008;40:703–706.

**13.** Cox A, Dunning AM, Garcia-Closas M et al. A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet 2007;39:352–358.

**14.** Dunning AM, Healey CS, Baynes C et al. Association of ESR1 gene tagging SNPs with breast cancer risk. Hum Mol Genet 2009;18:1131–1139.

**15.** Pharoah PD, Antoniou A, Bobrow M et al. Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 2002;31:33–36.

**16.** Peng S, Lu B, Ruan W et al. Genetic polymorphisms and breast cancer risk: Evidence from meta-analyses, pooled analyses, and genome-wide association studies. Breast Cancer Res Treat 2011;127:309–324.

**17.** Azzato EM, Lee AJ, Teschendorff A et al. Common germ-line polymorphism of C1QA and breast cancer survival. Br J Cancer 102:1294–1299.

**18.** Bachmann HS, Otterbach F, Callies R et al. The AA genotype of the regulatory BCL2 promoter polymorphism (938C>A) is associated with a favorable outcome in lymph node negative invasive breast cancer patients. Clin Cancer Res 2007;13:5790–5797.

**19.** Kasimir-Bauer S, Heubner M, Otterbach F et al. Prognostic relevance of the AQP5−1364C>A polymorphism in primary breast cancer. Mol Med Report 2009;2:645–650.

**20.** Brendle A, Brandt A, Johansson R et al. Single nucleotide polymorphisms in chromosomal instability genes and risk and clinical outcome of breast cancer: A Swedish prospective case-control study. Eur J Cancer 2009;45:435–442.

**21.** Gaudet MM, Hunter K, Pharoah P et al. Genetic variation in SIPA1 in relation to breast cancer risk and survival after breast cancer diagnosis. Int J Cancer 2009;124:1716–1720.

**22.** Hughes S, Agbaje O, Bowen RL et al. Matrix metalloproteinase single-nucleotide polymorphisms and haplotypes predict breast cancer progression. Clin Cancer Res 2007;13:6673–6680.

**23.** Long JR, Kataoka N, Shu XO et al. Genetic polymorphisms of the CYP19A1 gene and breast cancer survival. Cancer Epidemiol Biomarkers Prev 2006;15:2115–2122.

**24.** Otterbach F, Callies R, Frey UH et al. The T393C polymorphism in the gene GNAS1 of G protein is associated with survival of patients with invasive breast carcinoma. Breast Cancer Res Treat 2007;105:311–317.

**25.** Papadopoulou E, Tripsianis G, Anagnostopoulos K et al. Significance of serum tumor necrosis factor-alpha and its combination with HER-2 codon 655 polymorphism in the diagnosis and prognosis of breast cancer. Int J Biol Markers 2010;25:126–135.

**26.** Qu S, Long J, Cai Q et al. Genetic polymorphisms of metastasis suppressor gene NME1 and breast cancer survival. Clin Cancer Res 2008;14:4787–4793.

**27.** Udler MS, Azzato EM, Healey CS et al. Common germline polymorphisms in COMT, CYP19A1, ESR1,

CME

PGR, SULT1E1 and STS and survival after a diagnosis of breast cancer. Int J Cancer 2009;125:2687–2696.

28. Varadi V, Bevier M, Grzybowska E et al. Genetic variation in genes encoding for polymerase zeta subunits associates with breast cancer risk, tumour characteristics and survival. Breast Cancer Res Treat 2011;129:235–245.

29. Shu XO, Long J, Lu W et al. Novel genetic markers of breast cancer survival identified by a genome-wide association study. Cancer Res 2012;72:1182–1189.

30. Rebbeck TR, Lynch HT, Neuhausen SL et al. Prophylactic oophorectomy in carriers of BRCA1 or BRCA2 mutations. N Engl J Med 2002;346:1616–1622.

31. Brewster AM, Do KA, Thompson PA et al. Relationship between epidemiologic risk factors and breast cancer recurrence. J Clin Oncol 2007;25:4438–4444.

32. Chen MB WX, Shen W et al. . Association between polymorphisms of trinucleotide repeat containing 9 gene and breast cancer risk: Evidence from 62,005 subjects. Breast Cancer Res Treat 2011;126:177–183.

33. QIAamp DNA Mini and Blood Mini Handbook. Available at http://dna.uga.edu/docs/QIAamp_DNA_Mini_and_Blood_Mini_Handbook%20(Qiagen).pdf. Accessed February 27, 2013.

34. Applied Biosystems Arcturus PicoPure DNA Extraction Kit. Available at: http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_086335.pdf. Accessed April 25, 2013.

35. Bender R, Lange S. Adjusting for multiple testing: When and how? J Clin Epidemiol 2001;54:343–349.

36. Liang J, Chen P, Hu Z et al. Genetic variants in trinucleotide repeat-containing 9 (TNRC9) are associated with risk of estrogen receptor positive breast cancer in a Chinese population. Breast Cancer Res Treat 2010;124:237–241.

37. Huijts PE, Vreeswijk MP, Kroeze-Jansema KH et al. Clinical correlates of low-risk variants in FGFR2, TNRC9, MAP3K1, LSP1 and 8q24 in a Dutch cohort of incident breast cancer cases. Breast Cancer Res 2007;9:R78.

38. Reeves GK, Travis RC, Green J et al. Incidence of breast cancer and its subtypes in relation to individual and multiple low-penetrance genetic susceptibility loci. JAMA 2010;304:426–434.

39. Li L, Zhou X, Huang Z et al. TNRC9/LOC643714 polymorphisms are not associated with breast cancer risk in Chinese women. Eur J Cancer Prev 2009;18:285–290.

40. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. Lancet 2003;361:865–872.

41. Stranger BE, Nica AC, Forrest MS et al. Population genomics of human gene expression. Nat Genet 2007;39:1217–1224.

42. Broeks A, Schmidt MK, Sherman ME et al. Low penetrance breast cancer susceptibility loci are associated with specific breast tumor subtypes: Findings from the Breast Cancer Association Consortium. Hum Mol Genet 2011;20:3289–3303.

43. Ding S, Yu JC, Chen ST et al. Diverse associations between *ESR1* polymorphism and breast cancer development and progression. Clin Cancer Res 2010;16:3473–3484.

44. Hsieh S, Look MP, Sieuwerts AM et al. Distinct inherited metastasis susceptibility exists for different breast cancer subtypes: A prognosis study. Breast Cancer Research 2009;11:R75.

45. Lei H, Hemminki K, Johansson R et al. PAI-1–675 4G/5G polymorphism as a prognostic biomarker in breast cancer. Breast Cancer Res Treat 2008;109:165–175.

46. Chrisanthar R, Knappskog S, Lokkevik E et al. Predictive and prognostic impact of TP53 mutations and MDM2 promoter genotype in primary breast cancer patients treated with epirubicin or paclitaxel. PLoS One 6:e19249.

47. Bond GL, Hu W, Bond EE et al. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell 2004;119:591–602.

48. Colomer R, Monzo M, Tusquets I et al. A single-nucleotide polymorphism in the aromatase gene is associated with the efficacy of the aromatase inhibitor letrozole in advanced breast carcinoma. Clin Cancer Res 2008;14:811–816.

49. Smid M, Wang Y, Klijn JG et al. Genes associated with breast cancer metastatic to bone. J Clin Oncol 2006;24:2261–2267.

50. Andersen CL, Monni O, Wagner U et al. High-throughput copy number analysis of 17q23 in 3520 tissue specimens by fluorescence in situ hybridization to tissue microarrays. Am J Pathol 2002;161:73–79.

51. Barlund M, Forozan F, Kononen J et al. Detecting activation of ribosomal protein S6 kinase by complementary DNA and tissue microarray analysis. J Natl Cancer Inst 2000;92:1252–1259.

52. Isola JJ, Kallioniemi OP, Chu LW et al. Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer. Am J Pathol 1995;147:905–911.

53. Inaki K, Hillmer AM, Ukil L et al. Transcriptional consequences of genomic structural aberrations in breast cancer. Genome Res 2011;21:676–687.

54. Azzato EM, Pharoah PD, Harrington P et al. A genome-wide association study of prognosis in breast cancer. Cancer Epidemiol Biomarkers Prev 2010;19:1140–1143.

CME  This article is available for continuing medical education credit at CME.TheOncologist.com.