

Null Results in Brief

A Genome-Wide Association Study of Prognosis in Breast Cancer

Elizabeth M. Azzato^{1,4}, Paul D.P. Pharoah¹, Patricia Harrington¹, Douglas F. Easton², David Greenberg³, Neil E. Caporaso⁴, Stephen J. Chanock⁴, Robert N. Hoover⁴, Gilles Thomas⁴, David J. Hunter⁵, and Peter Kraft⁵

Abstract

Background: Traditional clinicopathologic features of breast cancer do not account for all the variation in survival. Germline genetic variation may provide additional prognostic information.

Materials and Methods: We conducted a genome-wide association study of survival after a diagnosis of breast cancer by obtaining follow-up data and genotyping information on 528,252 single-nucleotide polymorphisms for 1,145 postmenopausal women with invasive breast cancer (7,711 person-years at risk) from the Nurses' Health Study scanned in the Cancer Genetic Markers of Susceptibility initiative. We genotyped the 10 most statistically significant loci (most significant single-nucleotide polymorphism located in *ARHGAP10*; $P = 2.28 \times 10^{-7}$) in 4,335 women diagnosed with invasive breast cancer (38,148 years at risk) in the SEARCH (Studies of Epidemiology and Risk factors in Cancer Heredity) breast cancer study.

Results: None of the loci replicated in the SEARCH study (all $P > 0.10$). Assuming a minimum of 10 associated loci, the power to detect at least one with a minor allele frequency of 0.2 conferring a relative hazard of 2.0 at genome-wide significance ($P = 5 \times 10^{-8}$) was 99%.

Conclusion: We did not identify any common germline variants associated with breast cancer survival overall.

Impact: Our data suggest that it is unlikely that there are common germline variants with large effect sizes for breast cancer survival overall (hazard ratio >2). Instead, it is plausible that common variants associated with survival could be specific to tumor subtypes or treatment approaches. New studies, sufficiently powered, are needed to discover new regions associated with survival overall or by subtype or treatment subgroups. *Cancer Epidemiol Biomarkers Prev*; 19(4); 1140–3. ©2010 AACR.

Introduction

Traditional clinical and pathologic features related to breast cancer prognosis are not adequate predictors of survival (1), and it is possible that variation in germline DNA may provide additional information. Previous studies have focused on candidate genes, relying on our incomplete knowledge of tumor and host biology (2–8). In this study, we aimed to identify common germline variants associated with breast cancer–

specific survival after a diagnosis of breast cancer using a genome-wide approach. We obtained follow-up information on 1,145 women with invasive postmenopausal breast cancer from the Nurses' Health Study (NHS) cohort, genotyped using the Illumina HumanHap500 array, as part of the Cancer Genetic Markers of Susceptibility (CGEMS) initiative. We genotyped the most statistically significant associations in the SEARCH (Studies of Epidemiology and Risk factors in Cancer Heredity) breast cancer study.

Materials and Methods

Sample

The 1,145 women of the NHS/CGEMS sample used in this analysis and the genome-wide association study (GWAS) genotyping methods have been previously described (9, 10). Briefly, the NHS is a longitudinal study of 121,700 women enrolled in 1976. The CGEMS case-control study is derived from 32,826 participants who provided a blood sample in 1989–1990 and were followed for incident breast cancer until May 2004. Follow-up was conducted by personal mailings and searches of the National Death Index. The 1,145 women were genotyped

Authors' Affiliations: Departments of ¹Oncology and ²Public Health and Primary Care, Strangeways Research Laboratory, University of Cambridge; ³Eastern Cancer Registration and Information Centre, Unit C-Magog Court, Cambridge, United Kingdom; ⁴Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland; and ⁵Program in Molecular and Genetic Epidemiology, Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts

Corresponding Author: Elizabeth M. Azzato, 6120 Executive Plaza, Room 7106, Rockville, MD 20852. Phone: 301-435-7614; Fax: 301-402-4489. E-mail: azzatoe2@mail.nih.gov

doi: 10.1158/1055-9965.EPI-10-0085

©2010 American Association for Cancer Research.

using the Illumina HumanHap500 array as the first stage of a three-stage GWAS of breast cancer susceptibility. We removed single-nucleotide polymorphisms (SNP) with a minor allele frequency of <1% for a total of 528,252 SNPs.

The SEARCH breast cancer study ascertainment and follow-up is described elsewhere (4). Briefly, SEARCH is an ongoing population-based study of women diagnosed with breast cancer in the region of England included in the Eastern Cancer Registration and Information Centre. Study eligibility includes those diagnosed with invasive breast cancer at age <70 y at the start of the study in 1996, or those diagnosed at age 55 or younger since 1991 and alive at the start of the study (prevalent cases). Follow-up was obtained by death certificate flagging through the Office of National Statistics and the National Health Service Strategic Tracing Service. Genotyping sets 1 and 2 (4,335 women) were included in this analysis. Genotyping was done using optimized TaqMan assays (Applied Biosystems; ref. 11).

All participants in the studies provided informed consent. The NHS/CGEMS study protocol was approved by the Brigham and Women's Hospital Institutional Review Board. The SEARCH study was approved by the Eastern Multicentre Research Ethics Committee.

Statistical methods

We fitted Cox proportional hazards models to assess the association of genotype with breast cancer-specific mortality, adjusting for age category at diagnosis (44-59, 60-69, and 70-83 y). Follow-up for NHS ended at date of death from breast cancer or June 30, 2004. Statistical significance was based on a 1-degree-of-freedom trend test.

The 10 most statistically significant loci were genotyped in SEARCH and assessed for association with breast cancer-specific mortality, adjusting for age category at diagnosis (<44, 44-59, and 60-69 y), using Cox proportional hazards models allowing for left truncated (prevalent case) data (12). Follow-up ended for SEARCH at date of death from breast cancer or November 30, 2006; follow-up was censored at 10 y. Because the SEARCH population includes some premenopausal cases, analyses limited to individuals with ages at diagnosis of ≥ 44 and ≥ 55 y were also performed.

Results

The 1,145 women participating in the NHS/CGEMS study provided 7,711 person-years at risk (93 breast cancer deaths; Table 1). We tested the 528,252 SNPs for

Table 1. Study population characteristics

	NHS/CGEMS	SEARCH
Total no. of subjects	1,145	4,335
Total time at risk, y	7,711	28,148*
Median time at risk, y	6.00 (0.001-15.00) [†]	6.51 (0.036-9.81) [†]
No. of breast cancer deaths	93	587
Annual mortality rate	0.012 (0.010-0.014) [‡]	0.021 (0.019-0.023) [‡]
5-y survival rate	0.94 (0.944-0.969) [‡]	0.89 (0.883-0.905) [‡]
Median age at diagnosis, y	66 (44-83) [†]	51 (23-69) [†]
Histopathologic grade, <i>n</i> (%)		
Well differentiated	209 (18.25)	858 (19.79)
Moderately differentiated	389 (33.97)	1,652 (38.11)
Poorly differentiated	243 (21.22)	989 (22.81)
Unknown	304 (26.55)	836 (19.28)
Clinical stage, <i>n</i> (%)		
I	725 (63.32)	2,133 (49.20)
II	300 (26.20)	1,924 (44.38)
III	49 (4.28)	142 (3.28)
IV	0 (0.00)	44 (1.01)
Unknown	71 (6.20)	92 (2.12)
Estrogen receptor status, <i>n</i> (%)		
Positive	807 (70.48)	2,440 (56.29)
Negative	181 (15.81)	639 (14.74)
Unknown	157 (13.71)	1,256 (28.97)

*Follow-up censored at 10 y, analysis allowing for left-truncated data.

[†]Range of variable.

[‡]95% confidence interval.

Table 2. Hazard ratios, 95% confidence intervals, and *P* values for NHS/CGEMS and SEARCH studies for the top 10 most significant loci from NHS/CGEMS breast cancer survival GWAS

SNP	Location* Nearby gene(s)	Alleles [†]	NHS/CGEMS			SEARCH		
			MAF	HR [‡] (95% CI)	<i>P</i>	MAF	HR [§] (95% CI)	<i>P</i>
rs13124167	4q31.23 148894643 <i>ARHGAP10</i>	T, C	0.12	2.48 (2.13-2.82)	2.28 × 10 ⁻⁷	0.11	1.00 (0.83-1.19)	0.97
rs4529739	1p32.1 60477371	T, C	0.09	2.31 (1.96-2.65)	2.69 × 10 ⁻⁶	0.11	0.99 (0.83-1.19)	0.95
rs11591508	10p11.22 33324659	C, T	0.06	2.85 (2.41-3.29)	3.27 × 10 ⁻⁶	0.06	0.91 (0.71-1.16)	0.43
rs2571236	18q21.31 53607672	G, A	0.23	1.96 (1.67-2.25)	5.32 × 10 ⁻⁶	0.23	0.99 (0.86-1.13)	0.84
rs3094663	6p21.33 31215066 <i>PSORS1C1, CDSN, PSORS1C2, C6orf18</i>	G, A	0.30	1.94 (1.64-2.24)	1.10 × 10 ⁻⁵	0.28	0.98 (0.81-1.18)	0.82
rs352457	15q22.31 63564258 <i>DPP8</i>	G, A	0.06	2.46 (2.05-2.87)	1.82 × 10 ⁻⁵	0.06	1.06 (0.83-1.35)	0.64
rs936503	18q23 74788302	T, C	0.39	0.49 (0.15-0.82)	2.49 × 10 ⁻⁵	0.37	1.08 (0.95-1.21)	0.23
rs2282079	9p13.2 37026247 <i>PAX5, LOC401504</i>	G, A	0.05	2.03 (1.70-2.37)	2.76 × 10 ⁻⁵	0.04	0.84 (0.61-1.16)	0.28
rs17299684	15q25.2 82495353 <i>ADAMTSL3</i>	A, G	0.15	1.95 (1.64-2.27)	3.25 × 10 ⁻⁵	0.15	0.87 (0.74-1.03)	0.11
rs17296289	10p11.22 33300705 <i>ITGB1</i>	G, A	0.06	2.51 (2.06-2.96)	5.40 × 10 ⁻⁵	0.07	0.93 (0.73-1.18)	0.53

Abbreviations: MAF, minor allele frequency; HR, hazard ratio; 95% CI, 95% confidence interval.

*dbSNP build 130.

[†]Major allele, minor allele.

[‡]Adjusted for age at diagnosis categories (44-59, 60-69, and 70+ y).

[§]Adjusted for age at diagnosis categories (<44, 44-59, and 60-69 y).

an association with breast cancer survival, and no association reached nominal GWAS significance threshold ($P = 5 \times 10^{-8}$). The most significant locus was in the *ARHGAP10* gene ($P = 2.3 \times 10^{-7}$).

The 10 loci were genotyped in up to 4,335 women participating in the SEARCH study, who provided 28,148 years at risk (587 breast cancer deaths). None of the loci replicated in the SEARCH study (Table 2). There were no significant differences in the analyses when stratified by ages at diagnosis of ≥ 44 and ≥ 55 years (data not shown).

Discussion

The NHS/CGEMS study provides the unique opportunity for an agnostic search of the genome for common

genetic variants associated with breast cancer prognosis. To date, this is the first GWAS of breast cancer survival. However, we did not observe any SNP associations with a genome level of statistical significance ($P = 5 \times 10^{-8}$), nor did we replicate any of the 10 most statistically significant loci discovered in the GWAS in the SEARCH study.

Assuming a minimum of 10 associated loci, the power to detect at least one where the risk allele frequency is 0.2 conferring relative hazards of 1.6, 1.8, and 2.0 at genome-wide significance (5×10^{-8}), taking into account the staged study design, was 48%, 89%, and 99%, respectively. Because the power to detect larger magnitude effects and more prevalent alleles is correspondingly greater, it is unlikely that we have missed

common variants with large effect sizes. However, it is possible that variants with more modest effects were missed in the discovery analysis and were not carried forward to the replication phase. In addition, our power in the discovery GWAS is less favorable for rare variants or genes acting via a recessive mechanism.

Breast cancer is a heterogeneous disease and its prognosis varies significantly across tumor subtypes (13-18); additional factors such as patient characteristics (e.g., age, comorbidity, diet, etc.), treatment regimen, compliance, and individual pharmacogenetics also affect survival (19). It is plausible that germline genetic variation could be associated with survival by tumor subtype or treatment approaches. However, the power to investigate subgroups as well as interactions with environmental factors is limited with the current data set and will require larger consortial studies. Furthermore, the determination of common variants associated with survival is challenging in studies designed to discover common variants for etiology because of issues related to study design.

In conclusion, our study suggests that it is unlikely that there are many common germline variants with large effects (hazard ratio >2) on general breast cancer survival.

References

1. Tavassoli FA, Devilee P. Pathology and genetics of tumours of the breast and female genital organs. Lyon: International Agency for Research on Cancer, Oxford University Press; 2003.
2. Udler M, Azzato EM, Healey CS, et al. Common germline polymorphisms in COMT, CYP19A1, ESR1, PGR, SULT1E1 and STS and survival after a diagnosis of breast cancer. *IJC* 2009;125:2687-96.
3. Abraham JE, Harrington P, Driver KE, et al. Common polymorphisms in the prostaglandin pathway genes and their association with breast cancer susceptibility and survival. *Clin Cancer Res* 2009;15:2181-91.
4. Azzato EM, Driver KE, Lesueur F, et al. Effects of common germline genetic variation in cell cycle control genes on breast cancer survival: results from a population-based cohort. *Breast Cancer Res* 2008;10:R47.
5. Fagerholm R, Hofstetter B, Tommiska J, et al. NAD(P)H:quinone oxidoreductase 1 NQO1*2 genotype (P187S) is a strong prognostic and predictive factor in breast cancer. *Nat Genet* 2008;40:844-53.
6. Fasching PA, Loehberg CR, Strissel PL, et al. Single nucleotide polymorphisms of the aromatase gene (CYP19A1), HER2/neu status, and prognosis in breast cancer patients. *Breast Cancer Res Treat* 2008;112:89-98.
7. Udler M, Maia AT, Cebrian A, et al. Common germline genetic variation in antioxidant defense genes and survival after diagnosis of breast cancer. *J Clin Oncol* 2007;25:3015-23.
8. Goode EL, Dunning AM, Kuschel B, et al. Effect of germ-line genetic variation on breast cancer survival in a population-based study. *Cancer Res* 2002;62:3052-7.
9. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007;39:870-4.
10. 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;41:579-84.
11. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447:1087-93.
12. Azzato EM, Greenberg D, Shah M, et al. Prevalent cases in observational studies of cancer survival: do they bias hazard ratio estimates? *Br J Cancer* 2009;100:1806-11.
13. Anderson WF, Jatoi I, Devesa SS. Distinct breast cancer incidence and prognostic patterns in the NCI's SEER program: suggesting a possible link between etiology and outcome. *Breast Cancer Res Treat* 2005;90:127-37.
14. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999-2009.
15. Naderi A, Teschendorff AE, Barbosa-Morais NL, et al. A gene-expression signature to predict survival in breast cancer across independent data sets. *Oncogene* 2007;26:1507-16.
16. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530-6.
17. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869-74.
18. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003;100:8418-23.
19. Marsh S, McLeod HL. Pharmacogenetics and oncology treatment for breast cancer. *Expert Opin Pharmacother* 2007;8:119-27.

Further candidate gene and GWA studies powered for common variants with modest effects on survival, as well as tumor and treatment subgroups, are required.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Craig Luccarini, Don Conroy, and the SEARCH team.

Grant Support

The Nurses' Health Study is supported by U.S. NIH grant P01 CA087969. SEARCH is funded through a program grant from Cancer Research UK. E.M. Azzato is funded through the NIH-University of Cambridge Graduate Partnership Program and the Intramural Research Program of the National Cancer Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 01/19/2010; accepted 01/25/2010; published OnlineFirst 03/23/2010.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association
for Cancer Research

A Genome-Wide Association Study of Prognosis in Breast Cancer

Elizabeth M. Azzato, Paul D.P. Pharoah, Patricia Harrington, et al.

Cancer Epidemiol Biomarkers Prev 2010;19:1140-1143. Published OnlineFirst March 23, 2010.

Updated version Access the most recent version of this article at:
doi:[10.1158/1055-9965.EPI-10-0085](https://doi.org/10.1158/1055-9965.EPI-10-0085)

Cited articles This article cites 18 articles, 5 of which you can access for free at:
<http://cebp.aacrjournals.org/content/19/4/1140.full#ref-list-1>

Citing articles This article has been cited by 8 HighWire-hosted articles. Access the articles at:
<http://cebp.aacrjournals.org/content/19/4/1140.full#related-urls>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link
<http://cebp.aacrjournals.org/content/19/4/1140>.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.