The role of genetic breast cancer susceptibility variants as prognostic factors


1University Breast Center, Department of Gynecology and Obstetrics, University Hospital Erlangen, Comprehensive Cancer Center Erlangen Nuremberg and 2Institute of Human Genetics, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany, 3Division of Hematology and Oncology, Department of Medicine, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA, USA, 4Department of Oncology and 5Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK, 6Department of Oncology, CRUK/CRF Sheffield Cancer Research Centre, Institute for Cancer Studies, 7Department of Neuroscience, Academic Unit of Pathology and 8Department of Oncology, CRUK/CRF Sheffield Cancer Research Centre, Academic Unit of Surgical Oncology, University of Sheffield, Sheffield, UK, 9Department of Obstetrics and Gynecology and 10Department of Oncology, Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland, 11Copenhagen General Population Study, 12Department of Clinical Biochemistry and 13Department of Breast Surgery,

∗To whom correspondence should be addressed. Tel: +1 3105147700; Email: peter.fasching@uk-erlangen.de
†These authors contributed equally to this work.

© The Author 2012. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
Herlev Hospital, Copenhagen University Hospital, Copenhagen, Denmark, 14Department of Gynecology and Obstetrics, University Hospital, Johann-Wolfgang-Goethe University, Frankfurt a.M., Germany, 15Division of Molecular Pathology and 16Department of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Amsterdam, The Netherlands, 17Department of Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland, 18Genetic Epidemiology Laboratory, Department of Pathology and 19Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, Melbourne School of Population Health, The University of Melbourne, Melbourne, Australia, 20Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, UK, 21Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA, 22Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland, 23Department of Medical Oncology, Rotterdam Family Cancer Clinic, Erasmus Medical Center-Daniel den Hoed Cancer Center, Rotterdam, The Netherlands, 24Department of Human Genetics, 25Department of Surgical Oncology and 26Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands, 27UCI Center for Cancer Genetics Research & Prevention School of Medicine, University of California at Irvine, Irvine, CA, USA, 28School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine, 29Biocenter Kuopio, Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland, 30School of Medicine, Institute of Clinical Medicine, Oncology, 31Biocenter Kuopio, Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland, 32Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland 33Cancer Center, Kuopio University Hospital, Kuopio, Finland, 34Ontario Cancer Genetics Network, Cancer Care Ontario, Toronto, Canada, 35Department of Laboratory Medicine and Pathobiology, 36Department of Laboratory Medicine and Pathobiology, Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital and 37Department of Molecular Genetics, Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Canada, 38Department of Laboratory Medicine, Keenan Research Centre of the Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, Canada, 39Department of Medical Epidemiology and Biostatistics, 40Department of Molecular Medicine and Surgery and 41Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden, 42Human Genetics Laboratory, Genome Institute of Singapore, Singapore, Singapore, 43Division of Cancer Epidemiology and 44Molecular Epidemiology Unit, German Cancer Research Center (DKFZ), Heidelberg, Germany, 45Department of Obstetrics and Gynecology, University Hospital Ulm, Ulm University, Ulm, Germany, 46Clinics of Obstetrics and Gynaecology and 47Clinics of Radiation Oncology, Hannover Medical School, Hannover, Germany, 48Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia, 49Department of Medical Oncology, Josephine Nefkens Institute, Erasmus University Medical Center, Rotterdam, The Netherlands, 50Department of Laboratory Medicine and Pathology and 51Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, 52Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA, 53Cancer Prevention Institute of California, Fremont, CA, USA, 54Stanford University School of Medicine, Stanford, CA, USA, 55Stanford Cancer Institute, Stanford, CA, USA, 56Laboratory of Cancer Genetics, Department of Clinical Genetics and Biocenter Oulu, 57Department of Oncology and 58Department of Pathology, Oulu University Hospital, University of Oulu, Oulu, Finland, 59Department of Obstetrics and Gynecology, Heidelberg University Hospital, University of Heidelberg, Heidelberg, Germany, 60Section Gynecological Oncology, National Center for Tumor Diseases, University Hospital Heidelberg, Heidelberg, Germany, 61Queensland Institute of Medical Research, Brisbane, Australia, 62Peter MacCallum Cancer Centre, Melbourne, Australia, 63Vesalius Research Center, VIB and 64Leuven Multidisciplinary Breast Center, University Hospitals, K.U. Leuven, Leuven, Belgium, 65Human Genetics Group and 66Genetic & Molecular Epidemiology Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre [CNIO], Madrid, Spain, 67Servicio de Oncología Médica, Hospital Universitario La Paz, Madrid, Spain, 68Servicio de Cirugía General y Especialidades, Hospital Monte Naranco, Oviedo, Spain, 69Division of Cancer Prevention and Genetics and 70Department of Experimental Oncology, Istituto Europeo di Oncologia (IEO), Milan, Italy, 71Unit of Medical Genetics, Department of Preventive and Predictive Medicine and 72Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Italy, 73Fondazione Istituto FIRC di Oncologia Molecolare, IFOM, Milan, Italy, 74Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA, 75Division of Cancer Medicine, Peter MacCallum Cancer Centre, Melbourne, Australia, and 76Department of Medicine, St Vincent’s Hospital, The University of Melbourne, Victoria, Australia
Recent genome-wide association studies identified 11 single nucleotide polymorphisms (SNPs) associated with breast cancer (BC) risk. We investigated these and 62 other SNPs for their prognostic relevance. Confirmed BC risk SNPs rs17468277 (CASP8), rs1982073 (TGFB1), rs2981582 (FGFR2), rs13281615 (8q24), rs3817198 (LSP1), rs889312 (MAP3K1), rs3803662 (TOX3), rs13387042 (2q35), rs4973768 (SLC4A7), rs6504950 (COX11) and rs10941679 (5p12) were genotyped for 25,853 BC patients with the available follow-up; 62 other SNPs, which have been suggested as BC risk SNPs by a GWAS or as candidate SNPs from individual studies, were genotyped for replication purposes in subsets of these patients. Cox proportional hazard models were used to test the association of these SNPs with overall survival (OS) and BC-specific survival (BCS). For the confirmed loci, we performed an accessory analysis of publicly available gene expression data and the prognosis in a different patient group. One of the 11 SNPs, rs3803662 (TOX3) and none of the 62 candidate/GWAS SNPs were associated with OS and/or BCS at $P < 0.01$. The genotypic-specific survival for rs3803662 suggested a recessive mode of action [hazard ratio (HR) of rare homozygous carriers $= 1.21$; 95% CI: 1.09–1.35, $P = 0.0002$ and HR = 1.29; 95% CI: 1.12–1.47, $P = 0.0003$ for OS and BCS, respectively]. This association was seen similarly in all analyzed tumor subgroups defined by nodal status, tumor size, grade and estrogen receptor. Breast tumor expression of these genes was not associated with prognosis. With the exception of rs3803662 (TOX3), there was no evidence that any of the SNPs associated with BC susceptibility were associated with the BC survival. Survival may be influenced by a distinct set of germline variants from those influencing susceptibility.

INTRODUCTION

Over the past decade, genomic information has begun to be utilized to describe individual differences in prognosis between breast cancer (BC) patients (1). These data are mainly concerned with DNA expression or genetic variations within the tumor. Some of this information has already been translated into clinical practice or is under investigation in clinical trials (2–6). Additionally, there is growing evidence that an inherited component has an impact on prognosis. Daughters with BC, whose mother had died from BC within 10 years of diagnosis, were found to have a significantly worse survival (7). Apart from familial behavioral patterns, a variety of mechanisms might underlie associations between the inherited genetic profile of the host and prognosis. The genetic background could result in variation in drug–response phenotypes based on metabolism, transportation or elimination, referred to as pharmacogenetics (8,9) and affect both efficacy and toxicity of a drug. One pharmacogenetic example is the suspected association of genetic variants in the gene CYP2D6 and prognosis for tamoxifen-treated BC patients, although evidence is still conflicting (10–13). Another example is the observation that genetic variants of NQO1 could be associated with response to the drug epirubicin (14).

Alternatively, the genetic background of the host (including both high penetrance mutations as well as common variants) could result in different molecular profiles of a developing tumor. With respect to associations between genetic background and tumor subtypes, there is evidence from large collaborative studies that specific BC risk genotypes result in different molecular phenotypes of the tumor (15–17). This could result in a different molecular behavior and ultimately in a different prognosis for the BC patient. Finally, several studies have described associations of common germline genotypes with BC survival, without a specific functional explanation (18–25). For example, there is evidence that cancer risk factors can also act as prognostic factors after the onset of the disease; one study showed that hormone replacement therapy has an effect on the prognosis (26), and the timing and number of previous pregnancies might be associated with the prognosis as well (27).

The Breast Cancer Association Consortium (BCAC) has identified and validated several single nucleotide polymorphisms (SNPs) that are associated with the BC risk (28–33). Nine of these 11 SNPs were identified by validation of results from genome-wide association studies and two were identified through gene studies. It has previously been noted that rarer susceptibility variants for BC, notably the CHEK2 1100delC mutation, is associated with poorer survival (34,35). We hypothesized that these common susceptibility SNPs might also be associated with the prognosis. We therefore analyzed the association between these 11 SNPs and prognosis after BC investigating 25,853 patients from 23 studies participating in BCAC. In addition, we evaluated a further 62 SNPs, identified through GWAS or candidate studies but for which the evidence for susceptibility had not been confirmed, for association with prognosis.

RESULTS

General population information

A total of 23 BCAC studies were included for analyses of overall survival (OS) and 16 studies for BC-specific survival (BCS) analyses (Supplementary Material, Table S3). The data set comprised 25,853 BC patients, of whom 4076 died within the observation period. The mean follow-up time within individual studies ranged from 2.7 to 9.6 years (6.4
for the total data set). For BCS, data were available for 20,073 women, among whom there were 2282 BC-specific deaths (see Supplementary Material, Table S3). Patient and tumor characteristics by the study are shown in Supplementary Material, Tables S2a and b, and the total numbers of BC patients and events for each SNP are given in Supplementary Material, Table S4. The associations between morphological characteristics, estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) status, and SNP genotypes are shown in Tables 1 and 2. The distribution and the direction of the associations between SNPs and with tumor characteristics were similar to those found in previous analyses within BCAC (15,16).

**Association of SNPs with prognosis**

One of the 11 susceptibility SNPs, i.e. rs3803662 (TOX3), was significantly associated with OS and/or BCS in the complete data set at a nominal \( P < 0.01 \) (Table 3). The rare (T) allele of rs3803662 was associated with a poorer survival in all models; \( P \) heterogeneity was 0.0002 and 0.0009 for all-cause and BC-specific mortality, respectively (Table 3). The genotypic-specific results suggested a recessive mode of action (TT versus CC \( \text{HR}_{\text{adj}} = 1.21 \); 95% CI: 1.09–1.35, \( P = 0.0002 \) and \( \text{HR}_{\text{adj}} = 1.29 \); 95% CI: 1.12–1.47, \( P = 0.0003 \) for all-cause and BC-specific mortality, respectively). The risk in the rare CC homozygote genotype was associated with a lower survival compared with the common genotype (TT) (\( \text{HR}_{\text{adj}} = 1.31 \); 95% CI: 1.13–1.50, \( P = 0.0002 \) and \( \text{HR}_{\text{adj}} = 1.40 \); 95% CI: 1.15–1.70, \( P = 0.001 \) for all-cause and BC-specific mortality, respectively). The differences in the hazard ratio (HR) estimates was not statistically significant \( (P = 0.71) \). Additional analyses for this SNP by subgroups defined by grade, tumor size, nodal status, tumor biology, and survival for the outcome variable of OS (Table 4). The estimate of the association with prognosis was greater for ER-positive than ER-negative tumors (\( \text{HR}_{\text{adj}} = 1.81 \); 95% CI: 1.38–2.40, \( P = 0.001 \) and \( \text{HR}_{\text{adj}} = 1.40 \); 95% CI: 1.15–1.70, \( P = 0.001 \) for all-cause and BC-specific mortality, respectively).

### DISCUSSION

We analyzed 11 confirmed BC susceptibility SNPs and 62 candidate and GWAS-derived SNPs for the association with prognosis in up to 25,853 BC patients from 23 studies in BCAC. The strongest finding was that the risk allele of rs3803662 was associated with a poorer prognosis, consistent with a recessive model (HR = 1.21; 95% CI: 1.09–1.35, \( P = 0.0002 \) and \( \text{HR}_{\text{adj}} = 1.29 \); 95% CI: 1.12–1.47, \( P = 0.0003 \) for all-cause and BC-specific mortality, respectively). The distribution and the direction of the associations between SNPs and tumor characteristics were similar to those found in previous analyses within BCAC (15,16)

Breast tumor gene expression data for genes located close to the SNPs described above were analyzed for their association with prognosis. Probes with gene expression data in publicly available databases could be found for 9 out of the 11 confirmed BC risk SNPs (Supplementary Material, Table S8). We found no evidence of an association between TOX3 expression and prognosis in this data set (Supplementary Material, Table S9). However, the expression of RBL2 (retinoblastoma-like gene 2), which lies within 300 kb of the TOX3 locus, has previously shown an association with the rs3803662 genotype in lymphocytes (37). In the public data set, RBL2 expression was associated with prognosis only in ER-negative BC patients in one out of two analyzed probes for this gene (HR = 0.66 95% CI: 0.48–0.91). The most consistent evidence for an association with prognosis was found with probes in IGFBP2, which may be related to rs13387042 (four probes, minimum \( P = 0.01 \)) and FGFR2 (four probes, \( P = 0.003 \)). More limited evidence for an association with prognosis was found for probes in CASP8, TNS, LSP1 and COX11.

Breast tumor gene expression data for genes located close to the SNPs described above were analyzed for their association with prognosis. Probes with gene expression data in publicly available databases could be found for 9 out of the 11 confirmed BC risk SNPs (Supplementary Material, Table S8). We found no evidence of an association between TOX3 expression and prognosis in this data set (Supplementary Material, Table S9). However, the expression of RBL2 (retinoblastoma-like gene 2), which lies within 300 kb of the TOX3 locus, has previously shown an association with the rs3803662 genotype in lymphocytes (37). In the public data set, RBL2 expression was associated with prognosis only in ER-negative BC patients in one out of two analyzed probes for this gene (HR = 0.66 95% CI: 0.48–0.91). The most consistent evidence for an association with prognosis was found with probes in IGFBP2, which may be related to rs13387042 (four probes, minimum \( P = 0.01 \)) and FGFR2 (four probes, \( P = 0.003 \)). More limited evidence for an association with prognosis was found for probes in CASP8, TNS, LSP1 and COX11.

Breast tumor gene expression data for genes located close to the SNPs described above were analyzed for their association with prognosis. Probes with gene expression data in publicly available databases could be found for 9 out of the 11 confirmed BC risk SNPs (Supplementary Material, Table S8). We found no evidence of an association between TOX3 expression and prognosis in this data set (Supplementary Material, Table S9). However, the expression of RBL2 (retinoblastoma-like gene 2), which lies within 300 kb of the TOX3 locus, has previously shown an association with the rs3803662 genotype in lymphocytes (37). In the public data set, RBL2 expression was associated with prognosis only in ER-negative BC patients in one out of two analyzed probes for this gene (HR = 0.66 95% CI: 0.48–0.91). The most consistent evidence for an association with prognosis was found with probes in IGFBP2, which may be related to rs13387042 (four probes, minimum \( P = 0.01 \)) and FGFR2 (four probes, \( P = 0.003 \)). More limited evidence for an association with prognosis was found for probes in CASP8, TNS, LSP1 and COX11.

Breast tumor gene expression data for genes located close to the SNPs described above were analyzed for their association with prognosis. Probes with gene expression data in publicly available databases could be found for 9 out of the 11 confirmed BC risk SNPs (Supplementary Material, Table S8). We found no evidence of an association between TOX3 expression and prognosis in this data set (Supplementary Material, Table S9). However, the expression of RBL2 (retinoblastoma-like gene 2), which lies within 300 kb of the TOX3 locus, has previously shown an association with the rs3803662 genotype in lymphocytes (37). In the public data set, RBL2 expression was associated with prognosis only in ER-negative BC patients in one out of two analyzed probes for this gene (HR = 0.66 95% CI: 0.48–0.91). The most consistent evidence for an association with prognosis was found with probes in IGFBP2, which may be related to rs13387042 (four probes, minimum \( P = 0.01 \)) and FGFR2 (four probes, \( P = 0.003 \)). More limited evidence for an association with prognosis was found for probes in CASP8, TNS, LSP1 and COX11.

Breast tumor gene expression data for genes located close to the SNPs described above were analyzed for their association with prognosis. Probes with gene expression data in publicly available databases could be found for 9 out of the 11 confirmed BC risk SNPs (Supplementary Material, Table S8). We found no evidence of an association between TOX3 expression and prognosis in this data set (Supplementary Material, Table S9). However, the expression of RBL2 (retinoblastoma-like gene 2), which lies within 300 kb of the TOX3 locus, has previously shown an association with the rs3803662 genotype in lymphocytes (37). In the public data set, RBL2 expression was associated with prognosis only in ER-negative BC patients in one out of two analyzed probes for this gene (HR = 0.66 95% CI: 0.48–0.91). The most consistent evidence for an association with prognosis was found with probes in IGFBP2, which may be related to rs13387042 (four probes, minimum \( P = 0.01 \)) and FGFR2 (four probes, \( P = 0.003 \)). More limited evidence for an association with prognosis was found for probes in CASP8, TNS, LSP1 and COX11.

Breast tumor gene expression data for genes located close to the SNPs described above were analyzed for their association with prognosis. Probes with gene expression data in publicly available databases could be found for 9 out of the 11 confirmed BC risk SNPs (Supplementary Material, Table S8). We found no evidence of an association between TOX3 expression and prognosis in this data set (Supplementary Material, Table S9). However, the expression of RBL2 (retinoblastoma-like gene 2), which lies within 300 kb of the TOX3 locus, has previously shown an association with the rs3803662 genotype in lymphocytes (37). In the public data set, RBL2 expression was associated with prognosis only in ER-negative BC patients in one out of two analyzed probes for this gene (HR = 0.66 95% CI: 0.48–0.91). The most consistent evidence for an association with prognosis was found with probes in IGFBP2, which may be related to rs13387042 (four probes, minimum \( P = 0.01 \)) and FGFR2 (four probes, \( P = 0.003 \)). More limited evidence for an association with prognosis was found for probes in CASP8, TNS, LSP1 and COX11.
### Table 1. Patient and tumor characteristics by 11 BC susceptibility loci

<table>
<thead>
<tr>
<th>SNP</th>
<th>Variant</th>
<th>Age at diagnosis Mean</th>
<th>Age at diagnosis SD</th>
<th>Morphology (%)</th>
<th>Tumor size (%)</th>
<th>Lymph node status (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17468277</td>
<td>TT</td>
<td>54.8</td>
<td>10.7</td>
<td>72.2</td>
<td>13.0</td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>54.7</td>
<td>11.3</td>
<td>73.6</td>
<td>14.3</td>
<td>12.1</td>
</tr>
<tr>
<td>rs1982073</td>
<td>CC</td>
<td>54.6</td>
<td>11.6</td>
<td>75.5</td>
<td>13.4</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>20,372</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs13281615</td>
<td>AA</td>
<td>54.6</td>
<td>11.9</td>
<td>75.7</td>
<td>13.8</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>54.4</td>
<td>11.7</td>
<td>74.2</td>
<td>14.4</td>
<td>11.4</td>
</tr>
<tr>
<td>rs2981582</td>
<td>GG</td>
<td>54.3</td>
<td>11.3</td>
<td>73.5</td>
<td>14.8</td>
<td>11.8</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>22,361</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs3803662</td>
<td>CC</td>
<td>54.6</td>
<td>11.6</td>
<td>75.1</td>
<td>13.8</td>
<td>11.2</td>
</tr>
<tr>
<td>rs3817198</td>
<td>TT</td>
<td>54.3</td>
<td>11.5</td>
<td>74.9</td>
<td>13.9</td>
<td>11.2</td>
</tr>
<tr>
<td>rs889312</td>
<td>AA</td>
<td>54.6</td>
<td>11.7</td>
<td>74.8</td>
<td>14.2</td>
<td>11.0</td>
</tr>
<tr>
<td>rs13387042</td>
<td>AA</td>
<td>54.6</td>
<td>11.7</td>
<td>75.2</td>
<td>13.8</td>
<td>11.0</td>
</tr>
<tr>
<td>rs6504950</td>
<td>AA</td>
<td>54.1</td>
<td>11.5</td>
<td>74.6</td>
<td>13.1</td>
<td>12.4</td>
</tr>
<tr>
<td>rs10941679</td>
<td>AA</td>
<td>53.9</td>
<td>11.6</td>
<td>75.1</td>
<td>14.2</td>
<td>10.7</td>
</tr>
<tr>
<td>SNP</td>
<td>Variant</td>
<td>Grade (%)</td>
<td>ER (%)</td>
<td>PR (%)</td>
<td>HER2 (%)</td>
<td>PR (%)</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>-----------</td>
<td>--------</td>
<td>--------</td>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>rs17468277</td>
<td>TT</td>
<td>20.2</td>
<td>47.2</td>
<td>32.6</td>
<td>15.0</td>
<td>19.8</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>21.0</td>
<td>48.0</td>
<td>30.9</td>
<td>16.0</td>
<td>23.5</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>21.6</td>
<td>46.9</td>
<td>31.5</td>
<td>17.3</td>
<td>23.0</td>
</tr>
<tr>
<td>rs1982073</td>
<td>TT</td>
<td>21.5</td>
<td>47.8</td>
<td>30.7</td>
<td>15.7</td>
<td>22.6</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>21.3</td>
<td>47.7</td>
<td>31.0</td>
<td>14.9</td>
<td>21.3</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>21.7</td>
<td>46.9</td>
<td>31.4</td>
<td>15.3</td>
<td>23.8</td>
</tr>
<tr>
<td>rs13281615</td>
<td>AA</td>
<td>21.5</td>
<td>46.0</td>
<td>32.5</td>
<td>14.2</td>
<td>24.6</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>21.7</td>
<td>47.3</td>
<td>31.0</td>
<td>14.0</td>
<td>22.7</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>21.5</td>
<td>47.4</td>
<td>31.1</td>
<td>13.9</td>
<td>22.3</td>
</tr>
<tr>
<td>rs2981582</td>
<td>TT</td>
<td>21.7</td>
<td>45.5</td>
<td>34.0</td>
<td>16.1</td>
<td>26.8</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>22.4</td>
<td>47.7</td>
<td>29.8</td>
<td>16.2</td>
<td>22.1</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>23.5</td>
<td>47.9</td>
<td>28.6</td>
<td>17.0</td>
<td>19.6</td>
</tr>
<tr>
<td>rs3803662</td>
<td>CC</td>
<td>20.5</td>
<td>46.8</td>
<td>32.7</td>
<td>16.1</td>
<td>26.8</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>22.5</td>
<td>47.2</td>
<td>30.4</td>
<td>16.1</td>
<td>22.5</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>23.1</td>
<td>46.3</td>
<td>30.6</td>
<td>15.4</td>
<td>21.7</td>
</tr>
<tr>
<td>rs3817198</td>
<td>TT</td>
<td>21.9</td>
<td>46.6</td>
<td>31.5</td>
<td>16.2</td>
<td>23.4</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>22.2</td>
<td>45.6</td>
<td>32.2</td>
<td>16.5</td>
<td>22.7</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>23.1</td>
<td>47.2</td>
<td>29.7</td>
<td>17.2</td>
<td>24.2</td>
</tr>
<tr>
<td>rs889312</td>
<td>AA</td>
<td>21.9</td>
<td>47.1</td>
<td>31.0</td>
<td>15.7</td>
<td>23.2</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>21.7</td>
<td>46.9</td>
<td>31.4</td>
<td>16.8</td>
<td>22.8</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>21.5</td>
<td>48.3</td>
<td>30.3</td>
<td>17.0</td>
<td>22.8</td>
</tr>
<tr>
<td>rs13387042</td>
<td>GG</td>
<td>20.7</td>
<td>46.5</td>
<td>32.9</td>
<td>16.6</td>
<td>23.9</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>20.6</td>
<td>47.1</td>
<td>32.3</td>
<td>15.5</td>
<td>23.4</td>
</tr>
<tr>
<td>rs9473768</td>
<td>CC</td>
<td>21.1</td>
<td>47.9</td>
<td>30.9</td>
<td>12.5</td>
<td>22.9</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>21.9</td>
<td>46.5</td>
<td>31.7</td>
<td>12.9</td>
<td>22.1</td>
</tr>
<tr>
<td>rs6504950</td>
<td>AA</td>
<td>21.1</td>
<td>46.9</td>
<td>31.9</td>
<td>19.2</td>
<td>25.5</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>20.9</td>
<td>46.5</td>
<td>32.6</td>
<td>17.8</td>
<td>22.9</td>
</tr>
<tr>
<td>rs10941679</td>
<td>AA</td>
<td>21.0</td>
<td>47.8</td>
<td>31.3</td>
<td>17.6</td>
<td>22.3</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>22.7</td>
<td>48.5</td>
<td>28.8</td>
<td>18.1</td>
<td>22.6</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>22.8</td>
<td>50.7</td>
<td>26.5</td>
<td>16.4</td>
<td>20.4</td>
</tr>
</tbody>
</table>

aGenotype groups are represented in order of: homozygous non-risk, heterozygous, homozygous risk allele. Bold numbers indicate significant differences at $P < 0.05$ between SNP variant groups and other variables, excluding unknown categories; differences were tested by ANOVA for age at diagnosis and by nptrend (nptrend performs the non-parametric test for trend across ordered groups) for categorical variables. Percentages in unknown categories reflect the proportion of unknown information for each variable (percentages in the known categories sum up to 100%).
error rate of 0.001 (events = 4000, allele events = 8000); however, the power falls to 90% for a risk allele frequency of 0.1. Therefore, despite the large sample size, the power to detect HRs of the order of magnitude of relative risks detected for disease susceptibility phenotypes is less. We had 70% power to detect a risk allele of frequency 0.3 that is associated with an OS per-allele HR of 1.1, falling to 20% for an allele of frequency 0.1. Effect sizes for breast-specific mortality would be expected to be somewhat larger (since the genotypes are unlikely to be similarly associated with other causes of death), but our sample size was smaller for this endpoint (16 out of 23 studies with 2282 events). We had 98% power to detect a risk allele of frequency 0.3 that is associated with a BCS per-allele HR of 1.2, falling to 56% for an allele of frequency 0.1. Other strengths of this study include the high-quality genotyping with the stringent quality control (QC) criteria.

A weakness of the study is that the methods of clinical data collection varied across studies, although data were centrally checked and cleaned. There might therefore be heterogeneity in the assessment and completeness of tumor and patient characteristics such as tumor size, lymph node status and histopathological grade ER status, across the studies. Furthermore, inclusion of prevalent versus incident cases, the
assessment method and completeness of follow-up were different across the studies (Supplemental Material, Table S1). However, missing data and misclassification probabilities are likely to be independent of susceptibility genotype and will therefore result in an underestimation of true associations than generating false positive associations. Possible effects of the inclusion of prevalent cases were addressed by left truncation of the survival time of these patients (38). Another potential difficulty with these analyses presented is that the clinical course of disease depends on a variety of factors, and the association with any one common genetic variant might be diluted by the effect of other factors. In our study, we were able to adjust for ER status, grade, tumor size and lymph node status. Survival analyses that are additionally adjusted for other factors such as HER2/neu status might improve the estimates. However, even with more than 25,000 patients in our data set, the number of deaths is still too small to examine associations for most subgroups. Some associations may also be modified by treatment. We were unable to explore interactions with specific treatments due to limitations of available data. Collection of detailed treatment data is ongoing and should allow treatment-specific analyses in future.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Variant</th>
<th>ER-negative Unadjusted HR (95% CI)</th>
<th>Adjusted HR (95% CI)</th>
<th>P-value**</th>
<th>ER-positive Unadjusted HR (95% CI)</th>
<th>Adjusted HR (95% CI)</th>
<th>P-value**</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs17468277</td>
<td>TT</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td></td>
<td></td>
<td>1 (reference)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>0.86 (0.50–1.56)</td>
<td>0.94 (0.54–1.67)</td>
<td>0.69</td>
<td></td>
<td>0.99 (0.66–1.48)</td>
<td>0.92 (0.61–1.38)</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0.95 (0.55–1.64)</td>
<td>0.98 (0.57–1.71)</td>
<td>0.89</td>
<td></td>
<td>1.03 (0.69–1.53)</td>
<td>0.98 (0.66–1.46)</td>
</tr>
<tr>
<td>rs1982073</td>
<td>TT</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td></td>
<td></td>
<td>1 (reference)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>1.11 (0.95–1.30)</td>
<td>1.07 (0.91–1.25)</td>
<td>0.14</td>
<td></td>
<td>0.99 (0.88–1.11)</td>
<td>1.00 (0.89–1.12)</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>1.23 (1.00–1.51)</td>
<td>1.15 (0.94–1.42)</td>
<td>0.39</td>
<td></td>
<td>0.97 (0.83–1.14)</td>
<td>0.98 (0.83–1.15)</td>
</tr>
<tr>
<td>rs13281615</td>
<td>AA</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td></td>
<td></td>
<td>1 (reference)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>1.06 (0.92–1.21)</td>
<td>1.07 (0.93–1.22)</td>
<td></td>
<td></td>
<td>0.96 (0.87–1.07)</td>
<td>0.94 (0.85–1.04)</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>1.00 (0.83–1.20)</td>
<td>1.01 (0.84–1.22)</td>
<td></td>
<td></td>
<td>0.99 (0.87–1.13)</td>
<td>1.00 (0.87–1.14)</td>
</tr>
<tr>
<td>rs2981582</td>
<td>CC</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td></td>
<td></td>
<td>1 (reference)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>1.02 (0.89–1.17)</td>
<td>1.00 (0.88–1.16)</td>
<td>0.69</td>
<td></td>
<td>1.05 (0.94–1.17)</td>
<td>1.07 (0.96–1.19)</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>0.99 (0.82–1.20)</td>
<td>0.97 (0.80–1.17)</td>
<td></td>
<td></td>
<td>0.98 (0.86–1.13)</td>
<td>1.03 (0.90–1.18)</td>
</tr>
<tr>
<td>rs3803662</td>
<td>CC</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>0.93</td>
<td></td>
<td>1 (reference)</td>
<td>0.92 (0.85–1.04)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>0.98 (0.86–1.12)</td>
<td>0.97 (0.85–1.10)</td>
<td></td>
<td></td>
<td>0.94 (0.85–1.04)</td>
<td>0.98 (0.89–1.08)</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>1.14 (0.93–1.40)</td>
<td>1.11 (0.90–1.36)</td>
<td></td>
<td></td>
<td>1.23 (1.07–1.42)</td>
<td>1.31 (1.13–1.50)</td>
</tr>
<tr>
<td>rs3817198</td>
<td>TT</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>0.37</td>
<td></td>
<td>1 (reference)</td>
<td>0.45 (0.95–1.15)</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>0.99 (0.87–1.13)</td>
<td>0.92 (0.81–1.05)</td>
<td></td>
<td></td>
<td>1.04 (0.95–1.15)</td>
<td>1.04 (0.94–1.15)</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0.78 (0.61–0.98)</td>
<td>0.74 (0.59–0.93)</td>
<td></td>
<td></td>
<td>0.99 (0.84–1.18)</td>
<td>1.02 (0.86–1.21)</td>
</tr>
<tr>
<td>rs889312</td>
<td>AA</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>0.09</td>
<td></td>
<td>1 (reference)</td>
<td>0.66 (0.84–1.04)</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>1.10 (0.97–1.25)</td>
<td>1.09 (0.95–1.24)</td>
<td></td>
<td></td>
<td>0.95 (0.86–1.04)</td>
<td>0.93 (0.84–1.02)</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>1.15 (0.93–1.44)</td>
<td>1.07 (0.86–1.33)</td>
<td></td>
<td></td>
<td>0.84 (0.71–1.01)</td>
<td>0.86 (0.72–1.03)</td>
</tr>
<tr>
<td>rs13387042</td>
<td>GG</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>0.24</td>
<td></td>
<td>1 (reference)</td>
<td>0.44 (0.95–1.20)</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>1.02 (0.88–1.19)</td>
<td>1.03 (0.88–1.20)</td>
<td></td>
<td></td>
<td>1.07 (0.95–1.20)</td>
<td>1.10 (0.98–1.23)</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.87 (0.73–1.04)</td>
<td>0.89 (0.75–1.06)</td>
<td></td>
<td></td>
<td>1.00 (0.88–1.14)</td>
<td>1.03 (0.91–1.17)</td>
</tr>
<tr>
<td>rs4973768</td>
<td>CC</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>0.08</td>
<td></td>
<td>1 (reference)</td>
<td>0.17 (0.95–1.15)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>0.95 (0.81–1.11)</td>
<td>0.98 (0.84–1.15)</td>
<td></td>
<td></td>
<td>1.02 (0.91–1.15)</td>
<td>1.10 (0.98–1.23)</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>0.90 (0.74–1.09)</td>
<td>0.92 (0.76–1.12)</td>
<td></td>
<td></td>
<td>1.00 (0.87–1.15)</td>
<td>1.03 (0.91–1.17)</td>
</tr>
<tr>
<td>rs6504950</td>
<td>AA</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>0.56</td>
<td></td>
<td>1 (reference)</td>
<td>0.69 (0.86–1.28)</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>1.17 (0.90–1.52)</td>
<td>1.15 (0.88–1.49)</td>
<td></td>
<td></td>
<td>1.05 (0.86–1.28)</td>
<td>1.02 (0.83–1.24)</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>1.04 (0.80–1.35)</td>
<td>1.04 (0.80–1.35)</td>
<td></td>
<td></td>
<td>0.98 (0.80–1.19)</td>
<td>0.99 (0.81–1.20)</td>
</tr>
<tr>
<td>rs10941679</td>
<td>AA</td>
<td>1 (reference)</td>
<td>1 (reference)</td>
<td>0.19</td>
<td></td>
<td>1 (reference)</td>
<td>0.31 (0.95–1.24)</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>1.09 (0.95–1.25)</td>
<td>1.09 (0.95–1.25)</td>
<td></td>
<td></td>
<td>1.01 (0.91–1.12)</td>
<td>1.04 (0.94–1.15)</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>1.00 (0.76–1.31)</td>
<td>0.97 (0.74–1.28)</td>
<td></td>
<td></td>
<td>1.16 (0.98–1.39)</td>
<td>1.14 (0.95–1.36)</td>
</tr>
</tbody>
</table>

*Adjusted analyses included age at diagnosis, tumor size, nodal status and grade as co-variants, and were stratified by study; CI, confidence interval.
**P-values were derived from Wald statistics with 2df (overall test of heterogeneity). Unique Cox models per SNP and per outcome measure. All reference categories represent the homozygous non-risk allele (note: in earlier publications ORs for rs17468277 and rs6504950 for which the rare allele is protective for BC risk were shown in reference to the common allele).
We saw no clear evidence for an association between gene expression levels and prognosis that mirrored the associations with prognosis. Only IGFBP2, close to rs13387042, and FGFR2, showed any consistent association with prognosis, across several probes. The lack of consistency is not that surprising, since for most loci, causality is not known for certainty, nor is it known whether the association is mediated through gene expression. For example, no association between the rs3803662 genotype and TOX3 expression has been observed. Moreover, it is possible that variants in a locus influence gene expression during the course of pathogenesis, but might not necessarily have an effect in malignant tumor tissue. Generally, it has to be kept in mind that an alteration of gene expression is not the only way through which different genotypes can have an influence on the phenotype. Differences in gene expression caused by epigenetic changes or, for example, interactions with miRNAs might have an impact on BC pathogenesis and/or survival as well.

Multiple studies have confirmed that the rare allele of rs3803662 is associated with an increased risk of BC of 1.2-fold per allele (17,30,39,40). Although the rare allele increases the risk for both ER-positive and -negative cancer, the relative risk for ER-positive disease is higher (16). Similarly, the relative risk for a BC-related death was the higher in ER-positive BC patients. The association was in the same direction in ER-negative patients, but non-significant, perhaps reflecting the smaller sample size. The variant(s) in the region that are functionally related to BC risk remain unknown. Fine-scale mapping and association analyses of different ethnic groups have found rs4784227 to have the strongest association with the BC risk; however, 14 other candidates could not be excluded (41). An attempt to associate rs3803662 with gene expression of genes in close proximity revealed a possible association with the levels of Retinoblastoma like 2 protein (RBL2). RBL2 is part of the RB family and was found to play regulatory roles in G0-senescence (37). Genetically, its expression was also found to be associated with rs8050136, a diabetes and obesity susceptibility SNP, which is located in the intron 1 of FTO, showing that RBL1 might be regulated by cis-acting factors as much as 270 kb away (42). There is, however, no evidence to support an association between rs8050136 and BC risk. Alternatively, the association could be mediated through another gene, for which TOX3 is the most plausible candidate. SNP rs3803662 lies 8 kb upstream of TOX3, and is genetically linked with rs17271951, rs1362548, rs3095604 and rs4784227, which lie in the 5′ regulatory region of TOX3. TOX3 belongs to the diverse family of HMG-box proteins that function as architectural factors in the modification of chromatin structure by bending and unwinding DNA (43). TOX3 is differentially expressed in patients who experienced BC relapse to bone versus those who experienced relapse elsewhere in the body (44). TOX3 mediates the calcium-dependent transcription of c-fos (45). From our analyses, we found no conclusive evidence that either TOX3 or RBL2 expression were associated with BC prognosis.

Conclusion

BC-susceptibility SNPs have little or no association with survival for BC patients. One exception might be rs3803662. With an HR of 1.29; 95% CI: 1.12–1.47, lacking functional explanation this marker is not ready for clinical use. Further functional studies will be required as well as replication in studies focusing on treatment of different BC subtypes (e.g. triple negative). As well as known clinico-pathologic and patient factors, survival might be in part determined by a distinct set of germline variants from those influencing susceptibility, or the effects of other susceptibility variants may be too limited to be revealed even in a large data set. Nonetheless, the fact that some cancer susceptibility factors can have an association with prognosis might help link cancer prevention to cancer treatment.

MATERIALS AND METHODS

Patient selection

BCAC comprises 55 BC case–control studies in which susceptibility variants have been genotyped; these include both population- and hospital-based studies, and studies with both population-based and hospital-based ascertainment (http://www.srl.cam.ac.uk/consortia/bcac/). BCAC studies that had genotyped at least 1 of the 11 confirmed BC susceptibility variants analyzed and published by BCAC, and had follow-up data available at the time of analysis, were initially included (n = 27 studies, n = 36 436 cases; see the description of studies in Supplementary Material, Table S1). Cases from these studies were included on an individual basis following these criteria in hierarchical order: female (109 males excluded), follow-up and vital status available (3492 cases excluded), invasive breast tumor (1577 in situ cases, and 448 cases of unknown invasiveness excluded) and ascertained for their first tumor (411 cases ascertained for their second breast tumor and 360 where the basis of ascertainment was
unknown excluded). Of the 30,039 eligible cases, there were 28,259 of European origin, 720 Hispanic Americans, 958 Africans and 102 of unknown ethnicity. Owing to the small number of women of non-European origin, we based all analyses on women of European ancestry. Of 28,259 European cases, 26,306 remained after adjustment for study entry (see statistical analyses). Furthermore, four studies with <10 events for all-cause or BC-specific mortality were excluded from all the analyses leaving 25,853 cases in the analyses (Supplementary Material, Tables S1–S3).

Clinico-pathologic information and follow-up

Clinico-pathologic information and follow-up data were collected by each study individually through medical records and cancer registries. Data were pooled in the BCAC database according to a data dictionary, and checked for accuracy and consistency centrally. Data included were: age at diagnosis, tumor size (≤2 cm, >2 and ≤5 cm, >5 cm), lymph node status (negative or positive), differentiation grade (I, II, III), ER status, PR status, HER2 status, follow-up and vital status (all-cause mortality and BC-specific mortality) (Supplementary Material, Tables S2a and b, S3 and S4).

SNP selection

We selected nine SNPs identified through GWAS that had been genotyped through studies within BCAC (Supplementary Material, Table S5). Seven of these SNPs were discovered by a GWAS for which BCAC series were used as a replication stage: rs2981582, within intron 2 of FGFR2 (per allele OR = 1.26; 95% CI: 1.23–1.28); rs3803662, a synonymous coding SNP of LOC643714 which lies 8 kb upstream of TOX3 (per allele OR = 1.20; 95% CI: 1.16–1.24); rs889312, which lies in an LD block containing the MAP3K1 gene (per allele OR = 1.13; 95% CI: 1.10–1.16); rs3817198, which lies within intron 10 of LSP1 (per allele OR = 1.07; 95% CI: 1.04–1.11); rs13281615 at 8q24 (per allele OR = 1.08; 95% CI: 1.05–1.11) (30); rs4973768, which is located on the short arm of chromosome 3 near the potential causative genes SLC4A7 and NEK10 [OR = 1.11 (95% CI: 1.08–1.13; P = 4.1E – 23)] (32), and rs6504950, which lies within intron 1 of STXB4 [OR = 0.95 (95% CI: 0.92–0.97, P = 1.4 E – 08)] (32). Two further SNPs were identified in studies from Iceland: rs13387042 lies in 90 kb at 2q35 that contains neither known genes nor non-coding RNAs (46), OR = 1.12 (95% CI: 1.09–1.15; P = 1E – 19) (31); and rs10941679, located in the 5p12-11 region, which contains the genes FGFI0 and MRPS30 (OR = 1.19, P = 2.9E – 11) (33, 47).

Two validated SNPs from candidate gene studies were also included (28). rs1982073, a missense polymorphism within TGFBI (per allele OR = 1.08; 95% CI: 1.04–1.11, P = 1.5 E – 05); and rs1045485, non-synonymous change in CASP8 (per allele OR = 0.88; 95% CI: 0.84–0.92, P = 5.7 E – 07) (29).

In addition to the 11 SNPs described above, we also considered for completeness all other SNPs (n = 62), which had been genotyped by the BCAC (until 2009). These SNPs had been proposed for replication genotyping by participating groups within BCAC (Supplementary Material, Tables S6 and S7). They were selected either for replication of putative associations from a GWAS (n = 26) (30) or as candidate SNPs suggested from associations in other studies; nine of them reported in Cox et al. (29), five of them in Gaudet et al. (48) and the remainder in other case–control studies (36, 49–57).

Genotyping

Genotyping was performed in the framework of BCAC as described previously by Taqman and iPLEX assays (29–32). Genotype data were excluded from the analysis on a study-by-study basis according to the following BCAC QC guidelines: (i) any sample that consistently failed for >20% of the SNPs within a genotyping round, (ii) all samples on any one plate that had a call rate <90% for any one SNP, (iii) all genotype data for any SNP where overall call rate was <95%, (iv) all genotype data for any SNP where duplicate concordance was <98%. In addition, for any SNP where the P-value for departures from Hardy–Weinberg proportions for controls was <0.005, clustering of the intensity plots was reviewed manually and the data excluded if clustering was judged to be poor.

Susceptibility SNPs were genotyped by 22 studies for rs1045485/rs17468277 (CASP8), 18 studies for rs1982073 (TGFBI), 22 studies for rs2981582 (FGFR2), 23 studies for rs13281615 (8q24), 22 studies for rs3817198 (LSP1), 23 studies for rs889312 (MAP3K1), 24 studies for rs3803662 (TOX3), 23 studies for rs13387042 (2q35), 19 studies for rs4973768 (3p24), 21 studies for rs6504950 (17q23) and 23 studies for rs10941679 (5p12). Because of technical difficulties in genotyping rs1045485 in CASP8 reported in our original publication (29), a tightly linked SNP (rs17468277) was used as a surrogate in subsequent studies (r² = 1 in HapMap CEU).

Analysis of publicly available gene expression data

Independent from the BCAC data set and as an accessory analysis, we investigated the prognostic relevance of gene expression data of genes within close proximity of the 11 confirmed BC risk SNPs. Methods for utilizing gene expression data from publicly available databases have been described previously (58). Briefly, we combined a database of 3488 Affymetrix HGU133A microarrays from primary BC patients. Studies and patient characteristics are listed in Supplementary Material, Table S8. For comparability, only the ProbeSets available on the Affymetrix HGU-133A microarray were used from 29 data sets where HG-U133Plus2.0 microarrays were applied. Affymetrix expression data were analyzed by using the MAS5.0 (59) algorithm of the Affymetrix package (60) of the Bioconductor software project (61) (http://www.bioconductor.org/). Subsequently, data were log2 transformed and median centered across arrays. For further normalization, the expression values of all the genes on the array were multiplied by a scale factor S so that the magnitude (sum of the squares of the values) equals 1 (this method is similar to scaling by Z score transformation and uses the more robust median-centering). For stratification of samples according to
gene expression, conservative median splits of cohorts were used. Stratifications were performed separately in ER-positive and ER-negative subtypes of BC according to the respective medians on a data set by data set basis to avoid the introduction of a bias through different expression levels in these two subtypes.

Statistical analyses

Univariate analyses for each SNP (categorical variable with the common homozygous group as the reference) were carried out by fitting Cox proportional hazard (PH) regression models. All models were evaluated for the PH assumption by visual inspection of log–log plots and analytically using the Schoenfeld method. The primary test of significance was a (conservative) 2df Wald test for heterogeneity. All Cox models were stratified by study and left-truncated at the date of ascertainment (blood draw) to allow for inclusion of prevalent cases. Survival time was calculated from the date of diagnosis to the date of last follow-up or death, left-truncated for years to enter the study, if the entry was after diagnosis (54). Survival time was right censored at 10 years because some studies had a relatively short follow-up (Supplementary Material, Table S3), and to reduce the effects of competing risks in the overall mortality analysis, especially for older patients. Multivariate Cox models were fitted including additional covariates, i.e. age at diagnosis (continuous) and categorical variables for tumor size, lymph node status and grade. We also fitted additional models including ER status as a covariate, and fitted separate models for ER-positive and ER-negative cases. Missing values were included as a separate category (e.g. grade1, grade2, grade3, grade missing). In addition, univariate and multivariate Cox models were fitted using only those cases without missing information for the variables concerned (age, tumor size, lymph node status, grade and ER status). Results from the multivariate models adjusting for ER status and the complete case analyses were similar to the results shown in Table 3. For all Cox models including the 11 BC susceptibility SNPs, the none-risk allele was used as the reference; for the 62 other SNPs, the common allele was used as the reference. For all Cox models in-

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

We thank all the individuals who took part in the component studies contributing to this work and all the researchers, clinicians and administrative staff who have made possible those studies. In particular, we thank: Craig Lucarrini, Don Conroy and Caroline Baynes and the Studies of Epidemiology and Risk Factors in Cancer Heredity team (SEARCH); Kirsti-mari Aaltonen, Päivi Heikkilä, Karl von Smitten, Tuomas Heikkinen, Hanna Jäntti, Irja Erkkilä (HEBCS); Participants in CGPS (CPGS); Pei Chao, Michael Stagner, Stephen Chanock, Laurie Burdette, Amy Hutchinson, Jeff Yuenger, Louise Brinton, Beata Peplonska (PBS); Alice S. Whitemore and Dee W. West (NC-BCFR); Teresa Selander, Mona Gill, Lucine Collins, Naya Na Weerasooriya and members of the Ontario Familial Breast Cancer Registry (OFBCR); Margaret McCredie, Judi Maskiell, Kelly Aujard, the women and their families that participated in the Australian Breast Cancer Family Study (ABCFS); Reiner Strick, Silke Landrith, Janett Blom, Jan Molenaar (ORIGO); Natalia Bogdanova, Peter Schürrmann, Andreas Meyer, Johann Karstens, Peter Hillelmanns (HABC); Ian Brock, Graeme Elliott, Dan Connelly, Helen Cramp, Sue Higham (HSB); Jerry Reid, Diane Kampa, Allison Iwan, Richard Hoffbeck, Chris McClure, Laura Bowen (USRTS); Ans van den Ouweland, Jannet Blom, Petra Bos, Ellen Crepin, Annette Heemskerk-Gerritsen, Elisabeth Huijseks, Anja Nieuwlaat (RBCS); Tanya Koehler (GESBC); Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study (funded 2001–2009 by NHMRC and currently by the National Breast Cancer Foundation and Cancer Australia #628333) for their contributions to this resource, and the many families who contribute to kConFab/AOCS; Gillian Peuteman, Dominiek Smeets, Thomas Van Brussel and Kathleen Corthouts (LMBC); Mervi Grip, Kari Mononen and Meeri Otsukka (OBCS); Sandrine Tchatchou, Miriam Wiestler, Anne Langheinz (BSUCH); Eija Myöhänien and Helena Kemiläinen (KBCP); Paolo Radice, Siranoush Manoukian, Monica Barile, Marco A Pierotti and the Cancer Genetic Test Laboratory personnel (MBCG); Charo Alonso, Tais Moreno, Guillermo Pita, Primitiva Menendez, Anna González-Neira (CNIO-BCS).

Conflict of Interest statement. None declared.

FUNDING

This work was supported by funding from the European Community’s Seventh Framework Programme under grant agreement -223175 (HEALTH-F2-2009-223175) (CGPS), Co-ordination and genotyping in BCAC was funded by CR-UK (C1287/A10118 and C1287/A12014). Meetings of the BCAC have been funded by the European Union COST programme (BM0606). D.F.E. is a Principal Research
Fellow of CR-UK. The genotyping was supported by Cancer Research UK (C490/A10124) and (C1287/A12014). Funding of the constituent studies and data analysis was supported by the Cambridge Biomedical Research Centre, Cancer Research UK (C490/A10119). Other studies have been funded by the Dutch Cancer Society (grants NKI 2001-2423; NKI 2007-3839), the Dutch National Genomics Initiative, the Dutch Biobanking and Biomolecular Resources Research Infrastructure, the National Institutes of Health (NIH) National Cancer Institute (NCI) grant CA122340, NCI Recovery Act grant CA122340Z, NCI Specialized Program of Research Excellence (SPOR) in breast cancer CA116201, the ELAN research program, University Hospital Erlangen, the Deutsche Krebshilfe e. V. (70492), the state of Baden-Württemberg through the Medical Faculty of the University of Ulm (P.685), the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, the Associazione Italiana per la Ricerca sul Cancro (4017) and funds from Italian citizens who allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects ‘5x1000’), the ‘Stichting tegen Kanker’ (232-2008 and 196-2010), the special Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, the Academy of Finland and by the strategic funding of the University of Eastern Finland, the Chief Physician Johan Bosserup and Lise Bosserup Fund, the Danish Medical Research Council and Copenhagen University Hospital, Herlev Hospital, Rudolf Barling Foundation, the National Breast Cancer Foundation of Australia, Cancer Australia, the National Health and Medical Research Council of Australia, the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, the Cancer Foundation of Western Australia, the Victorian Health Promotion Foundation (Australia), the Breast Cancer Campaign (2004Nov49), Yorkshire Cancer Research (S295, S299 and core funding), the Dietmar-Hopp Foundation, the Helmholtz Society, the German Cancer Research Center (DKFZ), the Finnish Cancer Foundation, the Academy of Finland, the University of Oulu, the Oulu University Hospital, Biocenter Oulu; the Helsinki University Central Hospital Research Fund, Academy of Finland (132473), the Finnish Cancer Society, the Nordic Cancer Union and the Sigrid Jusélius Foundation; the Dutch Cancer Society, grants DDHK 2004-3124 and DDHK 2009-4318, the Swedish Cancer Society, the Stockholm Cancer Society, the Gustav V Jubilee foundation, the Bert von Kantzow foundation, the Agency for Science, Technology and Research of Singapore (A*STAR), the US National Institute of Health (NIH), the Susan G. Komen Breast Cancer Foundation and Märit and Hans Rausings Initiative Against Breast Cancer, the Genome Spain Foundation, the Red Temática de Investigación Cooperativa en Cáncer and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI081583 and PI081120). The ABCFS, NC-BCFR and OFBCR work was supported by the United States National Cancer Institute, National Institutes of Health (NIH) under RFA-CA-06-503 and through cooperative agreements with members of the Breast Cancer Family Registry (BCFR) and Principal Investigators, including Cancer Care Ontario (U01 CA69467), Cancer Prevention Institute of California (U01 CA69417) and University of Melbourne (U01 CA69638). Samples from the NC-BCFR were processed and distributed by the Coriell Institute for Medical Research. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the BCFR, nor does mention of trade names, commercial products or organizations imply endorsement by the US Government or the BCFR.

REFERENCES


