The role of genetic breast cancer susceptibility variants as prognostic factors

Peter A. Fasching[1,3,*,†], Paul D.P. Pharoah[4,5,†], Angela Cox[6,†], Heli Nevanlinna[9], Stig E. Bojesen[11,12], Thomas Karn[14], Annegien Broeks[15], Flora E. van Leeuwen[16], Laura J. van 't Veer[15], Renate Udo[16], Alison M. Dunning[4], Dario Greco[9], Kristiina Aittomäki[17], Carl Blomqvist[10], Mitul Shah[4], Børge G. Nordestgaard[11,12], Henrik Flyger[13], John L. Hopper[19], Melissa C. Southey[18], Carmel Apicella[19], Montserrat Garcia-Closas[20], Mark Sherman[21], Jolanta Lissowska[22], Caroline Seynaeve[23], Petra E.A. Huijts[24], Rob A.E.M. Tollenaar[25], Argyrios Ziogas[27], Arif B. Ekici[2], Claudia Rauh[1], Arto Mannermaa[28,29,32], Vesa Kataja[30,31,33], Veli-Matti Kosma[28,29,32], Jaana M. Hartikainen[28,29,32], Irene L. Andrulis[34,35,36,37], Hilmi Ozcelik[35,36], Anna-Marie Mulligan[35,37], Gord Glendon[33], Per Hall[39], Kamila Czene[39], Jianjun Liu[42], Jenny Chang-Claude[43], Shan Wang-Gohrke[45], Ursula Eilber[43], Stefan Nickels[43], Thilo Dörk[46], Maria Schiekel[46], Michael Bremer[47], Tjoung-Won Park-Simon[46], Graham G. Giles[16,19], Gianluca Severi[48,19], Laura Baglietto[48,19], Maartje J. Hooning[25], John W.M. Martens[49], Agnes Jager[23], Mieke Kriege[23], Annika Lindblom[40], Sara Margolin[41], Fergus J. Couch[50], Kristen N. Stevens[51], Janet E. Olson[51], Matthew Kosel[51], Simon S. Cross[7], Sabapathy P. Balasubramanian[8], Malcolm W.R. Reed[8], Alexander Miron[52], Esther M. John[53,54,55], Robert Winqvist[56], Katri Pylkäs[56], Arja Jukkola-Vuorinen[57], Saila Kauppila[58], Barbara Burwinkel[59,44], Frederik Marme[59,60], Andreas Schneeweiss[59,60], Christof Sohn[59], Georgia Chenevix-Trench[61], kConFab Investigators[62], Diether Lambrechts[63], Anne-Sophie Dieudonne[64], Sigrid Hatse[64], Erik van Limbergen[64], Javier Benitez[65], Roger L. Milne[66], M. Pilar Zamora[67], José Ignacio Arias Pérez[68], Bernardo Bonanni[69], Bernard Peissel[71], Bernard Loris[70], Paolo Peterlongo[72,73], Preetha Rajaraman[74], Sara J. Schonfeld[74], Hoda Anton-Culver[27], Peter Devilee[24,26], Matthias W. Beckmann[1], Dennis J. Slamon[3], Kelly-Anne Phillips[75,76], Jonine D. Figueroa[21], Manjeet K. Humphreys[5], Douglas F. Easton[4,5] and Marjanka K. Schmidt[15,16,†]

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

Herlev Hospital, Copenhagen University Hospital, Copenhagen, Denmark, [14]Department of Gynecology and Obstetrics, University Hospital, Johann-Wolfgang-Goethe University, Frankfurt a.M., Germany, [15]Division of Molecular Pathology and [16]Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Amsterdam, The Netherlands, [17]Department of Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland, [18]Genetic Epidemiology Laboratory, Department of Pathology and [19]Centre for Molecular, Enviromental, Genetic and Analytic Epidemiology, Melbourne School of Population Health, The University of Melbourne, Melbourne, Australia, [20]Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, UK, [21]Divsion of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA, [22]Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland, [23]Department of Medical Oncology, Rotterdam Family Cancer Clinic, Erasmus Medical Center-Daniel den Hoed Cancer Center, Rotterdam, The Netherlands, [24]Department of Human Genetics, [25]Department of Surgical Oncology and [26]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands, [27]UCI Center for Cancer Genetics Research & Prevention School of Medicine, University of California at Irvine, Irvine, CA, USA, [28]School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine, [29]Biocenter Kuopio, Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland, [30]School of Medicine, Institute of Clinical Medicine, Oncology, [31]Biocenter Kuopio, Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland, [32]Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland [33]Cancer Center, Kuopio University Hospital, Kuopio, Finland, [34]Ontario Cancer Genetics Network, Cancer Care Ontario, Toronto, Canada, [35]Department of Laboratory Medicine and Pathobiology, [36]Department of Laboratory Medicine and Pathobiology, Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital and [37]Department of Molecular Genetics, Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Canada, [38]Department of Laboratory Medicine, Keenan Research Centre of the Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, Canada, [39]Department of Medical Epidemiology and Biostatistics, [40]Department of Molecular Medicine and Surgery and [41]Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden, [42]Human Genetics Laboratory, Genome Institute of Singapore, Singapore, Singapore, [43]Division of Cancer Epidemiology and [44]Molecular Epidemiology Unit, German Cancer Research Center (DKFZ), Heidelberg, Germany, [45]Department of Obstetrics and Gynecology, University Hospital Ulm, Ulm University, Ulm, Germany, [46]Clinics of Obstetrics and Gynaecology and [47]Clinics of Radiation Oncology, Hannover Medical School, Hannover, Germany, [48]Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia, [49]Department of Medical Oncology, Josephine Nefkens Institute, Erasmus University Medical Center, Rotterdam, The Netherlands, [50]Department of Laboratory Medicine and Pathology and [51]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, [52]Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA, [53]Cancer Prevention Institute of California, Fremont, CA, USA, [54]Stanford University School of Medicine, Stanford, CA, USA, [55]Stanford Cancer Institute, Stanford, CA, USA, [56]Laboratory of Cancer Genetics, Department of Clinical Genetics and Biocenter Oulu, [57]Department of Oncology and [58]Department of Pathology, Oulu University Hospital, University of Oulu, Oulu, Finland, [59]Department of Obstetrics and Gynecology, Heidelberg University Hospital, University of Heidelberg, Heidelberg, Germany, [60]Section Gynecological Oncology, National Center for Tumor Diseases, University Hospital Heidelberg, Heidelberg, Germany, [61]Queensland Institute of Medical Research, Brisbane, Australia, [62]Peter MacCallum Cancer Centre, Melbourne, Australia, [63]Vesalius Research Center, VIB and [64]Leuven Multidisciplinary Breast Center, University Hospitals, K.U. Leuven, Leuven, Belgium, [65]Human Genetics Group and [66]Genetic & Molecular Epidemiology Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre [CNIO], Madrid, Spain, [67]Servicio de Oncología Médica, Hospital Universitario La Paz, Madrid, Spain, [68]Servicio de Cirugía General y Especialidades, Hospital Monte Naranco, Oviedo, Spain, [69]Division of Cancer Prevention and Genetics and [70]Department of Experimental Oncology, Istituto Europeo di Oncologia (IEO), Milan, Italy, [71]Unit of Medical Genetics, Department of Preventive and Predictive Medicine and [72]Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Italy, [73]Fondazione Istituto FIRC di Oncologia Molecolare, IFOM, Milan, Italy, [74]Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA, [75]Division of Cancer Medicine, Peter MacCallum Cancer Centre, Melbourne, Australia, and [76]Department 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 HR$_{adjusted}$ = 1.21; 95% CI: 1.09–1.35, $P = 0.0002$ and HR$_{adjusted}$ = 1.29; 95% CI: 1.12–1.47, $P = 0.0003$ for all-cause and BC-specific mortality, respectively), with no evidence of a difference in survival between TC heterozyotes and CC homozygotes. Because ER-positive and ER-negative tumors are known to have a different survival and different tumor biology, we performed Cox PH models for subgroups according to ER for the outcome variable of OS (Table 4). The estimate of the association with prognosis was greater for ER-positive than ER-negative tumors HR$_{adjusted}$ = 1.31; 95% CI: 1.13–1.50, $P = 0.0002$ and HR$_{adjusted}$ = 1.40; 95% CI: 1.15–1.70, $P = 0.001$ for all-cause and BC-specific mortality, respectively; however, the difference in the hazard ratio (HR) estimates was not statistically significant (P for SNPxER-status interaction = 0.33). Additional analyses for this SNP by subgroups defined by grade, tumor size and nodal status showed consistent associations in all subgroups (Fig. 1). One additional SNP [*LSP1* (rs3817198)] showed some evidence of an association for ER-negative disease ($P = 0.03$ test for heterogeneity); the rare CC homozygote genotype was associated with a lower mortality tumors compared with the common genotype (TT) (all-cause mortality HR$_{adjusted}$ = 0.74; 95% CI: 0.59–0.93, $P = 0.01$; BC-specific mortality HR$_{adjusted}$ = 0.74; 95% CI: 0.54–1.00, $P = 0.05$). Moreover, for all other BC susceptibility SNPs, there was no evidence of a consistent direction of worse survival in parallel with increased BC risks (survival analyses for the 11 BC susceptibility SNPs are shown comparing the BC risk allele with the non-risk allele; Tables 3 and 4).

Of the 62 candidate and GWAS-derived SNPs taken mostly from the GWAS of the Cambridge group (30) but some from other GWASs or candidates studies that had not been confirmed as being associated with BC by the BCAC, only six, i.e. rs144848, rs1318703, rs16998733, rs4666451, rs1042838 and rs2180341, showed evidence for the association with OS and/or BCS at $P < 0.05$ and none at $P < 0.01$ (Supplementary Material, Table S7). The strongest most consistent associations

were seen for a BCAC (Cambridge group) GWAS-derived SNP (28), rs1318703 (adjusted $P = 0.02$ for both overall and BCS) and a GWAS-derived SNP from a American Ashkenazi Jewish population rs2180341 (36) ($P = 0.03$ and $P = 0.01$ respectively); analyses according to the ER status showed effects at $P < 0.05$ for ER-positive tumors only (data not shown).

Analysis of gene expression on prognosis in publicly available data sets

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*.

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 HR = 1.29; 95% CI: 1.12–1.47, $P = 0.0003$ for all-cause and BC-specific mortality, respectively, after adjusting for other prognostic factors). A similar magnitude of association was observed in all subgroups defined by nodal status, tumor size, grading and ER status, although not significantly in all subgroups. Although reaching levels of significance of $P < 10^{-7}$ as for the effects of this SNP on the BC risk, the evidence for the association with survival is quite compelling given that only 11 susceptibility SNPs have been analyzed. Six of the 62 candidate/GWAS SNPs also showed a significant association with OS and/or BCS in the complete data set at a nominal $P < 0.05$, but none at $P < 0.01$ and we therefore did not consider these findings to be of significance. All these SNPs had been refuted by the BCAC to be associated with the BC risk, and we found no evidence that any of these SNPs plays a major role in BC survival in general.

A major strength of our study is the sample size. We had over 99% power to detect a risk allele of frequency 0.3 that is associated with a per-allele HR of 1.2 for OS at a type I

Table 1. Patient and tumor characteristics by 11 BC susceptibility loci[a]

SNP	Variant	N	Age at diagnosis Mean	SD	Morphology (%) IDC	ILC	Other	Unknown[a]	Tumor size (%) ≤2 cm	2–5 cm	>5 cm	Unknown[a]	Lymph node status (%) Negative	Positive	Unknown[a]
rs17468277	TT	314	54.8	10.7	**72.2**	**13.0**	**14.8**	14.0	59.4	36.7	3.9	27.1	62.9	37.1	15.0
	CT	4283	54.7	11.3	**73.6**	**14.3**	**12.1**	12.3	60.2	35.2	4.7	26.6	62.2	37.8	15.3
	CC	15 775	54.6	11.6	**75.5**	**13.4**	**11.1**	12.3	58.2	37.3	4.5	27.8	61.4	38.6	15.4
	Total	20 372													
rs1982073	TT	6761	54.1	11.6	74.9	13.7	11.4	6.8	58.6	36.8	4.6	27.4	**62.1**	**37.9**	16.2
	TC	8535	54.2	11.7	74.6	14.1	11.3	7.6	58.0	37.5	4.5	27.0	**61.8**	**38.2**	16.3
	CC	2795	53.6	11.5	75.9	13.0	11.1	9.0	58.5	36.8	4.7	29.1	**58.9**	**41.1**	17.4
	Total	18 091													
rs13281615	AA	7397	54.6	11.9	**75.7**	**13.8**	**10.6**	9.4	58.7	36.9	4.5	26.4	61.3	38.7	14.1
	AG	10 734	54.4	11.7	**74.2**	**14.4**	**11.4**	8.7	57.9	37.5	4.7	25.7	61.0	39.0	14.7
	GG	4230	54.5	11.3	**73.5**	**14.8**	**11.7**	9.3	60.3	35.3	4.4	25.3	60.0	40.0	15.2
	Total	22 361													
rs2981582	CC	6879	54.4	11.4	**76.2**	**13.3**	**10.5**	9.8	**57.2**	**38.3**	**4.5**	26.7	61.9	38.2	14.6
	CT	10 647	54.3	11.5	**74.1**	**14.5**	**11.5**	9.3	**58.5**	**36.9**	**4.6**	26.2	60.9	39.1	14.6
	TT	4167	54.3	11.5	**74.0**	**14.6**	**11.4**	10.0	**60.6**	**35.2**	**4.2**	28.2	60.3	39.7	15.3
	Total	21 693													
rs3803662	CC	10 957	54.6	11.6	75.1	13.8	11.2	11.1	58.0	37.3	4.7	27.8	61.1	38.9	14.7
	CT	9824	54.3	11.5	74.3	14.6	11.1	11.1	58.9	36.5	4.7	27.1	60.6	39.4	13.9
	TT	2369	54.1	11.8	74.5	14.9	10.6	10.9	59.2	36.9	3.9	26.0	60.5	39.5	13.0
	Total	23 150													
rs3817198	TT	9936	**54.3**	**11.5**	74.9	13.9	11.2	11.4	58.6	37.4	4.0	28.7	60.5	39.5	14.5
	TC	9195	**54.3**	**11.5**	75.0	14.0	11.0	11.7	58.0	37.1	4.9	28.9	60.5	39.5	15.1
	CC	2166	**53.7**	**11.2**	75.4	13.3	11.3	12.0	57.7	37.4	4.9	30.3	61.3	38.7	16.5
	Total	21 297													
rs889312	AA	11 146	**54.6**	**11.7**	74.8	14.2	11.0	9.0	58.7	36.7	4.6	25.6	61.3	38.7	14.2
	AC	9287	**54.5**	**11.5**	74.5	14.2	11.4	9.3	58.3	37.3	4.4	26.4	61.2	38.8	14.2
	CC	1994	**53.9**	**11.5**	73.9	14.6	11.5	10.4	58.4	36.7	4.9	26.5	61.2	38.8	15.2
	Total	22 427													
rs13387042	GG	5034	**54.1**	**11.6**	**75.7**	**14.1**	**10.3**	11.6	57.2	38.2	4.6	26.9	60.5	39.5	14.1
	GA	11 487	**54.6**	**11.7**	**75.2**	**13.8**	**11.0**	10.6	58.4	37.2	4.5	26.9	61.0	39.0	14.0
	AA	7256	**54.9**	**11.7**	**73.3**	**15.2**	**11.5**	11.0	59.3	35.9	4.8	26.5	60.8	39.2	13.8
	Total	23 777													
rs4973768	CC	5399	54.8	11.5	74.3	14.0	11.7	5.9	59.9	35.6	4.5	30.2	60.2	39.8	14.9
	CT	10 277	54.6	11.5	75.2	14.2	10.6	5.7	58.3	37.1	4.7	31.0	60.4	39.6	14.6
	TT	4930	54.5	11.5	73.9	15.1	11.0	6.3	59.3	36.8	3.9	31.6	61.3	38.7	15.6
	Total	20 606													
rs6504950	AA	1451	54.1	11.5	74.6	13.1	12.4	13.5	57.3	38.2	4.5	30.1	60.7	39.3	15.3
	GA	7801	54.2	11.7	76.8	13.2	10.0	12.7	58.7	37.1	4.2	30.7	61.5	38.5	16.0
	GG	10 807	54.3	11.7	76.8	13.6	9.7	12.7	57.0	37.9	5.1	30.4	61.9	38.1	15.5
	Total	20 059													
rs10941679	AA	10 632	53.9	11.6	75.1	14.2	10.7	12.9	39.0	26.4	3.6	31.1	60.4	39.7	16.2
	AG	8581	54.0	11.6	75.0	12.9	12.1	12.5	40.0	25.6	3.0	31.5	61.0	39.0	16.1
	GG	1547	54.3	11.6	73.5	14.3	12.2	11.0	39.0	26.7	3.1	31.1	61.3	38.7	15.9
	Total	20 760													

[a]Genotype 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%).

Table 2. Tumor characteristics by 11 BC susceptibility loci[a]

SNP	Variant	Grade (%)				ER (%)			PR (%)			HER2 (%)		
		I	II	III	Unknown[a]	Negative	Positive	Unknown[a]	Negative	Positive	Unknown[a]	Negative	Positive	Unknown[a]
rs17468277	TT	20.2	47.2	32.6	15.0	19.8	80.2	21.3	**34.7**	**65.4**	35.7	87.1	12.9	67.8
	CT	21.0	48.0	30.9	16.0	22.6	77.4	21.1	**34.1**	**65.9**	35.5	85.5	14.5	70.7
	CC	21.6	46.9	31.5	17.3	23.5	76.5	21.8	**36.5**	**63.5**	33.9	85.6	14.4	70.1
rs1982073	TT	21.5	47.8	30.7	15.7	23.0	77.0	23.3	35.8	64.2	37.8	86.5	13.5	70.2
	TC	21.3	47.7	31.0	14.9	22.6	77.4	23.1	35.3	64.7	37.4	85.6	14.4	70.0
	CC	21.7	46.9	31.4	15.3	23.8	76.2	22.1	38.7	61.3	34.6	87.3	12.8	67.4
rs13281615	AA	21.5	46.0	32.5	14.2	**24.6**	**75.5**	**20.5**	**37.2**	**62.8**	32.1	85.3	14.7	69.5
	AG	21.7	47.3	31.0	14.0	**22.7**	**77.3**	**20.2**	**36.1**	**63.9**	32.2	86.0	14.0	69.0
	GG	21.5	47.4	31.1	13.9	**22.3**	**77.7**	**21.4**	**34.9**	**65.1**	32.8	86.1	14.0	70.9
rs2981582	CC	**20.5**	**45.5**	**34.0**	16.1	**26.8**	**73.2**	20.8	**38.5**	**61.5**	33.2	85.2	14.8	70.4
	CT	**22.4**	**47.7**	**29.8**	16.2	**22.1**	**77.9**	20.9	**35.7**	**64.3**	32.8	86.6	13.4	70.4
	TT	**23.5**	**47.9**	**28.6**	17.0	**19.6**	**80.4**	21.0	**32.9**	**67.1**	32.9	87.1	13.0	70.5
rs3803662	CC	**20.5**	**46.8**	**32.7**	16.0	**24.1**	**75.9**	20.9	**37.3**	**62.7**	32.5	85.0	15.0	70.6
	CT	**22.5**	**47.2**	**30.4**	16.1	**22.5**	**77.5**	19.7	**35.0**	**65.0**	31.3	86.2	13.8	70.4
	TT	**23.1**	**46.3**	**30.6**	15.4	**21.7**	**78.3**	18.8	**33.4**	**66.6**	28.5	86.5	13.5	68.1
rs3817198	TT	21.9	46.6	31.5	16.2	23.4	76.7	19.9	36.7	63.3	31.7	**85.2**	**14.9**	70.1
	TC	22.2	45.6	32.2	16.5	22.7	77.3	21.0	35.3	64.7	33.8	**86.0**	**14.0**	71.1
	CC	23.1	47.2	29.7	17.2	24.2	75.8	21.7	37.5	62.5	34.2	**88.9**	**11.2**	72.3
rs889312	AA	21.9	47.1	31.0	15.7	23.2	76.8	20.2	36.0	64.0	31.9	86.8	13.3	71.4
	AC	21.7	46.9	31.4	16.8	22.8	77.2	21.0	35.9	64.1	32.8	85.9	14.1	71.6
	CC	21.5	48.3	30.3	17.0	22.8	77.2	22.0	34.1	65.9	33.5	85.8	14.2	71.0
rs13387042	GG	**20.7**	**46.5**	**32.9**	16.6	**23.9**	**76.1**	20.1	**36.8**	**63.3**	31.6	84.5	15.5	69.8
	GA	**20.6**	**47.1**	**32.3**	15.5	**23.4**	**76.6**	20.0	**36.6**	**63.4**	31.0	86.0	14.0	69.8
	AA	**23.3**	**47.6**	**29.0**	15.6	**22.2**	**77.8**	19.2	**34.3**	**65.7**	30.5	86.0	14.0	70.0
rs4973768	CC	**21.1**	**47.9**	**30.9**	12.5	22.9	77.2	21.0	35.7	64.3	33.7	85.1	15.0	70.4
	CT	**21.9**	**46.5**	**31.7**	12.9	22.1	77.9	21.1	35.8	64.3	34.0	85.9	14.1	70.4
	TT	**22.6**	**48.2**	**29.3**	13.0	21.4	78.6	22.2	35.5	64.5	35.2	87.3	12.7	69.6
rs6504950	AA	21.1	46.9	32.0	19.2	**25.5**	**74.5**	23.2	**37.1**	**63.0**	37.5	84.9	15.1	74.8
	GA	20.9	46.5	32.6	17.8	**22.9**	**77.1**	22.3	**35.4**	**64.6**	36.1	85.0	15.0	73.3
	GG	20.5	47.6	31.9	17.0	**22.3**	**77.8**	22.4	**34.1**	**65.9**	35.1	84.7	15.3	72.2
rs10941679	AA	**21.0**	**47.8**	**31.3**	17.6	**24.6**	**75.4**	20.0	**37.4**	**62.6**	32.0	85.3	14.7	67.8
	AG	**22.7**	**48.5**	**28.8**	18.1	**22.6**	**77.4**	20.0	**34.8**	**65.2**	33.0	86.4	13.7	67.6
	GG	**22.8**	**50.7**	**26.5**	16.4	**20.4**	**79.6**	20.2	**32.9**	**67.1**	33.7	85.7	14.3	67.0

[a]Genotype 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%).

Table 3. HR all-cause mortality and BC-specific mortality by the SNP[a]

SNP	Variant	All-cause mortality Unadjusted HR (95% CI)	*P*-value**	Adjusted HR (95% CI)	*P*-value**	Breast cancer-specific mortality Unadjusted HR (95% CI)	*P*-value**	Adjusted HR (95% CI)	*P*-value**
rs17468277	TT	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	CT	0.99 (0.73–1.32)		0.97 (0.72–1.30)		1.12 (0.75–1.67)		1.11 (0.74–1.67)	
	CC	1.01 (0.76–1.35)		0.99 (0.74–1.31)		1.13 (0.76–1.67)		1.11 (0.75–1.65)	
			0.83		0.92		0.83		0.87
rs1982073	TT	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	TC	1.06 (0.97–1.14)		1.05 (0.97–1.14)		1.08 (0.97–1.20)		1.05 (0.95–1.17)	
	CC	1.13 (1.01–1.26)		1.11 (1.00–1.24)		1.09 (0.94–1.26)		1.04 (0.90–1.21)	
			0.09		0.14		0.30		0.62
rs13281615	AA	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	AG	1.00 (0.93–1.07)		1.00 (0.93–1.07)		0.99 (0.90–1.10)		1.00 (0.91–1.10)	
	GG	1.00 (0.91–1.10)		1.01 (0.92–1.11)		0.95 (0.83–1.08)		0.98 (0.86–1.11)	
			1.00		0.95		0.67		0.93
rs2981582	CC	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	CT	1.01 (0.94–1.09)		1.05 (0.97–1.13)		1.06 (0.96–1.17)		1.09 (0.99–1.21)	
	TT	0.94 (0.85–1.03)		0.99 (0.90–1.09)		0.98 (0.86–1.11)		1.03 (0.90–1.17)	
			0.22		0.33		0.30		0.21
rs3803662	CC	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	CT	0.94 (0.88–1.01)		0.97 (0.91–1.04)		0.99 (0.90–1.08)		1.01 (0.92–1.11)	
	TT	**1.17 (1.05–1.30)**		**1.21 (1.09–1.35)**		**1.23 (1.07–1.41)**		**1.29 (1.12–1.47)**	
			0.0004		**0.0002**		**0.006**		**0.0009**
rs3817198	TT	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	TC	1.01 (0.94–1.08)		0.99 (0.93–1.07)		0.98 (0.89–1.07)		0.98 (0.89–1.07)	
	CC	0.91 (0.81–1.03)		0.92 (0.81–1.04)		0.97 (0.83–1.13)		0.97 (0.83–1.13)	
			0.28		0.41		0.88		0.84
rs889312	AA	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	AC	0.99 (0.93–1.06)		0.98 (0.92–1.05)		1.01 (0.92–1.10)		0.99 (0.91–1.09)	
	CC	0.97 (0.86–1.09)		0.97 (0.86–1.10)		1.05 (0.90–1.23)		1.04 (0.89–1.22)	
			0.87		0.80		0.81		0.85
rs13387042	GG	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	GA	1.02 (0.94–1.11)		1.04 (0.96–1.13)		1.02 (0.92–1.14)		1.05 (0.94–1.17)	
	AA	0.95 (0.86–1.04)		0.98 (0.89–1.07)		0.91 (0.80–1.03)		0.94 (0.83–1.06)	
			0.14		0.20		0.06		0.10
rs4973768	CC	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	CT	0.98 (0.91–1.07)		1.00 (0.92–1.08)		1.02 (0.91–1.14)		1.04 (0.93–1.16)	
	TT	0.97 (0.88–1.07)		1.01 (0.91–1.11)		0.96 (0.84–1.09)		1.01 (0.88–1.15)	
			0.85		0.98		0.55		0.75
rs6504950	AA	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	GA	1.01 (0.88–1.16)		1.00 (0.87–1.15)		0.97 (0.81–1.16)		0.93 (0.78–1.12)	
	GG	0.93 (0.81–1.06)		0.94 (0.82–1.08)		0.90 (0.75–1.07)		0.88 (0.73–1.05)	
			0.06		0.23		0.21		0.26
rs10941679	AA	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	AG	0.99 (0.92–1.06)		1.01 (0.94–1.08)		0.94 (0.85–1.04)		0.97 (0.88–1.07)	
	GG	1.03 (0.90–1.18)		1.04 (0.91–1.19)		1.00 (0.84–1.20)		1.04 (0.87–1.25)	
			0.83		0.82		0.47		0.71

[a]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).

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

Table 4. HR all-cause mortality by the SNP for ER-positive and ER-negative tumors separately[a]

SNP	Variant	ER-negative Unadjusted HR (95% CI)	*P*-value**	Adjusted HR (95% CI)	*P*-value**	ER-positive Adjusted HR (95% CI)	*P*-value**	Adjusted HR (95% CI)	*P*-value**
rs17468277	TT	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	CT	0.86 (0.50–1.56)		0.94 (0.54–1.67)		0.99 (0.66–1.48)		0.92 (0.61–1.38)	
	CC	0.95 (0.55–1.64)		0.98 (0.57–1.71)		1.03 (0.69–1.53)		0.98 (0.66–1.46)	
			0.69		0.89		0.79		0.59
rs1982073	TT	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	TC	1.11 (0.95–1.30)		1.07 (0.91–1.25)		0.99 (0.88–1.11)		1.00 (0.89–1.12)	
	CC	1.23 (1.00–1.51)		1.15 (0.94–1.42)		0.97 (0.83–1.14)		0.98 (0.83–1.15)	
			0.14		0.39		0.93		0.95
rs13281615	AA	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	AG	1.06 (0.92–1.21)		1.07 (0.93–1.22)		0.96 (0.87–1.07)		0.94 (0.85–1.04)	
	GG	1.00 (0.83–1.20)		1.01 (0.84–1.22)		0.99 (0.87–1.13)		1.00 (0.87–1.14)	
			0.69		0.66		0.74		0.44
rs2981582	CC	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	CT	1.02 (0.89–1.17)		1.00 (0.88–1.16)		1.05 (0.94–1.17)		1.07 (0.96–1.19)	
	TT	0.99 (0.82–1.20)		0.97 (0.80–1.17)		0.98 (0.86–1.13)		1.03 (0.90–1.18)	
			0.93		0.92		0.48		0.46
rs3803662	CC	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	CT	0.98 (0.86–1.12)		0.97 (0.85–1.10)		0.94 (0.85–1.04)		0.98 (0.89–1.08)	
	TT	1.14 (0.93–1.40)		1.11 (0.90–1.36)		**1.23 (1.07–1.42)**		**1.31 (1.13–1.50)**	
			0.37		0.45		**0.001**		**0.0002**
rs3817198	TT	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	TC	0.99 (0.87–1.13)		0.92 (0.81–1.05)		1.04 (0.95–1.15)		1.04 (0.94–1.15)	
	CC	0.78 (0.61–0.98)		**0.74 (0.59–0.93)**		0.99 (0.84–1.18)		1.02 (0.86–1.21)	
			0.09		**0.03**		0.66		0.72
rs889312	AA	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	AC	1.10 (0.97–1.25)		1.09 (0.95–1.24)		0.95 (0.86–1.04)		0.93 (0.84–1.02)	
	CC	1.15 (0.93–1.44)		1.07 (0.86–1.33)		0.84 (0.71–1.01)		0.86 (0.72–1.03)	
			0.24		0.44		0.13		0.12
rs13387042	GG	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	GA	1.02 (0.88–1.19)		1.03 (0.88–1.20)		1.07 (0.95–1.20)		1.10 (0.98–1.23)	
	AA	0.87 (0.73–1.04)		0.89 (0.75–1.06)		1.00 (0.88–1.14)		1.03 (0.91–1.17)	
			0.08		0.17		0.37		0.22
rs4973768	CC	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	CT	0.95 (0.81–1.11)		0.98 (0.84–1.15)		1.02 (0.91–1.15)		1.03 (0.91–1.15)	
	TT	0.90 (0.74–1.09)		0.92 (0.76–1.12)		1.00 (0.87–1.15)		1.03 (0.90–1.18)	
			0.56		0.69		0.93		0.91
rs6504950	AA	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	GA	1.17 (0.90–1.52)		1.15 (0.88–1.49)		1.05 (0.86–1.28)		1.02 (0.83–1.24)	
	GG	1.04 (0.80–1.35)		1.04 (0.80–1.35)		0.98 (0.80–1.19)		0.99 (0.81–1.20)	
			0.19		0.31		0.42		0.86
rs10941679	AA	1 (reference)		1 (reference)		1 (reference)		1 (reference)	
	AG	1.09 (0.95–1.25)		1.09 (0.95–1.25)		1.01 (0.91–1.12)		1.04 (0.94–1.15)	
	GG	1.00 (0.76–1.31)		0.97 (0.74–1.28)		1.16 (0.98–1.39)		1.14 (0.95–1.36)	
			0.42		0.40		0.23		0.33

[a]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).

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.
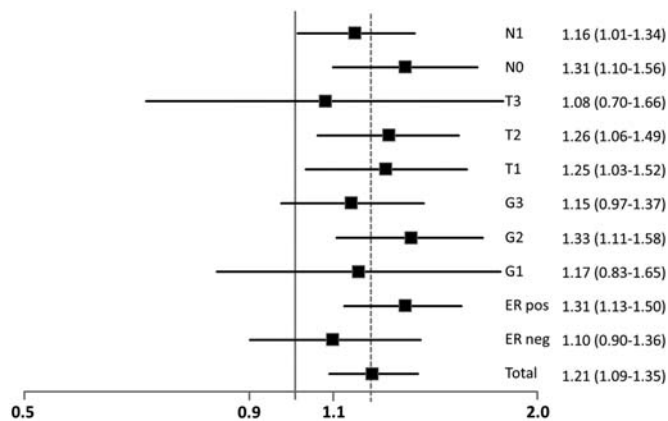
Figure 1. Adjusted HR (all-cause mortality) for the homozygous rare genotype of rs3803662 within clinical subgroups. Adjusted HR (95% CI) for the homozygous rare genotype of rs3803662 within clinical subgroups: N1, lymph node positive; N0, lymph node negative; T1, tumor size <2 cm; T2, tumor size 2–5 cm; T3, tumor size >5 cm; G1–3, differentiation grades 1–3; ER pos, estrogen receptor status positive; ER neg, estrogen receptor status negative. Analyses were performed with the common homozygous as reference group for each clinical subgroup and included age at diagnosis, tumor size, nodal status and grade as co-variants (excluding the covariant of the clinical subgroup of interest) and were stratified by study.

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 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 *STXBP4* [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 *FGF10* 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 *TGFB1* (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 (*TGFB1*), 22 studies for rs2981582 (*FGFR2*), 22 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 ($r^2 = 1$ 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 HG-U133A 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 $\log_2$ 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 nodes 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 category; for the 62 other SNPs, the common allele was used as the reference. All *P*-values reported are from two-sided tests. All analyses were performed using STATA version 11.1 (StataCorp, Texas, USA).

For the survival analysis of the expression data, multivariate Cox models were fitted—including age at diagnosis, tumor size, lymph node status, grade and hormone receptor status as covariates. The disease-free survival was used as an endpoint and survival time was right censored at 10 years. After fitting the models for the complete group of pooled BC patients, models for ER-positive and ER-negative patients were analyzed separately. All analyses for the expression data analysis were performed using R (http://www.r-project.org/) and SPSS version 17.0 (SPSS Inc., Chicago, IL, USA).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

REFERENCES

1. Sotiriou, C. and Pusztai, L. (2009) Gene-expression signatures in breast cancer. *N. Engl. J. Med.*, **360**, 790–800.

2. Cardoso, F., Van't Veer, L., Rutgers, E., Loi, S., Mook, S. and Piccart-Gebhart, M.J. (2008) Clinical application of the 70-gene profile: the MINDACT trial. *J. Clin. Oncol.*, **26**, 729–735.

3. Zujewski, J.A. and Kamin, L. (2008) Trial assessing individualized options for treatment for breast cancer: the TAILORx trial. *Future Oncol.*, **4**, 603–610.

4. Liedtke, C., Wolf, M.K. and Kiesel, L. (2010) New concepts for targeted systemic therapy in breast cancer. *Geburtsh. Frauenheilk.*, **70**, 625–633.

5. Harbeck, N., Salem, M., Nitz, U., Gluz, O. and Liedtke, C. (2010) Personalized treatment of early-stage breast cancer: present concepts and future directions. *Cancer Treat. Rev.*, **36**, 584–594.

6. Fehm, T., Janni, W., Kummel, S., Luftner, D., Lux, M.P., Schutz, F. and Fasching, P.A. (2011) Review: SABCS 2010-current treatment options for patients with breast cancer. *Geburtsh. Frauenheilk.*, **71**, 260–276.

7. Lindstrom, L.S., Hall, P., Hartman, M., Wiklund, F., Gronberg, H. and Czene, K. (2007) Familial concordance in cancer survival: a Swedish population-based study. *Lancet Oncol.*, **8**, 1001–1006.

8. Weinshilboum, R.M. and Wang, L. (2006) Pharmacogenetics and pharmacogenomics: development, science, and translation. *Annu. Rev. Genomics Hum. Genet.*, **7**, 223–245.

9. Wang, L., McLeod, H.L. and Weinshilboum, R.M. (2011) Genomics and drug response. *N. Engl. J. Med.*, **364**, 1144–1153.

10. Schroth, W., Goetz, M.P., Hamann, U., Fasching, P.A., Schmidt, M., Winter, S., Fritz, P., Simon, W., Suman, V.J., Ames, M.M. *et al.* (2009) Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. *JAMA*, **302**, 1429–1436.

11. Lash, T.L., Cronin-Fenton, D., Ahern, T.P., Rosenberg, C.L., Lunetta, K.L., Silliman, R.A., Garne, J.P., Sorensen, H.T., Hellberg, Y., Christensen, M. *et al.* (2011) CYP2D6 inhibition and breast cancer recurrence in a population-based study in Denmark. *J. Natl Cancer Inst.*, **103**, 489–500.

12. Abraham, J.E., Maranian, M.J., Driver, K.E., Platte, R., Kalmyrzaev, B., Baynes, C., Luccarini, C., Earl, H.M., Dunning, A.M., Pharoah, P.D. *et al.* (2011) CYP2D6 gene variants and their association with breast cancer susceptibility. *Cancer Epidemiol. Biomarkers Prev.*, **20**, 1255–1258.

13. Fasching, P.A., Fehm, T., Janni, W., Kummel, S., Luftner, D., Lux, M.P. and Maass, N. (2010) Breast cancer therapy—a state of the art review. *Geburtsh. Frauenheilk.*, **70**, 875–886.

14. Fagerholm, R., Hofstetter, B., Tommiska, J., Aaltonen, K., Vrtel, R., Syrjakoski, K., Kallioniemi, A., Kilpivaara, O., Mannermaa, A., Kosma, V.M. *et al.* (2008) NAD(P)H:quinone oxidoreductase 1 NQO1*2 genotype (P187S) is a strong prognostic and predictive factor in breast cancer. *Nat. Genet.*, **40**, 844–853.

15. Garcia-Closas, M., Hall, P., Nevanlinna, H., Pooley, K., Morrison, J., Richesson, D.A., Bojesen, S.E., Nordestgaard, B.G., Axelsson, C.K., Arias, J.I. *et al.* (2008) Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. *PLoS Genet.*, **4**, e1000054.

16. Broeks, A., Schmidt, M.K., Sherman, M.E., Couch, F.J., Hopper, J.L., Dite, G.S., Apicella, C., Smith, L.D., Hammet, F., Southey, M.C. *et al.* (2011) Low penetrance breast cancer susceptibility loci are associated with specific breast tumor subtypes: findings from the Breast Cancer Association Consortium. *Hum. Mol. Genet.*, **20**, 3289–3303.

17. Reeves, G.K., Travis, R.C., Green, J., Bull, D., Tipper, S., Baker, K., Beral, V., Peto, R., Bell, J., Zelenika, D. *et al.* (2010) Incidence of breast cancer and its subtypes in relation to individual and multiple low-penetrance genetic susceptibility loci. *JAMA*, **304**, 426–434.

18. Yao, S., Barlow, W.E., Albain, K.S., Choi, J.Y., Zhao, H., Livingston, R.B., Davis, W., Rae, J.M., Yeh, I.T., Hutchins, L.F. *et al.* (2010) Manganese superoxide dismutase polymorphism, treatment-related toxicity and disease-free survival in SWOG 8897 clinical trial for breast cancer. *Breast Cancer Res. Treat.*, **124**, 433–439.

19. Yu, C.P., Yu, J.C., Sun, C.A., Tzao, C., Ho, J.Y. and Yen, A.M. (2008) Tumor susceptibility and prognosis of breast cancer associated with the G870A polymorphism of CCND1. *Breast Cancer Res. Treat.*, **107**, 95–102.

20. Udler, M.S., Azzato, E.M., Healey, C.S., Ahmed, S., Pooley, K.A., Greenberg, D., Shah, M., Teschendorff, A.E., Caldas, C., Dunning, A.M. *et al.* (2009) Common germline polymorphisms in COMT, CYP19A1, ESR1, PGR, SULT1E1 and STS and survival after a diagnosis of breast cancer. *Int. J. Cancer*, **125**, 2687–2696.

21. Abraham, J.E., Harrington, P., Driver, K.E., Tyrer, J., Easton, D.F., Dunning, A.M. and Pharoah, P.D. (2009) Common polymorphisms in the prostaglandin pathway genes and their association with breast cancer susceptibility and survival. *Clin. Cancer Res.*, **15**, 2181–2191.

22. Fasching, P.A., Loehberg, C.R., Strissel, P.L., Lux, M.P., Bani, M.R., Schrauder, M., Geiler, S., Ringleff, K., Oeser, S., Weihbrecht, S. *et al.* (2008) Single nucleotide polymorphisms of the aromatase gene (CYP19A1), HER2/neu status, and prognosis in breast cancer patients. *Breast Cancer Res. Treat.*, **112**, 89–98.

23. Azzato, E.M., Lee, A.J., Teschendorff, A., Ponder, B.A., Pharoah, P., Caldas, C. and Maia, A.T. (2010) Common germ-line polymorphism of C1QA and breast cancer survival. *Br. J. Cancer*, **102**, 1294–1299.

24. Azzato, E.M., Pharoah, P.D., Harrington, P., Easton, D.F., Greenberg, D., Caporaso, N.E., Chanock, S.J., Hoover, R.N., Thomas, G., Hunter, D.J. *et al.* (2010) A genome-wide association study of prognosis in breast cancer. *Cancer Epidemiol. Biomarkers Prev.*, **19**, 1140–1143.

25. Azzato, E.M., Tyrer, J., Fasching, P.A., Beckmann, M.W., Ekici, A.B., Schulz-Wendtland, R., Bojesen, S.E., Nordestgaard, B.G., Flyger, H., Milne, R.L. *et al.* (2010) Association between a germline OCA2 polymorphism at chromosome 15q13.1 and estrogen receptor-negative breast cancer survival. *J. Natl Cancer Inst.*, **102**, 650–662.

26. Chlebowski, R.T., Anderson, G.L., Gass, M., Lane, D.S., Aragaki, A.K., Kuller, L.H., Manson, J.E., Stefanick, M.L., Ockene, J., Sarto, G.E. *et al.* (2010) Estrogen plus progestin and breast cancer incidence and mortality in postmenopausal women. *JAMA*, **304**, 1684–1692.

27. Barnett, G.C., Shah, M., Redman, K., Easton, D.F., Ponder, B.A. and Pharoah, P.D. (2008) Risk factors for the incidence of breast cancer: do they affect survival from the disease? *J. Clin. Oncol.*, **26**, 3310–3316.

28. Breast Cancer Association Consortium. (2006) Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. *J. Natl Cancer Inst.*, **98**, 1382–1396.

29. Cox, A., Dunning, A.M., Garcia-Closas, M., Balasubramanian, S., Reed, M.W., Pooley, K.A., Scollen, S., Baynes, C., Ponder, B.A., Chanock, S. *et al.* (2007) A common coding variant in CASP8 is associated with breast cancer risk. *Nat. Genet.*, **39**, 352–358.

30. Easton, D.F., Pooley, K.A., Dunning, A.M., Pharoah, P.D., Thompson, D., Ballinger, D.G., Struewing, J.P., Morrison, J., Field, H., Luben, R. *et al.* (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature*, **447**, 1087–1093.

31. Milne, R.L., Benitez, J., Nevanlinna, H., Heikkinen, T., Aittomaki, K., Blomqvist, C., Arias, J.I., Zamora, M.P., Burwinkel, B., Bartram, C.R. *et al.* (2009) Risk of estrogen receptor-positive and -negative breast cancer and single-nucleotide polymorphism 2q35-rs13387042. *J. Natl Cancer Inst.*, **101**, 1012–1018.

32. Ahmed, S., Thomas, G., Ghoussaini, M., Healey, C.S., Humphreys, M.K., Platte, R., Morrison, J., Maranian, M., Pooley, K.A., Luben, R. *et al.* (2009) Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat. Genet.*, **41**, 585–590.

33. Stacey, S.N., Manolescu, A., Sulem, P., Thorlacius, S., Gudjonsson, S.A., Jonsson, G.F., Jakobsdottir, M., Bergthorsson, J.T., Gudmundsson, J., Aben, K.K. *et al.* (2008) Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.*, **40**, 703–706.

34. CHEK2 Breast Cancer Case-Control Consortium. (2004) CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10 860 breast cancer cases and 9065 controls from 10 studies. *Am. J. Hum. Genet.*, **74**, 1175–1182.

35. Schmidt, M.K., Tollenaar, R.A., de Kemp, S.R., Broeks, A., Cornelisse, C.J., Smit, V.T., Peterse, J.L., van Leeuwen, F.E. and Van't Veer, L.J. (2007) Breast cancer survival and tumor characteristics in premenopausal women carrying the CHEK2*1100delC germline mutation. *J. Clin. Oncol.*, **25**, 64–69.

36. Gold, B., Kirchhoff, T., Stefanov, S., Lautenberger, J., Viale, A., Garber, J., Friedman, E., Narod, S., Olshen, A.B., Gregersen, P. *et al.* (2008) Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. *Proc. Natl Acad. Sci. USA*, **105**, 4340–4345.

37. Fiorentino, F.P., Symonds, C.E., Macaluso, M. and Giordano, A. (2009) Senescence and p130/Rbl2: a new beginning to the end. *Cell Res.*, **19**, 1044–1051.

38. Azzato, E.M., Greenberg, D., Shah, M., Blows, F., Driver, K.E., Caporaso, N.E. and Pharoah, P.D. (2009) Prevalent cases in observational studies of cancer survival: do they bias hazard ratio estimates? *Br. J. Cancer*, **100**, 1806–1811.

39. Travis, R.C., Reeves, G.K., Green, J., Bull, D., Tipper, S.J., Baker, K., Beral, V., Peto, R., Bell, J., Zelenika, D. *et al.* (2010) Gene-environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study. *Lancet*, **375**, 2143–2151.

40. Zheng, W., Wen, W., Gao, Y.T., Shyr, Y., Zheng, Y., Long, J., Li, G., Li, C., Gu, K., Cai, Q. *et al.* (2010) Genetic and clinical predictors for breast cancer risk assessment and stratification among Chinese women. *J. Natl Cancer Inst.*, **102**, 972–981.

41. Udler, M.S., Ahmed, S., Healey, C.S., Meyer, K., Struewing, J., Maranian, M., Kwon, E.M., Zhang, J., Tyrer, J., Karlins, E. *et al.* (2010) Fine scale mapping of the breast cancer 16q12 locus. *Hum. Mol. Genet.*, **19**, 2507–2515.

42. Jowett, J.B., Curran, J.E., Johnson, M.P., Carless, M.A., Goring, H.H., Dyer, T.D., Cole, S.A., Comuzzie, A.G., MacCluer, J.W., Moses, E.K. *et al.* (2010) Genetic variation at the FTO locus influences RBL2 gene expression. *Diabetes*, **59**, 726–732.

43. O'Flaherty, E. and Kaye, J. (2003) TOX defines a conserved subfamily of HMG-box proteins. *BMC Genomics*, **4**, 13.

44. Smid, M., Wang, Y., Klijn, J.G., Sieuwerts, A.M., Zhang, Y., Atkins, D., Martens, J.W. and Foekens, J.A. (2006) Genes associated with breast cancer metastatic to bone. *J. Clin. Oncol.*, **24**, 2261–2267.

45. Yuan, S.H., Qiu, Z. and Ghosh, A. (2009) TOX3 regulates calcium-dependent transcription in neurons. *Proc. Natl Acad. Sci. USA*, **106**, 2909–2914.

46. Stacey, S.N., Manolescu, A., Sulem, P., Rafnar, T., Gudmundsson, J., Gudjonsson, S.A., Masson, G., Jakobsdottir, M., Thorlacius, S., Helgason, A. *et al.* (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.*, **39**, 865–869.

47. Milne, R.L., Goode, E.L., García-Closas, M., Couch, F.J., Severi, G., Hein, R., Fredericksen, Z., Malats, N., Zamora, M.P., Pérez, J.I.A. *et al.* (2011) Confirmation of 5p12 as a susceptibility locus for progesterone-receptor-positive, lower grade breast cancer. *Cancer Epidemiol. Biomarkers Prev.*, **20**, 2222–2231.

48. Gaudet, M.M., Milne, R.L., Cox, A., Camp, N.J., Goode, E.L., Humphreys, M.K., Dunning, A.M., Morrison, J., Giles, G.G., Severi, G. *et al.* (2009) Five polymorphisms and breast cancer risk: results from the Breast Cancer Association Consortium. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 1610–1616.

49. Pooley, K.A., Healey, C.S., Smith, P.L., Pharoah, P.D., Thompson, D., Tee, L., West, J., Jordan, C., Easton, D.F., Ponder, B.A. *et al.* (2006) Association of the progesterone receptor gene with breast cancer risk: a single-nucleotide polymorphism tagging approach. *Cancer Epidemiol. Biomarkers Prev.*, **15**, 675–682.

50. Johnatty, S.E., Spurdle, A.B., Beesley, J., Chen, X., Hopper, J.L., Duffy, D.L. and Chenevix-Trench, G. (2008) Progesterone receptor polymorphisms and risk of breast cancer: results from two Australian breast cancer studies. *Breast Cancer Res. Treat.*, **109**, 91–99.

51. Toyama, T., Zhang, Z., Nishio, M., Hamaguchi, M., Kondo, N., Iwase, H., Iwata, H., Takahashi, S., Yamashita, H. and Fujii, Y. (2007) Association of TP53 codon 72 polymorphism and the outcome of adjuvant therapy in breast cancer patients. *Breast Cancer Res.*, **9**, R34.

52. Ramus, S.J., Vierkant, R.A., Johnatty, S.E., Pike, M.C., Van Den Berg, D.J., Wu, A.H., Pearce, C.L., Menon, U., Gentry-Maharaj, A., Gayther, S.A. *et al.* (2008) Consortium analysis of 7 candidate SNPs for ovarian cancer. *Int. J. Cancer*, **123**, 380–388.

53. Loizidou, M.A., Michael, T., Neuhausen, S.L., Newbold, R.F., Marcou, Y., Kakouri, E., Daniel, M., Papadopoulos, P., Malas, S., Kyriacou, K. *et al.* (2008) Genetic polymorphisms in the DNA repair genes XRCC1, XRCC2 and XRCC3 and risk of breast cancer in Cyprus. *Breast Cancer Res. Treat.*, **112**, 575–579.

54. Milne, R.L., Ribas, G., Gonzalez-Neira, A., Fagerholm, R., Salas, A., Gonzalez, E., Dopazo, J., Nevanlinna, H., Robledo, M. and Benitez, J. (2006) ERCC4 associated with breast cancer risk: a two-stage case-control study using high-throughput genotyping. *Cancer Res.*, **66**, 9420–9427.

55. Edvardsen, H., Tefre, T., Jansen, L., Vu, P., Haffty, B.G., Fossa, S.D., Kristensen, V.N. and Borresen-Dale, A.L. (2007) Linkage disequilibrium pattern of the ATM gene in breast cancer patients and controls; association of SNPs and haplotypes to radio-sensitivity and post-lumpectomy local recurrence. *Radiat. Oncol.*, **2**, 25.

56. Ye, C., Dai, Q., Lu, W., Cai, Q., Zheng, Y., Shu, X.O., Gu, K., Gao, Y.T. and Zheng, W. (2007) Two-stage case-control study of common ATM gene variants in relation to breast cancer risk. *Breast Cancer Res. Treat.*, **106**, 121–126.

57. Kirchhoff, T., Chen, Z.Q., Gold, B., Pal, P., Gaudet, M.M., Kosarin, K., Levine, D.A., Gregersen, P., Spencer, S., Harlan, M. *et al.* (2009) The 6q22.33 locus and breast cancer susceptibility. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 2468–2475.

58. Karn, T., Metzler, D., Ruckhaberle, E., Hanker, L., Gatje, R., Solbach, C., Ahr, A., Schmidt, M., Holtrich, U., Kaufmann, M. *et al.* (2010) Data-driven derivation of cutoffs from a pool of 3030 Affymetrix arrays to stratify distinct clinical types of breast cancer. *Breast Cancer Res. Treat.*, **120**, 567–579.

59. Affymetrix. (2001) Statistical algorithms reference guide, technical report.

60. Gautier, L., Cope, L., Bolstad, B.M. and Irizarry, R.A. (2004) affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*, **20**, 307–315.

61. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.*, **5**, R80.