Ghoussaini, Maya; Fletcher, Olivia; Michaillidou, Kyriaki; Turnbull, Clare; Schmidt, Marjanka K; Dicks, Ed; Dennis, Joe; Wang, Qin; Humphreys, Manjeet K; Luccarini, Craig; +177 more... Baynes, Caroline; Conroy, Don; Maranian, Melanie; Ahmed, Shahana; Driver, Kristy; Johnson, Nichola; Orr, Nicholas; dos Santos Silva, Isabel; Waifisz, Quinten; Meijers-Heijboer, Hanne; Uitterlinden, Andre G; Rivadeneira, Fernando; Netherlands Collaborative Group on Hereditary Breast and Ovarian; Hall, Per; Czene, Kamila; Irwanto, Astrid; Liu, Jianjun; Nevanlinna, Heli; Aittomäki, Kristiina; Blomqvist, Carl; Meindl, Alfons; Schmutzli, Ritz K; Müller-Myhsk, Bertram; Lichtner, Peter; Chang-Claude, Jenny; Hein, Rebecca; Nickel, Stefan; Flesch-Janys, Dieter; Tsimiklis, Helen; Makalic, Enes; Schmidt, Daniel; Bui, Minh; Hopper, John L; Apicella, Carmel; Park, Daniel J; Southey, Melissa; Hunter, David J; Chanock, Stephen J; Brooks, Annegien; Verhoeof, Semen; Hogervorst, Frans BL; Fasching, Peter A; Lux, Michael P; Beckmann, Matthias W; Ekici, Arif B; Sawyer, Elinor; Tomlinson, Ian; Kerin, Michael; Marne, Frederik; Schneeweiss, Andreas; Sohn, Christof; Burwinkel, Barbara; Guénél, Pascal; Truong, Thérèse; Cordina-Duverger, Emilie; Menegaux, Florence; Bojesen, Stig E; Nordestaard, Børge G; Nielsen, Sune F; Flyger, Henrik; Milne, Roger L; Alonso, M Rosario; González-Neira, Anna; Benitez, Javier; Anton-Culver, Hoda; Ziegos, Argyrios; Bernstein, Leslie; Dur, Christina Clarke; Brenner, Hermann; Müller, Heiko; Arndt, Volker; Stegmaier, Christa; Familial Breast Cancer Study (FBCS); Justen, Christina; Brauch, Hiltrud; Brümmer, Thomas; Gene Environment Interaction of Breast Cancer in Germany (GENICA); Wang-Gohrke, Shan; Eilber, Ursula; Dörk, Thilo; Schürmann, Peter; Bremer, Michael; Hillemanns, Peter; Bogdanova, Natalia V; Antonenkova, Natalia N; Rogov, Yuri I; Karstens, Johann H; Bermissheva, Marina; Prokofieva, Darya; Khusnutdinova, Elza; Lindblom, Annika; Margolin, Sara; Mannermaa, Arto; Kataja, Vesa; Kosma, Veli-Matti; Hartikainen, Jaana M; Lambrechts, Dieter; Yesilyurt, Betul T; Floris, Giuseppe; Leumen, Karin; Manoukian, Siranoush; Bonanni, Bernardo; Fortuzzi, Stefano; Peterlongo, Paolo; Couch, Fergus J; Wang, Xiaohu; Stevens, Kristen; Lee, Adam; Giles, Graham G; Baglietto, Laura; Severi, Gianluca; McLean, Catriona; Alnaes, Grethe Grenaker; Kristensen, Vessel; Borrensen-Dale, Anne-Lise; John, Esther M; Miron, Alexander; Winqvist, Robert; Pylkäs, Katri; Jukkola-Vuorinen, Arja; Kaupilla, Saila; Andruslis, Irene L; Glendon, Gerd; Molligan, Anna Marie; Devilee, Peter; van Asperen, Christie J; Tollenaar, Rob AEM; Seynaeve, Caroline; Figueroa, Jonine D; Garcia-Closas, Montserrat; Brinton, Louise; Lissowska, Jolanta; Hooning, Maartje J; Hollestelle, Antoinette; Oldenburg, Rogier A; van den Ouweland, Ans MW; Cox, Angela; Reed, Malcolm WR; Shah, Mitul; Jakubowska, Ania; Lubinski, Jan; Jaworska, Katarzyna; Durda, Katarzyna; Jones, Michael; Schoemaker, Minouk; Ashworth, Alan; Swerdlow, Anthony; Beesley, Jonathan; Chen, Xiaoping; ConFab Investigators; Australian Ovarian Cancer Study Group; Muir, Kenneth R; Lopatahanan, Arrtaita; Rattanamongkon, Suthee; Chaiwerawattana, Arkom; Kang, Dehee; Yoo, Koon-Young; Noh, Dong-Young; Shen, Chen-Yang; Wu, Jyh-Cherng; Wu, Pei-Ei; Hsiung, Chia-Ni; Perkins, Annie; Swann, Ruth; Velentzis, Louiza; Eccles, Diana M; Tapper, Will J; Gerty, Susan M; Graham, Nikki J; Ponder, Bruce AJ; Chenevix-Trench, Georgia; Pharoah,
Paul DP; Lathrop, Mark; Dunning, Alison M; Rahman, Nazneen; Peto, Julian; Easton, Douglas F; (2012) Genome-wide association analysis identifies three new breast cancer susceptibility loci. Nature genetics, 44 (3). pp. 312-318. ISSN 1061-4036 DOI: https://doi.org/10.1038/ng.1049

Downloaded from: http://researchonline.lshtm.ac.uk/id/eprint/57728/

DOI: https://doi.org/10.1038/ng.1049

Usage Guidelines:

Please refer to usage guidelines at https://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by-nc-nd/2.5/
Genome-wide association analysis identifies three new breast cancer susceptibility loci


Conflicts of interest The authors have no conflict of interest.


Published in final edited form as: Nat Genet. 44(3): 312–318. doi:10.1038/ng.1049.

1Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK 2Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, London, UK 3Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK 4Section of Cancer Genetics, Institute of Cancer Research, Sutton, UK 5Department of Experimental Therapy, Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands 6Department of Epidemiology, Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands 7Non-communicable Disease Epidemiology Department, London School of Hygiene and Tropical Medicine, London, UK 8Department of Clinical Genetics, VU University Medical Center, section Oncogenetics, Amsterdam, The Netherlands 9Department of Internal Medicine and Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands 10Netherlands Collaborative Group on Hereditary Breast and Ovarian Cancer 11Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm 17177, Sweden 12Human Genetics Division, Genome Institute of Singapore, Singapore 138672, Singapore 13Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland 14Department of Clinical Genetics, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland 15Department of Oncology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland 16Clinic of Gynaecology and Obstetrics, Division for Gynaecological Tumor-Genetics, Technische Universität München, München, Germany 17Department of Obstetrics and Gynaecology, Division of Molecular Gynaeco-Oncology, University of Cologne, Germany 18Max Planck Institute of Psychiatry, Munich, Germany 19Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany 20Division of Cancer Epidemiology, Deutsches Krebsforschungszentrum, Heidelberg, Germany 21PMV Research Group at the Department of Child and Adolescent Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany 22Department of Cancer Epidemiology/Clinical Cancer Registry, University Clinic Hamburg-Eppendorf, Hamburg, Germany 23Institute for Medical Biometrics and...
Epidemiology, University Clinic Hamburg-Eppendorf, Hamburg, Germany

Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Australia

Centre for Molecular, Environmental, Genetic, and Analytic Epidemiology, Melbourne School of Population Health, The University of Melbourne, Australia

Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA

Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland, USA

Department of Clinical Genetics, Family Cancer Clinic, Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands

Department of Molecular Pathology, Family Cancer Clinic, Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands

University Breast Center, Department of Gynecology and Obstetrics, University Hospital Erlangen, Erlangen, Germany

Institute of Human Genetics, Friedrich-Alexander University Erlangen Nuremberg, Erlangen, Germany

Division of Cancer Studies, NIHR Comprehensive Biomedical Research Centre, Guy’s & St. Thomas’ NHS Foundation Trust in partnership with King’s College London, London, UK

Welcome Trust Centre for Human Genetics and Oxford Biomedical Research Centre, University of Oxford, UK

Clinical Science Institute, University Hospital Galway, Galway, Ireland

Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg, Germany

National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany

Molecular Epidemiology Unit, German Cancer Research Center (DKFZ), Heidelberg, Germany

Inserm, CESP Centre for research in Epidemiology and Population Health, U1018, Environmental epidemiology of cancer, Villejuif, France

Université Paris-Sud, UMRS 1018, Villejuif, France

Copenhagen General Population Study, Herlev Hospital, Copenhagen University Hospital, Copenhagen, Denmark

Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, Copenhagen, Denmark

Department of Breast Surgery, Herlev Hospital, Copenhagen University Hospital, Copenhagen, Denmark

Genetic & Molecular Epidemiology Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Human Genotyping Unit, Human Cancer Genetics Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Department of Epidemiology, University of California Irvine, Irvine, California, USA

City of Hope Cancer Centre, Duarte, California, USA

Cancer Prevention Institute of California, Fremont, California, USA

Division of Clinical Epidemiology and Aging Research, German Cancer Research Center [DKFZ], Heidelberg, Germany

Saarland Cancer Registry, Saarbrücken, Germany

Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart

University of Tübingen, Germany

Institute for Prevention and Occupational Medicine of the German Social Accident Insurance (IFA), Bochum, Germany

Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany

Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany

Department of Obstetrics and Gynecology, University of Ulm, Ulm, Germany

Department of Obstetrics and Gynaecology, Hannover Medical School, Hannover, Germany

Department of Radiation Oncology, Hannover Medical School, Hannover, Germany

N.N. Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus

Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa, Russia

Department Molecular Medicine and Surgery, Karolinska Institutet, Stockholm

Department Onkologi-pathology, Karolinska Institutet, Stockholm

School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland

Biocenter Kuopio, Kuopio, Finland

Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland

School of Medicine, Institute of Clinical Medicine, Oncology, University of Eastern Finland, Kuopio, Finland

Department of Oncology, Kuopio University Hospital, Kuopio, Finland

Vesalius Research Center [VRC], VIB, Leuven, Belgium

Multidisciplinary Breast Center, University Hospital Gasthuisberg, Leuven, Belgium

Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto
Nazionale Tumori (INT), Milan, Italy
Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia (IEO), Milan, Italy
IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, Milan, Italy
Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Italy and IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, Milan, Italy
Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA
Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA
Department of Pharmacology, Mayo Clinic, Rochester, Minnesota, USA
Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Australia
Department of Anatomical Pathology, Alfred Hospital, Melbourne, Australia
Department of Genetics, Institute for Cancer Research, Oslo University Hospital, Radiumhospitalet, Oslo, Norway
Institute for Clinical Medicine, Faculty of Medicine, UiO, Norway
Cancer Prevention Institute of California, Fremont, California, USA
Stanford University School of Medicine, Stanford, California, USA
Dana-Farber Cancer Institute, Boston, Massachusetts, USA
Laboratory of Cancer Genetics, Department of Clinical Genetics and Biocenter Oulu, University of Oulu, Oulu University Hospital, Oulu, Finland.
Department of Oncology, University of Oulu, Oulu University Hospital, Oulu, Finland.
Department of Pathology, University of Oulu, Oulu University Hospital, Oulu, Finland.
Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital
Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada
Ontario Cancer Genetics Network, Cancer Care Ontario, Toronto, Ontario, Canada
Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada
Department of Laboratory Medicine, and the Keenan Research Centre of the Li Ka Shing Knowledge Institute, St Michael’s Hospital, Toronto, Ontario, Canada
Department of Human Genetics, Leiden University Medical Centre, Leiden, The Netherlands
Department of Pathology, Leiden University Medical Centre, Leiden, The Netherlands
Department of Clinical Genetics, Leiden University Medical Centre, Leiden, The Netherlands
Department of Surgery, Leiden University Medical Centre, Leiden, The Netherlands
Department of Medical Oncology, Rotterdam Family Cancer Clinic, Erasmus MC-Daniel den Hoed Cancer Center, Rotterdam, The Netherlands
Section of Epidemiology, Institute of Cancer Research, Sutton, United Kingdom
Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland
Department of Medical Oncology, Family Cancer Clinic, Erasmus University Medical Center, Rotterdam, The Netherlands
Department of Medical Oncology, Josephine Nefkens Institute, Erasmus University Medical Center, Rotterdam, The Netherlands
Department of Clinical Genetics, Family Cancer Clinic, Erasmus University Medical Center, Rotterdam, The Netherlands
Institute for Cancer Studies, Department of Oncology, University of Sheffield, Sheffield, UK
Academic Unit of Surgical Oncology, Department of Oncology, University of Sheffield, Sheffield, UK
International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland
Queensland Institute of Medical Research, Brisbane, Australia
Peter MacCallum Cancer Center, Melbourne, Australia
Warwick Medical School, Warwick University, Coventry, UK
Department of Preventive Medicine, Sirirajkrinwirot University, Ongkhara, Nakhon Nayok, Thailand
Department of Academic support, The National Cancer Institute of Thailand, Ministry of Public Health, Thailand
Seoul National University College of Medicine, Seoul, Korea
Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan
Department of Surgery, Tri-Service General Hospital, Taipei, Taiwan
Against Breast Cancer Research Group, Department of Molecular and Applied Biosciences, University of Westminster, London, UK
Faculty of Medicine, University of Southampton, Southampton, UK
Department of Oncology, University of Cambridge, Cambridge, UK
CRUK Cambridge Research Institute, Cambridge, UK
Centre National de Genotypage, Evry, France
Abstract

Breast cancer is the most common cancer among women. To date, 22 common breast cancer susceptibility loci have been identified accounting for ~8% of the heritability of the disease. We followed up 72 promising associations from two independent Genome Wide Association Studies (GWAS) in ~70,000 cases and ~68,000 controls from 41 case-control studies and nine breast cancer GWAS. We identified three new breast cancer risk loci on 12p11 (rs10771399; \(P=2.7 \times 10^{-35}\)), 12q24 (rs1292011; \(P=4.3 \times 10^{-19}\)) and 21q21 (rs2823093; \(P=1.1 \times 10^{-12}\)). SNP rs10771399 was associated with similar relative risks for both estrogen receptor (ER)-negative and ER-positive breast cancer, whereas the other two loci were associated only with ER-positive disease. Two of the loci lie in regions that contain strong plausible candidate genes: \(PTHLH\) (12p11) plays a crucial role in mammary gland development and the establishment of bone metastasis in breast cancer, while \(NRIP1\) (21q21) encodes an ER co-factor and has a role in the regulation of breast cancer cell growth.

Breast cancer is one of the most commonly occurring epithelial malignancies in women with an estimated one million new cases and over 400,000 deaths annually worldwide\(^1\). Familial aggregation and twin studies have demonstrated the substantial contribution of inherited susceptibility to breast cancer\(^2,3\). Over the last four years, we and others have conducted several genome-wide association studies (GWAS) and reported breast cancer susceptibility variants at 21 loci\(^4-14\) with an additional locus (\(CASP8\)) identified through a candidate gene approach\(^15\). These variants are associated with modest risks of the disease (per-allele odds ratios <1.3), and explain ~8% of the excess familial risk of breast cancer, while other rarer high and moderate risk loci contribute less than 20%, suggesting that other loci remain to be identified\(^16\).

To identify further breast cancer susceptibility loci, we selected 72 SNPs that were genotyped and found to be significantly associated with breast cancer at \(P<0.0001\) in either of two breast cancer GWAS in the UK (UK2 and BBCS)\(^17,18\). We attempted to genotype these SNPs in up to 41 case-control studies through the Breast Cancer Association Consortium (BCAC). After quality control (QC) exclusions (see Methods), we analysed data on 54,588 cases of invasive breast cancer, 2401 cases of Ductal Carcinoma in Situ (DCIS) and 58,098 controls. In addition, we utilised data from 7 additional breast cancer GWAS from which summary results had been obtained based on imputation to Hapmap 2 CEU. Results from the GWAS and BCAC replication were then combined to derive the overall evidence of association for each SNP based on 69,564 cases and 68,150 controls.

Three SNPs showed strong evidence for association in European women, consistent with the effect seen in the original GWAS (Table 1 and Figure 1). In each case, the genotype-specific odds ratios (ORs) were consistent with an allele dose (log-additive) model (Supplementary Table 1). SNP rs2823093 showed some evidence of heterogeneity in the per-allele ORs among studies in the replication stage \((P=0.0002)\), with particularly marked associations in two studies \((P=7.1 \times 10^{-7})\). The other two loci showed no evidence of heterogeneity among studies. Two additional SNPs on 17q21, rs2532348 and rs199523 (correlated at \(r^2=0.80\) in the UK2 GWAS), gave more limited evidence of replication \((P=0.00078\) and \(P=0.0063)\) and reached \(P=5.8 \times 10^{-7}\) and \(P=2.6 \times 10^{-6}\) respectively when combined with the GWAS data (Supplementary Table 2). These SNPs were only genotyped in the UK2 GWAS. They could not be imputed using HapMap, and were only successfully genotyped in 12 studies in the BCAC replication. Moreover, for SNP rs2532348 there was evidence of heterogeneity among studies in the per-allele ORs in BCAC \((P=0.001)\). Further data will be required to determine whether this SNP is associated with breast cancer risk. Three other SNPs \((rs10940235\) on 5q11, \(rs4403040\) on...
4q21 and rs6027564 on 20q13) showed evidence of replication at \( P < 0.01 \) but none reached genome-wide levels of statistical significance (Supplementary Table 2).

For women of Asian ancestry, SNP rs10771399 (12p11) was also associated with breast cancer risk, with the estimated OR being similar to that in women of European ancestry (Supplementary Table 3). There was no significant evidence of association for either SNPs rs1292011 (12q24) or rs2823093 (21q21) in women of Asian ancestry. For rs2823093, the estimated OR was in the opposite direction than that in women of European ancestry, but the estimates did not differ significantly (Supplementary Table 3).

SNP rs10771399 showed strong evidence of association with both estrogen receptor (ER)-positive and ER-negative breast cancer, with the estimated per-allele ORs being similar (based on 24,775 ER-positive and 7,122 ER-negative cases; Supplementary Table 4a). In contrast, for SNPs rs1292011 and rs2823093, the association was confined to ER-positive breast cancer, with no evidence of association for ER-negative disease (Supplementary Table 4a). These latter results conform to the general pattern of a preponderance of common susceptibility loci for ER-positive disease identified through GWAS based on cases unselected for disease subtype 19, 20. In terms of per-allele OR, SNP rs10771399 has one of the strongest effects identified to date for ER-negative breast cancer (OR 0.85, 95%CI 0.80-0.90). For all three SNPs, the per-allele OR for DCIS was similar to that for invasive disease (based on up to 2,148 DCIS cases; Supplementary Table 4b). For SNP rs10771399, the estimated OR was higher for 10 studies in which cases were selected for a positive family history and/or bilaterality, as would be expected under a polygenic model 21 (\( P = 0.027 \), Supplementary Table 5); however, exclusion of data from these studies made little difference to the estimated OR. There was no evidence for difference in the per-allele OR by age at diagnosis for any SNP (Supplementary Table 4c).

SNP rs10771399 lies in a ~300kb linkage disequilibrium (LD) block on 12p11 that contains one known gene, \( PTHLH \) (Parathyroid Hormone like Hormone isoform 1), also called \( PTHrP \) (Parathyroid hormone–related protein; Figure 2a). \( PTHrP \) is expressed in a wide variety of tissues and in many malignancies, including 60% of breast tumors and is required for normal mammary gland and bone development 22-25. During lactation it is released by the mammary gland to regulate the transfer of calcium from the skeleton to the milk 26, 27. Tumor secreted \( PTHrP \) mimics the action of parathyroid hormone (PTH) by binding to its receptor \( PTH1R \) 28 promoting humoral hypercalcemia as well as metastasis of breast cancer cells to the bone 23, 29-31. It has been suggested that \( PTHrP \) enhances tumorigenesis through its pro-proliferative and anti-apoptotic activity by promoting survival in cells subjected to apoptosis 32, 33. However, conflicting data regarding the correlation of \( PTHrP \) expression level and breast cancer survival have been found 24, 34-36. Moreover, a recent study reported that loss of \( PTHrP \) accelerates tumor incidence in DCIS and is associated with monocyte infiltration 37.

SNP rs1292011 on 12q24 lies in a ~100 kb LD block that contains no known genes (Figure 2b). SNPs in this region have been found to be associated with squamous esophageal carcinoma, renal cell carcinoma, liver adenoma, heart disease and type 1 diabetes as well as blood pressure and PSA levels 38-47. Two plausible cancer candidate genes, \( MAPKAPK5 \) (mitogen-activated protein kinase-activated protein kinase 5, also called \( MK5/PRAK \)) and \( TBX3 \) (T-box3), lie within 2 Mb of rs1292011. \( MAPKAPK5 \) is a member of the serine/threonine kinase family and is directly activated by \( Myc \) 48. \( TBX3 \) plays a role in mammary gland development 49 and its haplo-insufficiency is associated with Ulnar-Mammary disorder 50. \( TBX3 \) was found to be amplified and over-expressed in several cancers including breast cancer 51-54 and at high levels in plasma from breast and ovarian cancer patients 52. Recently, it has been shown that estrogen regulates the expansion of breast cancer stem cells
through the FGF/FGFR/TBX3 pathway\textsuperscript{52, 55} and that TBX3 is a direct downstream target of the Wnt/beta-catenin pathway\textsuperscript{56}. The expression of TBX3 was found to be significantly higher (P<0.0001) in ER-positive than in ER-negative breast cancer tumors in two independent datasets containing 781 tumors (with HGU-133A Affymetrix expression data)\textsuperscript{57} and 244 tumors (with 44k Agilent expression data)\textsuperscript{58}. These data suggest that the association of rs1292011 with ER-positive breast cancer could be mediated through its effect on TBX3.

SNP rs2823093 lies in a ~130 Kb LD block containing no known genes. The nearest gene, ~900 Kb downstream, is NRIP1 (Nuclear Receptor interacting protein 1) (Figure 2c) or also called RIP140 (Receptor-interacting protein 140). RIP140 acts as a strong transcriptional repressor for nuclear receptors\textsuperscript{59, 60}. It interacts with estrogen receptor α (ERα), represses the ER signalling and inhibits its mitogenic effects\textsuperscript{61}. This repression is mediated through interaction with FHL1, a protein involved in suppressing cancer cell growth and migration\textsuperscript{62}. Several lines of evidence suggest that RIP140 plays an important role in the regulation of breast cancer cell growth. Knockdown of RIP140 was found to induce growth promotion in an ER-positive breast cancer cell line\textsuperscript{61}. This protein was also highly induced following the treatment of human breast cancer cells with retinoids, known for their breast cancer growth suppression and their anti-estrogenic effects\textsuperscript{63-66}. A Spanish case-control study, which genotyped SNPs in 91 breast cancer candidate genes in ~700 cases and ~700 controls, identified a relatively rare SNP at this locus (rs926184 - MAF~2%), located 175 Kb upstream of rs2823093, which showed a modest association with breast cancer\textsuperscript{67}. These two SNPs are, however, not correlated (\(r^2=0\) in HapMap CEU). The expression of NRIP1 has been shown to be significantly higher in ER-positive than ER-negative tumors (p<0.0001)\textsuperscript{57, 58} suggesting that the association of rs28323093 with ER-positive breast cancer could be mediated through its effect on NRIP1 expression\textsuperscript{57, 58}.

The three novel susceptibility variants identified in this study are relatively common (MAF 0.11-0.41) and together explain ~0.7% of the familial risk of breast cancer, and bring the total contribution of common low-penetration breast cancer susceptibility loci to ~9%. The relative risks associated with these variants are modest, with the per-allele ORs for the risk allele ranging from 1.07 to 1.22 fold, but the causal variants underlying some of these loci might confer more substantial risks. The present work highlights the importance of combining GWAS and large-scale replication studies with tumor subtyping in the identification and characterisation of breast cancer susceptibility loci.

The genes in these regions (if proven to be the causal genes) underscore that diverse mechanisms are likely to be relevant to breast cancer pathogenesis. Re-sequencing of these loci, combined with fine-scale mapping and functional analyses will provide more insights into the genetic architecture of breast cancer and the pathogenesis of the disease.

\section*{Methods}

\subsection*{GWAS analysis}

Primary genotype data were obtained for nine breast cancer GWAS in populations of European ancestry (Supplementary Table 6). Standard QC was performed on all scans, as follows. We excluded all individuals with low call rate (<95%), extreme high or low heterozygosity (P<10\(^{-5}\)), and all individuals evaluated to be of non-European ancestry (>15% non-European component, by multidimensional scaling using the three Hapmap2 populations as a reference). We excluded SNPs with: call rate <95%; call rate <99% and MAF<5%, all SNPs with MAF<1%, and SNPs whose genotype frequencies departed from Hardy-Weinberg equilibrium at P<10\(^{-6}\) in controls or P<10\(^{-12}\) in cases. For highly

\begin{itemize}
  \item [Nat Genet. Author manuscript; available in PMC 2013 May 14.]}
significant SNPs the genotype intensity cluster plots were examined manually to judge reliability, either centrally or by contacting the original investigators.

Data were imputed for all scans for ~2.6M SNPs using HapMap version 2 CEU as a reference, using the program Mach v1.0. Estimated per-allele ORs and standard errors were generated from the imputed genotypes using Probabel69. For two studies (UK2 and HEBCS), estimates were adjusted by the first three principal components, since this was found to materially reduce the inflation. Residual inflation was then adjusted for by multiplying the variance by a genomic control adjustment factor, based on the ratio of the median chi-squared test statistic to its expected value. BBCS and UK2 used the same control data (WTCCC2) but different genotyping platforms. These studies were imputed separately. For the combined analysis, the control set was divided randomly between the two studies, in proportion to the size of case series, to provide disjoint strata. For a limited subset of SNPs that could not be imputed (including rs2532348 and rs199523 on 17q21), genotype data from the original scan(s) were used in the analysis.

Replication stage

SNPs for replication were genotyped in 46 studies, of which 4 were case-only studies that did not contribute to the current analysis (Supplementary Table 7). Data from BBCS were excluded as the same cases were included in the GWAS. Seven studies (HABCS, HMBCS, HUBCS, KARBAC, RBCS, SEARCH and SEBCS) were analysed by Fluidigm for 72 SNPs (Supplementary Table 2). We selected 63 SNPs selected from UK2: one replaced by a better surrogate, and one failed, so only data were available for 61 SNPs. Ten SNPs were selected from BBCS and one SNP was selected from both scans (The original SNP, rs1975930, also referred to as rs56003999, did not work by Fluidigm and in some iPLEX analyses and was replaced by a surrogate rs10771399, r²=0.95, which was typed in all studies). Samples from 27 studies were genotyped by iPLEX for 29 SNPs that showed the strongest associations. Seven additional studies (ABCFS, CGPS, MCCS, NC-BCFR, OFBCR, PBCS, UKBGS) were genotyped by Taqman for up to 4 SNPs that showed association after the Fluidigm and iPLEX genotyping, including all three 3 SNPs discussed in detail here. We restricted the analysis to individuals of European or East Asian ancestry, since the sample size for other ethnicities was too small to give meaningful results.

All studies complied with BCAC genotyping QC standards by including at least 2% of samples in duplicate and a common set of 93 CEPH DNAs used by the HapMap Consortium (HAPMAPPT01, Coriell Institute for Medical Research, Camden, NJ). Genotype data were excluded for: any sample that consistently failed genotyping for >20% of the SNPs typed; all samples on any one plate that had a SNP call rate <90%; all genotype data for any SNP where overall call rate was <95%; and all genotype data for any SNP where duplicate concordance was <98% (based on 2% of samples genotyped in duplicate). In addition, for any SNP for which the P-value for departure from Hardy-Weinberg equilibrium for controls was <0.005, clustering of the intensity plots was reviewed manually and the data excluded if clustering was judged to be poor. After QC exclusions we analysed data on 54,588 cases of invasive breast cancer, 2,401 cases of DCIS and 58,098 controls.

Per-allele and genotype-specific odds ratios for the replication stage were estimated using logistic regression, adjusted for study. Women of European and Asian ancestry were analysed separately. NC-BCFR contributed cases and controls to both European and Asian analyses; for the remaining studies the subjects were either predominantly European or predominantly Asian, and subjects from other minority ethnicities were excluded.

Statistical significance levels from the GWAS and BCAC replication phases were obtained by combining the logOR estimates and standard errors as in a fixed effect meta-analysis.
Heterogeneity in the OR association with each SNP by ER status was evaluated using a case-only analysis, by logistic regression. Heterogeneity by age was evaluated by fitting a linear age × genotype interaction term.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

We would thank the following individuals for their contribution to this project (study in brackets): Sten Cornelissen, Richard van Hien, Linde Braaf, Laura Van’t Veer, Bas Bueno-de-Mesquita and Sander Canisius (ABCS); Niall McInerney, Gabrielle Colleran, Andrew Rowan and Nicola Miller (BIGGS); Anne Langheinz (BSUCH); José Ignacio Arias Pérez, Pilar Zamora, Primitiva Menendez, Tais Moreno and Guillermo Pita (CNIO-BCS); Muriel Adank, Margreet Ausems and Senno Verhoef (DFBBCS); Ute Hamann, Yon-Dschun Ko, Christian Baisch, Hans-Peter Fischer, Beate Pesch, Sylvia Rabstein and Volker Harth (GENICA); Kirsimari Aaltonen, Päivi Heikkilä, Tuomas Heikkilänen, Dario Greco, RN Hanna Jäntti and Irja Erkkilä (HEBCS); Helena Kemiläinen, Eija Myöhänien and Aija Parkkinen (KBCP); Tracy Slanger, Elke Mutschelknauss, S. Behrens, R. Birr, M.Celik, U. Eisber, B. Kaspereit, N. Knese and K. Smitt (MARIE); Paolo Radice, Bernard Peissel, Monica Barile, Marco A. Pierotti (MBCSG); Teresa Selander, Mona Gill, Lucine Collins and Nayana Weerasooriya (OFBCR); Mervi Grip, Kari Mononen and Meeri Otsukka (OBCS); E. Krol-Warmerdam, and J. Blom (ORIGO); Dr. Prat Boonyawongviroj and Dr. Pornthep Siriwanarungsan (ACP). For full acknowledgements including funding see Supplementary Note.

**Reference List**


Figure 1.
Forest plots for the 3 SNPs showing evidence of association with breast cancer. Squares represent the estimated per-allele odds ratio (OR) for individual studies. The area of square is inversely proportional to the precise of the estimate. Diamonds represent the summary OR estimates for the subgroups indicated. Horizontal lines represent 95% confidence limits.
Figures 2a, b and c.
Association plots for the three new breast cancer susceptibility loci at (a) 12p11 (b) 12q24 and (c) 21q21 drawn using the SNAP software. Genotyped and imputed SNPs are plotted based on their chromosomal position in build 36 on the X axis and their overall $P$ values (as $-\log_{10}$ values) from the UK2 and BBCS GWAS on the Y axis. For each region, the most strongly associated SNP is represented by a diamond. The intensity of the red shading reflects the strength of correlation ($r^2$) between the best SNP and the other SNPs in the region. Genes present in the region (if any) are indicated in green.
<table>
<thead>
<tr>
<th>SNP</th>
<th>Chromosome Position</th>
<th>Alleles</th>
<th>MAF</th>
<th>Stage</th>
<th>Per-allele OR (95% CI)</th>
<th>P</th>
<th>Combined P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10771399</td>
<td>12p11 28046347</td>
<td>AG</td>
<td>0.12</td>
<td>UK2</td>
<td>0.79 (0.71-0.87)</td>
<td>3.1×10^{-6}</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.11 BBCS</td>
<td>0.84 (0.74-0.96)</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.10 Other GWAS</td>
<td>0.83 (0.75-0.91)</td>
<td>5.7×10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.12 BCAC replication</td>
<td>0.85 (0.83-0.88)</td>
<td>3.3×10^{-27} 2.7×10^{-35}</td>
</tr>
<tr>
<td>rs1292011</td>
<td>12q24 114320905</td>
<td>AG</td>
<td>0.41</td>
<td>UK2</td>
<td>0.88 (0.83-0.94)</td>
<td>5.8×10^{-5}</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.42 BBCS</td>
<td>0.95 (0.88-1.03)</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.40 Other GWAS</td>
<td>0.91 (0.86-0.96)</td>
<td>.0008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.41 BCAC replication</td>
<td>0.92 (0.91-0.94)</td>
<td>6.2×10^{-14} 4.3×10^{-19}</td>
</tr>
<tr>
<td>rs2823093</td>
<td>21q21 15442703</td>
<td>GA</td>
<td>0.26</td>
<td>UK2</td>
<td>0.96 (0.89-1.03)</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.26 BBCS</td>
<td>0.88 (0.76-0.92)</td>
<td>.00013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.26 Other GWAS</td>
<td>0.91 (0.85-0.97)</td>
<td>.0032</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.27 BCAC replication</td>
<td>0.94 (0.92-0.96)</td>
<td>1.7×10^{-9} 1.1×10^{-12}</td>
</tr>
</tbody>
</table>

1. Build 36
2. Minor allele listed second
3. Per copy of the minor allele

Nat Genet. Author manuscript; available in PMC 2013 May 14.