THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Identification of independent association signals and putative functional variants for breast cancer risk through fine-scale mapping of the 12p11 locus

Citation for published version:
EMBRACE, Zeng, C, Guo, X, Long, J, Kuchenbaecker, KB, Droit, A, Michailidou, K, Ghoussaini, M, Kar, S, Freeman, A, Hopper, JL, Milne, RL, Bolla, MK, Wang, Q, Dennis, J, Agata, S, Ahmed, S, Aittomäki, K, Andrulis, IL, Anton-Culver, H, Antonenkova, NN, Arason, A, Arndt, V, Arun, BK, Arver, B, Bacot, F, Barrowdale, D, Baynes, C, Beeghly-Fadiel, A, Benitez, J, Bermisheva, M, Blomqvist, C, Blot, WJ, Bogdanova, NV, Bojesen, SE, Bonanni, B, Borresen-Dale, A-L, Brand, JS, Brauch, H, Brennan, P, Brenner, H, Broeks, A, Brüning, T, Burwinkel, B, Buys, SS, Cai, Q, Caldes, T, Campbell, I, Carpenter, J, Chang-Claude, J & Figueroa, J 2016, 'Identification of independent association signals and putative functional variants for breast cancer risk through fine-scale mapping of the 12p11 locus', *Breast Cancer Research*, vol. 18, no. 1, pp. 64. https://doi.org/10.1186/s13058-016-0718-0

Digital Object Identifier (DOI):
10.1186/s13058-016-0718-0

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Breast Cancer Research

Publisher Rights Statement:
© 2016 The Author(s). Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0
International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and
reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to
the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver
(http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 19. Apr. 2024

Zeng et al. Breast Cancer Research (2016) 18:64
DOI 10.1186/s13058-016-0718-0

Breast Cancer Research

RESEARCH ARTICLE Open Access

Identification of independent association signals and putative functional variants for breast cancer risk through fine-scale mapping of the 12p11 locus

Chenjie Zeng[1], Xingyi Guo[1], Jirong Long[1], Karoline B. Kuchenbaecker[2], Arnaud Droit[3], Kyriaki Michailidou[2], Maya Ghoussaini[4], Siddhartha Kar[4], Adam Freeman[5], John L. Hopper[6], Roger L. Milne[6,7], Manjeet K. Bolla[2], Qin Wang[2], Joe Dennis[2], Simona Agata[9], Shahana Ahmed[10], Kristiina Aittomäki[11], Irene L. Andrulis[12,13], Hoda Anton-Culver[14], Natalia N. Antonenkova[15], Adalgeir Arason[16], Volker Arndt[17], Banu K. Arun[18], Brita Arver[19], Francois Bacot[20], Daniel Barrowdale[2], Caroline Baynes[10], Alicia Beeghly-Fadiel[1], Javier Benitez[21,22], Marina Bermisheva[23], Carl Blomqvist[24], William J. Blot[1,25], Natalia V. Bogdanova[26], Stig E. Bojesen[27,28,29], Bernardo Bonanni[30], Anne-Lise Borresen-Dale[31,32], Judith S. Brand[33], Hiltrud Brauch[34,35,36], Paul Brennan[37], Hermann Brenner[17,36,38], Annegien Broeks[39], Thomas Brüning[40], Barbara Burwinkel[41,42], Saundra S. Buys[43], Qiuyin Cai[1], Trinidad Caldes[44], Ian Campbell[45], Jane Carpenter[46], Jenny Chang-Claude[47,48], Ji-Yeob Choi[49,50,51], Kathleen B. M. Claes[52], Christine Clarke[53], Angela Cox[54], Simon S. Cross[55], Kamila Czene[33], Mary B. Daly[56], Miguel de la Hoya[44], Kim De Leeneer[52], Peter Devilee[57,58], Orland Diez[59], Susan M. Domchek[60], Michele Doody[61], Cecilia M. Dorfling[62], Thilo Dörk[63], Isabel dos-Santos-Silva[64], Martine Dumont[65], Miriam Dwek[66], Bernd Dworniczak[67], Kathleen Egan[68], Ursula Eilber[47], Zakaria Einbeigi[69], Bent Ejlertsen[70], Steve Ellis[2], Debra Frost[2], Fiona Lalloo[71], on behalf of EMBRACE[2], Peter A. Fasching[72,73], Jonine Figueroa[61], Henrik Flyger[74], Michael Friedlander[75], Eitan Friedman[76], Gaetana Gambino[77], Yu-Tang Gao[78], Judy Garber[79], Montserrat García-Closas[61,80], Andrea Gehrig[81], Francesca Damiola[82], Fabienne Lesueur[83], Sylvie Mazoyer[82], Dominique Stoppa-Lyonnet[84,85,86], behalf of GEMO Study Collaborators[87], Graham G. Giles[6,7], Andrew K. Godwin[88], David E. Goldgar[89], Anna González-Neira[21], Mark H. Greene[90], Pascal Guénel[91,92], Lothar Haeberle[72], Christopher A. Haiman[93], Emily Hallberg[94], Ute Hamann[95], Thomas V. O. Hansen[96], Steven Hart[94], Jaana M. Hartikainen[97,98,99], Mikael Hartman[100,101], Norhashimah Hassan[102,103], Sue Healey[104], Frans B. L. Hogervorst[105], Senno Verhoef[105], on behalf of HEBON[106], Carolyn B. Hendricks[107,108], Peter Hillemanns[63], Antoinette Hollestelle[109], Peter J. Hulick[110], David J. Hunter[111,112], Evgeny N. Imyanitov[113], Claudine Isaacs[114], Hidemi Ito[115], Anna Jakubowska[116], Ramunas Janavicius[117], Katarzyna Jaworska-Bieniek[116], Uffe Birk Jensen[118], Esther M. John[119,120,121], Charles Joly Beauparlant[122], Michael Jones[80], Maria Kabisch[95], Daehee Kang[49,50,51], Beth Y. Karlan[123], Saila Kauppila[124], Michael J. Kerin[125], Sofia Khan[126], Elza Khusnutdinova[23,127], Julia A. Knight[128,129], Irene Konstantopoulou[130], Peter Kraft[111,112], Ava Kwong[131,132], Yael Laitman[76], Diether Lambrechts[133,134], Conxi Lazaro[135], Loic Le Marchand[136], Chuen Neng Lee[100], Min Hyuk Lee[137], Jenny Lester[123], Jingmei Li[33], Annelie Liljegren[19], Annika Lindblom[138], Artitaya Lophatananon[139], Jan Lubinski[116], Phuong L. Mai[90], Arto Mannermaa[97,98,99], Siranoush Manoukian[140], Sara Margolin[141], Frederik Marme[41,142], Keitaro Matsuo[143], Lesley McGuffog[2], Alfons Meindl[144], Florence Menegaux[91,92], Marco Montagna[9], Kenneth Muir[139,145],

* Correspondence: wei.zheng@vanderbilt.edu
[1]Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, 2525 West End Avenue, 8th Floor, Nashville, TN 37203-1738, USA

BioMed Central

© 2016 The Author(s). **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Anna Marie Mulligan[146,147], Katherine L. Nathanson[60], Susan L. Neuhausen[148], Heli Nevanlinna[126], Polly A. Newcomb[149,150], Silje Nord[31], Robert L. Nussbaum[151], Kenneth Offit[152,197], Edith Olah[153], Olufunmilayo I. Olopade[154], Curtis Olswold[94], Ana Osorio[155,156], Laura Papi[157], Tjoung-Won Park-Simon[63], Ylva Paulsson-Karlsson[158], Stephanie Peeters[159], Bernard Peissel[160], Paolo Peterlongo[161], Julian Peto[64], Georg Pfeiler[162], Catherine M. Phelan[163], Nadege Presneau[66], Paolo Radice[164], Nazneen Rahman[165], Susan J. Ramus[166], Muhammad Usman Rashid[95,167], Gad Rennert[168], Kerstin Rhiem[169], Anja Rudolph[47], Ritu Salani[170], Suleeporn Sangrajrang[171], Elinor J. Sawyer[172], Marjanka K Schmidt[39], Rita K. Schmutzler[173,174,175,176], Minouk J. Schoemaker[80], Peter Schürmann[63], Caroline Seynaeve[109], Chen-Yang Shen[177,178], Martha J. Shrubsole[1], Xiao-Ou Shu[1], Alice Sigurdson[61], Christian F. Singer[179], Susan Slager[94], Penny Soucy[180], Melissa Southey[181], Doris Steinemann[182], Anthony Swerdlow[80,183], Csilla I. Szabo[184], Sandrine Tchatchou[185], Manuel R. Teixeira[186,187], Soo H. Teo[102,103], Mary Beth Terry[188], Daniel C. Tessier[20], Alex Teulé[189], Mads Thomassen[190], Laima Tihomirova[191], Marc Tischkowitz[192,193], Amanda E. Toland[194], Nadine Tung[195], Clare Turnbull[165], Ans M. W. van den Ouweland[196], Elizabeth J. van Rensburg[62], David ven den Berg[93], Joseph Vijai[152,197], Shan Wang-Gohrke[198], Jeffrey N. Weitzel[199], Alice S. Whittemore[120,121], Robert Winqvist[200,201], Tien Y. Wong[202], Anna H. Wu[93], Drakoulis Yannoukakos[203], Jyh-Cherng Yu[204], Paul D. P. Pharoah[2,10], Per Hall[33], Georgia Chenevix-Trench[205,206], on behalf of KConFab[205], AOCS Investigators[206], Alison M. Dunning[10], Jacques Simard[65], Fergus J. Couch[8], Antonis C. Antoniou[2], Douglas F. Easton[2,10] and Wei Zheng[1*]

Abstract

Background: Multiple recent genome-wide association studies (GWAS) have identified a single nucleotide polymorphism (SNP), rs10771399, at 12p11 that is associated with breast cancer risk.

Method: We performed a fine-scale mapping study of a 700 kb region including 441 genotyped and more than 1300 imputed genetic variants in 48,155 cases and 43,612 controls of European descent, 6269 cases and 6624 controls of East Asian descent and 1116 cases and 932 controls of African descent in the Breast Cancer Association Consortium (BCAC; http://bcac.ccge.medschl.cam.ac.uk/), and in 15,252 *BRCA1* mutation carriers in the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA). Stepwise regression analyses were performed to identify independent association signals. Data from the Encyclopedia of DNA Elements project (ENCODE) and the Cancer Genome Atlas (TCGA) were used for functional annotation.

Results: Analysis of data from European descendants found evidence for four independent association signals at 12p11, represented by rs7297051 (odds ratio (OR) = 1.09, 95 % confidence interval (CI) = 1.06–1.12; $P = 3 \times 10^{-9}$), rs805510 (OR = 1.08, 95 % CI = 1.04–1.12, $P = 2 \times 10^{-5}$), and rs1871152 (OR = 1.04, 95 % CI = 1.02–1.06; $P = 2 \times 10^{-4}$) identified in the general populations, and rs113824616 ($P = 7 \times 10^{-5}$) identified in the meta-analysis of BCAC ER-negative cases and *BRCA1* mutation carriers. SNPs rs7297051, rs805510 and rs113824616 were also associated with breast cancer risk at $P < 0.05$ in East Asians, but none of the associations were statistically significant in African descendants. Multiple candidate functional variants are located in putative enhancer sequences. Chromatin interaction data suggested that *PTHLH* was the likely target gene of these enhancers. Of the six variants with the strongest evidence of potential functionality, rs11049453 was statistically significantly associated with the expression of *PTHLH* and its nearby gene *CCDC91* at $P < 0.05$.

Conclusion: This study identified four independent association signals at 12p11 and revealed potentially functional variants, providing additional insights into the underlying biological mechanism(s) for the association observed between variants at 12p11 and breast cancer risk.

Keywords: Fine-scale mapping, Genetic risk factor, *PTHLH*, *CCDC91*, Breast cancer, *BRAC1* mutation carriers

Background

A previous genome-wide association study (GWAS) identified a common single nucleotide polymorphism (SNP), rs10771399 (termed the index SNP in this paper) at 12p11 to be associated with breast cancer risk in women of European descent [1]. This association, which did not vary by estrogen receptor (ER) status, was one of the most significant associations found for breast cancer risk in *Breast cancer 1* (*BRCA1*) mutation carriers so far, and the association was predominantly found in carriers with ER-negative (ER-(-)) breast cancer [2, 3]. This association was also replicated in East Asian women [4]. The index SNP lies in an approximately 300-kb linkage disequilibrium (LD) block, containing one known breast cancer associated gene that encodes parathyroid hormone-like hormone (*PTHLH*). This hormone has been shown to play a role in breast tumor initiation, progression, and metastasis in animal studies [5, 6] and was found to be associated with prognosis in breast cancer patients [7]. The index SNP, however, is located in a region with no evidence of functional significance [8]. The underlying biologic mechanisms and functional variants that drive the observed association have not yet been investigated. Furthermore, it is possible that additional independent risk signals may be present in the same region, as has been observed for other susceptibility regions [9–11]. In order to identify additional association signals at the12p11 locus with breast cancer risk, understand the underlying mechanisms and potential causal variants responsible for the association, we conducted a large fine-scale mapping study including data from 55,540 breast cancer cases and 51,168 controls in the Breast Cancer Association Consortium (BCAC) and 15,252 *BRCA1* mutation carriers in the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA).

Methods

Study population

The BCAC included 40 studies of women of European descent (48,155 cases and 43,612 controls), nine of Asian descent (6269 cases and 6624 controls), and two of African-American descent (1116 cases and 932 controls). The CIMBA included 45 studies of women of European descent (15,252 *BRCA1* mutation carriers), of whom 7797 had been diagnosed with breast cancer. Details on the study characteristics, participant characteristics and the methodology used by the BCAC and CIMBA have been published elsewhere [12–14]. Ethical approval of each study was given by the local institutional review boards. The full names of the institutional review boards that approved each study were listed in the Additional file 1.

SNP selection and genotyping

All SNPs within a 700-kb "fine mapping" interval at 12p11 (chr12: 27958733-28658733, hg19) were identified from the 1000 Genomes Project (1000G) (http://browser.1000genomes.org) CEU (April 2010) [15] and Hapmap III [16] (http://hapmap.ncbi.nlm.nih.gov/). The interval included all SNPs in LD ($r^2 > 0.1$) with the target SNP rs197593 ($r^2 = 0.95$ with the index SNP rs10771399) [1]. Tagging SNPs were selected to capture the remaining SNPs in the fine-mapping region at $r^2 > 0.9$. After quality control, genotypes for 441 SNPs were available for analysis. To improve the coverage, imputation was performed using data from the 1000G (March 2012) as the reference and the program IMPUTE2 [17] (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html). This was done separately for women of European, East Asian, and African descent and *BRCA1* mutation carriers. Using criteria of minor allele frequency (MAF) ≥2 % and an imputation quality $R^2 > 0.3$, genotype data were generated for a total of 1634 SNPs for studies of European women, 1360 for studies of East Asian women, 2508 for studies of African women in BCAC and 1646 for studies of *BRCA1* mutation carriers in CIMBA.

Statistical analysis

For BCAC studies, unconditional logistic regression models were used to estimate allelic odds ratios (OR) and their 95 % confidence intervals (CIs) of each of the SNPs included in the study. Analyses were performed separately for each ethnic group, and adjusted for study and principal components (seven for European studies and two each for Asian and African ancestry studies) [12]. Additional adjustment for age (age at diagnosis for cases and age at interview for controls) did not change the estimates, and thus age was not adjusted for in the main analyses. Tests of heterogeneity of the ORs across studies were conducted using Cochran's Q test. To identify independent association signals, we performed forward stepwise selection analyses with all SNPs associated with breast cancer risk at $P < 0.0001$ in BCAC European descendants or at $P < 0.005$ for East Asian descendants in the single-marker analysis. To reduce type 2 errors, we used a less stringent statistical significance threshold because of the smaller sample size for East Asian descendants than for European descendants in this study. Pairwise SNP-SNP interactions were evaluated using the likelihood ratio test for all SNPs selected from the forward stepwise regression analysis. Stratified analyses by ER status were performed, and the heterogeneity was assessed by case-only analysis. We estimated haplotype frequencies using the haplo.stats package under R with the expectation-maximum (EM) algorithm [18] and estimated the haplotype-specific ORs for women of European descent with adjustment for studies and principal components as described above. To evaluate whether the association varied by early-onset and late-onset cancer, stratified analyses by age at cancer diagnosis (≥45 or <45 years) were performed. The familial relative risk (FRR, λ) associated with independently

associated variants in this locus was calculated using the method described previously [19, 20].

For CIMBA studies, the associations between genetic variants and breast cancer risk were evaluated using a 1-degree of freedom (df) per allele trend test (*P*-trend), by modeling the retrospective likelihood of the observed genotypes conditional on breast cancer phenotypes [21]. To allow for the non-independence among related individuals, an adjusted test statistic was used, which took into account the correlation between study participants [22]. Per-allele hazard ratio (HR) estimates were obtained by maximizing the retrospective likelihood. All analyses were stratified by country of residence. To increase the statistical power to detect independent signals in *BRCA*1 mutation carriers, we conducted a meta-analysis of the BCAC and CIMBA studies [23]. Because approximately 80 % of breast tumors with known ER status in *BRCA1* mutation carriers were ER(-) [2], we only included the ER(-) breast cancer cases for BCAC studies. We combined the logarithm of the per-allele HR estimated in *BRCA1* mutation carriers and the logarithm of the per-allele OR estimated in BCAC using a fixed-effects model. We further determined whether there is evidence for independent association signals through a serial of conditional meta-analyses. We performed a conditional analysis on the top variant identified in the meta-analysis mentioned above in each consortium, and carried out the meta-analysis on the conditional *P* value for each variant to identify the most significant variant after conditioning on the top variant in the whole region. We continued to perform the conditional meta-analyses until the most significant association found had a *P* value >0.0001.

Functional annotation

We used the Encyclopedia of DNA Elements (ENCODE) chromatin states (chromHMM) annotation, DNase I hypersensitive, transcription factor binding sites, histone modifications of epigenetic markers (H3K4Me1, H3K4Me3 and H3K27Ac) data from ENCODE [24] (http://genome.ucsc.edu/ENCODE/) to determine the likely regulatory elements. We used chromatin interaction analysis by paired end tag (ChIA-PET), genome conformation capture (Hi-C) data from ENCODE and enhancer-promoter interaction data predicted by He et al. [25] to identify putative gene targets in mammary cell lines (human mammary epithelial cells (HMEC) and Michigan Cancer Foundation-7 (MCF7)). We used maps of enhancers as defined in Corradin et al. [8] and Hnisz et al. [26] to identify the locations of potential enhancers. We obtained RNA-seq data from ENCODE, respectively, to evaluate the expression of protein-coding genes in mammary cell lines at this locus. We also used the same data in the chronic myeloid leukemia cell line (K562) as a comparison if available.

To predict the most likely functional variants, we mapped all candidates to the transcription factor binding maps generated by ENCODE [24], based on the hypothesis that causal variants alter the binding affinity of transcription factors. We prioritized variants that were located in binding sites of master transcription factors of breast cancer and disrupted binding motif of transcription factors. We also prioritized variants that were located in active promoter regions in mammary cell lines. Two publicly available tools, RegulomeDB [27] (see http://regulome.stanford.edu/) and HaploReg V3 [28] (see http://www.broadinstitute.org/mammals/haploreg/haploreg.php), were also used to evaluate those candidate functional variants.

Expression quantitative trait loci (eQTL) analysis

The eQTL analyses in tumor tissues were performed as previously described [29, 30]. Briefly, we downloaded RNA-Seq V2, DNA methylation and SNP genotype data of 1006 breast cancer tumor tissues from The Cancer Genome Atlas (TCGA) data portal [26] (see http://cancergenome.nih.gov/). We log2-transformed the RNA-Seq by expectation-maximization (RSEM) value of each gene, and performed principal component adjustment of gene expression data to remove potential batch effects. Residual linear regression analysis was used to detect eQTLs while adjusting for methylation and copy number alterations (CNA), according to the approach proposed by Li et al. [29].

The eQTL analyses in 135 tumor-adjacent normal breast tissues were performed using data from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) [31] as previously described [32]. Briefly, gene expression levels were measured by the Illumina HT12 v3 microarray platform. Genotyping was performed using the Affymetrix SNP 6.0 array. Imputation was performed using data from the 1000G (CEU, March 2012) as the reference. Linear regression was performed to evaluate the association between genotypes and gene expression levels using the R (http://www.r-project.org/) package Matrix eQTL [32].

Results

Association results among women of European ancestry

Of the 2075 SNPs evaluated, 833 were associated with breast cancer risk in women of European descent at $P < 0.0001$ (Fig. 1). Using forward stepwise selection, we identified two SNPs that were independently associated with breast cancer risk with conditional $P < 0.0001$, tagging two independent signals (Table 1, Fig. 1). The index SNP is located in signal 2, approximately 30 kb upstream of the *PTHLH* gene and was in strong LD with

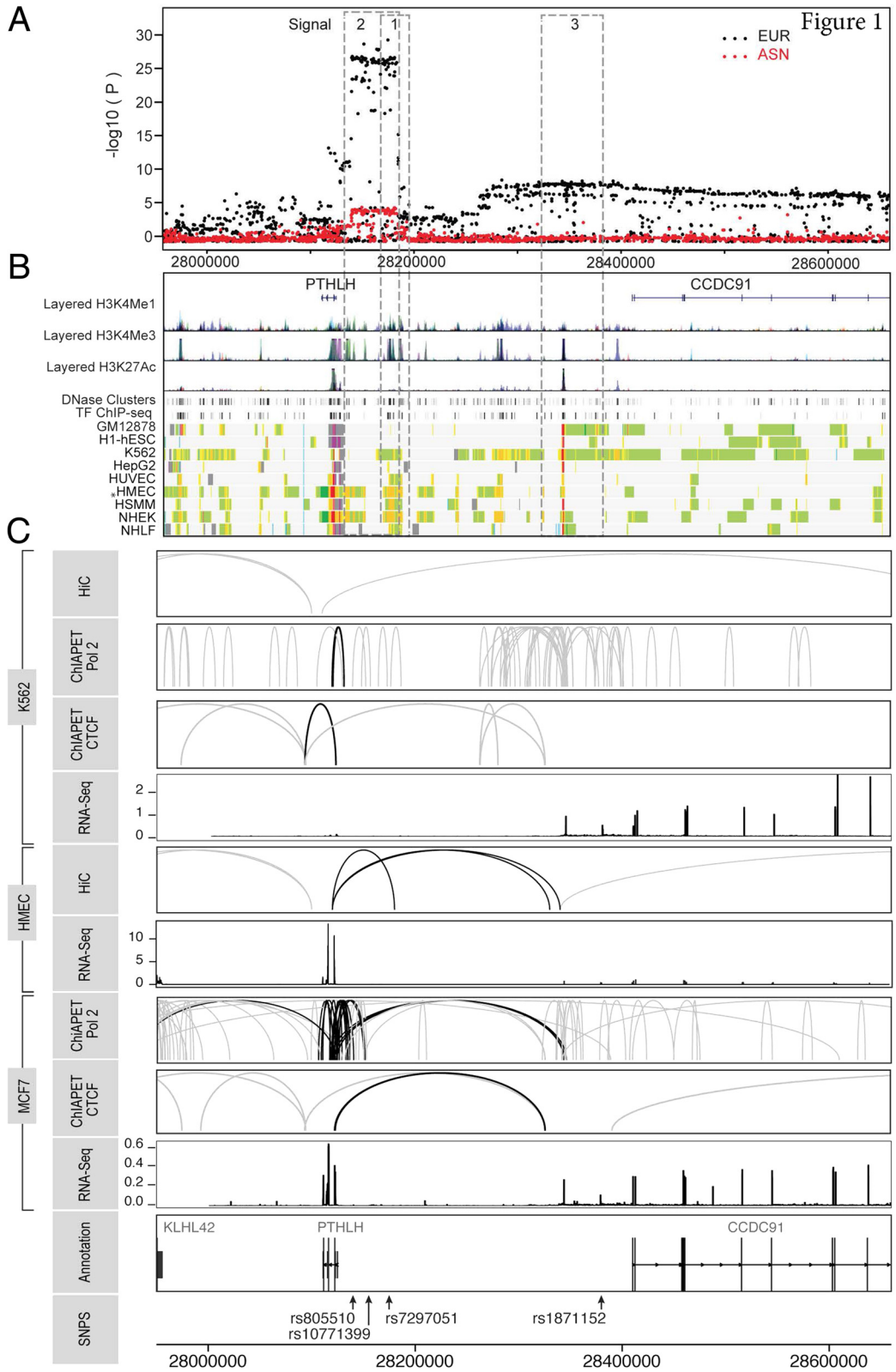

Fig. 1 (See legend on next page.)

(See figure on previous page.)
Fig. 1 Genetic mapping and epigenetic landscape of the 12p11 locus (**a**). Regional association plot of the genotyped and imputed Illumina iSelect genotyping array of the Collaborative Oncological Gene-environment Study (iCOGS) genotype data. Three independent signals were identified, marked as signal 1, 2 and 3. **b** Functional annotations using data from the Encyclopedia of DNA Elements (ENCODE) project. From *top* to *bottom*, the epigenetic signals evaluated include histone modifications, DNase clusters, transcription factor ChIP-seq clusters, and ENCODE chromatin states (ChromHMM) in the ENCODE cell lines. The signals of different layered histone modifications from the same ENCODE cell line are shown in the same color (the detailed color scheme for each ENCODE cell line is described in the UCSC genome browser; http://genome.ucsc.edu). *Red* and *orange* in chromatin states represent active promoter and strong enhancer regions, respectively (the detailed color scheme of the chromatin states was described in the previous study [45]). All tracks were generated by the UCSC genome browser (hg 19). **c** Long-range chromatin interactions. From *top* to *bottom*, genome conformation capture (*Hi-C*), chromatin interaction analysis by paired end tag (*ChIA-PET*) and RNA-Seq data from K562 cell lines, Hi-C and RNA-Seq from human mammary epithelial cells (*HMEC*), ChIA-PET and RNA-Seq from MCF7 cell lines, gene annotations and single nucleotide polymorphism (*SNP*) annotations. *Black lines* represent interactions with the promoter region (-1500/+500) of *Parathyroid hormone-like hormone* (*PTHLH*), and *gray lines* represent chromatin interaction that did not involve the *PTHLH* promoter region. The value of the RNA-Seq analysis corresponds to the mean reads per million (RPM) value for *PTHLH* from 65 K562, 4 HMEC and 19 MCF7 datasets, respectively. The annotation has been obtained through the Bioconductor annotation package TxDb.Hsapiens.UCSC.hg19.knownGene. The Hi-C and ChIA-PET raw data, available in the Gene Expression Omnibus (GEO) [GSE63525.K56, GSE33664, GSE39495], were processed using the GenomicRanges package. The tracks have been generated using ggplot2 and ggbio libraries in R

the lead SNP (rs805510) for this signal ($r^2 = 0.92$). The lead SNP in signal 1, rs7297051, is located approximately 50 kb upstream of the *PTHLH* gene, and was in moderate LD with the index SNP ($r^2 = 0.42$). The lead SNPs for signals 1 and 2 were moderately correlated ($r^2 = 0.36$). After adjusting for the lead SNPs in signals 1 and 2, we found evidence of the presence of a third independent association signal (lead SNP rs1871152; conditional $P = 2 \times 10^{-4}$, Table 1, Fig. 1). Signal 3 lies approximately 60 kb upstream of another gene, *coiled-coil domain containing 91* (*CCDC91*). SNP rs1871152 was not correlated with the lead SNP in signal 1 or signal 2 ($r^2 = 0.01$ for rs7297051 and $r^2 = 0.03$ for rs805510). All lead SNPs for these three signals were associated with breast cancer risk at $P < 5 \times 10^{-8}$ in single-marker analyses (rs7297051 OR = 0.88, $P = 4 \times 10^{-28}$; rs805510 OR = 0.85, $P = 10^{-25}$; rs1871152 OR = 0.94, $P = 3 \times 10^{-8}$). No apparent heterogeneity in the ORs of the identified SNPs across the 40 studies in BCAC was found (all $P_{\text{heterogeneity}} > 0.75$). No statistically significant interactions between any pair of these three lead SNPs were found (all $P > 0.05$).

Using the lead SNP from each signal, rs805510, rs7297051 and rs1871152, we identified seven haplotypes with a frequency greater than 1 % (Table 2). The most common haplotype (frequency 51 %), carrying the major allele of each SNP, was used as the reference in the association analysis. The most statistically significant association was observed for the haplotype carrying the minor alleles at both signals 1 and 2 (TTA and TTG), while less pronounced yet significant associations were observed for individuals carrying the minor allele for signal 1 but not signal 2 (CTA and CTG), consistent with results for the independent association signals from the regression analyses. The evidence for signal 3 comes largely from the observation that the CCG haplotype, which carries the rare allele for signal 3 alone, was associated with reduced risk. The haplotype carrying only the minor allele in the lead SNP for signal 2 was too rare to evaluate. Stratified analyses revealed no evidence of any apparent heterogeneity in the association of these haplotypes with breast cancer risk by age at breast cancer diagnosis (age at diagnosis <45 vs ≥45 years).

The associations of the three SNPs did not vary appreciably by ER status (Additional file 2: Table S3). In an attempt to identify potential independent association signals that might have been missed in the analysis

Table 1 Independent association signals identified for breast cancer risk in the 12p11 locus in women of European ancestry

Signal	SNPs	Position (hg 19)	Alleles	EAF	LD (r^2)[b]	Univariate analysis		Conditional analysis		SNPs retained for functional annotation[e]
						Per-allele OR (95 % CI)[c]	*P*-trend	Per-allele OR (95 % CI)[d]	*P*-trend	
2	Index[a] rs10771399	28155080	G*/A	0.12	-	0.85 (0.83–0.88)	5×10^{-25}	-	-	-
1	rs7297051	28174817	T*/C	0.24	0.42	0.88 (0.86–0.90)	4×10^{-28}	0.92 (0.89–0.94)	3×10^{-9}	rs812020, chr12:28164044, rs2619434, rs2590275
2	rs805510	28139846	T*/C	0.12	0.88	0.85 (0.82–0.88)	10^{-25}	0.93 (0.89–0.96)	2×10^{-5}	74 SNPs[f]
3	rs1871152	28379826	G*/A	0.31	0.04	0.94 (0.92–0.96)	3×10^{-8}	0.96 (0.94–0.98)	2×10^{-4}	376 SNPs[g]

*Effect alleles. [a]Identified in the initial genome-wide association study conducted in women of European descent [1]. [b]Linkage disequilibrium (LD) with rs10771399 for women of European descent. [c]Adjusted for studies, and the top principal components and an additional principal component accounting for the Leuven Multidisciplinary Breast Centre (LMBC) study. [d]Included all three variants, and was adjusted for studies, and the top eight principal components as well as an additional principal component accounting for the LMBC study. [e]Associated single nucleotide polymorphisms (SNPs) with a likelihood ratio >1/100 relative to the lead SNP in each signal. [f]See Table S2 in Additional file 5. [g]See Table S2 in Additional file 5. *EAF* effect allele frequency in controls, *OR* odds ratio, *CI* confidence interval

Table 2 Associations between common haplotypes derived using lead single nucleotide polymorphisms and breast cancer risk in women of European ancestry

Haplotype	Overall breast cancer			Breast cancer (age at diagnosis <45 years)			Breast cancer (age at diagnosis ≥45 years)			$P_{heterogeneity}$ [b]
rs805510 - rs7297051-rs1871152	Frequency	OR (95 % CI)[a]	*P* value	Frequency	OR (95 % CI)[a]	*P* value	Frequency	OR (95 % CI)[a]	*P* value	
C-C-A	0.51	1.00 (Ref)	Ref	0.52	1.00 (Ref)	Ref	0.51	1.00 (Ref)	Ref	-
C-C-G	0.24	0.92 (0.89–0.95)	7×10^{-8}	0.22	0.94 (0.89–1.00)	0.04	0.24	0.92 (0.89–0.95)	4×10^{-7}	0.24
C-T-A	0.09	0.90 (0.87–0.95)	3×10^{-6}	0.09	0.96 (0.89–1.03)	0.28	0.09	0.90 (0.86–0.94)	4×10^{-7}	0.09
C-T-G	0.03	0.89 (0.82–0.96)	2×10^{-3}	0.03	0.85 (0.73–0.98)	0.02	0.03	0.89 (0.82–0.96)	3×10^{-3}	0.37
T-T-A	0.04	0.82 (0.77–0.88)	9×10^{-9}	0.04	0.76 (0.67–0.87)	5×10^{-5}	0.04	0.83 (0.76–0.85)	5×10^{-8}	0.19
T-T-G	0.07	0.79 (0.76–0.83)	3×10^{-23}	0.06	0.78 (0.71–0.85)	5×10^{-8}	0.07	0.81 (0.77–0.85)	3×10^{-18}	0.45
Rare	0.01	0.88 (0.79–0.99)	0.04	0.01	0.90 (0.72–1.13)	0.37	0.01	0.88 (0.78–0.99)	0.04	0.45

[a]Adjusted for studies and the top principal components. [b]*P* for heterogeneity between cases with age at diagnosis <45 years and ≥45 years. *Ref* reference

described above that included all breast cancer cases (Table 1), we conducted forward stepwise regression analyses separately for ER(+) and ER(-) cases. For the ER(+) breast cancer, the lead SNPs for signals 1 and 2 were identical to those found for all cases combined. For signal 3, however, a different lead SNP (rs7959641) was identified, which was moderately correlated with rs1871152, the lead SNP identified in the overall analysis ($r^2 = 0.28$) (Additional file 2: Table S3). The lead SNP for signal 3 in ER(-) cases is different from the SNP identified in all cases combined, but these two SNPs were highly correlated ($r^2 = 0.86$) (Additional file 2: Table S3).

Association results for *BRCA1* mutation carriers of European descent

Of the 2087 SNPs evaluated in the CIMBA among *BRCA1* mutation carriers of European descent, 234 were associated with breast cancer risk at $P < 0.0001$. The most significant association was found with rs113824616 (per-C allele HR 0.73, 95 % CI 0.64–0.82, $P = 1 \times 10^{-7}$; Table 3). The three lead SNPs identified in BCAC had similar associations, although the association was statistically significant at $P < 0.05$ in conditional analyses only for the lead SNPs of signals 1 and 3 (rs7297051 and rs1871152, respectively) (Additional file 3: Table S4). Meta-analysis of data from BCAC ER(-) cases and CIMBA showed that rs113824616 was associated with breast cancer risk after adjusting for rs7297051 (conditional $P = 7 \times 10^{-5}$, r^2 with rs10773199 = 0.40; Table 3). No additional independent signals were identified. We defined the association signal represented by SNP rs113824616 as signal 4.

Association results among women of East Asian ancestry

Of the 1801 SNPs evaluated, 118 were associated with breast cancer risk in women of East Asian ancestry ($P < 0.005$) (Fig. 1). The four lead SNPs in European descendants had a similar association with breast cancer risk in East Asian women, although the association was statistically significant at $P < 0.005$ only for the lead SNPs of signals 1 and 2 (rs7297051 and rs805510, respectively) (Additional file 4: Table S5). The MAFs for the lead SNPs of signals 1, 2 and 4 were similar to those in Europeans, but the MAF for signal 3 (rs1871152) was markedly lower in East Asians. In conditional regression analyses, only the association with signal 1 was independently statistically significant, perhaps due to the small sample size. The per-allele ORs did not differ materially from those in Europeans in the conditional analysis (data not shown).

The most significant association in Asians was with SNP rs2737455 (MAF = 0.17, per-major (T) allele OR = 1.16, 95 % CI 1.09–1.25, $P = 10^{-5}$). Among women of East Asian descent, this SNP was in high LD with the two lead SNPs for signals 1 and 2 identified in populations of European ancestry, rs7297051 ($r^2 = 0.67$) and rs805510 ($r^2 = 0.84$). This variant was also associated with breast cancer in women of European descent (per T-allele OR = 1.17, 95 % CI 1.14–1.21, $P = 5 \times 10^{-25}$). No additional independent signal was found on stepwise regression.

Association results for women of African ancestry

Of the 2949 SNPs evaluated in African descendants, 116 were statistically significantly associated with breast cancer risk at $P < 0.05$. The most significant association was with rs10843021 (MAF = 0.38, per-C allele OR = 1.22, 95 % CI 1.08–1.39, $P = 0.001$), which is located 60 kb downstream of the gene *PTHLH*. This SNP is not in LD with any of the lead SNPs identified for women of European or East Asian descent (all $r^2 < 0.02$). There was some evidence of association of this SNP with breast cancer risk in women of European descent ($P = 8 \times 10^{-5}$) but not in women of Asian descent ($P = 0.23$). None of the lead SNPs identified for women of European or East Asian descent were associated with breast cancer risk at

Table 3 Independent association signals in the meta-analysis of BCAC (ER-) and *BRCA1* mutation carriers from CIMBA

	SNPs	Position (hg 19)	Alleles	EAF	LD (r^2)[§]	Univariate analysis		Conditional analysis	
						Per-allele effect (95 % CI)[a]	*P*-trend	Per-allele effect (95 % CI)[b]	*P*-trend
Index[‡]	rs10771399	28155080	G*/A	0.10	-	0.86 (0.80–0.91)	3×10^{-6}	-	-
Meta-analysis of ER-negative cancer (BCAC + CIMBA)									
BCAC ER-									
Signal 1	rs7297051	28174817	T*/C	0.24	0.42	0.87 (0.83–0.91)	3×10^{-10}	0.89 (0.85–0.94)	1×10^{-5}
Signal 4	rs113824616	28184905	C*/T	0.05	0.40	0.75 (0.67–0.84)	5×10^{-7}	0.86 (0.76–0.98)	0.02
CIMBA *BRCA1* mutation carriers									
Signal 1	rs7297051	28174817	T*/C	0.23	0.37	0.89 (0.85–0.93)	3×10^{-7}	0.94 (0.90–0.98)	0.003
Signal 4	rs113824616	28184905	C*/T	0.04	0.49	0.73 (0.64–0.82)	1×10^{-7}	0.83 (0.74–0.93)	0.001

Effect for Breast Cancer Association Consortium (BCAC): odds ratio; effect for Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA) cohort: hazard ratio.
*Effect alleles. [a]Adjusted for studies, and the top principal components. [b]Included both variants, and adjusted for studies and the top principal components.
SNPs single nucleotide polymorphisms, *EAF* effect allele frequency in the or (BCAC) controls, *LD* linkage disequilibrium, *CI* confidence interval, *ER* estrogen receptor. [§]represents LD with the index SNP rs10771399. [‡]represented the index SNP, Identified in the initial genome-wide association study conducted in women of European descent [1]

$P < 0.05$ in African descendants, although the directions of the associations were consistent and the effect sizes did not differ significantly (Additional file 4: Table S5). The MAF of the index SNP rs10771399 (MAF = 0.04) was much lower in African descendants than that in Asian and European descendants ($P < 0.001$).

Functional annotation

To identify putative causal variants, we used data from European descendants to exclude any variants that had a likelihood ratio <1/100 relative to the most significantly associated SNP in each signal (33). Based on this threshold, four variants in signal 1, 74 variants in signal 2, 376 variants in signal 3, and 2 variants in signal 4 were retained as candidates for causal variants (Fig. 1a and Additional file 5: Table S2).

Using data from ENCODE, we found that the histone markers (H3K27Ac and H3K4Me3) were enriched in each signal (Fig. 1b). Using both ChIA-PET chromatin interaction data and Hi-C data from ENCODE, we identified multiple and dense chromosomal interactions of variants at signals 1 and 2 with the promoter region of *PTHLH* in MCF7 cells (Fig. 1c). There was some evidence of interaction of variants at signal 3 with the promoter of *PTHLH* (Fig. 1c).

Using maps of predicted enhancer regions produced by Hnisz et al. [26] and Corradin et al. [8], we found that multiple candidate variants were located in enhancer regions in mammary cell lines (Fig. 2). Using predicted enhancer-promoter interaction data in HMEC and MCF7 cell lines generated by He et al. [25] (Fig. 2), we identified two interacting genes of these enhancers, *CCDC91* and *PTHLH*.

We next overlaid these candidate variants to the transcription factor binding site maps generated from ENCODE. We identified rs812020 within signal 1, rs788463 and rs10843066 within signal 2, and rs10843110, rs56318627 and rs11049453 within signal 3 to be the most likely functional variants (Fig. 3a and b; Additional file 6: Table S6). These SNPs were within or close to binding sites of multiple breast cancer-related transcription factors. Furthermore, these SNPs were predicted to disrupt

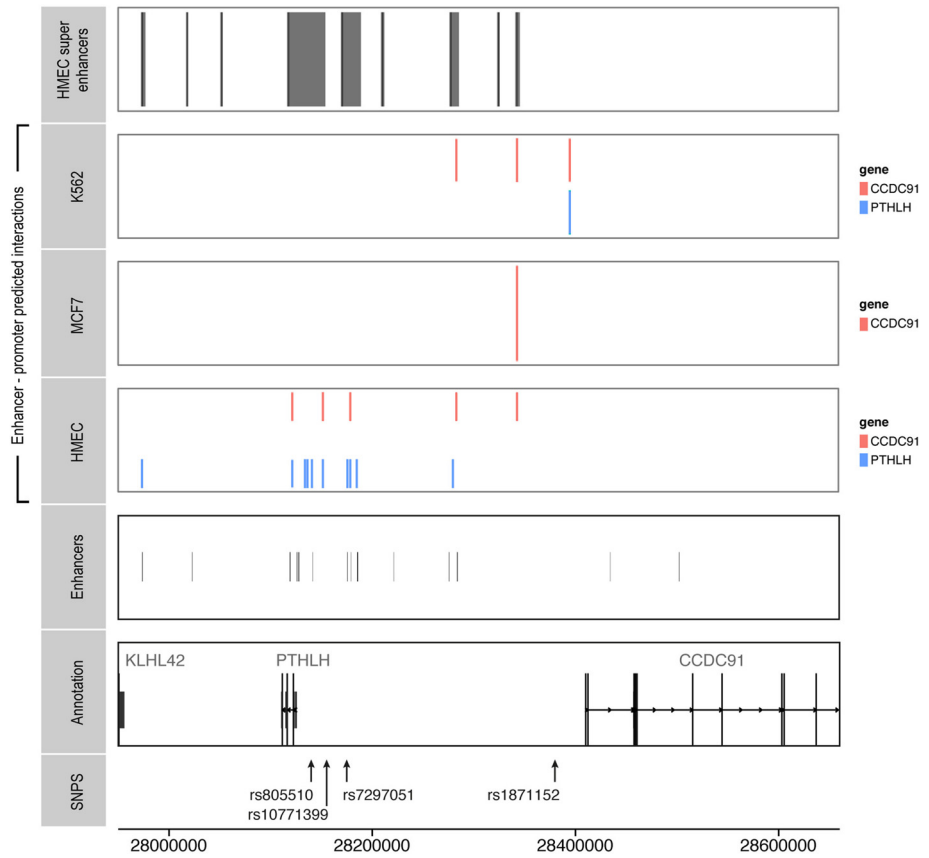

Fig. 2 Enhancer-promoter interaction data at 12p11. From *top* to *bottom*, enhancer locations as defined by Corradin et al. [8] and Hnisz et al. [26] are shown in human mammary epithelial cells (*HMEC*) cell lines. Enhancer-promoter (EP)-predicted interactions as defined by He et al. [25] are shown in K562, MCF7 and HMEC cells. Gene annotations and single nucleotide polymorphism (*SNP*) annotations. *Orange* EP interactions are those with the *coiled-coil domain containing 91* (*CCDC91*) gene; *blue* EP are those with *Parathyroid hormone-like hormone* (*PTHLH*)

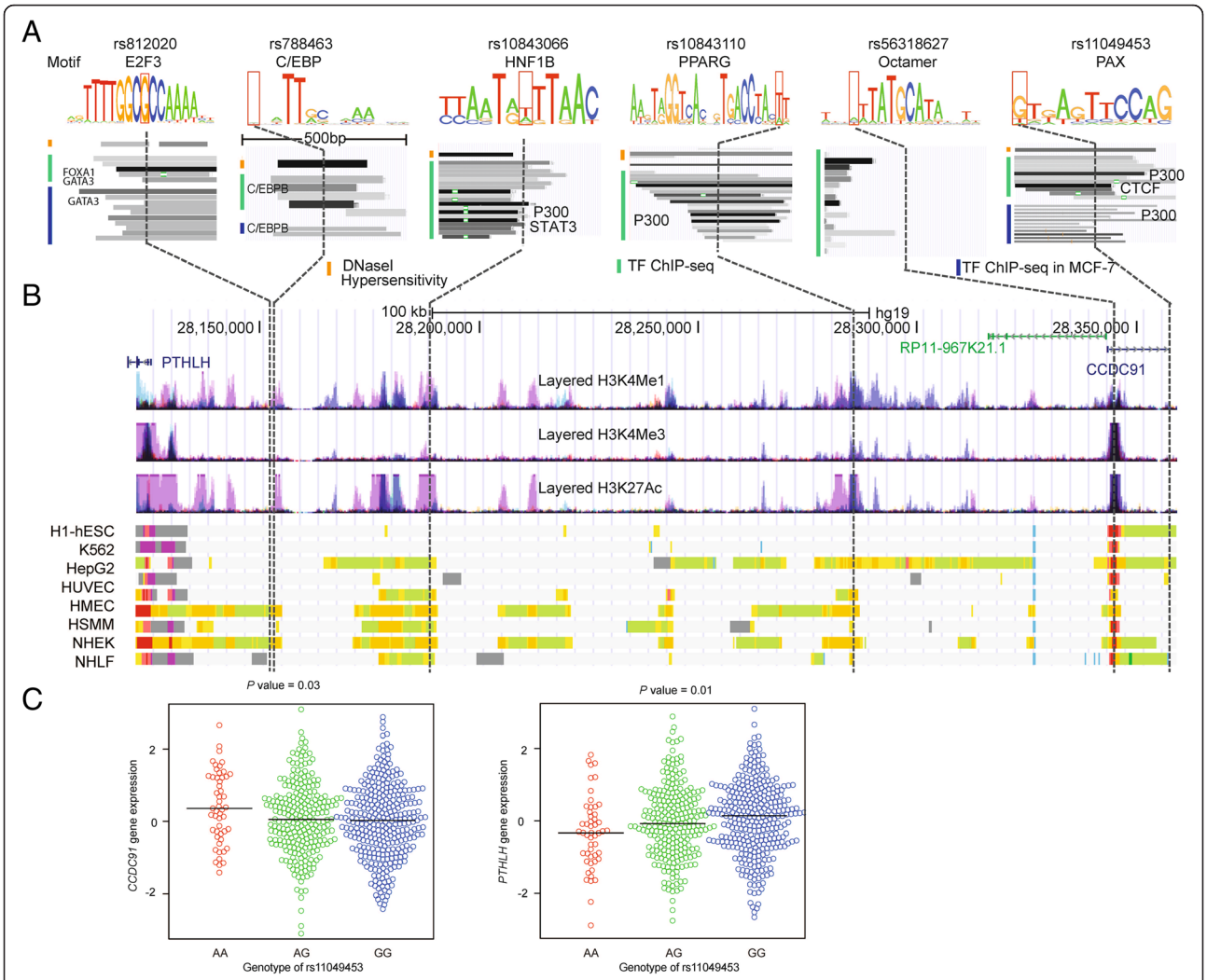

Fig. 3 Putative functional variants and association of rs11049453 with gene expression in breast tumor tissues. **a** Epigenetic signals of five potential functional variants. From *top* to *bottom*, lanes showing that those variants mapped to transcription factors predicted binding motifs, DNase I hypersensitivity sites and transcription factor ChIP-Seq binding peaks in the Encyclopedia of DNA Elements (ENCODE) cell lines and MCF7. The corresponding location of each variant is indicated by a dashed line. **b** Epigenetic landscape at the 12p11 locus for breast cancer risk. From *top* to *bottom*, RefSeq genes (*PTHLH* and *CCDC91*), layered H3K4Me1, H3K4Me3 and H3K27Ac histone modifications and annotation using chromatin states on the ENCODE cell lines. The signals of different layered histone modifications from the same ENCODE cell line are shown in the same color (the detailed color scheme for each ENCODE cell line is described in the UCSC genome browser). *Red* and *orange* in the chromatin states represent the active promoter and strong enhancer regions, respectively (the detailed color scheme of the chromatin states was described in the previous study [45]). **c** rs11049453 and the expression of *coiled-coil domain containing 91* (*CCDC91*) and *parathyroid hormone-like hormone* (*PTHLH*). The association of the genotypes and the expression level of each gene was evaluated by residual linear regression [29]. *bp* base pairs, *C/EBP* CCAAT/enhancer-binding protein, *E2F3* E2F transcription factor 3, *HNF1B* HNF1 homeobox B, *PPARG* peroxisome proliferator-activated receptor gamma, *PAX* paired box

the binding motifs recognized by transcription factors (Fig. 3a and b), suggesting a regulatory role. For example, in signal 1, rs812020 (per C-allele OR = 0.89, 95 % CI 0.87–0.91, $P = 2 \times 10^{-27}$) was annotated to a region bound by multiple key transcription factors for breast cancer, including GATA3 and FOXA1 (Fig. 3a and b). This SNP is predicted to disrupt the binding motif recognized by the transcription factor E2F3 and may change its binding affinity [32]. E2F3 has been found to increase centrosome amplification in mammary epithelial cells and regulate breast tumor development and metastasis [33]. In signal 3, SNP rs11049453 (per G-allele OR = 1.06, 95 % CI 1.04–1.08, $P = 9 \times 10^{-8}$) was in the binding site of transcription factors P300 and CTCF in MCF7 cell lines [31] (Fig. 3). It was also predicted to disrupt the binding motif of paired box (PAX) [33], which has been associated with the progression of breast cancer [34, 35]. No functional significance of the candidate variants in signal 4 was found.

To further explore the potential target genes, we performed eQTL analysis in both breast tumor and normal tissues. Using data on tumor tissues from TCGA, we found that rs10843110, rs56318627 and rs11049453 within signal 3 were associated with the expression of *PTHLH* at $P < 0.05$ and *CCDC91* at $P < 0.10$ (Additional file 7: Table S7). Among these highly correlated SNPs, the most significant association was found for rs11049453: the risk allele G of rs11049453 was associated with increased expression of *PTHLH* ($P = 0.01$) and decreased expression of *CCDC91* ($P = 0.03$, Fig. 3c). However, we did not find any statistically significant association for these six variants using data from adjacent normal breast tissues from METABRIC (all $P > 0.05$).

Discussion

Through a fine-scale mapping study at 12p11, we identified four independent association signals for breast cancer risk in women of European descent. It is of interest that the fourth signal was identified only through the meta-analysis of ER(-) breast cancer and *BRCA1* mutation carriers, suggesting that this signal may be more specific to ER(-) cancers. The associations of these signals were in general consistent in women of European and East Asian descent.

Multiple genetic studies have confirmed that a locus at 12p11 is associated with breast cancer risk [2, 4]. However, it remained unknown whether the observed association was due to a single or multiple causal variants at this locus. In this study, we demonstrated that there were at least four independent signals at 12p11, three 100 kb upstream of the gene *PTHLH* (signals 1, 2 and 4), and one 60 kbp upstream from the gene *CCDC91* (signal 3), suggesting that there may be multiple causal variants and multiple underlying mechanisms for the observed association at the 12p11 locus. Furthermore, we identified multiple candidate causal variants at each signal: four in signal 1, 74 in signal 2, 376 in signal 3 and 2 in signal 4. Using functional genomic data from ENCODE, we observed that multiple candidate functional variants located in enhancer regions, and identified *PTHLH* and *CCDC91* as the likely target genes for these enhancers. Using data on transcription factor binding, we identified six putative functional variants with strong evidence of regulation of gene expression. Among these six variants, we observed that the rs11049453 was significantly associated with the expression of *PTHLH* and *CCDC91*. However, we could not exclude the possibility that there were other functional variants and other target genes at this locus.

PTHLH encodes the protein PTHrP, which has intracrine, autocrine or paracrine action in most normal tissues; its downstream effects include promotion of growth and anti-apoptotic effects [36]. It is a cause of humoral hypercalcemia of malignancy [37], and is expressed in more than two thirds of breast tumor tissue samples [7, 38]. It has been shown to affect the regulation of tumor-related genes, and is thought to affect the proliferation and migration of breast cancer cells [39]. PTHrP plays an important role in the formation of osteolytic bone metastases in breast cancer through its action on osteoblasts to increase RANK-ligand and promote osteoclast formation [40]. It has been proposed that PTHrP may promote breast cancer tumorigenesis; however, previous studies had conflicting results [41]. Less is known about the function of the *CCDC91* gene, which is located approximately 232 kb from the *PTHLH* gene. *CCDC91* encodes a protein known as p56 accessory protein or GGA binding partner, which binds proteins, and facilitates the transportation of secreted proteins through the trans-Golgi network [42]. *CCDC91* is also expressed in a variety of cancer cell lines including MCF7 [43]. Using cBioPortal (http://www.cbioportal.org/public-portal/), we found that both *PTHLH* and *CCDC91* genes were altered in breast tumors and that there was a statistically significant co-occurrence of alternations (including mutations and copy number aberrations) in both genes (P for tendency towards co-occurrence <0.001). Together with our findings, these results suggest that there might be correlation between these two genes and that alterations in both genes might contribute concurrently to breast cancer susceptibility. Future studies evaluating both genes and their interrelationship are needed to elucidate the underlying mechanism.

Functional annotation data suggested that the functional variants underlying the observed association, mainly those in signal 2, are located in enhancer regions involved in the transcriptional regulation of *PTHLH* and *CCDC91* in the MCF7 and HMEC cells. Moreover, we did not find similar functional evidence for the same region in the K562 cells, which suggests that the regulatory effects might be context-specific. We identified multiple putative functional variants associated with transcriptional factors that have been found to be important for breast cancer, including GATA3, FOXA1, C/EBP, P300 and STAT3, and overlapped with binding motifs of transcriptional factors, including E2F3, C/EBP, HNF1B, PPARG and PAX. Despite strong evidence for altering the binding of transcription factor and regulating gene transcription, we found only one eQTL among these putative functional variants, which lies in signal 3, suggesting that the underlying functional variants might exert a more subtle regulatory effect on gene expressions than expected. Although we found strong genetic and epigenetic evidence for potential functional variants in

signals 1 and 2, we did not observe statistically significant association between these variants and the expression of *PTHLH* or *CCDC91*, or any other protein-coding genes within a flanking region of 500 kb for each variant. It is possible that the causal variants in these two signals might be involved in regulating non-coding genes or more distant genes. Future functional studies that comprehensively investigate the regulatory elements at these loci and their target genes will be needed to elucidate the molecular mechanisms.

The top risk variants identified in women of Asian and European ancestry were not associated with breast cancer risk in African descendants. It is possible that these top risk variants might not be correlated with the causal variants in African descendants due to their different LD structures. For example, the effect allele frequencies (EAFs) for the index SNP rs10773199 and the top risk variant rs805510 in African descendants were 0.04 and 0.45, respectively, and the EAFs for these two SNPs were similar in European descendants (EAF = 0.12 for both SNPs) and in East Asian descendants (EAF = 0.17 and 0.15, respectively), suggesting a distinct LD structure at this locus in African descendants. Similarly, the EAF for the SNP rs113824616 in African descendants (EAF = 0.01) was substantially lower than that in European descendants (EAF = 0.05). In addition, the sample size for African descendants included in this study was small and the power to detect the association of these variants was low. A previous fine-mapping study in African Americans with a larger sample size (3016 cases/2745 controls) than our study (1116 cases/932 controls) showed that rs10773199 is marginally associated with breast cancer risk (OR = 0.84, $P = 0.089$) [44], suggesting that there might be an association of the 12p11 locus with breast cancer risk in African descendants. Studies with a large sample size are needed to elucidate the association between this locus and breast cancer risk in African descendants.

To date this is the largest and most comprehensive fine-mapping study of the 12p11 region in relation to breast cancer risk. By using densely genotyped data from a very large number of cases and controls of European descent, we derived highly reliable estimates of the association between each common SNP and breast cancer risk in women of European descent. The sample size was relatively small for East Asian and African descendants, and associations with risk of overall breast cancer and molecular subtypes in these populations should be further evaluated in future larger studies.

Conclusions

Through fine-mapping of the 12p11 locus, we identified multiple independent association signals for breast cancer risk. We estimate that the four independent signals identified by this study explain approximately 1 % of the familial relative risk of breast cancer in populations of European ancestry, more than doubling the risk explained by the index SNP (0.4 %). Bioinformatics analyses revealed that these signals are mapped to enhancer regions that interact with the gene *PTHLH and CCDC91.* We identified putative functional variants that might contribute to the observed association. Our findings also suggest a possible interrelation between *PTHLH* and *CCDC91* in the etiology of breast cancer. Our study has expanded the knowledge of genetic risk associated with breast cancer at the 12p11 locus and provided clues for future functional characterization.

Additional files

Additional file 1: Table S1. Ethical committees that approved each study. (PDF 94 kb)

Additional file 2: Table S3. Independent association signals for risk of estrogen (ER)-positive and ER-negative breast cancer in European descendants. (PDF 47 kb)

Additional file 3: Table S4. Associations of independent signals for breast cancer risk for BRCA1 mutation carriers. (PDF 64 kb)

Additional file 4: Table S5. Associations of independent signals for breast cancer risk in women of East Asian and African descent. (PDF 66 kb)

Additional file 5: Table S2. List of the variants that were retained for further functional annotation in European descendants. (PDF 54 kb)

Additional file 6: Table S6. Putative functional SNPs identified using the ENCODE data. (PDF 50 kb)

Additional file 7: Table S7. Gene expression analysis for putative functional SNPs using 1,006 breast tumor samples in TCGA. (PDF 46 kb)

Abbreviations

BCAC, Breast Cancer Association Consortium; *BRCA1*, Breast cancer 1; C/EBP, CCAAT/enhancer-binding protein; *CCDC91*, Coiled-coil domain containing 91; ChIA-PET, chromatin interaction analysis by paired end tag; CI, confidence interval; CIMBA, Consortium of Investigators of Modifiers of *BRCA1/2*; CNA, copy number alterations; E2F3, E2F transcription factor 3; EAF, effect allele frequency; EM, expectation-maximum; ENCODE, Encyclopedia of DNA Elements; eQTL, expression quantitative trait loci; ER, estrogen receptor; FOXA1, forkhead box A1; GATA3, trans-acting T-cell-specific transcription factor GATA-3; GWAS, genome-wide association study; Hi-C, genome conformation capture; HMEC, human mammary epithelial cells; HNF1B, HNF1 homeobox B; HR, hazard ratio; iCOGS, Illumina iSelect genotyping array of the Collaborative Oncological Gene-environment Study; IMPUTEv2, IMPUTE version 2; LD, linkage disequilibrium; MAF, minor allele frequency; MCF7, Michigan Cancer Foundation-7; METABRIC, Molecular Taxonomy of Breast Cancer International Consortium; OR, odds ratio; PAX, paired box; PPARG, peroxisome proliferator-activated receptor gamma; PTHLH, parathyroid hormone-like hormone; QC, quality control; SNP, single nucleotide polymorphism; STAT3, signal transducer and activator of transcription 3; TCGA, The Cancer Genome Atlas

Acknowledgements

We thank all the individuals who took part in these studies and all the researchers, study staff, clinicians and other healthcare providers, technicians and administrative staff who have enabled this work to be carried out. In particular, we would like to thank Terence 'Jack' Martin (St Vincents Institute, Melbourne, Australia) for carefully reviewing this manuscript. COGS would not have been possible without the contributions of the following: Andrew Berchuck (OCAC), Rosalind A. Eeles, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), Lesley McGuffog, Andrew Lee, and Ed Dicks,

Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory, Javier Benitez, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit, and Daniel C. Tessier, Francois Bacot, Daniel Vincent, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, Sune F. Nielsen, Borge G. Nordestgaard, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility. ABCFS wishes to thank Maggie Angelakos, Judi Maskiell and Gillian Dite. ABCS wishes to thank Sten Cornelissen, Richard van Hien, Linde Braaf, Frans Hogervorst, Senno Verhoef, Laura van 't Veer, Emiel Rutgers, C Ellen van der Schoot and Femke Atsma. ABCTB wishes to thank Christine Clarke, Rosemary Balleine, Robert Baxter, Stephen Braye, Jane Carpenter, Jane Dahlstrom, John Forbes, Soon Lee, Debbie Marsh, Adrienne Morey, Nirmala Pathmanathan, Rodney Scott, Allan Spigelman, Nicholas Wilcken and Desmond Yip. Samples are made available to researchers on a non-exclusive basis. The ACP study wishes to thank the participants in the Thai Breast Cancer study. Special thanks also go to the Thai Ministry of Public Health (MOPH), doctors and nurses who helped with the data collection process. Finally, the ACP study would like to thank Dr. Prat Boonyawongviroj, the former Permanent Secretary of MOPH and Dr. Pornthep Siriwanarungsan, the Department Director-General of Disease Control who have supported the study throughout. BBCS wishes to thank Eileen Williams, Elaine Ryder-Mills and Kara Sargus. BIGGS wishes to thank Niall McInerney, Gabrielle Colleran, Andrew Rowan and Angela Jones. BOCS wishes to thank The Wellcome Trust Case Control Consortium (see the WTCCC website for a full list of contributing investigators). BSUCH wishes to thank Peter Bugert and Medical Faculty Mannheim. CGPS wishes to thank staff and participants of the Copenhagen General Population Study, and Dorthe Uldall Andersen, Maria Birna Arnadottir, Anne Bank and Dorthe Kjeldgård Hansenor for the excellent technical assistance. CNIO-BCS thanks Guillermo Pita, Charo Alonso, Daniel Herrero, Nuria Álvarez, Pilar Zamora, Primitiva Menendez and the Human Genotyping-CEGEN Unit (CNIO). The CTS Steering Committee includes Leslie Bernstein, Susan Neuhausen, James Lacey, Sophia Wang, Huiyan Ma, Yani Lu and Jessica Clague DeHart at the Beckman Research Institute of City of Hope, Dennis Deapen, Rich Pinder, Eunjung Lee, and Fred Schumacher at the University of Southern California, Pam Horn-Ross, Peggy Reynolds, Christina Clarke Dur and David Nelson at the Cancer Prevention Institute of California, and Hoda Anton-Culver, Argyrios Ziogas and Hannah Park at the University of California Irvine. DIETCOMPLY thanks the patients, nurses and clinical staff involved in the study. The University of Westminster Against Breast Cancer Research Unit acknowledges funding from the charity Against Breast Cancer (Registered Charity Number 1121258). ESTHER thanks Hartwig Ziegler, Sonja Wolf and Volker Hermann. GC-HBOC thanks Heide Hellebrand, Stefanie Engert and GC-HBOC (Supported by Deutsche Krebshilfe). GENICA thanks Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany (HB, Wing-Yee Lo, Christina Justenhoven), German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ) (HB), Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany (Yon-Dschun Ko, Christian Baisch), Institute of Pathology, University of Bonn, Germany (Hans-Peter Fischer), Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany (Ute Hamann), Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany (TB, Beate Pesch, Sylvia Rabstein, Anne Lotz) and Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany (Volker Harth). GESBC thanks Ursula Eilber. HABCS wishes to thank Michael Bremer. HEBCS wishes to thank Kirsimari Aaltonen, Karl von Smitten, Sofia Khan, Tuomas Heikkinen and Irja Erkkilä. HMBCS wishes to thank Peter Hillemanns, Hans Christiansen and Johann H. Karstens. KBCP wishes to thank Eija Myöhänen and Helena Kemiläinen. kConFab/AOCS wishes to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics and the clinical follow up study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia and the National Institute of Health (USA)) for their contributions to this resource, and the many families who contribute to kConFab. LAABC thanks all the study participants and the entire data collection team, especially Annie Fung and June Yashiki. LMBC wishes to thank Gilian Peuteman, Dominiek Smeets, Thomas Van Brussel and Kathleen Corthouts. MARIE wishes to thank Petra Seibold, Dieter Flesch-Janys, Judith Heinz, Nadia Obi, Alina Vrieling, Sabine Behrens, Ursula Eilber, Muhabbet Celik, Til Olchers and Stefan Nickels. MBCSG wishes to thank Daniela Zaffaroni of the Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Monica Barile and Irene Feroce of the Istituto Europeo di Oncologia (IEO) and the personnel of the Cogentech Cancer Genetic Test Laboratory. MSKCC thanks Marina Corines and Lauren Jacobs. MTLGEBCS would like to thank Martine Tranchant (CHU de Québec Research Center), Marie-France Valois, Annie Turgeon and Lea Heguy (McGill University Health Center, Royal Victoria Hospital, McGill University) for DNA extraction, sample management and skillful technical assistance. J.S. is Chairholder of the Canada Research Chair in Oncogenetics. MYBRCA wishes to thank Phuah Sze Yee, Peter Kang, Kang In Nee, Kavitta Sivanandan, Shivaani Mariapun, Yoon Sook-Yee, Daphne Lee, Teh Yew Ching and Nur Aishah Mohd Taib for DNA extraction and patient recruitment. NBCS wishes to thank Dr. Kristine Kleivi, PhD (K.G. Jebsen Centre for Breast Cancer Research, Institute of Clinical Medicine, University of Oslo, Oslo, Norway and Department of Research, Vestre Viken, Drammen, Norway), Dr. Lars Ottestad, MD (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Prof. Em. Rolf Kåresen, MD (Department of Oncology, Oslo University Hospital and Faculty of Medicine, University of Oslo, Oslo, Norway), Dr. Anita Langerød, PhD (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Ellen Schlichting, MD (Department for Breast and Endocrine Surgery, Oslo University Hospital Ullevaal, Oslo, Norway), Dr. Marit Muri Holmen, MD (Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway), Prof. Toril Sauer, MD (Department of Pathology at Akershus University hospital, Lørenskog, Norway), Dr. Vilde Haakensen, MD (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Olav Engebråten, MD (Institute for Clinical Medicine, Faculty of Medicine, University of Oslo and Department of Oncology, Oslo University Hospital, Oslo, Norway), Prof. Bjørn Naume, MD (Division of Cancer Medicine and Radiotherapy, Department of Oncology, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Cecile E. Kiserud, MD (National Advisory Unit on Late Effects after Cancer Treatment, Department of Oncology, Oslo University Hospital, Oslo, Norway and Department of Oncology, Oslo University Hospital, Oslo, Norway), Dr. Kristin V. Reinertsen, MD (National Advisory Unit on Late Effects after Cancer Treatment, Department of Oncology, Oslo University Hospital, Oslo, Norway and Department of Oncology, Oslo University Hospital, Oslo, Norway), Assoc. Prof. Åslaug Helland, MD (Department of Genetics, Institute for Cancer Research and Department of Oncology, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Margit Riis, MD (Dept of Breast- and Endocrine Surgery, Oslo University Hospital, Ullevål, Oslo, Norway), Dr. Ida Bukholm, MD (Department of Breast-Endocrine Surgery, Akershus University Hospital, Oslo, Norway and Department of Oncology, Division of Cancer Medicine, Surgery and Transplantation, Oslo University Hospital, Oslo, Norway), Prof. Per Eystein Lønning, MD (Section of Oncology, Institute of Medicine, University of Bergen and Department of Oncology, Haukeland University Hospital, Bergen, Norway), and Grethe I. Grenaker Alnæs, M.Sc. (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway). NBHS wishes to thank study participants and research staff for their contributions and commitment to this study. OBCS thanks Meeri Otsukka and Kari Mononen. OFBCR wishes to thank Teresa Selander and Nayana Weerasooriya. ORIGO wishes to thank E. Krol-Warmerdam and J. Blom for patient accrual, administering questionnaires and managing clinical information. The LUMC survival data were retrieved from the Leiden hospital-based cancer registry system (ONCDOC) with the help of Dr. J. Molenaar. PBCS wishes to thank Louise Brinton, Mark Sherman, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao and Michael Stagner. pKARMA wishes to thank the Swedish Medical Research Counsel. RBCS wishes to thank Petra Bos, Jannet Blom, Ellen Crepin, Elisabeth Huijskens, Annette Heemskerk and the Erasmus MC Family Cancer Clinic. SASBAC thanks the Swedish Medical Research Counsel. SBCGS wishes to thank the study participants and research staff for their contributions and commitment to this study. SBCS wishes to thank Sue Higham, Helen Cramp, Ian Brock, Sabapathy Balasubramanian and Dan Connley. SEARCH wishes to thank the SEARCH and EPIC teams. SGBCC wishes to thank the participants and research coordinator Kimberley Chua. SKKDKFZS wishes to thank all study participants, clinicians, family doctors, researchers and technicians for their contributions and commitment to this study. TNBCC wishes to thank Robert Pilarski and Charles Shapiro who were instrumental in the formation

of the OSU Breast Cancer Tissue Bank. We thank the Human Genetics Sample Bank for processing of samples and providing OSU Columbus area control samples. UCIBCS wishes to thank Irene Masunaka. UKBGS wishes to thank Breakthrough Breast Cancer and the Institute of Cancer Research for support and funding of the Breakthrough Generations Study, and the study participants, study staff, and the doctors, nurses and other health care providers and health information sources who have contributed to the study. We acknowledge NHS funding to the Royal Marsden/ICR NIHR Biomedical Research Centre.

CIMBA studies acknowledge the following. BCFR-AU wishes to thank Maggie Angelakos, Judi Maskiell, Gillian Dite and Helen Tsimiklis. BCFR-NY wishes to thank members and participants in the New York site of the Breast Cancer Family Registry for their contributions to the study. BCFR-ON wishes to thank members and participants in the Ontario Familial Breast Cancer Registry for their contributions to the study. BFBOCC-LT acknowledges Vilius Rudaitis, Laimonas Griškevičius, Ramūnas Janavičius. BFBOCC-LV acknowledges Drs. Janis Eglitis, Anna Krilova and Aivars Stengrevics. BMBSA wishes to thank the families who contribute to the BMBSA study. BRICOH wishes to thank Yuan Chun Ding and Linda Steele for their work in participant enrollment and biospecimen and data management. CNIO wishes to thank Alicia Barroso, Rosario Alonso and Guillermo Pita for their assistance. The CONSIT TEAM wishes to thank Maria Grazia Tibiletti of the Ospedale di Circolo-Università dell'Insubria,Varese, Italy, Giulietta Scuvera of the Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Gabriele Capone of the University of Florence, Florence, Italy, Alessandra Viel and Riccardo Dolcetti of the CRO Aviano National Cancer Institute, Aviano, Italy, Aline Martayan of the Istituto Nazionale Tumori Regina Elena, Rome, Italy, Stefania Tommasi e Brunella Pilato of the Istituto Nazionale Tumori "Giovanni Paolo II", Bari, Italy, Liliana Varesco of the IRCCS, AOU San Martino - IST Istituto Nazionale per la Ricerca sul Cancro, Genoa, Italy, Laura Cortesi of the University of Modena and Reggio Emilia, Modena, Italy and Laura Ottini of the University La Sapienza, Rome, Italy. FCCC thanks Ms. JoEllen Weaver and Dr. Betsy Bove for their technical support. Genetic Modifiers of Cancer Risk in BRCA1/2 Mutation Carriers (GEMO) study: National Cancer Genetics Network (UNICANCER Genetic Group), France. We wish to thank all the GEMO collaborating groups for their contribution to this study. GEMO Collaborating Centers are: Coordinating Centres, Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Hospices Civils de Lyon - Centre Léon Bérard & Equipe, Génétique du cancer du sein, Centre de Recherche en Cancérologie de Lyon: Olga Sinilnikova†, Sylvie Mazoyer, Francesca Damiola, Laure Barjhoux, Carole Verny-Pierre, Mélanie Léone, Nadia Boutry-Kryza, Alain Calender, Sophie Giraud; and Service de Génétique Oncologique, Institut Curie, Paris: Dominique Stoppa-Lyonnet, Marion Gauthier-Villars, Bruno Buecher, Claude Houdayer, Etienne Rouleau, Lisa Golmard, Agnès Collet, Virginie Moncoutier, Cédrick Lefol, Muriel Belotti, Antoine de Pauw, Camille Elan, Catherine Nogues, Emmanuelle Fourme, Anne-Marie Birot. Institut Gustave Roussy, Villejuif: Brigitte Bressac-de-Paillerets, Olivier Caron, Marine Guillaud-Bataille. Centre Jean Perrin, Clermont–Ferrand: Yves-Jean Bignon, Nancy Uhrhammer. Centre Léon Bérard, Lyon: Christine Lasset, Valérie Bonadona, Sandrine Handallou. Centre François Baclesse, Caen: Agnès Hardouin, Pascaline Berthet, Dominique Vaur, Laurent Castera. Institut Paoli Calmettes, Marseille: Hagay Sobol, Violaine Bourdon, Tetsuro Noguchi, Audrey Remenieras, François Eisinger. CHU Arnaud-de-Villeneuve, Montpellier: Isabelle Coupier, Pascal Pujol. Centre Oscar Lambret, Lille: Jean-Philippe Peyrat, Joëlle Fournier, Françoise Révillion, Philippe Vennin†, Claude Adenis. Centre Paul Strauss, Strasbourg: Danièle Muller, Jean-Pierre Fricker. Institut Bergonié, Bordeaux: Emmanuelle Barouk-Simonet, Françoise Bonnet, Virginie Bubien, Nicolas Sevenet, Michel Longy. Institut Claudius Regaud, Toulouse: Christine Toulas, Rosine Guimbaud, Laurence Gladieff, Viviane Feillel. CHU Grenoble: Dominique Leroux, Hélène Dreyfus, Christine Rebischung, Magalie Peysselon. CHU Dijon: Fanny Coron, Laurence Faivre. CHU St-Etienne: Fabienne Prieur, Marine Lebrun, Caroline Kientz. Hôtel Dieu Centre Hospitalier, Chambéry: Sandra Fert Ferrer. Centre Antoine Lacassagne, Nice: Marc Frénay. CHU Limoges: Laurence Vénat-Bouvet. CHU Nantes: Capucine Delnatte. CHU Bretonneau, Tours: Isabelle Mortemousque. Groupe Hospitalier Pitié-Salpétrière, Paris: Florence Coulet, Chrystelle Colas, Florent Soubrier, Mathilde Warcoin. CHU Vandoeuvre-les-Nancy: Johanna Sokolowska, Myriam Bronner. CHU Besançon: Marie-Agnès Collonge-Rame, Alexandre Damette. Creighton University, Omaha, USA: Henry T. Lynch, Carrie L. Snyder. G-FAST wishes to thank the technical support of Ilse Coene en Brecht Crombez. GOG: this study was supported by National Cancer Institute grants to the Gynecologic Oncology Group (GOG) Administrative Office and Tissue Bank (CA 27469), the GOG Statistical and Data Center (CA 37517), and the GOG Cancer Prevention and Control Committee (CA 101165). Drs. Greene, Mai and Savage were supported by funding from the Intramural Research Program, NCI. HCSC was supported by a grant RD12/00369/0006 and 12/00539 from ISCIII (Spain), partially supported by European Regional Development FEDER funds. We acknowledge Alicia Tosar for her technical assistance. HCSC wishes to thank the technical support of Ilse Coene en Brecht Crombez. HEBCS would like to thank Dr. Kristiina Aittomäki, Taru A. Muranen, Drs. Carl Blomqvist and Kirsimari Aaltonen and RNs Irja Erkkilä and Virpi Palola for their help with the HEBCS data and samples. The Hereditary Breast and Ovarian Cancer Research Group Netherlands (HEBON) consists of the following Collaborating Centers: Coordinating center: Netherlands Cancer Institute, Amsterdam, NL: M.A. Rookus, F.B.L. Hogervorst, F.E. van Leeuwen, S. Verhoef, M.K. Schmidt, N.S. Russell, J.L. de Lange, R. Wijnands; Erasmus Medical Center, Rotterdam, NL: J.M. Collée, A.M.W. van den Ouweland, M.J. Hooning, C. Seynaeve, C.H.M. van Deurzen, I.M. Obdeijn; Leiden University Medical Center, NL: C.J. van Asperen, J.T. Wijnen, R.A.E.M. Tollenaar, P. Devilee, T.C.T.E.F. van Cronenburg; Radboud University Nijmegen Medical Center, NL: C.M. Kets, A.R. Mensenkamp; University Medical Center Utrecht, NL: M.G.E.M. Ausems, R.B. van der Luijt, C.C. van der Pol; Amsterdam Medical Center, NL: C.M. Aalfs, T.A.M. van Os; VU University Medical Center, Amsterdam, NL: J.J.P. Gille, Q. Waisfisz, H.E.J. Meijers-Heijboer; University Hospital Maastricht, NL: E.B. Gómez-Garcia, M.J. Blok; University Medical Center Groningen, NL: J.C. Oosterwijk, A.H. van der Hout, M.J. Mourits, G.H. de Bock; The Netherlands Foundation for the detection of hereditary tumours, Leiden, NL: H.F. Vasen; The Netherlands Comprehensive Cancer Organization (IKNL): S. Siesling, J.Verloop; The Dutch Pathology Registry (PALGA): L.I.H. Overbeek. The HEBON study is supported by the Dutch Cancer Society grants NKI1998-1854, NKI2004-3088, NKI2007-3756, the Netherlands Organization of Scientific Research grant NWO 91109024, the Pink Ribbon grants 110005 and 2014-187.WO76, the BBMRI grant NWO 184.021.007/CP46 and the Transcan grant JTC 2012 Cancer 12-054. HEBON thanks the registration teams of IKNL and PALGA for part of the data collection. HRBCP wishes to thank Hong Kong Sanatoriuma and Hospital for their continual support. HUNBOCS wishes to thank the Hungarian Breast and Ovarian Cancer Study Group members (Janos Papp, Tibor Vaszko, Aniko Bozsik, Timea Pocza, Judit Franko, Maria Balogh, Gabriella Domokos and Judit Ferenczi, Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary) and the clinicians and patients for their contributions to this study. HVH wishes to thank the Oncogenetics Group (VHIO), and the High Risk and Cancer Prevention Unit of the University Hospital Vall d'Hebron. ICO wishes to thank the ICO Hereditary Cancer Program team led by Dr. Gabriel Capella. INHERIT would like to thank Dr. Martine Dumont, Martine Tranchant for sample management and skillful technical assistance. J.S. is Chairholder of the Canada Research Chair in Oncogenetics. J.S. and P.S. were part of the QC and Genotyping coordinating group of iCOGS (http://ccge.medschl.cam.ac.uk/research/consortia/icogs/) (BCAC and CIMBA). IPOBCS wishes to thank Drs. Ana Peixoto, Catarina Santos, Patrícia Rocha and Pedro Pinto for their skillful contribution to the study. KCONFAB wishes to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the clinical follow up study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia, and the National Institute of Health (USA)) for their contributions to this resource, and the many families who contribute to kConFab. The Modifier Study of Quantitative Effects on Disease (MODSQUAD) acknowledges ModSQuaD members Csilla Szabo (National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA); Lenka Foretova and Eva Machackova (Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute and MF MU, Brno, Czech Republic) and Michal Zikan, Petr Pohlreich and Zdenek Kleibl (Oncogynecologic Center and Department of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, Prague, Czech Republic). MSKCC wishes to thank Anne Lincoln and Lauren Jacobs. NICCC wishes to thank the NICCC National Familial Cancer Consultation Service team led by Sara Dishon, the laboratory team led by Dr. Flavio Lejbkowicz, and the research field operations team led by Dr. Mila Pinchev. NRG Oncology thanks the investigators of the Australia New Zealand NRG Oncology group. OCGN wishes to thank members and participants in the Ontario Cancer Genetics Network for their contributions to the study. SEABASS would like to thank Yip Cheng Har, Nur Aishah Mohd Taib, Phuah Sze Yee, Norhashimah Hassan and all the research nurses, research

assistants and doctors involved in the MyBrCa Study for assistance in patient recruitment, data collection and sample preparation. In addition, we thank Philip Iau, Sng Jen-Hwei and Sharifah Nor Akmal for contributing samples from the Singapore Breast Cancer Study and the HUKM-HKL study respectively. The Malaysian Breast Cancer Genetic Study is funded by research grants from the Malaysian Ministry of Science, Technology and Innovation, Ministry of Higher Education (UM.C/HIR/MOHE/06) and charitable funding from the Cancer Research Initiatives Foundation. The SMC team wishes to acknowledge the assistance of the Meirav Comprehensive breast cancer center team at the Sheba Medical Center for assistance in this study. Swedish scientists participating as SWE-BRCA collaborators are: Åke Borg, Håkan Olsson, Helena Jernström, Karin Henriksson, Katja Harbst, Maria Soller and Ulf Kristoffersson from Lund University and University Hospital; Anna Öfverholm, Margareta Nordling, Per Karlsson and Zakaria Einbeigi from Gothenburg Sahlgrenska University Hospital; Anna von Wachenfeldt, Annelie Liljegren, Annika Lindblom, Brita Arver, Gisela Barbany Bustinza and Johanna Rantala from Stockholm and Karolinska University Hospital; Beatrice Melin, Christina Edwinsdotter Ardnor and Monica Emanuelsson from Umeå University Hospital; Hans Ehrencrona, Maritta Hellström Pigg and Richard Rosenquist from Uppsala University; and Marie Stenmark-Askmalm and Sigrun Liedgren from Linköping University Hospital. UCHICAGO wishes to thank Cecilia Zvocec, Qun Niu, the physicians, genetic counselors, research nurses and staff of the Cancer Risk Clinic for their contributions to this resource and the many families who contribute to our program. UCLA thanks Joyce Seldon, MSGC and Lorna Kwan, MPH, for assembling the data for this study. UCSF would like to thank Dr. Robert Nussbaum and the following genetic counselors for participant recruitment: Beth Crawford, Kate Loranger, Julie Mak, Nicola Stewart, Robin Lee, Amie Blanco and Peggy Conrad, and Ms. Salina Chan for data management. UKFOCR thanks Paul Pharoah, Simon Gayther, Susan Ramus, Carole Pye, Patricia Harrington and Eva Wozniak for their contributions towards the UKFOCR. UPENN wishes to thank the Breast Cancer Research Foundation, Susan G. Komen Foundation for the cure and Basser Research Center for BRCA. VFCTG wishes to thank Geoffrey Lindeman, Marion Harris and Martin Delatycki of the Victorian Familial Cancer Trials Group. VFCTG also thanks Sarah Sawyer and Rebecca Driessen for assembling these data and Ella Thompson for performing all DNA amplification. Funding was as follows: the work conducted for this project at Vanderbilt Epidemiology Center is supported in part by NIH grant R37CA070867 and endowment funds for the Ingram Professorship and Anne Potter Wilson Chair. BCAC is funded by Cancer Research UK (C1287/A10118, C1287/A12014) and by the European Community Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2–2009-223175) (COGS). Funding for the iCOGS infrastructure came from the European Community Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund. The Australian Breast Cancer Family Study (ABCFS) was supported by grant UM1 CA164920 from the National Cancer Institute (USA). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia) and the Victorian Breast Cancer Research Consortium. J.L.H. is a National Health and Medical Research Council (NHMRC) Australia Fellow and a Victorian Breast Cancer Research Consortium Group Leader. M.C.S. is a NHMRC Senior Research Fellow and a Victorian Breast Cancer Research Consortium Group Leader. The ABCS study was supported by the Dutch Cancer Society (grants NKI 2007-3839; 2009 4363); BBMRI-NL, which is a research infrastructure financed by the Dutch government (NWO 184.021.007); and the Dutch National Genomics Initiative. The Australian Breast Cancer Tissue Bank is generously supported by the National Health and Medical Research Council of Australia, The Cancer Institute NSW and the National Breast Cancer Foundation. The ACP study is funded by the Breast Cancer Research Trust, UK. The work of the BBCC was partly funded by ELAN-Fond of the University Hospital of Erlangen. The BBCS is funded by Cancer Research UK and Breakthrough Breast Cancer and acknowledges NHS funding to the NIHR Biomedical Research Centre, and the National Cancer Research Network (NCRN). ES is supported by NIHR Comprehensive Biomedical Research Centre, Guy's & St. Thomas' NHS Foundation Trust in partnership with King's College London, UK. IT is supported by the Oxford Biomedical Research Centre. BOCS is supported by funds from Cancer Research UK (C8620/A8372 and C8620/A8857), a US Military Acquisition (ACQ) Activity, Era of Hope Award (W81XWH-05-1-0204) and the Institute of Cancer Research (UK). C.T. is funded by a Medical Research Council (UK) Clinical Research Fellowship. BOCS acknowledges NHS funding to the Royal Marsden/Institute of Cancer Research NIHR Specialist Cancer Biomedical Research Centre. The BSUCH study was supported by the Dietmar-Hopp Foundation, the Helmholtz Society and the German Cancer Research Center (DKFZ). The CECILE study was funded by Fondation de France, Institut National du Cancer (INCa), Ligue Nationale contre le Cancer, Ligue contre le Cancer Grand Ouest, Agence Nationale de Sécurité Sanitaire (ANSES) and Agence Nationale de la Recherche (ANR). The CGPS was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council and Herlev Hospital. The CNIO-BCS was supported by the Instituto de Salud Carlos III, the Red Temática de Investigación Cooperativa en Cáncer and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923 and PI12/00070). The CTS was initially supported by the California Breast Cancer Act of 1993 and the California Breast Cancer Research Fund (contract 97-10500) and is currently funded through the National Institutes of Health (R01 CA77398). Collection of cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885. HAC receives support from the Lon V Smith Foundation (LVS39420). The University of Westminster curates the DietCompLyf database created by and funded by Against Breast Cancer Registered Charity No. 1121258. The ESTHER study was supported by a grant from the Baden Württemberg Ministry of Science, Research and Arts. Additional cases were recruited in the context of the VERDI study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe). The GC-HBOC was supported by Deutsche Krebshilfe (107 352). The GENICA was funded by the Federal Ministry of Education and Research (BMBF) Germany grants 01KW9975/5, 01KW9976/8, 01KW9977/0 and 01KW0114, the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, and the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, and Johanniter Krankenhaus, Bonn, Germany. The GESBC was supported by the Deutsche Krebshilfe e. V. [70492] and the German Cancer Research Center (DKFZ). The HABCS study was supported by an intramural grant from Hannover Medical School. The HEBCS was financially supported by the Helsinki University Central Hospital Research Fund, Academy of Finland (266528), the Finnish Cancer Society, The Nordic Cancer Union and the Sigrid Juselius Foundation. The HERPACC was supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports, Culture and Technology of Japan, by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from Ministry of Health, Labor and Welfare of Japan, by Health and Labor Sciences Research Grants for Research on Applying Health Technology from the Ministry Health, Labor and Welfare of Japan, National Cancer Center Research and Development Fund and Grant form Takeda Health Foundation. The HMBCS was supported by a grant from the Friends of Hannover Medical School and by the Rudolf Bartling Foundation. The HUBCS was supported by a grant from the German Federal Ministry of Research and Education (RUS08/017). Financial support for KARBAC was provided through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Cancer Society, The Gustav V Jubilee foundation and and Bert von Kantzows foundation. The KBCP was financially supported by the special Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations and strategic funding of the University of Eastern Finland. kConFab is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. Financial support for the AOCS was provided by the United States Army Medical Research and Materiel Command (DAMD17-01-1-0729), Cancer Council Victoria, Queensland Cancer Fund, Cancer Council New South Wales, Cancer Council South Australia, The Cancer

Foundation of Western Australia, Cancer Council Tasmania and the National Health and Medical Research Council of Australia (NHMRC; 400413, 400281, 199600). G.C.T. and P.W. are supported by the NHMRC. RB was a Cancer Institute NSW Clinical Research Fellow. LAABC is supported by grants (1RB-0287, 3 PB-0102, 5 PB-0018, 10 PB-0098) from the California Breast Cancer Research Program. Incident breast cancer cases were collected by the USC Cancer Surveillance Program (CSP) which is supported under subcontract by the California Department of Health. The CSP is also part of the National Cancer Institute Division of Cancer Prevention and Control Surveillance, Epidemiology, and End Results Program, under contract number N01CN25403. LMBC is supported by the Stichting tegen Kanker (232-2008 and 196-2010). Diether Lambrechts is supported by the FWO and the KULPFV/10/016-SymBioSysII. The MARIE study was supported by the Deutsche Krebshilfe e.V. (70-2892-BR I, 106332, 108253, 108419), the Hamburg Cancer Society, the German Cancer Research Center (DKFZ) and the Federal Ministry of Education and Research (BMBF) Germany (01KH0402). MBCSG is supported by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects "5x1000"). The MCBCS was supported by the NIH grants CA128978, CA116167, CA176785 an NIH Specialized Program of Research Excellence (SPORE) in Breast Cancer [CA116201], and the Breast Cancer Research Foundation and a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation. MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR). The MEC was supported by NIH grants CA63464, CA54281, CA098758 and CA132839. MSKCC is supported by grants from the Breast Cancer Research Foundation and Robert and Kate Niehaus Clinical Cancer Genetics Initiative. The work of MTLGEBCS was supported by the Quebec Breast Cancer Foundation, the Canadian Institutes of Health Research for the CIHR Team in Familial Risks of Breast Cancer program, grant number CRN-87521 and the Ministry of Economic Development, Innovation and Export Trade, grant number PSR-SIIRI-701. MYBRCA is funded by research grants from the Malaysian Ministry of Science, Technology and Innovation (MOSTI), Malaysian Ministry of Higher Education (UM.C/HIR/MOHE/06) and Cancer Research Initiatives Foundation (CARIF). Additional controls were recruited by the Singapore Eye Research Institute, which was supported by a grant from the Biomedical Research Council (BMRC08/1/35/19/550), Singapore and the National medical Research Council, Singapore (NMRC/CG/SERI/2010). The NBCS has received funding from the K.G. Jebsen Centre for Breast Cancer Research, the Research Council of Norway grant 193387/V50 (to A-L Børresen-Dale and V.N. Kristensen) and grant 193387/H10 (to A-L Børresen-Dale and V.N. Kristensen), South Eastern Norway Health Authority (grant 39346 to A-L Børresen-Dale) and the Norwegian Cancer Society (to A-L Børresen-Dale and V.N. Kristensen). The NBHS was supported by NIH grant R01CA100374. Biological sample preparation was conducted using the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The Northern California Breast Cancer Family Registry (NC-BCFR) was supported by grant UM1 CA164920 from the National Cancer Institute (USA). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centres in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR. The NHS was funded by NIH grant CA87969. The OBCS was supported by research grants from the Finnish Cancer Foundation, the Academy of Finland (grant number 250083, 122715 and Center of Excellence grant number 251314), the Finnish Cancer Foundation, the Sigrid Juselius Foundation, the University of Oulu, the University of Oulu Support Foundation and the special Governmental EVO funds for Oulu University Hospital-based research activities. The Ontario Familial Breast Cancer Registry (OFBCR) was supported by grant UM1 CA164920 from the National Cancer Institute (USA). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centres in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR. The ORIGO study was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16). The PBCS was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The pKARMA study was supported by Märit and Hans Rausings Initiative Against Breast Cancer. The RBCS was funded by the Dutch Cancer Society (DDHK 2004-3124, DDHK 2009-4318). The SASBAC study was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the US National Institute of Health (NIH) and the Susan G. Komen Breast Cancer Foundation. The SBCGS was supported primarily by NIH grants R01CA64277, R01CA148667 and R37CA70867. Biological sample preparation was conducted using the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The scientific development and funding of this project were, in part, supported by the Genetic Associations and Mechanisms in Oncology (GAME-ON) Network U19 CA148065. The SBCS was supported by Yorkshire Cancer Research S295, S299, S305PA and Sheffield Experimental Cancer Medicine Centre. The SCCS is supported by a grant from the National Institutes of Health (R01 CA092447). Data on SCCS cancer cases used in this publication were provided by the Alabama Statewide Cancer Registry; Kentucky Cancer Registry, Lexington, KY; Tennessee Department of Health, Office of Cancer Surveillance; Florida Cancer Data System; North Carolina Central Cancer Registry, North Carolina Division of Public Health; Georgia Comprehensive Cancer Registry; Louisiana Tumor Registry; Mississippi Cancer Registry; South Carolina Central Cancer Registry; Virginia Department of Health, Virginia Cancer Registry; Arkansas Department of Health, Cancer Registry, 4815 W. Markham, Little Rock, AR 72205. The Arkansas Central Cancer Registry is fully funded by a grant from National Program of Cancer Registries, Centers for Disease Control and Prevention (CDC). Data on SCCS cancer cases from Mississippi were collected by the Mississippi Cancer Registry which participates in the National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention (CDC). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the CDC or the Mississippi Cancer Registry. SEARCH is funded by a programme grant from Cancer Research UK (C490/A10124) and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. SEBCS was supported by the BRL (Basic Research Laboratory) programme through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2012-0000347). SGBCC is funded by the NUS start-up Grant, National University Cancer Institute Singapore (NCIS) Centre Grant and the NMRC Clinician Scientist Award. Additional controls were recruited by the Singapore Consortium of Cohort Studies-Multi-ethnic cohort (SCCS-MEC), which was funded by the Biomedical Research Council, grant number 05/1/21/19/425. SKKDKFZS is supported by the DKFZ. The SZBCS was supported by grant PBZ_KBN_122/P05/2004. The TBCS was funded by The National Cancer Institute Thailand. The TNBCC was supported by a Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), a grant from the Breast Cancer Research Foundation, a generous gift from the David F. and Margaret T. Grohne Family Foundation, the Stefanie Spielman Breast Cancer fund and the OSU Comprehensive Cancer Center, the Hellenic Cooperative Oncology Group research grant (HR R_BG/04) and the Greek General Secretary for Research and Technology (GSRT) Program, Research Excellence II, the European Union (European Social Fund – ESF), and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - ARISTEIA. The TWBCS is supported by the Taiwan Biobank project of the Institute of Biomedical Sciences, Academia Sinica, Taiwan. The UCIBCS component of this research was supported by the NIH (CA58860, CA92044) and the Lon V Smith Foundation (LVS39420). The UKBGS is funded by Breakthrough Breast Cancer and the Institute of Cancer Research (ICR), London. ICR acknowledges NHS funding to the NIHR Biomedical Research Centre. The US3SS study was supported by Massachusetts (K.M.E., R01CA47305), Wisconsin (P.A.N., R01 CA47147) and New Hampshire (L.T.-E., R01CA69664) centres, and Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The USRT study was funded by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services. Support for CIMBA studies: BCFR was supported by grant UM1 CA164920 from the National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names and commercial products, or organizations imply endorsement by the US Government or the BCFR. BFBOCC is partly supported by Lithuania (BFBOCC-LT), Research Council of Lithuania grant LIG-07/2012; BIDMC is supported by the Breast Cancer Research Foundation. BRCA-gene mutations

and breast cancer in South African women (BMBSA) was supported by grants from the Cancer Association of South Africa (CANSA) to Elizabeth J. van Rensburg; SLN was partially supported by the Morris and Horowitz Families Endowed Professorship. The CBCS was supported by the NEYE Foundation. CNIO was partially supported by Spanish Association against Cancer (AECC08), RTICC 06/0020/1060, FISPI08/1120, Mutua Madrileña Foundation (FMMA) and SAF2010-20493. City of Hope Clinical Cancer Genetics Community Network and the Hereditary Cancer Research Registry, supported in part by Award Number RC4CA153828 (PI: J. Weitzel) from the National Cancer Institute and the Office of the Director, National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Funds for CONSIT TEAM were from Italian citizens who allocated the 5x1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects '5x1000') to SM and from FiorGen Foundation for Pharmacogenomics to LP. The CIMBA data management and data analysis were supported by Cancer Research UK grants C12292/A11174 and C1287/A10118. SH is supported by an NHMRC Program Grant to GCT. ACA is a Cancer Research UK Senior Cancer Research Fellow. GCT is an NHMRC Senior Principal Research Fellow. The DEMOKRITOS has been co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program of the General Secretariat for Research & Technology: SYN11_10_19 NBCA. Investing in knowledge society through the European Social Fund. The DKFZ study was supported by the DKFZ. EMBRACE is supported by Cancer Research UK Grants C1287/A10118 and C1287/A11990. D. Gareth Evans and Fiona Lalloo are supported by an NIHR grant to the Biomedical Research Centre, Manchester. The Investigators at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust are supported by an NIHR grant to the Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. Ros Eeles and Elizabeth Bancroft are supported by Cancer Research UK Grant C5047/A8385. Ros Eelses is also supported by NIHR support to the Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. FCCC was supported by The University of Kansas Cancer Center (P30 CA168524) and the Kansas Bioscience Authority Eminent Scholar Program. A.K.G. was funded by 5U01CA113916, R01CA140323, and by the Chancellors Distinguished Chair in Biomedical Sciences Professorship. The German Consortium of Hereditary Breast and Ovarian Cancer (GC-HBOC) is supported by the German Cancer Aid (grant no 109076, Rita K. Schmutzler) and by the Center for Molecular Medicine Cologne (CMMC). The GEMO study was supported by the Ligue Nationale Contre le Cancer; the Association "Le cancer du sein, parlons-en!" Award; the Canadian Institutes of Health Research for the "CIHR Team in Familial Risks of Breast Cancer" program and the French National Institute of Cancer (INCa). GEORGETOWN (CI) received support from the Non-Therapeutic Subject Registry Shared Resource at Georgetown University (NIH/NCI grant P30-CA051008), the Fisher Center for Familial Cancer Research, and Swing For the Cure. Kim De Leeneer (the G-FAST study) is supported by GOA grant BOF10/GOA/019 (Ghent University) and spearhead financing of Ghent University Hospital. The HCSC was supported by a grant RD12/00369/0006 and 12/00539 from ISCIII (Spain), partially supported by European Regional Development FEDER funds. HEBCS was financially supported by the Helsinki University Hospital Research Fund, Academy of Finland (266528), the Finnish Cancer Society and the Sigrid Juselius Foundation. The HEBON study is supported by the Dutch Cancer Society grants NKI1998-1854, NKI2004-3088, NKI2007-3756, the Netherlands Organization of Scientific Research grant NWO 91109024, the Pink Ribbon grant 110005 and the BBMRI grant NWO 184.021.007/CP46. HEBON thanks the registration teams of the Comprehensive Cancer Centre Netherlands and Comprehensive Centre South (together the Netherlands Cancer Registry) and PALGA (Dutch Pathology Registry) for part of the data collection. HRBCP is supported by The Hong Kong Hereditary Breast Cancer Family Registry and the Dr. Ellen Li Charitable Foundation, Hong Kong. The Hungarian Breast and Ovarian Cancer Study was supported by Hungarian Research Grants KTIA-OTKA CK-80745, OTKA K-112228 and the Norwegian EEA Financial Mechanism Hu0115/NA/2008-3/OP-9. ICO: contract grant sponsor, Asociación Española Contra el Cáncer, Spanish Health Research Fund; Carlos III Health Institute; Catalan Health Institute and Autonomous Government of Catalonia. Contract grant numbers: ISCIIIRETIC RD06/0020/1051, RD12/0036/008, PI10/01422, PI10/00748, PI13/00285, PIE13/00022, 2009SGR290 and 2014SGR364. The IHCC was supported by Grant PBZ_KBN_122/P05/2004. The ILUH group was supported by the Icelandic Association "Walking for Breast Cancer Research" and by the Landspitali University Hospital Research Fund. INHERIT was supported by the Canadian Institutes of Health Research for the "CIHR Team in Familial Risks of Breast Cancer" programme, the Canadian Breast Cancer Research Alliance-grant 019511 and the Ministry of Economic Development, Innovation and Export Trade – grant PSR-SIIRI-701. IOVHBOCS is supported by Ministero della Salute and "5x1000" Istituto Oncologico Veneto grant. The IPOBCS study was in part supported by Liga Portuguesa Contra o Cancro. kConFab is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. KOHBRA is supported by a grant from the National R&D Program for Cancer Control, Ministry for Health, Welfare and Family Affairs, Republic of Korea (1020350). MAYO is supported by NIH grants CA116167, CA128978 and CA176785, an NCI Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), a U.S. Department of Defence Ovarian Cancer Idea award (W81XWH-10-1-0341), a grant from the Breast Cancer Research Foundation, a generous gift from the David F. and Margaret T. Grohne Family Foundation. Jewish General Hospital Weekend to End Breast Cancer, Quebec Ministry of Economic Development, Innovation and Export Trade. MODSQUAD was supported by MH CZ - DRO (MMCI, 00209805) and by the European Regional Development Fund and the State Budget of the Czech Republic (RECAMO, CZ.1.05/2.1.00/03.0101) to LF, and by Charles University in Prague project UNCE204024 (MZ). MSKCC is supported by grants from the Breast Cancer Research Foundation, the Robert and Kate Niehaus Clinical Cancer Genetics Initiative, and the Andrew Sabin Research Fund. NAROD was supported by NIH grant:1R01 CA149429-01. The research of Drs. MH Greene and PL Mai was supported by the Intramural Research Program of the US National Cancer Institute, NIH, and by support services contracts NO2-CP-11019-50 and N02-CP-65504 with Westat Inc, Rockville, MD. NICCC is supported by Clalit Health Services in Israel. Some of its activities are supported by the Israel Cancer Association and the Breast Cancer Research Foundation (BCRF), NY. NNPIO has been supported by the Russian Federation for Basic Research (grants 13-04-92613, 14-04-93959 and 15-04-01744). NRG Oncology was supported by National Cancer Institute grants to the NRG Oncology Administrative Office and Tissue Bank (CA 27469), the NRG Oncology Statistical and Data Center (CA 37517), and NRG Oncology's Cancer Prevention and Control Committee (CA 101165). Drs. Greene, Mai and Savage were supported by funding from the Intramural Research Program, NCI. OSUCCG is supported by the Ohio State University Comprehensive Cancer Center. PBCS was supported by the ITT (Istituto Toscano Tumori) grants 2011-2013. SEABASS was supported by the Ministry of Science, Technology and Innovation, Ministry of Higher Education (UM.C/HIR/MOHE/06) and Cancer Research Initiatives Foundation. SMC was partially funded through a grant by the Israel cancer association and the funding for the Israeli Inherited Breast Cancer Consortium. SWE-BRCA collaborators are supported by the Swedish Cancer Society. UCHICAGO is supported by NCI Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA125183), R01 CA142996, 1U01CA161032 and by the Ralph and Marion Falk Medical Research Trust, the Entertainment Industry Fund National Women's Cancer Research Alliance and the Breast Cancer research Foundation. OIO is an ACS Clinical Research Professor. UCLA was supported by the Jonsson Comprehensive Cancer Center Foundation, Breast Cancer Research Foundation. UCSF was supported by the UCSF Cancer Risk Program and Helen Diller Family Comprehensive Cancer Center. UKFOCR was supported by a project grant from CRUK to Paul Pharoah. UPENN was supported by National Institutes of Health (NIH) (R01-CA102776 and R01-CA083855; Breast Cancer Research Foundation; Susan G. Komen Foundation for the cure, Basser Research Center for BRCA. UPITT/MWH was supported by the Frieda G. and Saul F. Shapira BRCA-Associated Cancer Research Program, Hackers for Hope Pittsburgh. Kate Lawrenson is funded by Ovarian Cancer Research Fund (OCRF) grant number 258807 and an Ann Schreiber Program of Excellence award from the Ovarian Cancer Research Fund (POE/USC/01.12). Janet Lee and Howard Shen are funded by National Institute of Health grant number 5 U19 CA148112-02. Tassja Spindler is funded by National Institute of Health grant number CA173531-01. Work was performed within the USC Norris Comprehensive Cancer Center which is supported by a Cancer Center Support Grant (award number P30 CA014089) from the National Cancer Institute. VFCTG was supported by the Victorian Cancer Agency, Cancer Australia, National Breast Cancer Foundation. Dr. Karlan is funded by the American Cancer Society Early Detection Professorship (SIOP-06-258-01-COUN) and the National Center for Advancing Translational Sciences (NCATS), Grant UL1TR000124.

Authors' contributions
CZ and WZ wrote the manuscript. DFE directed the BCAC. ACA, GCT and FJC coordinated the CIMBA studies. C Z and KK conducted the statistical analyses. AD and XG conducted the bioinformatics analyses. All authors made substantial contributions toward the conception and design, or acquisition of data, or analysis and interpretation of data. All authors read and approved the final manuscript.

Authors' information
† means prematurely passed away.

Competing interests
The study sponsors had no role in the design of the study, the collection, analysis or interpretation of the data, the writing of the manuscript or the decision to submit the manuscript for publication. All authors declare that they have no conflict of interest.

Ethics approval and consent to participate
Ethical approval of each study was given by the local institutional review boards. The full names of the institutional review boards that approved each study are listed in Additional file 1.

Author details
[1]Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, 2525 West End Avenue, 8th Floor, Nashville, TN 37203-1738, USA. [2]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [3]Proteomics Center, CHU de Québec Research Center and Department of Molecular Medicine, Laval University, Quebec, Canada. [4]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge CB1 8RN, UK. [5]Department of Surgery, St Vincent's Hospital, Melbourne, VIC, Australia. [6]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global health, The University of Melbourne, Melbourne, Australia. [7]Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia. [8]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA. [9]Immunology and Molecular Oncology Unit, Istituto Oncologico Veneto IOV - IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico), Padua, Italy. [10]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK. [11]Department of Clinical Genetics, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. [12]Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada. [13]Department of Molecular Genetics, University of Toronto, Toronto, Canada. [14]Department of Epidemiology, University of California Irvine, Irvine, CA, USA. [15]N.N. Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus. [16]Department of Pathology, Landspitali University Hospital and BMC (Biomedical Centre), Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [17]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany. [18]University of Texas MD Anderson Cancer Center, Houston, TX, USA. [19]Department of Oncology, Karolinska University Hospital, Stockholm, Sweden. [20]McGill University and Génome Québec Innovation Centre, Montréal, Canada. [21]Human Cancer Genetics Program, Spanish National Cancer Research Centre, Madrid, Spain. [22]Centro de Investigación en Red de Enfermedades Raras, Valencia, Spain. [23]Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa, Russia. [24]Department of Oncology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. [25]International Epidemiology Institute, Rockville, MD, USA. [26]Department of Radiation Oncology, Hannover Medical School, Hannover, Germany. [27]Copenhagen General Population Study, Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark. [28]Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark. [29]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [30]Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia, Milan, Italy. [31]Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway. [32]K.G. Jebsen Center for Breast Cancer Research, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [33]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [34]Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany. [35]University of Tübingen, Tübingen, Germany. [36]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany. [37]International Agency for Research on Cancer, Lyon, France. [38]Division of Preventive Oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [39]Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands. [40]Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum, Bochum, Germany. [41]Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg, Germany. [42]Molecular Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany. [43]Department of Medicine, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT, USA. [44]Molecular Oncology Laboratory, Hospital Clinico San Carlos, IdISSC (El Instituto de Investigación Sanitaria del Hospital Clínico San Carlos), Madrid, Spain. [45]Cancer Genetics Laboratory, Peter MacCallum Cancer Centre, Melbourne, Australia. [46]Australian Breast Cancer Tissue Bank, Westmead Millennium Institute, University of Sydney, Sydney, Australia. [47]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [48]University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [49]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, South Korea. [50]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, South Korea. [51]Cancer Research Institute, Seoul National University, Seoul, South Korea. [52]Center for Medical Genetics, Ghent University, Ghent, Belgium. [53]Westmead Millenium Institute for Medical Research, University of Sydney, Sydney, Australia. [54]Sheffield Cancer Research, Department of Oncology, University of Sheffield, Sheffield, UK. [55]Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield, UK. [56]Department of Clinical Genetics, Fox Chase Cancer Center, Philadelphia, PA, USA. [57]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands. [58]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands. [59]Oncogenetics Group, University Hospital Vall d'Hebron, Vall d'Hebron Institute of Oncology (VHIO) and Universitat Autònoma de Barcelona, Barcelona, Spain. [60]Department of Medicine, Abramson Cancer Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [61]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA. [62]Department of Genetics, University of Pretoria, Pretoria, South Africa. [63]Clinics of Obstetrics and Gynaecology, Hannover Medical School, Hannover, Germany. [64]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK. [65]Genomics Center, Centre Hospitalier Universitaire de Québec Research Center, Laval University, Québec City, Canada. [66]Department of Biomedical Sciences, Faculty of Science and Technology, University of Westminster, London, UK. [67]Institute of Human Genetics, Muenster, Germany. [68]Division of Population Sciences, Moffitt Cancer Center & Research Institute, Tampa, FL, USA. [69]Department of Oncology, Sahlgrenska University Hospital, Gothenburg, Sweden. [70]Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [71]Genetic Medicine, Manchester Academic Health Sciences Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK. [72]Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany. [73]David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, CA, USA. [74]Department of Breast Surgery, Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark. [75]ANZ GOTG Coordinating Centre, Australia New Zealand GOG, Camperdown, NSW, Australia. [76]Susanne Levy Gertner Oncogenetics Unit, Sheba Medical Center, Tel-Hashomer, Israel. [77]Section of Genetic Oncology, Deparment of Laboratory Medicine, University and University Hospital of Pisa, Pisa, Italy. [78]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China. [79]Cancer Risk and Prevention Clinic, Dana-Farber Cancer Institute, Boston, MA, USA. [80]Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK. [81]Institute of Human Genetics, University Würzburg, Wurzburg, Germany. [82]INSERM U1052, CNRS UMR5286, Université Lyon, Centre de Recherche en Cancérologie de Lyon, Lyon, France. [83]Genetic Epidemiology of Cancer team, Inserm, U900, Institut Curie, Mines ParisTech, 75248 Paris, France. [84]Department of Tumour Biology, Institut Curie, Paris, France. [85]Institut Curie, INSERM U830, Paris, France. [86]Université Paris Descartes, Sorbonne Paris Cité, Paris, France. [87]GEMO study: National Cancer Genetics Network, UNICANCER Genetic Group, , France. [88]Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, KS, USA. [89]Department of Dermatology, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT, USA.

[90]Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA. [91]Environmental Epidemiology of Cancer, Center for Research in Epidemiology and Population Health, INSERM, Villejuif, France. [92]University Paris-Sud, Villejuif, France. [93]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [94]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA. [95]Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany. [96]Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [97]Cancer Center, Kuopio University Hospital, Kuopio, Finland. [98]Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland. [99]Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland. [100]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore. [101]Department of Surgery, National University Health System, Singapore, Singapore. [102]Cancer Research Initiatives Foundation, Subang Jaya, Selangor, Malaysia. [103]Breast Cancer Research Unit, Cancer Research Institute, University Malaya Medical Centre, Kuala Lumpur, Malaysia. [104]Cancer Division, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia. [105]Family Cancer Clinic, Netherlands Cancer Institute, Amsterdam, The Netherlands. [106]The Hereditary Breast and Ovarian Cancer Research Group Netherlands (HEBON), Coordinating center: Netherlands Cancer Institute, Amsterdam, The Netherlands. [107]Suburban Hospital, Bethesda, MD, USA. [108]Care of City of Hope Clinical Cancer Genetics Community Research Network, Duarte, CA, USA. [109]Department of Medical Oncology, Family Cancer Clinic, Erasmus MC Cancer Institute, Rotterdam, The Netherlands. [110]Center for Medical Genetics, NorthShore University HealthSystem, Evanston, IL, USA. [111]Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston, MA, USA. [112]Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA. [113]N.N. Petrov Institute of Oncology, St. Petersburg, Russia. [114]Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC, USA. [115]Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Aichi, Japan. [116]Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland. [117]State Research Institute Centre for Innovative Medicine, Vilnius, Lithuania. [118]Department of Clinical Genetics, Aarhus University Hospital, Aarhus, N, Denmark. [119]Department of Epidemiology, Cancer Prevention Institute of California, Fremont, CA, USA. [120]Department of Health Research and Policy - Epidemiology, Stanford University School of Medicine, Stanford, CA, USA. [121]Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA. [122]Genomics Center, Centre Hospitalier Universitaire de Québec Research Center and Laval University, Quebec City, QC, Canada. [123]Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA. [124]Department of Pathology, Oulu University Hospital, University of Oulu, Oulu, Finland. [125]School of Medicine, National University of Ireland, Galway, Ireland. [126]Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. [127]Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Russia. [128]Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada. [129]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada. [130]Molecular Diagnostics Laboratory, IRRP, National Centre for Scientific Research "Demokritos", Athens, Greece. [131]The Hong Kong Hereditary Breast Cancer Family Registry, Cancer Genetics Center, Hong Kong Sanatorium and Hospital, Hong Kong, China. [132]Department of Surgery, The University of Hong Kong, Hong Kong, China. [133]Vesalius Research Center, Leuven, Belgium. [134]Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium. [135]Molecular Diagnostic Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain. [136]University of Hawaii Cancer Center, Honolulu, HI, USA. [137]Department of Surgery, Soonchunhyang University and Hospital, Seoul, South Korea. [138]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [139]Division of Health Sciences, Warwick Medical School, Warwick University, Coventry, UK. [140]Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale Tumori (INT), Milan, Italy. [141]Department of Oncology - Pathology, Karolinska Institutet, Stockholm, Sweden. [142]National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany. [143]Department of Preventive Medicine, Kyushu University Faculty of Medical Sciences, Fukuoka, Japan. [144]Division of Gynaecology and Obstetrics, Technische Universität München, Munich, Germany. [145]Institute of Population Health, University of Manchester, Manchester, UK. [146]Laboratory Medicine Program, University Health Network, Toronto, ON, Canada. [147]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada. [148]Beckman Research Institute of City of Hope, Duarte, CA, USA. [149]Carbone Cancer Center, University of Wisconsin, Madison, WI, USA. [150]Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [151]Department of Medicine and Genetics, University of California, San Francisco, San Francisco, CA, USA. [152]Clinical Genetics Research Lab, Department of Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA. [153]Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary. [154]Center for Clinical Cancer Genetics and Global Health, University of Chicago Medical Center, Chicago, IL, USA. [155]Human Genetics Group, Human Cancer Genetics Program, Spanish National Cancer Centre (CNIO), Madrid, Spain. [156]Biomedical Network on Rare Diseases (CIBERER), Madrid, Spain. [157]Unit of Medical Genetics, Department of Biomedical, Experimental and Clinical Sciences, University of Florence, Florence, Italy. [158]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [159]University Hospital Gashuisberg, Leuven, Belgium. [160]Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale Tumori (INT), Milan, Italy. [161]IFOM, Fondazione Istituto FIRC (Italian Foundation of Cancer Research) di Oncologia Molecolare, Milan, Italy. [162]Department of Obstetrics and Gynecology, and Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria. [163]Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL, USA. [164]Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale Tumori (INT), Milan, Italy. [165]Section of Cancer Genetics, The Institute of Cancer Research, London, UK. [166]Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, CA, USA. [167]Department of Basic Sciences, Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH & RC), Lahore, Pakistan. [168]Clalit National Israeli Cancer Control Center and Department of Community Medicine and Epidemiology, Carmel Medical Center and B. Rappaport Faculty of Medicine, Haifa, Israel. [169]Centre of Familial Breast and Ovarian Cancer, Department of Gynaecology and Obstetrics and Centre for Integrated Oncology (CIO), Center for Molecular Medicine Cologne (CMMC), University Hospital of Cologne, Cologne, Germany. [170]Obstetrics and Gynecology, Ohio State University College of Medicine, Columbus, OH, USA. [171]National Cancer Institute, Bangkok, Thailand. [172]Research Oncology, Guy's Hospital, King's College London, London, UK. [173]Division of Molecular Gyneco-Oncology, Department of Gynaecology and Obstetrics, University Hospital of Cologne, Cologne, Germany. [174]Center of Familial Breast and Ovarian Cancer, University Hospital of Cologne, Cologne, Germany. [175]Center for Integrated Oncology, University Hospital of Cologne, Cologne, Germany. [176]Center for Molecular Medicine, University Hospital of Cologne, Cologne, Germany. [177]School of Public Health, China Medical University, Taichung, Taiwan. [178]Taiwan Biobank, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. [179]Department of Obstetrics and Gynecology, Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria. [180]Centre Hospitalier Universitaire de Québec Research Center and Laval University, Quebec City, QC, Canada. [181]Genetic Epidemiology Laboratory, Department of Pathology, University of Melbourne, Parkville, VIC, Australia. [182]Hannover Medical School, Hannover, Germany. [183]Division of Breast Cancer Research, The Institute of Cancer Research, London, UK. [184]National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA. [185]Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, Canada. [186]Department of Genetics, Portuguese Oncology Institute, Porto, Portugal. [187]Biomedical Sciences Institute (ICBAS), Porto University, Porto, Portugal. [188]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA. [189]Genetic Counseling Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain. [190]Department of Clinical Genetics, Odense University Hospital, Odense, C, Denmark. [191]Latvian Biomedical Research and Study Centre, Riga, Latvia. [192]Program in Cancer Genetics, Departments of Human Genetics and Oncology, McGill University,

Montreal, QC, Canada. [193]Currently at Medical School Cambridge University, Cambridge, UK. [194]Department of Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH, USA. [195]Department of Medical Oncology, Beth Israel Deaconess Medical Center, Boston, MA, USA. [196]Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands. [197]Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA. [198]Department of Obstetrics and Gynecology, University of Ulm, Ulm, Germany. [199]Clinical Cancer Genetics, for the City of Hope Clinical Cancer Genetics Community Research Network, Duarte, CA, USA. [200]Laboratory of Cancer Genetics and Tumor Biology, Department of Clinical Chemistry and Biocenter Oulu, University of Oulu, Oulu, Finland. [201]Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre NordLab, Oulu, Finland. [202]Singapore Eye Research Institute, National University of Singapore, Singapore, Singapore. [203]Department of Medical Oncology, Papageorgiou Hospital, Aristotle University of Thessaloniki School of Medicine, Thessaloniki, Greece. [204]Department of Surgery, National Taiwan University Hospital, Taipei, Taiwan. [205]Department of Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia. [206]Peter MacCallum Cancer Center, The University of Melbourne, Melbourne, Australia.

Received: 23 December 2015 Accepted: 18 May 2016
Published online: 21 June 2016

References

1. Ghoussaini M, Fletcher O, Michailidou K, Turnbull C, Schmidt MK, Dicks E, et al. Genome-wide association analysis identifies three new breast cancer susceptibility loci. Nat Genet. 2012;44(3):312–8.
2. Antoniou AC, Kuchenbaecker KB, Soucy P, Beesley J, Chen X, McGuffog L, et al. Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers. Breast Cancer Res. 2012;14(1):R33.
3. Kuchenbaecker KB, Neuhausen SL, Robson M, Barrowdale D, McGuffog L, Mulligan AM, et al. Associations of common breast cancer susceptibility alleles with risk of breast cancer subtypes in BRCA1 and BRCA2 mutation carriers. Breast Cancer Res. 2014;16(6):3416.
4. Zheng W, Zhang B, Cai Q, Sung H, Michailidou K, Shi J, et al. Common genetic determinants of breast-cancer risk in East Asian women: a collaborative study of 23 637 breast cancer cases and 25 579 controls. Hum Mol Genet. 2013;22(12):2539–50.
5. Li J, Karaplis AC, Huang DC, Siegel PM, Camirand A, Yang XF, et al. PTHrP drives breast tumor initiation, progression, and metastasis in mice and is a potential therapy target. J Clin Invest. 2011;121(12):4655–69.
6. Fleming NI, Trivett MK, George J, Slavin JL, Murray WK, Moseley JM, et al. Parathyroid hormone-related protein protects against mammary tumor emergence and is associated with monocyte infiltration in ductal carcinoma in situ. Cancer Res. 2009;69(18):7473–9.
7. Henderson MA, Danks JA, Slavin JL, Byrnes GB, Choong PF, Spillane JB, et al. Parathyroid hormone-related protein localization in breast cancers predict improved prognosis. Cancer Res. 2006;66(4):2250–6.
8. Corradin O, Saiakhova A, Akhtar-Zaidi B, Myeroff L, Willis J, Cowper-Sal lari R, et al. Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. Genome Res. 2014;24(1):1–13.
9. Edwards SL, Beesley J, French JD, Dunning AM. Beyond GWASs: illuminating the dark road from association to function. Am J Hum Genet. 2013;93(5):779–97.
10. Bojesen SE, Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet. 2013;45(4):371–84. 384e371-372.
11. French JD, Ghoussaini M, Edwards SL, Meyer KB, Michailidou K, Ahmed S, et al. Functional variants at the 11q13 risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers. Am J Hum Genet. 2013;92(4):489–503.
12. Meyer KB, O'Reilly M, Michailidou K, Carlebur S, Edwards SL, French JD, et al. Fine-scale mapping of the FGFR2 breast cancer risk locus: putative functional variants differentially bind FOXA1 and E2F1. Am J Hum Genet. 2013;93(6):1046–60.
13. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet. 2013;45(4):353–61. 361e351-352.
14. Chenevix-Trench G, Milne RL, Antoniou AC, Couch FJ, Easton DF, Goldgar DE. Cimba: An international initiative to identify genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA). Breast Cancer Res. 2007;9(2):104.
15. Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. Nature. 2010;467(7319):1061–73.
16. International HapMap C. The International HapMap Project. Nature. 2003;426(6968):789–96.
17. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. G3. 2011;1(6):457–70.
18. Long JC, Williams RC, Urbanek M. An E-M algorithm and testing strategy for multiple-locus haplotypes. Am J Hum Genet. 1995;56(3):799–810.
19. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007;447(7148):1087–93.
20. Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet. 2009;41(3):324–8.
21. Barnes DR, Antoniou AC. Unravelling modifiers of breast and ovarian cancer risk for BRCA1 and BRCA2 mutation carriers: update on genetic modifiers. J Intern Med. 2012;271(4):331–43.
22. Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, et al. A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. Nat Genet. 2010;42(10):885–92.
23. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26(17):2190–1.
24. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57–74.
25. He B, Chen C, Teng L, Tan K. Global view of enhancer-promoter interactome in human cells. Proc Natl Acad Sci USA. 2014;111(21):E2191–9.
26. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. Cell. 2013;155(4):934–47.
27. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;22(9):1790–7.
28. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012;40(Database issue):D930–4.
29. Cline MS, Craft B, Swatloski T, Goldman M, Ma S, Haussler D, et al. Exploring TCGA pan-cancer data at the UCSC Cancer Genomics Browser. Sci Rep. 2013;3:2652.
30. Cai Q, Zhang B, Sung H, Low SK, Kweon SS, Lu W, et al. Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1. Nat Genet. 2014;46(8):886–90.
31. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486(7403):346–52.
32. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics. 2012;28(10):1353–8.
33. Pique-Regi R, Degner JF, Pai AA, Gaffney DJ, Gilad Y, Pritchard JK. Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Res. 2011;21(3):447–55.
34. Robson EJ, He SJ, Eccles MR. A PANorama of PAX genes in cancer and development. Nat Rev Cancer. 2006;6(1):52–62.
35. Ballestar E, Paz MF, Valle L, Wei S, Fraga MF, Espada J, et al. Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. EMBO J. 2003;22(23):6335–45.
36. Miao D, Su H, He B, Gao J, Xia Q, Zhu M, et al. Severe growth retardation and early lethality in mice lacking the nuclear localization sequence and C-terminus of PTH-related protein. Proc Natl Acad Sci USA. 2008;105(51):20309–14.
37. Suva LJ, Winslow GA, Wettenhall RE, Hammonds RG, Moseley JM, Diefenbach-Jagger H, et al. A parathyroid hormone-related protein implicated in malignant hypercalcemia: cloning and expression. Science. 1987;237(4817):893–6.

38. Southby J, Kissin MW, Danks JA, Hayman JA, Moseley JM, Henderson MA, et al. Immunohistochemical localization of parathyroid hormone-related protein in human breast cancer. Cancer Res. 1990;50(23):7710–6.
39. Dittmer A, Vetter M, Schunke D, Span PN, Sweep F, Thomssen C, et al. Parathyroid hormone-related protein regulates tumor-relevant genes in breast cancer cells. J Biol Chem. 2006;281(21):14563–72.
40. Akhtari M, Mansuri J, Newman KA, Guise TM, Seth P. Biology of breast cancer bone metastasis. Cancer Biol Ther. 2008;7(1):3–9.
41. Okoumassoun LE, Russo C, Denizeau F, Averill-Bates D, Henderson JE. Parathyroid hormone-related protein (PTHrP) inhibits mitochondrial-dependent apoptosis through CK2. J Cell Physiol. 2007;212(3):591–9.
42. Mardones GA, Burgos PV, Brooks DA, Parkinson-Lawrence E, Mattera R, Bonifacino JS. The trans-Golgi network accessory protein p56 promotes long-range movement of GGA/clathrin-containing transport carriers and lysosomal enzyme sorting. Mol Biol Cell. 2007;18(9):3486–501.
43. Kolker E, Higdon R, Haynes W, Welch D, Broomall W, Lancet D, et al. MOPED: Model Organism Protein Expression Database. Nucleic Acids Res. 2012;40(Database issue):D1093–9.
44. Feng Y, Stram DO, Rhie SK, Millikan RC, Ambrosone CB, John EM, et al. A comprehensive examination of breast cancer risk loci in African American women. Hum Mol Genet. 2014;23(20):5518–26.
45. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 2012;9(3):215–6.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central