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

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:27822369

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Identification of independent association signals and putative functional variants for breast cancer risk through fine-scale mapping of the 12p11 locus


* Correspondence: wei.zheng@vanderbilt.edu
1Division 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
© 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.
Multiple recent genome-wide association studies (GWAS) have identified a single nucleotide polymorphism (SNP), rs10771399, at 12p11 that is associated with breast cancer risk. We performed a fine-scale mapping study of a 700 kb region including 441 genotyped and more than 4,200 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 (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; \( P = 3 \times 10^{-5} \)), 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, BRCA1 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 (1116 cases and 932 controls). The CIMBA included 45 studies of women of European descent (1116 cases and 932 controls), nine of Asian descent (6269 cases and 6624 controls), and two of African ancestry studies (seven 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 BRCA1 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.)
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, $P = 4 \times 10^{-38}$; rs805510 OR = 0.85, $P = 10^{-35}$; rs1871152 OR = 0.94, $P = 3 \times 10^{-40}$). 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

<table>
<thead>
<tr>
<th>Signal</th>
<th>SNPs</th>
<th>Position (hg 19)</th>
<th>Alleles</th>
<th>EAF</th>
<th>LD ($r^2$)</th>
<th>Univariate analysis</th>
<th>Conditional analysis</th>
<th>SNPs retained for functional annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per-allele OR (95 % CI)</td>
<td>P-trend</td>
<td>Per-allele OR (95 % CI)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>28155080</td>
<td>G*/A</td>
<td>0.12</td>
<td>-</td>
<td>0.85 (0.83-0.88)</td>
<td>$5 \times 10^{-5}$</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>28174817</td>
<td>T*/C</td>
<td>0.24</td>
<td>0.42</td>
<td>0.88 (0.86-0.90)</td>
<td>$4 \times 10^{-38}$</td>
<td>0.92 (0.89-0.94)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>28139846</td>
<td>T*/C</td>
<td>0.12</td>
<td>0.88</td>
<td>0.85 (0.82-0.88)</td>
<td>$10^{-25}$</td>
<td>0.93 (0.89-0.96)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>28378926</td>
<td>T*/C</td>
<td>0.31</td>
<td>0.04</td>
<td>0.94 (0.92-0.96)</td>
<td>$3 \times 10^{-6}$</td>
<td>0.96 (0.94-0.98)</td>
</tr>
</tbody>
</table>

*Effect alleles. *Identified in the initial genome-wide association study conducted in women of European descent [1]. *Linkage disequilibrium (LD) with rs10771399 for women of European descent. *Adjusted for studies, and the top principal components and an additional principal component accounting for the Leuven Multidisciplinary Breast Centre (LMBC) study. *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. *Associated single nucleotide polymorphisms (SNPs) with a likelihood ratio >1/100 relative to the lead SNP in each signal. *See Table S2 in Additional file 5. *See Table S2 in Additional file 5. EAF effect allele frequency in controls, OR odds ratio, CI confidence interval.
<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Overall breast cancer</th>
<th></th>
<th>Breast cancer (age at diagnosis &lt;45 years)</th>
<th></th>
<th>Breast cancer (age at diagnosis ≥45 years)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rs805510 - rs7297051 - rs1871152</td>
<td>frequency</td>
<td>OR (95 % CI)</td>
<td>P value</td>
<td>frequency</td>
<td>OR (95 % CI)</td>
<td>P value</td>
</tr>
<tr>
<td>C-C-A</td>
<td>0.51</td>
<td>1.00 (Ref)</td>
<td>Ref</td>
<td>0.52</td>
<td>1.00 (Ref)</td>
<td>Ref</td>
</tr>
<tr>
<td>C-C-G</td>
<td>0.24</td>
<td>0.92 (0.89–0.95)</td>
<td>2 × 10^{-8}</td>
<td>0.22</td>
<td>0.94 (0.89–1.00)</td>
<td>0.04</td>
</tr>
<tr>
<td>C-T-A</td>
<td>0.09</td>
<td>0.90 (0.87–0.95)</td>
<td>3 × 10^{-6}</td>
<td>0.09</td>
<td>0.96 (0.89–1.03)</td>
<td>0.28</td>
</tr>
<tr>
<td>C-T-G</td>
<td>0.03</td>
<td>0.89 (0.82–0.96)</td>
<td>2 × 10^{-3}</td>
<td>0.03</td>
<td>0.85 (0.73–0.98)</td>
<td>0.02</td>
</tr>
<tr>
<td>T-T-A</td>
<td>0.04</td>
<td>0.82 (0.77–0.88)</td>
<td>1 × 10^{-5}</td>
<td>0.04</td>
<td>0.76 (0.67–0.87)</td>
<td>5 × 10^{-5}</td>
</tr>
<tr>
<td>T-T-G</td>
<td>0.07</td>
<td>0.79 (0.76–0.83)</td>
<td>3 × 10^{-33}</td>
<td>0.06</td>
<td>0.78 (0.71–0.85)</td>
<td>5 × 10^{-8}</td>
</tr>
<tr>
<td>Rare</td>
<td>0.01</td>
<td>0.88 (0.79–0.99)</td>
<td>0.04</td>
<td>0.01</td>
<td>0.90 (0.72–1.13)</td>
<td>0.37</td>
</tr>
</tbody>
</table>

*Adjusted for studies and the top principal components. *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 rs1877152, 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 rs1877152, 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 descent 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 (rs1877152) 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, $P = 1 \times 10^{-5}$). 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 East Asian 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 ($r^2 < 0.02$). 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

<table>
<thead>
<tr>
<th>Index</th>
<th>SNP</th>
<th>Position (hg 19)</th>
<th>Alleles</th>
<th>EAF</th>
<th>LD ($r^2$)</th>
<th>Univariate analysis</th>
<th>Conditional analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>SNPs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per-allele effect (95 % CI)</td>
<td>$P$-trend</td>
</tr>
<tr>
<td>Meta-analysis of ER-negative cancer (BCAC + CIMBA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCAC ER-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Signal 1</td>
<td>rs7297051</td>
<td>28174817</td>
<td>T*/C</td>
<td>0.24</td>
<td>0.42</td>
<td>0.87 (0.83–0.91)</td>
<td>3 $\times 10^{-10}$</td>
</tr>
<tr>
<td>Signal 4</td>
<td>rs113824616</td>
<td>28184905</td>
<td>C*/T</td>
<td>0.05</td>
<td>0.40</td>
<td>0.75 (0.67–0.84)</td>
<td>5 $\times 10^{-7}$</td>
</tr>
<tr>
<td>CIMBA BRCA1 mutation carriers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Signal 1</td>
<td>rs7297051</td>
<td>28174817</td>
<td>T*/C</td>
<td>0.23</td>
<td>0.37</td>
<td>0.89 (0.85–0.93)</td>
<td>3 $\times 10^{-7}$</td>
</tr>
<tr>
<td>Signal 4</td>
<td>rs113824616</td>
<td>28184905</td>
<td>C*/T</td>
<td>0.04</td>
<td>0.49</td>
<td>0.73 (0.64–0.82)</td>
<td>1 $\times 10^{-7}$</td>
</tr>
</tbody>
</table>


*Effect alleles. *Adjusted for studies, and the top principal components. §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).

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...
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.

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 (PTHHLH 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 (PTHHLH). 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.
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 \( (p = 0.03, \text{Fig. 3c}) \). 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 kb 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 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) 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, CAAT/enhancer-binding protein; CCDC91, Coiled-coil domain containing 91; CHA-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; HMBC, human mammary epithelial cells; HNF1B, HNF1 homeobox B; HR, hazard ratio; iCOGS, Illumina Select 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. Tessler, Francois Bacot, Daniel Vincent, Sylvie Laboisserière and Frédéric Robidoux and the staff of the McGill University and Génome Québec Innovat’om Centre, Sune F. Nielsen, Borge G. Nordestgaard, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hillier, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility. ABCFS wishes to thank Maggie Angelakos, Judi Maskill and Gillian Dite. ABCS wishes to thank Steen Cornelissen, Richard van Hen, Linde Braaf, Frans Hogervorst, Svenno Verhoef, Laura van ’t Veer, Emil 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 ACSF 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 ACSF study would like to thank Dr. Prat Boonyawongvej, the former Permanent Secretary of MOPH and Dr. Pornthep Siriwaranusaran, the Department Director-General of Disease Control who have supported the study throughout. ABCFS wishes to thank Eileen Williams, Elaine Ryder-Mills and Kara Sargus. BIGGS wishes to thank Niall McNiceney, Gabriella Collieran, Andrew Rowan and Angela Jones. BOCs wishes to thank the Welcome Trust Case Control Consortium (WTCCC) website for a full list of contributing investigators. BSUICH 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 Arndottir, Anne Bank and Dorthe Kjeldgaard Hansen for the excellent technical assistance. CNIO-BCS thanks Guillermo Pita, Charo Alonso, Daniel Herreb, Nuria Alvarez, 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. ABCFS wishes to thank the National Breast Cancer Foundation, Cancer Australia and the National Institute of Health (USA) for which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia and the National Institute of Health (USA). ABCFS wishes to thank Daniel Zaffaroni of the Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Monia Barile and Irene Feroce of the Istituto Europeo di Oncologia (IEO) and the personnel of the Cogentech Cancer Genetic Test Laboratory. MSICCC thanks Marina Corines and Lauren Jacobs. MTLGBCS would like to thank Martin Tranchant (CHU de Québec Research Centre), 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. MYBRCAS wishes to thank Phuah Sae Yue, Peter Kang, Kang Hoon, Kavitta Sivanandan, Sivaiah 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 Ottstedt, 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 Langrãed, 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 Ulleval, Oslo, Norway), Dr. Mari Muri Holmen, MD (Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway), Prof. Tori Sauer, MD (Department of Pathology at Akershus University Hospital, Lærangen, Norway), Dr. Vilde Haakensen, MD (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Olav Engebretsen, 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), Kristin V. Reintensd, 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), Associate Prof. Åslaug Helland, MD (Department of Genetics, Institute for Cancer Research and Department of Oncology, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Margit Rils, MD (Dept of Breast- and Endocrine Surgery, Oslo University Hospital, Ulleval, Oslo, Norway), Dr. Ida Buhholm, MD (Department of Breast-Endocrine Surgery, Akershus University Hospital, Oslo, Norway and Department of Oncology, Division of Cancer Medicine and 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. Gremaker Alvnes, 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 Otsuuka and Kari Mononen. OBFIR wishes to thank Teresa Selander and Nayana Weerasooriya. ORGO 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 (ONCOLOG) with the help of Dr. J. Molenaar. PBSCS wishes to thank Louise Brinton, Mark Sherman, Neolonia Szeszenia-Dabrowsa, Beata Pepłonska, Witold Zatorski, Pei Chao and Michael Stagner. pKARMA wishes to thank the Swedish Medical Research Council. PBSCS wishes to thank Petra Bos, Jannet Blom, Ellen Crepin, Elisabeth Huijser, Annette Heemspek and the Erasmus MC Family Cancer Clinic. SASBC thanks the Swedish Medical Research Council. 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 Connelly. SEARCH wishes to thank the SEARCH and EPIC teams. SGBCS wishes to thank the participants and research coordinator Kimberly Cha. SKDKFCS wishes to thank all study participants, clinicians, family doctors, researchers and technicians for their contributions and commitment to this study. TNBC wishes to thank Robert Pilarksi and Charles Shapiro who were instrumental in the formation of Zeng et al. Breast Cancer Research (2016) 18:64
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. UICBBS wishes to thank Irene Masunaka. UKBBS 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 NHIR Biomedical Research Centre.

CIMBA studies acknowledge the following. BCFR-AU wishes to thank Maggie Angelakos, Judi Maskell, 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 Vilnis Rudaitis, Laimonas Griskvičius, Ramūnas Janavičius. BFBOCC-LV acknowledges Drs. Janis Eglitis, Anna Krivola and Aivars Stengrevics. BMBSA wishes to thank the families who contribute to the BMBSA study. BRCOSH wishes to thank Yuan Chau Dinh and Linda Steele for their work in participant enrollment and biospecimen and data management. CNOIO wishes to thank Alicia Barroso, Rosario Alonso and Guillermo Pita for their assistance. The CONSIST TEAM wishes to thank Maria Grazia Tibiletti of the Ospedale di Circolo-Università dell’Insubria,Varese, Italy, Giulietta Scucner of the Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Gabriele Capone of the University of Florence, Florence, Italy, Vincenza Vel and Riccarda Ronchetti of the Insubria National Cancer Institute, Aviano, Italy, Aline Martayan of the Istituto Nazionale Tumori Regina Elena, Rome, Italy, Stefania Tommasi and Brunella Pilato of the Istituto Nazionale Tumori “Giovanni Paolo II”, Bari, Italy, Liliana Varesco of the IRCCS, ADU 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. FCCLS thanks Ms. JillEileen Weaver and Dr. Betty 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 Centres 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 Damilola, Laura Barjhoux, Carole Verny-Pierre, Mélène Léone, Nadia Boutez-Krzyza, Alain Calender, Sophie Geraud; and Service de Génétique Oncologique, Institut Curie, Paris: Dominique Stoppa-Lyonnet, Marion Gauthier-Villars, Bruno Buecher, Claude Houdayer, Ève Pinel, Lisa Golmard, Agnès Hardouin, Pascaline Berthet, Dominique Vaur, Laurent Castera. Institut Paoli Calmettes, Marseille: Hagay Solob, Violaine Bourdon, Nadia Boutez-Krzyza, Alain Calender, Sophie Geraud.

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. UICBBS wishes to thank Irene Masunaka. UKBBS 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 NHIR Biomedical Research Centre.

CIMBA studies acknowledge the following. BCFR-AU wishes to thank Maggie Angelakos, Judi Maskell, 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. BFBOCC-LT acknowledges Vilnis Rudaitis, Laimonas Griskvičius, Ramūnas Janavičius. BFBOCC-LV acknowledges Drs. Janis Eglitis, Anna Krivola and Aivars Stengrevics. BMBSA wishes to thank the families who contribute to the BMBSA study. BRCOSH wishes to thank Yuan Chau Dinh and Linda Steele for their work in participant enrollment and biospecimen and data management. CNOIO wishes to thank Alicia Barroso, Rosario Alonso and Guillermo Pita for their assistance. The CONSIST TEAM wishes to thank Maria Grazia Tibiletti of the Ospedale di Circolo-Università dell’Insubria,Varese, Italy, Giulietta Scucner of the Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Gabriele Capone of the University of Florence, Florence, Italy, Vincenza Vel and Riccarda Ronchetti of the Insubria National Cancer Institute, Aviano, Italy, Aline Martayan of the Istituto Nazionale Tumori Regina Elena, Rome, Italy, Stefania Tommasi and Brunella Pilato of the Istituto Nazionale Tumori “Giovanni Paolo II”, Bari, Italy, Liliana Varesco of the IRCCS, ADU 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. FCCLS thanks Ms. JillEileen Weaver and Dr. Betty 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 Centres 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 Damilola, Laura Barjhoux, Carole Verny-Pierre, Mélène Léone, Nadia Boutez-Krzyza, Alain Calender, Sophie Geraud; and Service de Génétique Oncologique, Institut Curie, Paris: Dominique Stoppa-Lyonnet, Marion Gauthier-Villars, Bruno Buecher, Claude Houdayer, Ève Pinel, Lisa Golmard, Agnès Hardouin, Pascaline Berthet, Dominique Vaur, Laurent Castera. Institut Paoli Calmettes, Marseille: Hagay Solob, Violaine Bourdon, Nadia Boutez-Krzyza, Alain Calender, Sophie Geraud.
assistants and doctors involved in the MyRCa Study for assistance in patient recruitment, data collection and sample preparation. In addition, we thank Philip lau, Sng Jen-Hwee and Shafishah 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/CHIR/MOHE/06) and charitable funding from the Cancer Research Initiatives Foundation. The SMC team wishes to acknowledge the assistance of the Malaysian 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, Häkan Olsson, Helena Jernstrom, Karin Henriksson, Katja Harbst, Maria Soller and UK Kristoffersson from Lund University and University Hospital; Anna Överholm, Margareta Nordling, Per Karlsson and Zakaria Einbeigi from Gothenburg Sahlgrenska University Hospital; Anna von Wachenfeldt, Annelie Liljeberg, Annika Lindblom, Brita Arver, Gisela Barbany Bustinha and Johanna Rantala from Stockholm and Karolinska University Hospital; Beatrice Melin, Christina Edwinsdotter Arndor and Monica Emanuelsson from Umeå University Hospital; Hans Ehrencrona, Marita Hellström Pegg and Richard Rosenquist from Uppsala University; and Marie Stenmark-Askmalm and Signur Liedgren from Linköping University Hospital. UCHICAGO wishes to thank Cecilia Zvosec, 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, NSGIC and Lorna Kwan, MPH, for assembling the data for this study. UICSF wishes to thank Dr. Vardavas and Dr. Nussibaum and the following genetic counselors for participant recruitment: Beth Crawford, Kate Loranger, Julie Mak, Nicola Stewart, Robin Lee, Arnie Blanco and Peggy Conrad, and Ms. Salina Chan for data management. UKFOCORS thanks Paul Pharaoh, Simon Gayther, Susan Ramus, Carole Pye, Patricia Harrington and Eva Wozniak for their contributions towards the UKFOCR. UPENH wishes to thank the Breast Cancer Research Foundation, Susan G. Komen Foundation for the cure and Basier Research Center for BRCA. VFTCG wishes to thank Geoffrey Lindeman, Marion Harris and Martin Delaytcky of the Victorian Familial Cancer Trials Group. VFTCG also thanks Sarah Sawyer and Rebecca Driesen 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/A10214) and by the European Community Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2-2009-223175) (ICOGS). Funding for the ICGOS infrastructure came from the European Community Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2-2009-223175) (ICOGS). Cancer Research UK (C1287/A10118, C1287/A10710, C12922/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C1917/A16656), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA48065 and 1U19 CA48112 - 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 U101 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 ABCFS 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 184.02.007); 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 Health and Medical Research Council of Australia (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. R.B. was a Cancer Institute NSW Clinical Research Fellow. LAAC is supported by grants (1R8-0287, 3 PB-0102, 5 PB-0018, 10 PB-0089) from the California Breast Cancer Research Program. Incident breast cancer cases were collected by the US National 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 N01-CN-25403. LMBC is supported by the Stichting tegen Kanker (232-2008 and 196-2010). Dether Lambrechts is supported by the FWO and the KULPfV/10/016-SymBioSys. The MIARES study was supported by the Deutsche Krebsforschung e.V. (70-2892-BI, 106332, 108253, 108419), the Hamburg Cancer Society, the German Cancer Research Center (DKFZ) and the Federal Ministry of Education and Research (BMBF) Germany (01KH0402). MBCS5 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, CA111617, 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. NCCS was supported by grants from the Missouri Department of Health, Cancer Registry. 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 Ontario Cancer Institute Research Centre (BFBOCC-LT), Research Council of Lithuania grant LIG-07-2012, BIDMC is supported by the Breast Cancer Research Foundation. BRCA-gene mutations were funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The pKARMA study was supported by Marit and Hans Rausing Initiative Against Breast Cancer. The RBCS was funded by the Dutch Cancer Society (DDH 2004-3124, DDH 2009-4318). The SABASC 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 CA148665. The SBCS was supported by Yorkshire Cancer Research S295, S299, S305PA and Sheffield Experimental Cancer Medicine Centre. The SCCS was supported by a grant from the National Institutes of Health (R01 CA922447). 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 were collected by the Arkansas Central 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 grant to the Cancer Research UK (C490/ 101/24) 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 (2012-000347). SGCCS is funded by the NSW start-up Grant, National University Cancer Institute Singapore (NCIS) Centre Grant and the MRC 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 05/121/10/425. SKDKF2S is supported by the DKFZ. The SB2S was supported by grant PB2_122/PO5/2004. The TBCS was funded by The National Cancer Institute Thailand. The TNBBC 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 Secretariat 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 TNBBCS is supported by the Taiwan Biobank project of the Institute of Biomedical Sciences, Academia Sinica, Taiwan. The UCBICS component of this research was supported by the NIH (CA58860, CA92044) and the Lon V Smith Foundation (LVS39420). The UNBICS is funded by Breakfast 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 CA74147) 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 US3SS study was supported 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 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 Ontario Cancer Institute Research Centre (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 (6002002016), FSP0811320, Mutua Madrideria Foundation (FMM) 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 Institute on Cancer 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 CONST NT team were from Italian citizens who allocated the Sx1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects ’Sx1000’) to SM and from ForeGen Foundation for Pharmacogenomics to LP. The CIMBA data management and data analysis were supported by Cancer Research UK grants C12292/A111147 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: SYIN11_10_19 NBCA. Investing in knowledge society through the European Union was supported by the DIFKZ. EMBRACE is supported by Cancer Research UK Grants C1287/A10118 and C1287/A11900. D. Gareth Evans and Fiona Laloo are supported by an NHRI 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 NHRI grant to the Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. Ros Eales and Elizabeth Bancroft are supported by Cancer Research UK Grant CS047/A8385. Ros Eales is also supported by NHRI grant to the Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. FCCCC was supported by The University of Kansas Cancer Center (P30 CA18524) and the Kansas Bioscience Authority Eminent Scholar Program. A.K.G. was funded by SU01CA13916, 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 100976), Rita K. Schmutzer) 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, parons-en" Award; the Canadian Institutes of Health Research for the “OHR Team in Familiar 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 (NH/NCI grant P30-CA051000), the Fisher Center for Familial Cancer Research, and Swing For the Cure. Kim De Leenheer (the G-FAST study) is supported by GOA grant ROF10/GOA619 (Ghent University) and spearhead financing of Ghent University Hospital. The HCS was supported by a grant R01/1200369/0006 and 12/00539/15/ISC (Spain), partially supported by European Regional Development FEDER funds. HEBCS was financially supported by the Helsinki University Hospital Research Fund, Academy of Finland (265628), the Finnish Cancer Society and the Sigrid Juselius Foundation. The HEBON study is supported by the Dutch Cancer Society grants NK1999-1854, NK2004-3088, NK2007-3756, the Netherlands Organization of Scientific Research grant NWO 9110024, the Pink Ribbon grant 110005 and the BBMR grant NWO 184020147/CP46. HEBON thanks the registration teams of the Comprehensive Cancer Centre Netherlands and Comprehensive Cancer 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. HKG was partially supported by Portuguese Cancer Research Fund (FCT 2013-02174, 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 MI CZ - DRO (MIW, 20095805) and by the European Regional Development Fund and the State Budget of the Czech Republic (RECAMO, CZ.05.1.01/0.3.01/01.0011) to LF, and by Charles University in Prague project UAE204024 (M2). NSCC is supported by an NHRI grant to the Biomedical 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, NH, and by support services contracts NO2-CP-11019-50 and NO2-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, NNP0 has been supported by the Russian Federation for Basic Research (grants 13-04-9263, 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, NICI OS160131 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/H 1/MOHE/06) and Cancer Research Initiatives Foundation. SMC was supported by the Spanish Association against Cancer (AECC08), 5x1000 (together the Netherlands Cancer Registry) and PALGA (Dutch Pathology Registry) for part of the data collection. HRBCP is supported by the Intramural Research Program, NICI OS160131 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/H 1/MOHE/06) and Cancer Research Initiatives Foundation. SMC was supported by the Spanish Association against Cancer (AECC08), 5x1000 (together the Netherlands Cancer Registry) and PALGA (Dutch Pathology Registry) for part of the data collection. HRBCP is supported by the Intramural Research Program, NICI OS160131 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/H 1/MOHE/06) and Cancer Research Initiatives Foundation. SMC was supported by the Spanish Association against Cancer (AECC08), 5x1000 (together the Netherlands Cancer Registry) and PALGA (Dutch Pathology Registry) for part of the data collection. HRBCP is supported by the Intramural Research Program, NICI OS160131 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/H 1/MOHE/06) and Cancer Research Initiatives Foundation. SMC.
Author details

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

Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Cancer Institute, Rotterdam, The Netherlands.

Malaya Medical Centre, Kuala Lumpur, Malaysia.

Epidemiology, Stanford University School of Medicine, Stanford, CA, USA.

Aichi Cancer Center Research Institute, Aichi, Japan.

Berghofer Medical Research Institute, Brisbane, QLD, Australia.

Genetic Epidemiology and Statistical Genetics, Harvard School of Public USA.

Hope Clinical Cancer Genetics Community Research Network, Duarte, CA, USA.

University of Helsinki, Helsinki, Finland.

University of California, Fremont, CA, USA.

Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC, USA.

Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Aichi, Japan. 110Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland.

State Research Institute Centre for Innovative Medicine, Vilnius, Lithuania.

Department of Clinical Genetics, Aarhus University Hospital, Aarhus, N, Denmark. 111Department of Epidemiology, Cancer Prevention Institute of California, Freemont, CA, USA.

Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA. 112Genomics Center, Centre Hospitalier Universitaire de Quebec Research Center and Laval University, Quebec City, QC, Canada. 113Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA.

Department of Pathology, Oulu University Hospital, University of Oulu, Oulu, Finland.

School of Medicine, National University of Ireland, Galway, Ireland.

Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. 114Department of Genetics and Fundamental Medicine, Bashke State University, Ufa, Russia. 115Prosman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada. 116Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada.

Molecular Diagnostics Laboratory, IRRP, National Centre for Scientific Research "Demokritos", Athens, Greece. 117The Hong Kong Hereditary Breast Cancer Family Registry, Cancer Genetics Center, Hong Kong Sanatorium and Hospital, Hong Kong, China. 118Department of Surgery, The University of Hong Kong, Hong Kong, China. 119Vesalius Research Center, Leuven, Belgium.

Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium. 120Molecular Diagnostic Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain. 121University of Hawaii Cancer Center, Honolulu, HI, USA. 122Department of Surgery, Soonchunhyang University and Hospital, Seoul, South Korea. 123Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. 124Division of Health Sciences, Warwick Medical School, Warwick University, Coventry, UK. 125Unit 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. 126Division of Biomedical Research and Centre for Rare Diseases, University of Washington, Seattle, WA, USA.

National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. 127Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada. 128Molecular Diagnostics Laboratory, IRRP, National Centre for Scientific Research "Demokritos", Athens, Greece. 117The Hong Kong Hereditary Breast Cancer Family Registry, Cancer Genetics Center, Hong Kong Sanatorium and Hospital, Hong Kong, China. 118Department of Surgery, The University of Hong Kong, Hong Kong, China. 119Vesalius Research Center, Leuven, Belgium.

Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium. 120Molecular Diagnostic Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain. 121University of Hawaii Cancer Center, Honolulu, HI, USA. 122Department of Surgery, Soonchunhyang University and Hospital, Seoul, South Korea. 123Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. 124Division of Health Sciences, Warwick Medical School, Warwick University, Coventry, UK. 125Unit 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. 126Division of Biomedical Research and Centre for Rare Diseases, University of Washington, Seattle, WA, USA.

National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. 127Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada. 128Molecular Diagnostics Laboratory, IRRP, National Centre for Scientific Research "Demokritos", Athens, Greece. 117The Hong Kong Hereditary Breast Cancer Family Registry, Cancer Genetics Center, Hong Kong Sanatorium and Hospital, Hong Kong, China. 118Department of Surgery, The University of Hong Kong, Hong Kong, China. 119Vesalius Research Center, Leuven, Belgium.

Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium. 120Molecular Diagnostic Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain. 121University of Hawaii Cancer Center, Honolulu, HI, USA. 122Department of Surgery, Soonchunhyang University and Hospital, Seoul, South Korea. 123Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. 124Division of Health Sciences, Warwick Medical School, Warwick University, Coventry, UK. 125Unit 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. 126Division of Biomedical Research and Centre for Rare Diseases, University of Washington, Seattle, WA, USA.
Montreal, QC, Canada. 109Currently at Medical School Cambridge University, Cambridge, UK. 109Department of Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH, USA. 109Department of Medical Oncology, Beth Israel Deaconess Medical Center, Boston, MA, USA. 109Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands. 109Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA. 109Department of Obstetrics and Gynecology, University of Ulm, Ulm, Germany. 109Clinical Genetics, for the City of Hope Clinical Genetics Community Research Network, Duarte, CA, USA. 109Laboratory of Cancer Genetics and Tumor Biology, Department of Clinical Chemistry and Biocenter Oulu, University of Oulu, Oulu, Finland. 109Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre Nordlab, Oulu, Finland. 109Singapore Eye Research Institute, National University of Singapore, Singapore, Singapore. 109Department of Medical Oncology, Papageorgiou Hospital, Aristote University of Thessaloniki School of Medicine, Thessaloniki, Greece. 109Department of Surgery, National Taiwan University Hospital, Taipei, Taiwan. 109Department of Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia. 109Peter MacCallum Cancer Centre, The University of Melbourne, Melbourne, Australia.

Received: 23 December 2015 Accepted: 18 May 2016

Published online: 21 June 2016

References


