Common Genetic Variants and Modification of Penetrance of BRCA2-Associated Breast Cancer

Gaudet, Mia M.; Kirchhoff, Tomas; Green, Todd; Vijai, Joseph; Korn, Joshua M.; Guiducci, Candace; Segre, Ayellet V.; McGee, Kate; McGuffog, Lesley; Kartsonaki, Christiana; Morrison, Jonathan; Healey, Sue; Sinilnikova, Olga M.; Stoppa-Lyonnet, Dominique; Mazoyer, Sylvie; Gauthier-Villars, Marion; Sobol, Hagay; Longy, Michel; Frenay, Marc; Hogervorst, Frans B. L.; Rookus, Matti A.; Collee, J. Margriet; Hoogerbrugge, Nicoline; van Roozendaal, Kees E. P.; Piedmonte, Marion; Rubinstein, Wendy; Nerenstone, Stacy; Van Le, Linda; Blank, Stephanie V.; Caldes, Trinidad; de la Hoya, Miguel; Nevanlinna, Heli; Aittomaki, Kristiina; Lazaro, Conxi; Blanco, Ignacio; Arason, Adalgeir; Johannsson, Oskar T.; Barkardottir, Rosa B.; Devilee, Peter; Olopade, Olofenmilayo I.; Neuhausen, Susan L.; Wang, Xianshu; Fredericksen, Zachary S.; Peterlongo, Paolo; Manoukian, Siranoush; Barile, Monica; Viel, Alessandra; Radice, Paolo; Phelan, Catherine M.; Narod, Steven

Published in:
PLoS Genetics

DOI:
10.1371/journal.pgen.1001183

2010

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 17. Nov. 2019
Common Genetic Variants and Modification of Penetration of BRCA2-Associated Breast Cancer


1 Department of Epidemiology and Population Health and Department of Obstetrics and Gynecology and Women’s Health, Albert Einstein College of Medicine, New York, New York, United States of America, 2 Clinical Genetics Service, Department of Medicine, Sloan-Kettering Institute, Memorial Sloan-Kettering Cancer Center, New York, New York, United States of America, 3 Cancer Biology and Genetics Program, Sloan-Kettering Institute, Memorial Sloan-Kettering Cancer Center, New York, New York, United States of America, 4 Broad Institute of Harvard and Massachusetts Institute of Technology, Harvard Medical School, Boston, Massachusetts, United States of America, 5 Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, 6 Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, United States of America, 7 Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, United States of America, 8 Center for Cancer Research, Cancer Inflammation Program, Human Genetics Section, National Cancer Institute – Frederick, Frederick, Maryland, United States of America, 9 Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom, 10 Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Centre Hospitalier Universitaire de Lyon/Centre Léon Bérard, Lyon, France, 11 Équipe labellisée LIGUE 2008, UMR2501 CNRS, Centre Léon Bérard, Université de Lyon, Lyon, France, 12 Institut Curie, Service de Génétique, INSERM U830, F-75248, Université Paris Descartes, Paris, France, 13 Service de Génétique Oncologique, Institut Curie, Paris, France, 14 Département Oncologie génétique, Prévention et Dépistage, INSERM CIC-P9502, Institut Paoli-Calmettes/Université d’Aix-Marseille II, Marseille, France, 15 Institut Bergonié, Bordeaux, France, 16 Centre Antoine Lacassagne, Nice, France, 17 GEMO Study - Cancer Genetics Network “Groupe Génétique et Cancer”, Fédération Nationale des Centres de Lutte Contre le Cancer, Paris, France, 18 Family Cancer Clinic, Netherlands Cancer Institute, Amsterdam, The Netherlands, 19 Department of Epidemiology, Netherlands Cancer Institute, Amsterdam, The Netherlands, 20 Department of Medical Oncology, Rotterdam Family Cancer Clinic, Erasmus University Medical Center, Rotterdam, The Netherlands, 21 Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, 22 Department of Clinical Genetics, University Medical Center, Maastricht, The Netherlands, 23 Gynecologic Oncology Group Statistical and Data Center, Roswell Park Cancer Institute, Buffalo, New York, United States of America, 24 NorthShore University Health System, Evanston, Illinois, United States of America, 25 Central Connecticut Cancer Consortium, Hartford Hospital, Hartford, Connecticut, United States of America, 26 University of North Carolina, Chapel Hill, North Carolina, United States of America, 27 New York University School of Medicine, New York, New York, United States of America, 28 Molecular Oncology Laboratory, Hospital Clinico San Carlos, Madrid, Spain, 29 Department of Obstetrics and Gynecology, Helsinki University Central Hospital, Helsinki, Finland, 30 Department of Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland, 31 Hereditary Cancer Program, Catalan Institute of Oncology, Barcelona, Spain, 32 Department of Oncology, Landspítali–LSH, Reykjavik, Iceland, 33 Department of Pathology, Landspítali–LSH, Reykjavik, Iceland, 34 Faculty of Medicine, University of Iceland, Reykjavik, Iceland, 35 Department of Human Genetics and Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands, 36 Center for Clinical Cancer Genetics and Global Health, Department of Medicine, University of Chicago Medical Center, Chicago, Illinois, United States of America, 37 Department of Population Sciences, the Beckman Research Institute of the City of Hope, Duarte, California, United States of America, 38 Laboratory of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, United States of America, 39 Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, United States of America, 40 Unit of Genetic Susceptibility to Cancer, Department of Experimental Oncology and Molecular Medicine, Fondazione IRCCS Istituto Nazionale Tumori (INT), Milan, Italy, 41 IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, Milan, Italy, 42 Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Milan, Italy, 43 Division of Cancer...
BRCA2 GWAS

Prevention and Genetics, Istituto Europeo di Oncologia (IEO), Milan, Italy, 44 Division of Experimental Oncology 1, Centro di Riferimento Oncologico (CRO), IRCCS, Aviano (PN), Italy, 45 Moffitt Cancer Center, Tampa, Florida, United States of America, 46 Women’s College Research Institute, Toronto, Canada, 47 CHS National Cancer Control Center and Department of Community Medicine and Epidemiology, Carmel Medical Center, Haifa, Israel, 48 Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada, 49 Cancer Care Ontario, Ontario Cancer Genetics Network, University of Toronto, Toronto, Canada, 50 Departments of Molecular Virology, Immunology, and Medical Genetics and Internal Medicine, Ohio State University, Columbus, Ohio, United States of America, 51 Immunology and Molecular Oncology Unit, Istituto Oncologico Veneto, IRCCS, Padua, Italy, 52 Department of Oncology and Surgical Sciences, University of Padua, Padua, Italy, 53 The Susan Leyt Gertner Oncogenetics Unit, Institute of Genetics, Sheba Medical Center, Tel Hashomer, Israel, 54 Department of Oncology, Lund University, Lund, Sweden, 55 Division of General Internal Medicine, Department of Medicine, University of California San Francisco, San Francisco, California, United States of America, 56 Gynaecological Oncology Unit, UCL EGA Institute for Women’s Health, University College London, United Kingdom, 57 Department of Oncology, The Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, 58 Department of Cell and Molecular Biology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, United States of America, 59 Center for Clinical Epidemiology and Biostatistics, Department of Biostatistics and Epidemiology, The University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, United States of America, 60 Genetics and Population Health Division, Queensland Institute of Medical Research, Brisbane, Australia, 61 Peter MacCallum Cancer Centre, Melbourne, Australia, 62 Cancer Prevention Institute of California, Fremont, California, United States of America, 63 Centre for Genetic Epidemiology, University of Melbourne, Melbourne, Australia, 64 Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, United States of America, 65 Fox Chase Cancer Center, Philadelphia, Pennsylvania, United States of America, 66 Centre for Genetic Epidemiology, University of Melbourne, Melbourne, Australia, 67 Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, United States of America, 68 Division of Hematology-Oncology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, United States of America, 69 Department of Clinical Biochemistry, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark, 70 Clinical Genetics Branch, National Cancer Institute, Rockville, Maryland, United States of America, 71 Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Research Centre, Madrid, Spain, 72 Institute of Biology and Molecular Genetics, Universidad de Valladolid (IBGM-UVA), Valladolid, Spain, 73 Oncology Service, Hospital Clínico Universitario Lozano Blesa, Zaragoza, Spain, 74 Human Genetics Group and Genotyping Unit, Human Cancer Genetics Programme, Spanish National Cancer Research Centre, Madrid, Spain, 75 City of Hope Cancer Center, Duarte, California, United States of America, 76 Dana Farber Cancer Institute, Harvard University, Boston, Massachusetts, United States of America, 77 Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany, 78 Genetic Medicine, Manchester Academic Health Sciences Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester, United Kingdom, 79 Oncogenetics Team, The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, United Kingdom, 80 Clinical Genetics, Guy’s and St. Thomas’ NHS Foundation Trust, London, United Kingdom, 81 Oxford Regional Genetics Service, Churchill Hospital, Oxford, United Kingdom, 82 Nottingham Clinical Genetics Service, Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom, 83 Leicestershire Clinical Genetics Service, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom, 84 Women’s Cancer Program, Department of Medical Oncology, Fox Chase Cancer Center, Philadelphia, Pennsylvania, United States of America, 85 Centre of Familial Breast and Ovarian Cancer, Department of Gynaecology and Obstetrics and Centre for Integrated Oncology (CIO), University Hospital of Cologne, Cologne, Germany, 86 Department of Gynaecology and Obstetrics, Division of Tumor Genetics, Klinikum rechts der Isar, Technical University Munich, Munich, Germany, 87 Department of Gynaecology and Obstetrics, University Hospital of Schleswig-Holstein, Christian-Albrechts University Kiel, Kiel, Germany, 88 Institute of Cell and Molecular Pathology, Hannover Medical School, Hannover, Germany, 89 Program in Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, New York, New York, United States of America

Abstract

The considerable uncertainty regarding cancer risks associated with inherited mutations of BRCA2 is due to unknown factors. To investigate whether common genetic variants modify penetrance for BRCA2 mutation carriers, we undertook a two-staged genome-wide association study in BRCA2 mutation carriers. In stage 1 using the Affymetrix 6.0 platform, 592,163 filtered SNPs genotyped were available on 899 young (<40 years) affected and 804 unaffected carriers of European ancestry. Associations were evaluated using a survival-based score test adjusted for familial correlations and stratified by country of the study and BRCA2*6174delT mutation status. The genomic inflation factor (λ) was 1.011. The stage 1 association analysis revealed multiple variants associated with breast cancer risk: 3 SNPs had p-values <10−5 and 39 SNPs had p-values<10−4. These variants included several previously associated with sporadic breast cancer risk and two novel loci on chromosome 20 (rs311499) and chromosome 10 (rs16917302). The chromosome 10 locus was in linkage disequilibrium with BRCA2*6174delT. The combined per allele HR of the minor allele for the novel locus rs16917302 was 0.75 (95% CI 0.66–0.86, p = 3.8×10−5) and for rs311499 was 0.72 (95% CI 0.61–0.85, p = 6.6×10−5). FGR2 rs2981575 had the strongest association with breast cancer risk (per allele HR = 1.28, 95% CI 1.18–1.39, p = 1.2×10−5). These results indicate that SNPs that modify BRCA2 penetrance identified by an agnostic approach thus far are limited to variants that also modify risk of sporadic BRCA2 wild-type breast cancer.


Editor: James M. Ford, Stanford University School of Medicine, United States of America

Received June 8, 2010; Accepted September 28, 2010; Published October 28, 2010

This is an open-access article distributed under the terms of the Creative Commons Public Domain declaration which stipulates that, once placed in the public domain, this work may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose.

Funding: This study was supported by the Starr Cancer Research Consortium and the Breast Cancer Research Foundation, as well as by NIH NCI: P20CA103694-3 and the Lymphoma Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: olftik@msccc.org

† These authors contributed equally to this work.

‡ Membership of each consortium is provided in the Acknowledgements.
Author Summary

The risk of breast cancer associated with BRCA2 mutations varies widely. To determine whether common genetic variants modify the penetrance of BRCA2 mutations, we conducted the first genome-wide association study of breast cancer among women with BRCA2 mutations using a two-stage approach. The major finding of the study is that only those loci known to be associated with breast cancer risk in the general population, including FGR2 (rs2981575), modified BRCA2-associated risk in our high-risk population. Two novel loci, on chromosomes 10 in ZNF365 (rs16917302) and chromosome 20 (rs311499), were shown to modify risk in BRCA2 mutation carriers, although not at a genome-wide level of significance. However, the ZNF365 locus has recently independently been associated with breast cancer risk in sporadic tumors, highlighting the potential significance of this zinc-finger-containing gene in breast cancer pathogenesis. Our results indicate that it is unlikely that other common variants have a strong modifying effect on BRCA2 penetrance.

Introduction

After more than a decade of clinical testing for mutations of BRCA1 and BRCA2, there remains considerable uncertainty regarding cancer risks associated with inherited mutations of these genes. This variable penetrance is most striking for BRCA2 [1–4], and it affects medical management [5]. Women with the same BRCA2 mutation may develop breast, ovarian or other cancers at different ages or not at all [6]. In a segregation analysis of families identified through breast cancer cases diagnosed before age 55, the residual familial clustering after accounting for BRCA1 and BRCA2 mutations could be explained by a large number of low penetrance genes with multiplicative effects on breast cancer risk [7,8]. A candidate gene approach in BRCA2 mutation carriers led to the discovery of loci that modify the penetrance of BRCA2 mutations, such as RAD51 135 G>C [9] and perhaps CASP8 [10,11] and IGFBP2 [12], if replicated. To investigate whether other common single nucleotide polymorphisms (SNP), copy number variants (CNV), or copy number polymorphisms (CNP) modify penetrance for BRCA2 mutation carriers, we undertook a two-staged genome-wide association study (GWAS) in BRCA2 mutation carriers from the international Consortium for Investigators of Modifiers of BRCA1/2 (CIMBA) and other international studies. We hypothesized that an agnostic search for breast cancer loci in an enriched population of BRCA2 mutation carriers, the first among this high risk population, would provide greater power than a sporadic population of equal number, and would yield associations specific to BRCA2 carriers and/or the general population.

Results

Stage 1 and Stage 2 Genotyping

In stage 1, genotype data were available for 899 young (<40 years) affected and 804 older (>40 years) unaffected carriers of European ancestry after quality control filtering and removal of ethnic outliers (Figure S1). A total of 592,163 filtered SNPs genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 platform passed quality control assessment. In stage 1, comparison of the observed and expected distributions (quantile-quantile plot: Figure S2) showed little evidence for an inflation of the test statistics (genomic inflation factor λ = 1.01), thereby excluding the possibility of significant hidden population substructure, cryptic relatedness among subjects or differential genotype calling between BRCA2 affected and BRCA2 unaffected carriers. Multiple variants were found to be associated with breast cancer risk (Figure S3): 3 SNPs had p<10^-7 and 39 SNPs had p<10^-4. The most significant association (p=3.6×10^-6) was observed for FGR2 rs2981582 (Table 1), a variant previously shown to be associated with increased risk of BRCA2-related breast cancer [13]. A positive association was also observed with rs3003662 (Table 1), near TOX3, which has also been associated with sporadic breast cancer risk [13].

Using the stage 1 data, we also performed a GSEA as implemented in MAGENTA [14] to evaluate whether a functionally-related set of genes relevant to BRCA2 function (Table S1) was enriched for relative risk associations (see Statistical Methods). The 59 genes selected are related to the Fanconi anemia pathway [15] as well as other pathways reported in the literature to regulate or interact with BRCA1/2 [16]. These showed no enrichment of associations with the breast cancer risk (p = 0.56). In addition, eight of 125 known cancer susceptibility alleles identified by previous GWAS of other cancers [17] were associated with BRCA2 modification in the current study, a number not greater than expected (Kolmogorov-Smirnov p = 0.60) by chance alone. Of the 113 most significantly associated SNPs (p<10^-5) in our study, three showed significant association (p<0.05) with BRCA2-associated breast cancer risk in a complimentary GWAS [18].

In the combined stage 1 and stage 2 results, four independent SNPs (pairwise r^2<0.80) were associated with increased risk of breast cancer risk with p-values<10^-4 (Table 1). Previously identified breast cancer susceptibility loci [13,19,20] had the most significant associations among BRCA2 mutation carriers (FGR2: per allele p-value = 1.2×10^-3 and TOX3: per allele p-value = 4.9×10^-5). Novel loci, rs16917302 on chromosome 10 and rs311499 on chromosome 20, had HRs in stage 2 that were in the same direction as those observed for stage 1 (Figure 1, Table 1), but were smaller in magnitude (HR = 0.67 (95% CI:0.56-0.80) vs. 0.83 (95% CI: 0.70–1.04) for rs16917302; HR = 0.60 (95%CI:0.50-0.70) vs. 0.84 (95%CI: 0.67–1.06) for rs311499) perhaps reflecting a “winner’s curse” effect [21]. The associations for these SNPs were not statistically significant in stage 2 (Table 1). In the combined stage 1 and stage 2 dataset, the C allele of rs16917302 was associated with lower risk of breast cancer (per allele HR = 0.75, 95% CI 0.66–0.86; p = 2.7×10^-5; Table 1), and the C allele of rs311499 was associated with a reduced risk (per allele HR = 0.72, 95% CI 0.61–0.85; p = 7.9×10^-5; Table 1). A full list of stage 2 results can be found in Table S2. Using the combined stage 1 and stage 2 data, there was no evidence that the HR for SNP rs16917302 changes with age (p=0.63), but there was some evidence that the per-allele HR for rs311499 may increase with age (p = 0.034).

Copy Number Variant Analysis

We also examined the association of both high-frequency CNPs and low-frequency CNVs to case-control status using the stage 1 data. After performing standard quality control measures including a minor allele frequency (MAF) threshold of 5%, we identified 191 polymorphisms with reliable genotypes. No associations were found between CNVs and the phenotype; there was no inflation or deflation of the test statistic, and the best p-value was 4×10^-3. We similarly assessed less common CNPs, and found neither the overall burden of events (or any subclass thereof, such as large deletions overlapping genes) nor any specific locus associated with breast cancer risk (Figure S4).
Excess Sharing in Genetic Isolates and Outbred Populations Analyzed

Because of the prior evidence of significant LD extent around the 6174delT (c.5946delT) founder mutation in the Ashkenazi Jewish population [22], we explored the potential excess sharing of the genome compared to the BRCA2 region in both Ashkenazi Jewish and non-Jewish European ancestries. Using GERMLINE [23], shared segments of greater than 5 cM were computed based on the imputed genotype dataset. In the BRCA2 region, we observed a significant excess of sharing amongst both Ashkenazi (n = 304) and non-Jewish (n = 1331) individuals compared to samples from an autism study (n = 808) suggesting common and non-Jewish (n = 1331) individuals compared to observed a significant excess of sharing amongst both Ashkenazi Jewish and non-Jewish European ancestries. Using GERMLINE [23], shared segments of greater than 5 cM were computed based on the imputed genotype dataset. In the BRCA2 region, we observed a significant excess of sharing amongst both Ashkenazi (n = 304) and non-Jewish (n = 1331) individuals compared to samples from an autism study (n = 808) suggesting common founders for BRCA2 mutations. Examining sites across the genome every 2.5 cM (excluding telomere and centromere regions), we observed possible pairs share segments greater than 5 cM that on average 0.005% (n = 50,177, s.d = 55.3, max = 491) for non-Jewish individuals and 0.12% (n = 141.11, s.d = 57.32, max = 525) for Ashkenazi Jewish individuals. Comparing cases and controls, we did not observe a significant difference in number of pairs of samples sharing segments greater than 5cM across the genome excluding chromosome 13. That is, there was no evidence of overall excess sharing across the genome other than for the BRCA2 locus within the Ashkenazi Jewish and non-Ashkenazi Jewish populations in the study.

Discussion

In this GWAS of BRCA2 mutation carriers, the first in this high risk population, we found previously identified breast cancer susceptibility loci modified risk of BRCA2-associated breast cancer with similar magnitude of association. Although FGR2 (rs2981575) was the only locus to reach genome-wide statistical significance, novel loci, rs16917302 and rs10509168 were each (rs2981575) was the only locus to reach genome-wide statistical significance, novel loci, rs16917302 and rs10509168 were each associated with breast cancer risk.

rs16917302 is located on chromosome 10, in the zinc finger protein 363 gene (ZNF363). A recent multistage GWAS of 15,992 sporadic breast cancer cases and 16,891 controls also observed an inverse association (per allele OR = 0.82, 95% CI 0.82-0.91, p = 5.1 x 10^-15) between breast cancer risk and rs10509168, a SNP 18kb from rs16917302 (pairwise r^2 = 0.1) and located in intron 4 of ZNF363 [24]. Of the 3,659 cases and 4,897 controls in phase 1 of that study, imputation revealed that the locus identified in our BRCA2 study, rs10509168, was significantly associated with risk for breast cancer (p = 0.02) (Easton DF, personal communication). The second novel SNP in the current study, rs311499, is located on chromosome 20, within a region containing several possible candidate genes including GMEB2, SRMS, PTK6, STMN3, and TNFRSF6 among others.

Table 1. Estimates of breast cancer association for loci (two confirmatory loci at FGFR2 and TOX3, and two novel loci with stage 1 and 2 combined of p<10^-4) among BRCA2 mutation carriers in a two-staged genome-wide association study.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Chr.</th>
<th>Stage 1 N (Controls/ Cases)</th>
<th>Stage 2 N (Controls/ Cases)</th>
<th>Stage 1 and 2 Combined N (Controls/ Cases)</th>
<th>HR (95% CI) p-value^2</th>
<th>HR (95% CI) p-value^2</th>
<th>HR (95% CI) p-value^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>FGFR2</td>
<td>rs2981575</td>
<td>10</td>
<td>794/892</td>
<td>1,222/1,263</td>
<td>2,016/2,155</td>
<td>1.30 (1.16–1.45) 1.28 (1.18–1.39) 1.2 x 10^-8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOX3</td>
<td>rs3803662</td>
<td>16</td>
<td>804/899</td>
<td>1,222/1,263</td>
<td>2,026/2,162</td>
<td>1.19 (1.05–1.34) 1.22 (1.07–1.39) 2.8 x 10^-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZNF365</td>
<td>rs16917302</td>
<td>10</td>
<td>804/898</td>
<td>1,222/1,264</td>
<td>2,026/2,162</td>
<td>0.67 (0.56–0.80) 0.85 (0.70–0.94) 0.14 (0.07–0.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMEB2, etc.</td>
<td>rs311499</td>
<td>20</td>
<td>792/882</td>
<td>1,209/1,255</td>
<td>2,001/2,137</td>
<td>0.60 (0.47–0.78) 0.84 (0.67–0.96) 0.13 (0.07–0.22)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^1 p-value was calculated based on the 1-degree of freedom score test statistic stratified by country of study and 6174delT (c.5946delT) mutation status, and modified to allow for the non-independence among related individuals.

^2Per allele hazard ratios (HR) (i.e., multiplicative model) were estimated on the log scale, assuming independence of age, using the retrospective likelihood. All analyses were stratified by country of residence and 6174delT (c.5946delT) mutation status, and used calendar-year- and cohort-specific breast cancer incidence rates for BRCA2.

The combined stage 1 and stage 2 analyses were also stratified by stage.

^3The region also includes other possible genes including SRMS, PTK6, STMN3, and TNFRSF6 among others.

doi:10.1371/journal.pgen.1001183.t001

In this GWAS of BRCA2 mutation carriers, the first in this high risk population, we found previously identified breast cancer susceptibility loci modified risk of BRCA2-associated breast cancer with similar magnitude of association. Although FGR2 (rs2981575) was the only locus to reach genome-wide statistical significance, novel loci, rs16917302 and rs10509168 were each associated with breast cancer risk.

rs16917302 is located on chromosome 10, in the zinc finger protein 363 gene (ZNF363). A recent multistage GWAS of 15,992 sporadic breast cancer cases and 16,891 controls also observed an inverse association (per allele OR = 0.82, 95% CI 0.82-0.91, p = 5.1 x 10^-15) between breast cancer risk and rs10509168, a SNP 18kb from rs16917302 (pairwise r^2 = 0.1) and located in intron 4 of ZNF363 [24]. Of the 3,659 cases and 4,897 controls in phase 1 of that study, imputation revealed that the locus identified in our BRCA2 study, rs10509168, was significantly associated with risk for breast cancer (p = 0.02) (Easton DF, personal communication). The second novel SNP in the current study, rs311499, is located on chromosome 20, within a region containing several possible candidate genes including GMEB2, SRMS, PTK6, STMN3, and TNFRSF6 among others.
Figure 1. Association signals, genetic structure, and linkage disequilibrium of the novel modifier loci of BRCA2 penetrance in the regions surrounding rs1691730 on chromosome 10 and rs311499 on chromosome 20. The color of the dots indicates linkage disequilibrium (LD; based on r2 values) in the CEU population (as per scale). Triangle plots below represent LD from actual data of BRCA2 carriers in the study.
doi:10.1371/journal.pgen.1001183.g001
replicated, the two additional SNPs identified here would only explain about 1.7% of the variance in breast cancer risk among BRCA2 mutation carriers. Taken together, the combined effects of all the common and putative risk modifiers in this study only account for ~4% of the variance of BRCA2 mutations, compared with 1.1% for the single RAD51 135 G>C variant, which is rare and biologically-linked to BRCA2 function, as shown by candidate gene studies [9]. Thus, the common alleles that modify risk in BRCA1 and BRCA2 backgrounds appear to have comparable associated risks in sporadic ER-positive and ER-negative tumors, respectively [18]. While individual SNPs are unlikely to be used to guide radiographic screening and risk-reducing surgical strategies, the combined effect of these SNPs may ultimately be used for the tailor management of subsets of BRCA1 mutation carriers [5].

While we took great efforts to collect all of the possible known BRCA2 mutation carriers, there were insufficient numbers to stratify by race and BRCA2 mutations with the exception of BRCA2*6174delT mutations. Due to the small numbers of women of non-European ancestry who have participated in the individual studies represented here, the current analysis was based only on women who had genetic backgrounds consistent with HapMap CEU samples. While we expect that SNPs identified among women of European ancestry might also be applicable to women of other genetic backgrounds, additional research in these populations will be needed. Similarly, the observed associations represented across all types of mutations, and specifically a weighted average of BRCA2*6174delT and non-delT mutations. It is possible that the observed associations may only modify the penetrance of specific BRCA2 mutations due to differential effects on function or differences in genetic background. Our analysis was stratified on the basis of the most common BRCA2 mutation, BRCA2*6174delT, which is prevalent in individuals with an Ashkenazi Jewish ancestry. Large numbers of mutation carriers will be necessary to calculate mutation-specific estimates. In addition, there was a drop-out of SNPs in the two phases of this study. While we were able to achieve a representative coverage of the genome, it is also possible that additional studies using denser arrays may provide further information.

As expected, we observed associations with some of the major common genetic variants seen in genome-wide scans of breast cancer in a non-BRCA1/2 mutation background. However, we found no evidence for loci with stronger effects than FGFR2. Although we observed an association with a novel locus at \( \Xi_{NF365} \) that appears also to be a risk factor for sporadic breast cancer, overall, our results suggest that there are no common variants with major effects (i.e., OR>2.0) that are specific in BRCA2 carriers. Similarly, in a recent report of SNPs from sporadic breast cancer GWAS genotyped in a restricted set of BRCA1/2 carriers [27], loci in \( LOC134997 \) (rs9393597: per allele HR = 1.55, 95% CI 1.25–1.92, \( p = 6.0 \times 10^{-5} \)) and \( FBXL7 \) (rs12652447: \( HR = 1.37, 95\% \) CI 1.16–1.62, \( p = 1.7 \times 10^{-4} \)) were associated with BRCA2 breast cancer risk with p-values weaker than FGFR2 reported here (per allele \( p-value = 1.2 \times 10^{-8} \)), although the magnitudes of the associations were slightly stronger than FGFR2 (HR = 1.28). Although these SNPs were not in our genotyped panel of SNPs at stage 1, imputation results indicate that SNP rs9393597 has a p-value of 0.008 and SNP rs12652447 a p-value of 0.04 for association with breast cancer risk for the BRCA2 mutation carriers in our stage1. However, there is substantial overlap between our study and the study of Wang et al. [27].

Replication in larger datasets will be necessary to precisely estimate the magnitude of the associations of suspected loci identified from our study, candidate gene analysis [10–12], and other selection approaches [27]. It is of interest, however, that when utilizing an agnostic approach in BRCA2 mutation carriers in this study, the major determinants of risk variation in mutation carriers are those that also modify risk in subsets of sporadic, BRCA1/2 wild type, breast cancer. However, it remains possible that unique variants with smaller effects, or rarer variants (not evaluated in this experiment), may be specific modifiers of breast cancer risk in BRCA2 carriers. Their detection would require study populations much larger than the current analysis, which is presently the largest such cohort assembled.

Materials and Methods

Study Subjects

Ethics statement. All carriers were recruited to studies (Table 2) at the host institutions under IRB-approved protocols.

Selection of affected individuals and controls. A total of 6,272 BRCA2 carriers from 39 studies (Table 2) and 14 countries contributed DNA samples for this project. With the exception of NICC, all studies are members of the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) [28]. Recruitment of carriers was conducted predominantly through cancer genetics clinics, and enrolled through national or regional efforts. Other studies were recruited through population-based or community-based ascertainment. All subjects provided written informed consent. Eligible female carriers were aged 18 years or older, were self-reported ‘white’, and had mutations in BRCA2. Data were available on age at study recruitment, age at cancer diagnosis, age of bilateral prophylactic mastectomy, BRCA1/2 mutation description, and self-reported ethnicity. Only a limited number of cases had detailed information on tumor characteristics (e.g., estrogen and progesterone receptor status); therefore, subtype analyses were not performed at this stage.

Genotyping and Quality Control

Stage 1 Affymetrix genotyping. All eligible DNA samples provided by participating centers were subjected to a rigorous quality control assessment, including measures of overall DNA quality and quantity. A total of 1,156 young (≤50 years) affected women and 1,038 unaffected women with high quality DNA samples were selected (Table 2). For time efficiency, stage 1 genotyping occurred in two phases: phase 1 included 421 cases and 404 controls and phase 2 included 735 cases and 634 controls.

Prior to the genome-wide scan, we genotyped five SNPs previously genotyped by the CIMBA study centers as a pre-filter for sample identification. Thirty-one samples (Figure S1) were discordant in the two genotyping rounds and were excluded from further analysis.

The genotyping for the stage 1 GWAS was performed on 2,163 eligible carriers using the Affymetrix 6.0 GeneChip array that included 906,622 SNPs (Figure S1). To further monitor the identity of the DNA samples, a fingerprinting panel of 14 SNPs with a minor allele frequency >10% in HapMap European individuals were genotyped on all samples, using Sequenom iPLEX, before and after Affymetrix genotyping. The AMG gender assay was used for gender assessment. As an additional quality control measure, cases and controls were interleaved on each plate to eliminate technical bias. Each plate also included one HapMap CEU DNA sample.

The DNA samples and genotyping calls for both phases of stage 1 were filtered through a series of data quality control parameters using the Birdseed module of the Birdsuite software developed at Broad Institute [29]. Among the 2,163 samples genotyped in the stage 1 GWAS, 253 failed to hybridize to the chip due to poor DNA quality and were excluded (Figure S1). Fifty-five samples were dropped with call rates <95%. Three samples were
Table 2. Description of affected and unaffected carriers selected for BRCA2 GWAS Stage 1 and 2.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Stage 1</th>
<th></th>
<th>Stage 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Affected (n = 1,156)</td>
<td>Unaffected (n = 1,038)</td>
<td>Affected (n = 1,524)</td>
<td>Unaffected (n = 1,508)</td>
</tr>
<tr>
<td><strong>Age at Censoring</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 40</td>
<td>763</td>
<td>66.7</td>
<td>11</td>
<td>1.1</td>
</tr>
<tr>
<td>40–44</td>
<td>308</td>
<td>26.9</td>
<td>230</td>
<td>22.2</td>
</tr>
<tr>
<td>45–49</td>
<td>72</td>
<td>6.3</td>
<td>223</td>
<td>22.4</td>
</tr>
<tr>
<td>50–54</td>
<td>1</td>
<td>0.1</td>
<td>176</td>
<td>17.0</td>
</tr>
<tr>
<td>55–59</td>
<td>0</td>
<td>0.0</td>
<td>138</td>
<td>13.3</td>
</tr>
<tr>
<td>60+</td>
<td>0</td>
<td>0.0</td>
<td>248</td>
<td>24.0</td>
</tr>
<tr>
<td><strong>Self-reported Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>125</td>
<td>10.9</td>
<td>80</td>
<td>7.7</td>
</tr>
<tr>
<td>Caucasian</td>
<td>873</td>
<td>76.3</td>
<td>723</td>
<td>69.9</td>
</tr>
<tr>
<td>Ashkenazi Jewish</td>
<td>146</td>
<td>12.8</td>
<td>232</td>
<td>22.4</td>
</tr>
<tr>
<td><strong>DELF Mutation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carrier</td>
<td>161</td>
<td>14.1</td>
<td>271</td>
<td>26.2</td>
</tr>
<tr>
<td>Non-carrier</td>
<td>983</td>
<td>85.9</td>
<td>764</td>
<td>73.8</td>
</tr>
<tr>
<td><strong>Country of Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>109</td>
<td>9.5</td>
<td>82</td>
<td>7.9</td>
</tr>
<tr>
<td>Canada</td>
<td>98</td>
<td>8.6</td>
<td>172</td>
<td>16.6</td>
</tr>
<tr>
<td>Denmark</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>France</td>
<td>52</td>
<td>4.5</td>
<td>25</td>
<td>2.4</td>
</tr>
<tr>
<td>Finland</td>
<td>27</td>
<td>2.4</td>
<td>27</td>
<td>2.6</td>
</tr>
<tr>
<td>Germany</td>
<td>68</td>
<td>5.9</td>
<td>31</td>
<td>3.0</td>
</tr>
<tr>
<td>Iceland</td>
<td>25</td>
<td>2.2</td>
<td>9</td>
<td>0.9</td>
</tr>
<tr>
<td>Israel</td>
<td>49</td>
<td>4.3</td>
<td>87</td>
<td>8.4</td>
</tr>
<tr>
<td>Italy</td>
<td>110</td>
<td>9.6</td>
<td>44</td>
<td>4.3</td>
</tr>
<tr>
<td>Spain</td>
<td>107</td>
<td>9.4</td>
<td>71</td>
<td>6.9</td>
</tr>
<tr>
<td>Sweden</td>
<td>13</td>
<td>1.1</td>
<td>13</td>
<td>1.3</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>15</td>
<td>1.3</td>
<td>26</td>
<td>2.5</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>181</td>
<td>15.8</td>
<td>179</td>
<td>17.3</td>
</tr>
<tr>
<td>USA</td>
<td>290</td>
<td>25.4</td>
<td>290</td>
<td>28.0</td>
</tr>
<tr>
<td><strong>Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCFR-Australia</td>
<td>19</td>
<td>1.7</td>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>BCFR-NCCC</td>
<td>12</td>
<td>1.0</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>BCFR-Ontario</td>
<td>29</td>
<td>2.5</td>
<td>28</td>
<td>2.7</td>
</tr>
<tr>
<td>BCFR-UT</td>
<td>18</td>
<td>1.6</td>
<td>18</td>
<td>1.7</td>
</tr>
<tr>
<td>BCFR-FCCC</td>
<td>2</td>
<td>0.2</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>BCFR-NY</td>
<td>4</td>
<td>0.3</td>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>BIDMC</td>
<td>10</td>
<td>0.9</td>
<td>20</td>
<td>1.9</td>
</tr>
<tr>
<td>CBCS</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>CGB_NCI</td>
<td>7</td>
<td>0.6</td>
<td>15</td>
<td>1.4</td>
</tr>
<tr>
<td>CNIO</td>
<td>49</td>
<td>4.3</td>
<td>33</td>
<td>3.2</td>
</tr>
<tr>
<td>COH</td>
<td>30</td>
<td>2.6</td>
<td>13</td>
<td>1.3</td>
</tr>
<tr>
<td>DFCI</td>
<td>14</td>
<td>1.2</td>
<td>22</td>
<td>2.1</td>
</tr>
<tr>
<td>DKFZ</td>
<td>7</td>
<td>0.6</td>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>EMBRACE</td>
<td>178</td>
<td>15.6</td>
<td>173</td>
<td>16.7</td>
</tr>
<tr>
<td>FCCC</td>
<td>14</td>
<td>1.2</td>
<td>10</td>
<td>1.0</td>
</tr>
<tr>
<td>GC-HBOC</td>
<td>61</td>
<td>5.3</td>
<td>26</td>
<td>2.5</td>
</tr>
</tbody>
</table>
contaminated, 43 were identified by genotyping to be duplicates, and 4 were male; all were dropped from analyses.

SNPs were also filtered using Birdseed and were removed if monomorphic or 10% missing (n = 38,962), genotype call rates 95% (n = 50,810), minor allele frequencies 1% (n = 104,792), departures from Hardy-Weinberg Equilibrium (p < 10^-6; n = 1,090), differential missingness with respect to phenotype (p < 10^-3; n = 275), and differential missingness with respect to nearby SNPs (p < 10^-10; n = 22,065). A total of 6,212 SNPs had different missingness patterns in phase 1 compared to phase 2, and were excluded. Since we found that significant missingness correlated to SNPs mapping to longer fragments of Affymetrix 6.0 digestion products, we also removed the SNPs on fragments longer than 1000bp (n = 85,990).

With the remaining 1,805 carriers and 596,426 SNPs, an iterative process proceeded to drop all individuals with low call rates (>95%), high autosomal heterozygosity rates (false discovery rate >0.1%), and high identity by descent scores (≥0.95) and to drop all SNPs with minor allele frequencies <1% and SNP call rates <95% until the final run contained individuals above the individual and SNP filter thresholds (n = 1,747 samples and 592,566 SNPs). A more stringent HWE filter (p < 10^-7) was then applied and 403 additional SNPs were dropped. Nine individuals with missing mutation descriptions were removed. Finally, principal components analysis was used to identify the ethnic outliers (Figure S5). A total of 1,743 BRCA2 mutation carriers and the HapMap3 data for 210 individuals of European (CEU), Han Chinese (CHB), and Yoruba (YRI) African descent were available for multidimensional scaling using the genomic kinship matrix estimated using a set of 53,641 autosomal and uncorrelated SNPs. A cut-off of >0.1% was used to exclude samples with non-CEU ancestry (n = 35). Genotype-phenotype association analyses were based on 1,703 (899 young affected and 804 unaffected) BRCA2 mutation carriers and 592,163 SNPs, covering 85% of the common HapMap 3 SNPs (imputed with r^2 ≥ 0.8 (see below), including 64% of the markers that were removed in the QC process).

Where directly genotyped data were not available, probabilities were imputed with Beagle 3.0.2 (using the default parameters) using CEU + TSI samples on HapMap3 release2 B36 as the reference panel (410 chromosomes, 1.4 M SNPs).

**Stage 2 Sequenom iPLEX genotyping.** The primary SNP selection strategy was based on the results of the kinship-adjusted score test of 592,163 GWAS genotyped SNPs. From stage 1, a total of 79 top independent regions (p ≤ 1.5 × 10^-4) with pairwise r^2 values <0.80 were selected for genotyping in stage 2 (Figure S6). For the top 10 SNPs if available, an additional correlated SNP (pairwise r^2 values >0.8; n = 5) was selected to serve as genotyping
back. The remaining SNPs for stage 2 were selected based on two alternate strategies. First, we added the 14 (as well as FGFR2 counted in the top 10 SNPs above) confirmed breast cancer SNPs from prior independent GWAS of sporadic breast cancer. Second, we also selected the 15 top independent regions (pairwise $R^2<0.50$) based on the ranking of the p-values from a logistic regression analysis of 1.5 million imputed SNPs. In total for stage 2 replication phase, we selected 113 SNPs and 1,524 breast cancer carriers and 1,508 control carriers (Table 2) for genotyping using the Sequenom iPLEX platform.

Samples were excluded for call rates $\leq 95\%$ ($n = 476$), duplication in stage 2 (identity by state (IBS) $\leq 1.0$; $n = 43$), duplication in stage 1 and 2 (IBS; $n = 25$), lack of complete phenotype data ($n = 1$), and insufficient country-specific numbers ($n = 1$; Figure S6). A total of 100 SNPs were successfully multiplexed into three pools; the remaining 13 SNPs were not genotyped. Genotyping QC filters excluded 15 SNPs due to call rates $\leq 90\%$ ($n = 14$) and MAF $<1\%$ ($n = 1$). In summary, the final association analyses in stage 2 were based on 2,486 carriers (1,264 affected and 1,222 unaffected carriers) and 85 SNPs.

**Statistical Methods**

**Defining time at risk.** Carriers were censored at the first breast or ovarian cancer or bilateral prophylactic mastectomy, whichever occurred first. Carriers who developed any cancer were censored at time of bilateral prophylactic mastectomy if it occurred more than a year prior to the cancer diagnosis (to avoid censoring at bilateral mastectomies related to diagnosis in which rounded ages were used). The remaining carriers were censored at the age of last observation. This was defined either by the age/date at interview or age at follow-up depending on the information provided by the participating center. Carriers censored at diagnosis of breast cancer were considered cases in the analysis. Mutation carriers censored at ovarian cancer diagnosis were considered unaffected. Carriers with a censoring/last follow-up age older than age 80 were censored at age 80 because there are no reliable cancer incidence rates for **BRCA1/2** carriers beyond age 80.

**Genotype–phenotype associations.** Analyses, based on 1,703 **BRCA2** mutation carriers and 592,163 SNPs, were performed within a survival analysis framework. Since the mutation carriers were not selected at random with respect to their disease status, standard methods of survival (e.g., Cox regression) may lead to biased estimates of relative risk [30]. Therefore, analyses were conducted by modeling the retrospective likelihood of the observed genotypes conditional on the disease phenotypes. The associations between genotype and breast cancer risk at both stages were assessed using the 1-degree of freedom score test statistic based on this retrospective likelihood, as previously described [9,18]. All models were stratified by country of study and 6174delIT (c.594delIT) mutation status, the most common **BRCA2** mutation in this study and a marker of the Ashkenazi Jewish population among Ashkenazi Jewish women [31–33]. Since the linkage disequilibrium structure among Ashkenazi Jewish people may differ from other mutation carriers [34], stratifying by the 6174delIT provides additional control for population stratification. To allow for the non-independence among related individuals, an adjusted version of the score test was used in which the variance of the score was derived by taking into account the correlation between the genotypes [35,36]. Analyses were performed in R using the GenABEL libraries [37] and custom written software.

To estimate the magnitude of the associations, the effect of each SNP was modelled either as a per allele hazard ratio (HR) (i.e., multiplicative model) or as separate HRs for heterozygotes and homozygotes, and these were estimated on the log scale. The HRs were assumed to be independent of age (i.e. we used a Cox proportional-hazards model). For the most significant novel associations this assumption was verified by adding a genotype-by-age interaction term to the model to fit models in which the HR changed with age. The retrospective likelihood was implemented in the pedigree-analysis software MENDEL [38] as previously described [9]. All analyses were stratified by country of residence and 6174delIT (c.594delIT) mutation status, and used calendar-year- and cohort-specific breast cancer incidence rates for **BRCA2** [25]. The combined stage 1 and stage 2 analyses were also stratified by stage. Parameter estimates were obtained by maximising the retrospective likelihood. To allow for the non-independence among related mutation carriers, we used a robust variance estimation approach in order to obtain standard errors for the parameters [39,40]. Related individuals were identified through a unique family identifier.

**Copy number variant analysis.** We also examined the association of both high-frequency and low-frequency copy number variants (CNV) to the age of diagnosis of breast cancer as a dichotomous trait using the stage 1 data [29]. We called known, common variants (copy number polymorphisms, CNPs) with Canary [29]. CNP alleles lower than 1% in frequency were removed, to maximize the number of the CNPs that were bi-allelic instead of multi-allelic. CNPs were removed that had for call rate $<95\%$, differential missingness by genotype ($p<10^{-5}$), or departure from Hardy-Weinberg proportions ($p<10^{-5}$). Post-QC, we had 191 high-quality genotype polymorphisms. We used PLINK to assess association using logistic regression and the same ancestry covariates of no interest as with SNPs. We similarly assessed less common CNVs discovered by Birdseye [29] for association with age at diagnosis using PLINK [41]. Finally, we also looked specifically at CNVs overlapping the **BRCA2** gene itself using LOD scores and Birdseye.

**Haplotype sharing analysis.** We looked for evidence of excess sharing across the genome and the **BRCA2** region. Using GERMILINE [23], shared segments of greater than 5 cM were computed based on the imputed genotype dataset among both Ashkenazi ($n = 304$) and non-Jewish ($n = 1,331$) samples compared to samples from an autism study ($n = 808$) (Figure S3). Examining sites across the genome every 2.5 cM (excluding telomere and centromere regions), we computed the mean of the proportion, standard deviation, and the maximum values for non-Jewish and Ashkenazi women, respectively.

**Gene Set Enrichment Analysis.** We tested whether 59 genes known to regulate or interact with **BRCA2** [16] (Table S1) were enriched for associations with age of onset of breast cancer in **BRCA2** mutation carriers, using a new implementation of Gene Set Enrichment Analysis (GSEA) called Meta-Analysis Gene-Set Enrichment of variaNT Associations (MAGENTA) [14]. The 59 genes were compiled using a Pubmed abstract mining software, Chilibot [42], and were selected if they were related to the Fanconi anemia pathway [15] as well as others reported from literature to regulate or interact with **BRCA1/2** [43]. An association p-value was calculated for each gene in the genome, defined as the most-significant association p-value of all genotyped SNPs that lie within 110 kb upstream and 40 kb downstream to the gene’s most extreme transcript boundaries, followed by correction for gene score confounders (gene size, number of SNPs per gene and linkage disequilibrium related properties). SNP association p-values were taken from the stage 1 GWAS. To compute a GSEA p-value for the **BRCA1** gene set, the fraction of genes with an association p-value more significant than the 95 percentile of all
gene p-values in the genome was compared to a null distribution, generated by randomly sampling gene-sets of identical size from the genome 10,000 times. Of the 59 BRC4 interactors, two genes were assigned the same most significant SNP due to physical proximity in the genome. To prevent potential over-estimation of gene set enrichment due to physical clustering of genes in a gene set, we retained only one gene of each subset of genes assigned the same best SNP (the gene with the most significant gene p-value) for the analysis of both the real and permuted gene sets.

Supporting Information

Figure S1 Data filtering of stage 1 BRC4 GWAS. Found at: doi:10.1371/journal.pgen.1001183.s001 (0.57 MB TIF)

Figure S2 Quantile-quantile plot comparing expected distribution of chi-square values and observed chi-square values from a genome-wide scan of breast cancer cases and unaffected BRC4 carriers. Found at: doi:10.1371/journal.pgen.1001183.s002 (0.31 MB TIF)

Figure S3 Manhattan plot of p-values by chromosomal position from a genome-wide scan of breast cancer cases and unaffected BRC4 carriers [Visualized using SVS7 (Goldenhelix)]. Found at: doi:10.1371/journal.pgen.1001183.s003 (0.51 MB TIF)

Figure S4 Quantile-quantile plot comparing expected distribution of p-values and observed p-values of association of common copy number polymorphisms (CNPs) from a genome-wide scan of breast cancer cases and unaffected BRC4 carriers. Found at: doi:10.1371/journal.pgen.1001183.s004 (0.32 MB TIF)

Figure S5 Principal components analysis, including all eligible (after filtering) BRC4 stage 1 samples and HapMap samples. Found at: doi:10.1371/journal.pgen.1001183.s005 (2.39 MB TIF)

Figure S6 Data filtering of stage 2 BRC4 GWAS. Found at: doi:10.1371/journal.pgen.1001183.s006 (0.35 MB TIF)

Table S1 List of 59 BRC4 interactors or regulators and their gene association p-values to breast cancer age of onset in BRC4 mutation carriers. Found at: doi:10.1371/journal.pgen.1001183.s007 (0.13 MB DOC)

Table S2 Ranked results for the 85 SNPs successfully genotyped in stage 2, BRC4 GWAS. Found at: doi:10.1371/journal.pgen.1001183.s008 (0.13 MB DOC)

Acknowledgments

HEBON, kConFab, and GEMO were listed authors. The GEMO study is represented by the named authors: Olga M. Sinilnikova, Dominique Stoppa-Lyonnet, Sylvie Mazoyer, Marion Gauthier-Villars, Hagay Sobol, Michel Longy, and Marc Frenay. We wish to thank all the GEMO collaborating groups for their contribution to this study. GEMO Collaborating Centers are: Coordinating Centers, Unite´ Mixte de Ge´ne´tique Constitutionnelle des Cancers Fre´quents, Centre Hospitalier Universitaire de Lyon/Centre Leon Bérard, & UMR5291 CNRS, Universite´ de Lyon, Lyon: Olga Sinilnikova, Laure Barjhoux, Sophie Giraud, Melanie Le`one, Sylvie Mazoyer; and INSERM U509, Service de Ge´ne´tique Oncologique, Institut Curie, Paris: Dominique Stoppa-Lyonnet, Marion Gauthier-Villars, Claude Houdayer, Virginie Moncoutier, Muriel Beloti, Antoine de Pauw. Institut Gustave Roussey, Villejuif: Brigitte Bressac-de-Pailleres, Audrey Remenieras, Veronique Byrde, Olivier Caron, Gilbert Lenoir. Centre Jean Perrin, Clermont-Ferrand: Yves-Jean Bigon, Nancy Uhrhammer. Centre Leon Bérard, Lyon: Christine Lasset, Valerie Bonadona. Centre Francois Baclesse, Caen: Aignes Hardouin, Pascaline Berthet. Institut Paoli Calmettes, Marseille: Hagay Sobol, Violaine Bourdon, Tetsuo Noguchi, Francois Eisinger. Groupe Hospitalier Pitié-Salpêtrière, Paris: Florence Coulet, Chrysatelle Colas, Florent Soulouire. CHU de Arnaud-de-Villeneuve, Montpellier: Isabelle Coupiere. CRBEP, Oscar Lambert, Poitiers: Philippe Pouillot, Fran¸coise Revillon, Philippe Vennin, Claude Adenis. Centre René Huguenin, St Cloud: Etienne Rouleau, Rosette Liderace, Liliane Demange, Catherine Nougues. Centre Paul Strauss, strasbourg: Danielle Muller, Jean-Pierre Fricker. Institut Bergonié, Bordeaux: Michel Longy, Nicolas Sevenet. Institut Claudius Regaud, toulouse: Christine Toulas, Rosine Guimbaud, Laurence Gladieff, Viviane Felliel. CHU de Grenoble: Dominique Leroux, Heline Dreyvus, Christine Bischung. CHU de Dijon: Cecile Cassini, Laurence Faivre. CHU de St-Etienne: Fabienne Prieur. H´otel Dieu Centre Hospitalier, Chambéry: Sandra Ferr Ferrer. Centre Antoine Lacassagne, Nice: Marc Frenay. CHU de L´imoges: Laurence Vénat-Bouvet. Creighton University, Omaha, USA: Henry T. Lynch. The HEBON study is represented by the named authors: Frans B. L. Hogervorst, Matti A. Rookus, J. Margriet Collée, Nicoole Hoogerbrugge; Van Rosendothal. The kConFab study is represented by the named authors: Coordinating center: Netherlands Cancer Institute, Amsterdam: Frans B. L. Hogervorst, Senno Verhoef, Martijn Verheus, Laura J. van ’t Veer, Flora E. Van Leewen, Matti A. Rookus; Erasmus Medical Center, Rotterdam; Margriet Collée, Ans M.W. van den Ouweland, Agnes Jager, Maartje J. Hooning, Madeleine M.A. Tilanus-Linthorst, Caroline Seynaeve; Leiden University Medical Center, Leiden: Christi J. van Asperen, Jaul T. Wijnen, Maaike P. Vreeswijk, Rob A. Tollenar, Peter Van den Ouweland, Radboud Universiteit Nijmegen Medical Center, Nijmegen; Marjolijn J. Ligtelenberg, Nicoole Hoogerbrugge; University Medical Center Utrecht, Utrecht: Margreet G. Ausems, Rob B. van der Luitj; Amsterdam Medical Center: Cora M. Aalfs, Theo A. van Osi; VU University Medical Center, Amsterdam: Johan J.P. Gilles, Quinter Waifsfas, Hanne Meijers-Heijboer; University Hospital Maastricht, Maastricht: Escarina B. Gomez-Garcia, Gees E. van Rosendothal, Marinus J. Blok; University Medical Center Gresty’s University; Jan G. Oosterwijk, Annemarie H van der Hout, Marian J. Mourits; The Netherlands Foundation for the detection of hereditary tumours, Leiden, the Netherlands: Hans F. Vanes. For the kConFab study, Amanda B. Spurdle and Georgia Chenevis-Trench take responsibility for the manuscript on behalf of the kConFab members (listed at http://www. kconfab.org/Organisation/Members.aspx). The authors also wish to thank Heather Thor, Eveline Niedernary, all the kConFab study staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study for their contributions to this resource, and the many families who contribute to kConFab. ABS and GCT are Senior and Senior Principal Research Fellows of the NHMRC respectively. From the OG CN study, we wish to thank Teresa Selander, Navaya Weerasooriya and members of the Ontario Cancer Genetics Network for their contributions to the study. The LUMC study would like to thank hans Vazen, Inge van Leeuwen, and Hanne Meijers for patient accrual. For the CNIO study, the authors would like to thank R.M. Alonso, G. Pita and R.M. Måne for their assistance. We thank IBGM, Universidad de Valladolid and Consejeria de Sanidad Junta de Castilla y Le´on. For the SWE-BRCA, the authors would like to also acknowledge the contribution of Per Karlsson, Margareta Nordling, Aninka Bergman and Zakari Einbeig, Gothenburg, Sahlgrenska University Hospital; Marie Stemmer-Hansen, University of Linköping University Hospital; Åke Borg, Niklas Loman, Håkan Olsson, Ulf Kristoffersson, Helena Jernström, Katja Harbst and Karin Henriksson, Land University Hospital; Annika Lindblom, Brita Arver, Anna von Wachenfeldt, Annele Liijegren, Gisela Barbany-Bustina and Johanna Rantalak, Stockholm, Karolinska University Hospital; Beatrice Melin, Henrik Grönborg, Eva-Lena Stattin and Monica Emmelshausen, Umeå University Hospital; Hans Ehrenrother, Richard Rosenquist Brandelli and Niklas Dahl, Uppsala University Hospital; Douglas F. Easton is the PI of the EMBRACE study. EMBRACE Collaborating Centers are: Coordinating Center, Cambridge: Susan Pook, Margaret Cook, Clare Oliver, Debra Gregor; North of Scotland Regional Genetics Service, Aberdeen: Helen Gregory, Zosia Medzhybodzka. Northern Ireland Regional Genetics Service, Belfast: Patrick J. Morrison, LisaJeffers. West

References


