PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/137733

Please be advised that this information was generated on 2017-11-27 and may be subject to change.
DNA Glycosylases Involved in Base Excision Repair May Be Associated with Cancer Risk in BRCA1 and BRCA2 Mutation Carriers


PLOS Genetics | www.plosgenetics.org 1 April 2014 | Volume 10 | Issue 4 | e1004256
Base Excision Repair Genes Are BRCA1/2 Modifiers

Abstract

Single Nucleotide Polymorphisms (SNPs) in genes involved in the DNA Base Excision Repair (BER) pathway could be associated with cancer risk in carriers of mutations in the high-penetration susceptibility genes BRCA1 and BRCA2, given the relation of synthetic lethality that exists between one of the components of the BER pathway, PARP1 (poly ADP ribose polymerase), and both BRCA1 and BRCA2. In the present study, we have performed a comprehensive analysis of 18 genes involved in BER using a tagging SNP approach in a large series of BRCA1 and BRCA2 mutation carriers. 144 SNPs were analyzed in a two stage study involving 23,463 carriers from the CIMBA consortium (the Consortium of Investigators of Mutation Carriers). 144 SNPs were analyzed in a two stage study involving 23,463 carriers from the CIMBA consortium (the Consortium of Investigators of Mutation Carriers). Eleven SNPs showed evidence of association with breast and/or ovarian cancer at p<0.05 in the combined analysis. Four of the five genes for which strongest evidence of association was observed were DNA glycosylases. The strongest evidence was for rs1466785 in the NEIL2 (endonuclease VIII-like 2) gene (HR: 1.09, 95% CI (1.03–1.16), p = 2.7×10^{-5}) for association with breast cancer risk in BRCA2 mutation carriers, and rs2304277 in the OGG1 (8-guanine DNA glycosylase) gene, with ovarian cancer risk in BRCA1 mutation carriers (HR: 1.12 95%CI: 1.03–1.21, p = 4.8×10^{-5}). DNA glycosylases involved in the first steps of the BER pathway may be associated with cancer risk in BRCA1/2 mutation carriers and should be more comprehensively studied.
Introduction

Carrying an inherited mutation in the BRCA1 or BRCA2 gene increases a woman’s lifetime risk of developing breast, ovarian, and other cancers. The estimated cumulative risk of developing breast cancer by the age of 70 in BRCA1 and BRCA2 mutation carriers varies between 43% to 86%; similarly, between 11% to 59% of mutation carriers will develop ovarian cancer by the age of 70 [1-3]. These considerable differences in disease manifestation suggest the existence of other genetic or environmental factors that modify the risk of cancer development. The Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA), was established in 2006 [4] and with more than 40,000 mutation carriers currently provides the largest sample size for reliable evaluation of even modest associations between single-nucleotide polymorphisms (SNPs) and cancer risk. CIMBA studies have so far demonstrated that more than 25 SNPs are associated with the risk of developing breast or ovarian cancer for BRCA1 or BRCA2 carriers. These were identified through genome-wide association studies (GWAS) of breast or ovarian cancer in the general population or through BRCA1- and BRCA2-specific GWAS [5-8]. Cells harboring mutations in BRCA1 or BRCA2 show impaired homologous recombination (HR) [9-11] and are thus critically dependent on other processes of the DNA repair machinery such as poly ADP ribose polymerase (PARP1) involved in the Base Excision Repair (BER) pathway. The BER pathway is crucial for the replacement of aberrant bases generated by different causes [12]. A deficiency in BER can give rise to a further accumulation of double-strand DNA breaks which, in the presence of a defective BRCA1 or BRCA2 background, could persist and lead to cell cycle arrest or cell death; this makes BRCA-deficient cells extremely sensitive to PARP inhibitors, as previously demonstrated [13]. We hypothesize that SNPs in PARP1 and other members of BER may be associated with cancer risk in BRCA1 and BRCA2 mutation carriers. SNPs in XRCC1, one of the main components of BER, have been recently evaluated within the CIMBA consortium [14], however a comprehensive study has not yet been performed of either XRCC1 or the other genes participating in BER.

In the present study, we used a tagging SNP approach to evaluate whether the common genetic variation in the genes involved in the BER pathway could be associated with cancer risk in a large series of BRCA1/2 mutation carriers using a two-stage approach. The first stage involved an analysis of 144 tag SNPs in 1,787 Spanish and Italian BRCA1/2 mutation carriers. In stage II, the 36 SNPs showing the strongest evidence of association in stage I, were evaluated in a further 23,463 CIMBA mutation carriers included in the Collaborative Oncological Gene-environment Study (COGS) and genotyped using the iCOGS custom genotyping array.

Results

Breast cancer association

In stage I, 144 selected Tag SNPs covering the 18 selected BER genes were genotyped in 968 BRCA1 and 819 BRCA2 mutation carriers from five CIMBA centres (Spanish National Cancer ResearchCentre (CNIO), Hospital Clinico San Carlos (HCSG), Catalan Institute of Oncology (ICO), Demokritos and Milan Breast Cancer Study Group (MBCSG)). Of those, 30 were excluded because of low call-rates, minor allele frequency (MAF)<0.05, evidence of deviation from Hardy Weinberg Equilibrium (p-value<10-3) or monomorphism. Associations with...
breast cancer risk were assessed for 94 SNPs, as summarized in Table S1. The 36 SNPs that showed evidence of association at p<0.05 were selected for analysis in stage II. Of the 36 SNPs successfully genotyped in the whole CIMBA series comprising 15,252 BRCA1 and 8211 BRCA2 mutation carriers, consistent evidence of association with breast cancer risk (p-trend<0.05) was observed for six SNPs (Table 1). The strongest evidence of association was observed for rs1466785 in the NEIL2 gene (HR: 1.09, 95% CI: 1.03–1.16, p = 6×10^{-5}). The SNP was primarily associated with ER-negative breast cancer (HR: 1.20, 95% CI: 1.06–1.37, p = 4×10^{-5}), although the difference in HRs for ER-positive and ER-negative disease was not statistically significant. The evidence of association in Stage II was somewhat stronger when considering the genotype-specific models, with the dominant being the best fitting (HR: 1.20, 95% CI: 1.09–1.37, p = 1.0×10^{-5}). The associations remained significant and the estimated effect sizes remained consistent with the overall analysis when the data were reanalyzed excluding samples used in stage I of the study (data not shown). Imputation using the 1000 genomes data showed that there were several SNPs in strong linkage disequilibrium (LD) with rs1466785 showing more significant associations (p<10^{-5}) (Figure 1).

Ovarian cancer association
Due to lack of power we did not perform analysis of associations with ovarian cancer in stage I. However, we performed this analysis for the 36 SNPs tested in stage II. Although they had been selected based on their evidence of association with breast cancer risk, under the initial hypothesis they are also plausible modifiers of ovarian cancer risk for BRCA1 and BRCA2 mutations carriers. We found four SNPs associated with ovarian cancer risk with a p-trend<0.01 in BRCA1 or BRCA2 mutation carriers (Table 1). The strongest association was found for rs2304277 in OGG1 in BRCA1 mutation carriers (HR: 1.12, 95% CI: 1.03–1.21, p = 4.8×10^{-5}). The association was somewhat stronger under the dominant model (HR: 1.19, 95% CI: 1.08–1.3, p = 6×10^{-5}). Although three other SNPs were found to be associated with ovarian cancer risk in BRCA2 mutation carriers (p-trend<10^{-5}), these results were based on a relatively small number of ovarian cancer cases. Imputed data did not show any SNPs with substantially more significant associations with ovarian cancer risk except for rs3093926 in PARP2, associated with ovarian cancer risk in BRCA2 mutation carriers for which there was a SNP, rs61993542, with a stronger association (HR: 0.67, p = 4.6×10^{-5}) (Figure S1).

Discussion
Based on the interaction of synthetic lethality that has been described between PARP1 and both BRCA1 and BRCA2, we hypothesize that this and other genes involved in the BER pathway could potentially be associated with cancer risk in BRCA1/2 mutation carriers. Several studies have recently investigated the association of some of the BER genes with breast cancer, however, no definitive conclusions can be drawn, given that some publications suggest that SNPs in these genes can be associated with breast cancer risk with marginal p-values while others rule out a major role of these genes in the disease [15–21]. There is only one study from the CIMBA consortium which has evaluated the role of three of the most studied SNPs in the XRCC1 gene, c.-77C>T (rs3213245) p.Arg280His (rs25489) and p.Gln399Arg (rs25487), ruling out associations of these variants with cancer risk in BRCA1 and BRCA2 mutation carriers [14]. However, a comprehensive analysis of neither XRCC1 nor the other genes involved in the pathway in the context of BRCA mutation carriers has been performed. In the present study we have assessed the common genetic variation of 18 genes participating in BER by using a two stage strategy.

Eleven SNPs showed evidence of association with breast and/or ovarian cancer at p<0.05 in stage II of the experiment (Table 1). Of those, six showed a p-trend value<0.01 and were therefore considered the best candidates for further evaluation. Only one of those six, rs1466785 in the NEIL2 gene (endonuclease VIII-like gene) showed an association with breast cancer risk while the other five, rs2304277 in OGG1 (beta-guanine DNA glycosylase), rs167715 and rs4135087 in Tdg (thymine-DNA glycosylase), rs3093926 in PARP2 (Poly(ADP-ribose) polymerase 2) and rs34259 in Ung (uracil-DNA glycosylase) were associated with ovarian cancer risk. The minor allele of NEIL2-rs1466785 was associated with increased breast cancer risk in BRCA2 mutation carriers; moreover, when considering the genotype-specific risks observed that the best fitting model was the dominant one. NEIL2 is one of the oxidized base-specific DNA glycosylases that participate in the initial steps of BER and specifically removes oxidized bases from transcribing genes [22]. By imputing using the 1000 genome data we found six correlated SNPs in strong LD with rs1466785 (r2>0.8), located closer or inside the gene and showing slightly stronger and more significant associations with the disease and therefore being better candidate causal variants. From those, we considered rs804276 and rs804271 as the best candidates given that they showed the most significant associations (p = 6×10^{-5} and p = 8×10^{-5} respectively) and there were available epidemiological or functional data supporting their putative role in cancer. SNP rs804276 has been associated with disease recurrence in patients with bladder cancer treated with Bacillus Calmette-Guérin (BCG) (HR: 2.71, 95% CI: 1.75–4.20, p = 9.9×10^{-5}) [23]. SNP rs804271 is located in a positive regulatory region in the promoter of the gene, between two potential cis-binding sites for reactive oxygen species responsive transcription factors in which sequence variation has
### Table 1. Associations with breast and ovarian cancer risk for SNPs observed at p-trend<0.05 in stage II of the experiment.

<table>
<thead>
<tr>
<th>BRCA1 carriers SNP name</th>
<th>Gene</th>
<th>Unaffected (Number)</th>
<th>Affected (Number)</th>
<th>HR per allele(^a)</th>
<th>HR heterozygote(^b)</th>
<th>HR homozygote(^b)</th>
<th>p-trend(^c)</th>
<th>p-het(^c)</th>
<th>p-hom(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs3847954(^d)</td>
<td>UNG</td>
<td>7455</td>
<td>7797</td>
<td>0.18</td>
<td>0.19</td>
<td>1.05 (1.00–1.11)</td>
<td>1.09 (1.02–1.16)</td>
<td>0.99 (0.84–1.16)</td>
<td>0.04</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2073668</td>
<td>OGG1</td>
<td>12786</td>
<td>2461</td>
<td>0.22</td>
<td>0.23</td>
<td>1.09 (1.01–1.18)</td>
<td>1.16 (1.05–1.27)</td>
<td>1.03 (0.82–1.28)</td>
<td>0.016</td>
</tr>
<tr>
<td>rs2269112</td>
<td>OGG1</td>
<td>12789</td>
<td>2461</td>
<td>0.17</td>
<td>0.18</td>
<td>1.11 (1.02–1.21)</td>
<td>1.11 (1.01–1.23)</td>
<td>1.21 (0.92–1.58)</td>
<td>0.013</td>
</tr>
<tr>
<td>rs2304277</td>
<td>OGG1</td>
<td>12783</td>
<td>2462</td>
<td>0.21</td>
<td></td>
<td>1.12 (1.03–1.21)</td>
<td>1.19 (1.08–1.3)</td>
<td>1.01 (0.79–1.30)</td>
<td>4.8 × 10(^{-3})</td>
</tr>
<tr>
<td>rs10161263</td>
<td>SMUG1</td>
<td>12790</td>
<td>2462</td>
<td>0.34</td>
<td>0.32</td>
<td>0.92 (0.86–0.99)</td>
<td>0.88 (0.80–0.97)</td>
<td>0.90 (0.78–1.04)</td>
<td>0.024</td>
</tr>
<tr>
<td>BRCA2 carriers SNP name</td>
<td>Gene</td>
<td>Unaffected (Number)</td>
<td>Affected (Number)</td>
<td>HR per allele(^a)</td>
<td>HR heterozygote(^b)</td>
<td>HR homozygote(^b)</td>
<td>p-trend(^c)</td>
<td>p-het(^c)</td>
<td>p-hom(^c)</td>
</tr>
<tr>
<td>Breast cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2073668(^e)</td>
<td>OGG1</td>
<td>3879</td>
<td>4328</td>
<td>0.23</td>
<td>0.21</td>
<td>0.91 (0.85–0.98)</td>
<td>0.95 (0.87–1.04)</td>
<td>0.75 (0.62–0.91)</td>
<td>0.018</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2269112(^f)</td>
<td>OGG1</td>
<td>3880</td>
<td>4329</td>
<td>0.17</td>
<td>0.16</td>
<td>0.91 (0.84–0.99)</td>
<td>0.93 (0.85–1.03)</td>
<td>0.76 (0.58–0.99)</td>
<td>0.035</td>
</tr>
<tr>
<td>rs3136811(^g)</td>
<td>POLB</td>
<td>3873</td>
<td>4321</td>
<td>0.06</td>
<td>0.07</td>
<td>1.12 (1.005–1.25)</td>
<td>1.17 (1.03–1.32)</td>
<td>0.86 (0.49–1.48)</td>
<td>0.032</td>
</tr>
<tr>
<td>rs2304277(^h)</td>
<td>OGG1</td>
<td>3880</td>
<td>4330</td>
<td>0.21</td>
<td>0.19</td>
<td>0.91 (0.84–0.97)</td>
<td>0.94 (0.85–1.03)</td>
<td>0.74 (0.60–0.91)</td>
<td>0.013</td>
</tr>
<tr>
<td>rs1466785(^i)</td>
<td>NEIL2</td>
<td>3879</td>
<td>4330</td>
<td>0.43</td>
<td>0.43</td>
<td>1.09 (1.03–1.16)</td>
<td>1.20 (1.09–1.37)</td>
<td>1.16 (1.03–1.31)</td>
<td>2.7 × 10(^{-3})</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs167715</td>
<td>TDG</td>
<td>7577</td>
<td>631</td>
<td>0.12</td>
<td>0.09</td>
<td>0.76 (0.62–0.94)</td>
<td>0.72 (0.58–0.90)</td>
<td>0.89 (0.41–1.89)</td>
<td>7.4 × 10(^{-3})</td>
</tr>
<tr>
<td>rs3093926(^j)</td>
<td>PARP2</td>
<td>7580</td>
<td>631</td>
<td>0.07</td>
<td>0.05</td>
<td>0.64 (0.49–0.84)</td>
<td>_</td>
<td>_</td>
<td>1.5 × 10(^{-2})</td>
</tr>
<tr>
<td>rs4135087</td>
<td>TDG</td>
<td>7580</td>
<td>631</td>
<td>0.09</td>
<td>0.11</td>
<td>1.32 (1.09–1.59)</td>
<td>1.33 (1.07–1.65)</td>
<td>1.67 (0.84–3.28)</td>
<td>2.8 × 10(^{-3})</td>
</tr>
<tr>
<td>rs34259</td>
<td>UNG</td>
<td>7580</td>
<td>631</td>
<td>0.2</td>
<td>0.17</td>
<td>0.80 (0.69–0.94)</td>
<td>0.84 (0.70–1.01)</td>
<td>0.51 (0.29–0.90)</td>
<td>7.6 × 10(^{-3})</td>
</tr>
</tbody>
</table>

\(^a\)Hazard Ratio per allele (1 df) estimated from the retrospective likelihood analysis.  
\(^b\)Hazard Ratio under the genotype specific models (2df) estimated from the retrospective likelihood analysis.  
\(^c\)p-values were based on the score test.  
\(^d\)HR per allele of 1.69 and p-trend of 1 × 10\(^{-4}\) for BRCA2 mutation carriers in stage I of the study.  
\(^e\)HR per allele of 1.43 and p-trend of 0.01 for BRCA1 mutation carriers in stage I of the study.  
\(^f\)HR per allele of 1.69 and p-trend of 0.03 for BRCA2 mutation carriers in stage I of the study.  
\(^g\)HR per allele of 0.64 and p-trend of 0.057 for BRCA2 mutation carriers in stage I of the study.  
\(^h\)HR per allele of 1.25 and p-trend of 0.04 for BRCA1 mutation carriers in stage I of the study.  
\(^i\)HR per allele of 1.25 and p-trend of 0.058 for BRCA2 mutation carriers in stage I of the study.  
\(^j\)rs3093926 did not yield results under the genotype specific model due to the low minor allele frequency.  

Highlighted in bold are those SNPs showing strongest associations with breast or ovarian cancer risk (p<0.01).  
Complete description of results from stage I are included in Supplementary Table S1.
been proven to alter the transcriptional response to oxidative stress [24]. Moreover, this SNP has been proposed to partly explain the inter-individual variability observed in NEIL2 expression levels in the general population and has been proposed as a potential risk modifier of disease susceptibility [25].

Several studies have been published showing associations between SNPs in NEIL2 and lung or oropharyngeal cancer risk [26,27] but to our knowledge, no association with breast cancer risk has been reported. We hypothesize that the potential association observed in the present study could be explained by the interaction between NEIL2 and BRCA2, each of them causing a deficiency in the BER and HR DNA repair pathways, respectively. This would explain why the breast cancer risk modification due to rs1466785 would only be detected in the context of BRCA2 mutation carriers and not in the general population.

The strongest evidence of association found in BRCA1 carriers was between rs2304277 in the OGG1 gene and ovarian cancer risk. The association was more significant when considering the dominant model. OGG1 removes 8-oxodeoxyguanosine which is generated by oxidative stress and is highly mutagenic, and it has been suggested that SNPs in the gene could be associated with cancer risk [28–31]. This is an interesting result, given that to date only one SNP, rs4691139 in the 4q35.3 region, also identified through the iCOGS effort, did not show better results for a more plausible causal SNP.

We have identified four SNPs associated with ovarian cancer risk in BRCA2 mutation carriers, rs167715 and rs4135087 in the TDG gene, rs34259 in the UNG gene and rs3093926 in PARP2. However, these last results should be interpreted with caution given that the number of BRCA2 carriers affected with ovarian cancer is four-fold lower than for BRCA1 carriers and the statistical power was therefore more limited, increasing the possibility of false-positives. In the case of PARP2, imputed data showed a lower p-value of association (4×10^{-4}) for another SNP, rs61995542, that had a slightly higher MAF than rs3093926 (0.074 vs. 0.067) (Figure S1). However, it must still be interpreted with caution due to small number of ovarian cancer cases in the BRCA2 group.

It is worth noting that, four of the five genes for which strongest evidence of association was observed, are all DNA glycosylases participating in the initiation of BER by removing damaged or mismatched bases. Apart from the already mentioned NEIL2 and OGG1, TDG initiates repair of G/T and G/U mismatches commonly associated with CpG islands, while UNG removes uracil in DNA resulting from deamination of cytosine or replicative incorporation of dUMP. We have not found strong associations with SNPs in genes involved in any other parts of the pathway, such as strand incision, trimming of ends, gap filling or ligation. It has been suggested that at least in the case of uracil repair, base removal is the major rate-limiting step of BER [33]. This is consistent with our findings, suggesting that SNPs causing impairment in the function of these specific DNA glycosylases through the 1000 Genome did not show better results for a more plausible causal SNP.

Figure 1. p-values of association (−log10 scale) with breast cancer risk in BRCA2 carriers for genotyped and imputed SNPs in the NEIL2 gene. SNP rs1466785 is indicated with a purple arrow and the best causal imputed SNPs, rs804276 and rs804271 are indicated with a red arrow. Colors represent the pairwise r². Plot generated with LocusZoom [42] (http://csg.sph.umich.edu/locuszoom/).

doi:10.1371/journal.pgen.1004256.g001
could give rise to accumulation of single strand breaks and subsequently DNA double strand breaks that, in the HR defective context of \textit{BRCA1/2} mutation carriers would increase breast and ovarian cancer risk.

The fact that the SNPs tested are located in genes participating in the same DNA repair pathway as PARP1, make them especially interesting, not only as risk modifiers but also because they could have an impact on patients' response to treatment with PARP inhibitors. \textit{BRCA1/2} mutation carriers harboring a potential modifier SNP in DNA glycosylases could be even more sensitive to PARP1 due to a constitutional slight impairment of the BER activity. This is a hypothesis that should be confirmed in further studies. The design of this study in two stages, the hypothesis-based approach adopted to select genes, and that it is based on the largest possible series of \textit{BRCA1} and \textit{BRCA2} carriers available nowadays, mean that the results obtained are quite solid. However, the study still has some limitations such as the possible existence of residual confounding due to environmental risk factors for which we did not have information.

In summary, we have identified at least two SNPs, rs1466785 and rs2304277, in the DNA glycolylases \textit{NEIL2} and \textit{OGG1}, potentially associated with increased breast and ovarian cancer risks in \textit{BRCA2} and \textit{BRCA1} mutation carriers, respectively. Our results suggest that glycosylases involved in the first steps of the BER pathway may be cancer risk modifiers in \textit{BRCA1/2} mutation carriers and should be more comprehensively studied. If confirmed, these findings could have implications not only for risk assessment, but also for treatment of \textit{BRCA1/2} mutation carriers with PARP inhibitors.

\section*{Materials and Methods}

\subsection*{Subjects}

Eligible subjects were female carriers of deleterious mutations in \textit{BRCA1} or \textit{BRCA2} aged 18 years or older [6]. A total of 55 collaborating CIMBA studies contributed genotypes for the study. Numbers of samples included from each are provided in Table S2. A total of 1,787 mutation carriers (968 with mutations in \textit{BRCA1} and 819 with mutations in \textit{BRCA2}) from the CNIO, HCSC, ICO, Demokritos and MBCSG were genotyped in the first stage of the study. Stage II included 23,463 CIMBA samples (15,252 with mutations in \textit{BRCA1} and 8,211 with mutations in \textit{BRCA2}). All carriers participated in clinical and/or research studies at the host institution under IRB-approved protocols.

\subsection*{Methods stage I}

\textbf{Selection and genotyping of SNPs.} Eighteen genes (\textit{UNG, SMUG1, MBD4, TDG, OGG1, MUTYH, NTHL1, MGP, NEIL1, NEIL2, APEX1, APEX2, LIG3, XRCC1, PNKP, POLB, PARP1 and PARP2}) involved in the BER pathway were selected, based on the information available at http://www.cgd.icnet.uk/DNA_Repair_Genes.html as at the 31st December, 2009. Tag SNPs for the selected genes were defined using Haploview v.4.0 (http://www.broad.mit.edu/mpg/haploview) with an $r^2$ threshold of 0.85 and a minimum minor allele frequency of 0.05. In addition, SNPs with potentially functional effects already described in the literature were selected. A final number of 144 SNPs was included in an oligonucleotide pool assay for genotyping using the Illumina Veracode technology (Illumina Inc., San Diego, CA). Three hundred nanograms of DNA from each sample were genotyped using the GoldenGate Genotyping Assay with Veracode technology (Illumina Inc., San Diego, CA). Three hundred nanograms of DNA from each sample were genotyped

\textbf{Statistics analysis.} Duplicate samples and CEPH trios (Coriell Cell Repository, Camden, NJ) were genotyped across the plates. SNPs showing Mendelian allele-transmission errors or showing discordant genotypes across duplicates were excluded. In stage II, 23,463 CIMBA samples (15,252 with mutations in \textit{BRCA1} and 8,211 with mutations in \textit{BRCA2}) were nominated for inclusion on iCOGS based on statistical evidence of association in stage I of the present study (p < 0.05). Of these, 36 were successfully manufactured and genotyped in CIMBA mutation carriers.

\textbf{iCOGS genotyping and quality control.} Genotyping was performed at Mayo Clinic and the McGill University and Génomique Québec Innovation Centre (Montreal, Canada). Genotypes were called using Illumina's GenCall algorithm. Sample and quality control process have been described in detail elsewhere [32,35]. After the quality control process a total of 23,463 carriers were genotyped for the 36 selected SNPs.

\subsection*{Methods stage II}

\textbf{iCOGS SNP array.} Stage II of the experiment was performed as part of the iCOGS genotyping experiment. The iCOGS custom array was designed in collaboration between the Breast Cancer Association Consortium (BCAC), the Ovarian Cancer Association Consortium (OCAC), the Prostate Cancer Association Group to Investigate Cancer Associated in the Genome (PRACTICAL) and CIMBA. The final design comprised 211,155 successfully manufactured SNPs of which approximately 17.5% had been proposed by CIMBA. A total of 43 SNPs were nominated for inclusion on iCOGS based on statistical evidence of association in stage I of the present study (p<0.05). Of these, 36 were successfully manufactured and genotyped in CIMBA mutation carriers.

\textbf{iCOGS genotyping and quality control.} Genotyping was performed at Mayo Clinic and the McGill University and Génomique Québec Innovation Centre (Montreal, Canada). Genotypes were called using Illumina’s GenCall algorithm. Sample and quality control process have been described in detail elsewhere [32,35]. After the quality control process a total of 23,463 carriers were genotyped for the 36 selected SNPs.

\textbf{Statistical analysis.} Both breast and ovarian cancer associations were evaluated in stage II. Censoring for breast cancer followed the same approach as in stage I. Censoring for ovarian cancer risk occurred at risk-reducing salpingo-oophorectomy or last follow-up.

The genotype-disease associations were evaluated within a survival analysis framework, by modelling the retrospective likelihood of the observed genotypes conditional on the disease phenotypes [9,34,36,37]. The associations between genotype and breast or ovarian cancer risk were assessed using the 1 d.f. score test statistic based on this retrospective likelihood. To allow for the
non-independence among related individuals, we accounted for the correlation between the genotypes by estimating the kinship coefficient for each pair of individuals using the available genomic data [34,38,39]. These analyses were performed in R using the GenABELibraries and custom-written functions in FORTRAN and Python.

To estimate the magnitude of the associations (HRs), the effect of each SNP was modeled either as a per-allele HR (multiplicative model) or as genotype-specific HRs, and was estimated on the log-scale by maximizing the retrospective likelihood. The retrospective likelihood was fitted using the pedigree-analysis software MENDEL. The variances of the parameter estimates were obtained by robust variance estimation based on reported family membership. All analyses were stratified by country of residence and based on calendar-year and cohort-specific breast cancer incidence rates for mutation carriers. Countries with small number of mutation carriers were combined with neighbouring countries to ensure sufficiently large numbers within each stratum. USA and Canada were further stratified by reported Ashkenazi Jewish (AJ) ancestry.

**Imputation.** Genotypes were imputed separately for \(BRCA1\) and \(BRCA2\) mutation carriers using the v3 April 2012 release (Genomes Project et al., 2012) as reference panel. To improve computation efficiency we used a two-step procedure which involved pre-phasing in the first step and imputation of the phased data in the second. Pre-phasing was carried out using the SHAPEIT software [40]. The IMPUTE version 2 software was used for the subsequent imputation [41]. SNPs were excluded from the association analysis if their imputation accuracy was \(r^2<0.3\) or MAF <0.005 in any of the data sets. For the final analysis we only took in account those SNPs with an imputation accuracy \(r^2>0.7\), MAF >0.01 and being located in the region comprised within 15 kilo bases (kb) downstream and upstream the gene where the genotyped SNP showing an association was located (Table 1). Associations between imputed genotypes and breast cancer risk were evaluated using a version of the score test as described above but with the posterior genotype probabilities replacing the genotypes.

**Supporting Information**

Figure S1 p-values of association (~log10 scale) with breast and ovarian cancer risk in \(BRCA1\) and \(BRCA2\) carriers for genotyped and imputed SNPs considering 15 kb upstream and downstream the genes in which SNPs described in Table 1 were located. rs numbers of SNPs from Table 1 are indicated at the top of each panel and in the graph with a purple arrow. For \(PARP2\) gene, the imputed SNP with the strongest association, rs61995542 is indicated with a red arrow. Colors represent the pairwise \(r^2\).

Table S1 Association with breast cancer for the 94 SNPs selected for analysis in stage I.

Table S2 number of \(BRCA1\) and \(BRCA2\) carriers by study.

**Acknowledgments**

CNIO acknowledgments: Genotyping was provided by ‘Centro Nacional de Genotipado - Instituto de Salud Carlos III’ (CeGen-ISCIII; www.cegen.org). COGS acknowledgments: this study would not have been possible without the contributions of the following: Per Hall (COGS); Douglas F. Easton (BCAC), Andrew Berchuck (OCAC), Rosalind A. Eeles, Douglas F. Easton, Ali Amin Al Olama, Zsofia Kote-Jarai (PRACTICAL), Georgia Chenevix-Trench, Antonis Antoniou, Fergus Couch and Ken Oliff (CIMBA), Joe Dennis, Alison M. Dunning, Andrew Lee, and Ed Dicks (Cambridge), Javier Benitez, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit, Jacques Simard and Daniel C. Tessler, Francois Bacot, Daniel Vincent, Sylvie LaBoissiere and Frederic Robidoux and the staff of the McGill University and Genome Quebec Innovation Centre, Stig E. Bojesen, Sune F. Nielsen, Borge G. Nordestgaard, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hiller, Jefriet Meyer and the staff of Mayo Clinic Genotyping Core Facility.

Swedish Breast Cancer Study (SWE-BRCA): SWE-BRCA members are Ake Borg, Department of Oncology, Lund University, Lund. Anna Olverholm, Department of Clinical Genetics, Sahlgrenska University Hospital, Gothenburg, Sweden. Anna von Wachenfeldt, Department of Oncology, Karolinska University Hospital, Stockholm, Sweden. Annelie Liljegren, Department of Oncology, Karolinska University Hospital, Stockholm, Sweden. Annika Lindholm, Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. Beatrice Melin, Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden. Brita Arver, Department of Oncology, Karolinska University Hospital, Stockholm, Sweden. Christina Edwindsdotter Ardnor, Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden. Gisela Barbany Bustinza, Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. Håkan Olsson, Department of Oncology, Lund University Hospital, Lund, Sweden. Hans Ehrencreuna, Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala and Department of Clinical Genetics, Lund University Hospital, Lund, Sweden. Helena Jerström, Department of Oncology, Lund University, Lund, Sweden. Johanna Rantalaa, Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. Karin Henriksson, Oncologie Centre, Regional Tumour Registry, Lund University Hospital, Lund, Sweden. Katja Harbst, Department of Oncology, Lund University, Lund, Sweden. Margareta Nordling, Department of Clinical Genetics, Sahlgrenska University Hospital, Gothenburg, Sweden. Maria Soller, Department of Clinical Genetics, Lund University Hospital, Lund, Sweden. Marie Stenmark-Axkman, Division of Clinical Genetics, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden. Maritta Hellström Pigg, Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Uppsala. Monica Emanuelhs, Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden. Niklas Loman, Department of Oncology, Lund University Hospital, Lund, Sweden. Per Karlsson, Department of Oncology, Sahlgrenska University Hospital, Gothenburg, Sweden. Richard Rosenquist, Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden. Sigrun Lidgren, Division of Clinical Genetics, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden. SWE-BRCA, Ulf Kristoffersson, Department of Clinical Genetics, Lund University Hospital, Lund, Sweden. Zakaria Einhaeji, Department of Oncology, Sahlgrenska University Hospital, Gothenburg, Sweden.

The University of Chicago Center for Clinical Cancer Genetics and Global Health (UCHICAGO) wish to thank Cecilia Zvoce, Qun Niu, musicians, 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.

UCSF Cancer Risk Program and Helen Diller Family Comprehensive Cancer Center: We would like to thank Ms. Salina Chan for her data management and the following genetic counselors for participant recruitment: Beth Crawford, Nicola Stewart, Julie Mak, and Kate Lamvik.

Baltic Familial Breast Ovarian Cancer Consortium (BFBOCC) study: we acknowledge Vilis Rudaitis, Laimonas Grisikevičius, BFBOCC-LV acknowledge Drs. Janis Eglitis, Anna Krilova and Aivars Stengrevics. \(BRCA\)-gene mutations and breast cancer in South African women (BMBSA) wish to thank the families who contribute to the BMBSA study.

Beckman Research Institute of the City of Hope (BRICOH) study: we wish to thank Greg Wilhoite, Linda Steele, and Marie Plato for their work in participant enrollment and biospecimen and data management.

The City of Hope Clinical Genetics Community Research Network: we wish to thank the collaborating clinicians in the Clinical Cancer Genetics Community Research Network for patient recruitment and follow up; Kai Yang for masterful assistance with the database, and Hazel Mariveles for network coordination.
COINoorzio Studi Italiani su Tumori Ereditari Alla Mamella (CONSIT TEAM) acknowledged: Monica Barile and Irene Forcato of the Istituto Europeo di Oncologia, Milan, Italy; Antonella Savarese and Aline Martayan of the Istituto Nazionale Tumori Regina Elena, Rome, Italy; Stefania Tommasi, Brunella Pilato and Rossana Lambo of the Istituto Nazionale Tumori “Giovanni Paolo II” - Bari, Italy and the personnel of the Cogenetics Cancer Genetic Test Laboratory, Milan, Italy.


The German Consortium of Hereditary Breast and Ovarian Cancer (GC-HBOC): we are very thankful to all family members who participated in this study, Wolfram Heinritz, Center Leipzig, and Dieter Schäfer, Center Frankfurt, for providing DNA samples and Juliane Kohler for excellent technical assistance.

Author Contributions

Conceived and designed the experiments: AO RLM GJP. Performed the experiments: AO TV RA BH LTJ JS KO FJC. Analyzed the data: RLM KKu ACA. Contributed reagents/materials/analysis tools: AO RLM Kku GJP. Wrote the first draft of the article: AO RLM GJP. Contributed to the writing of the manuscript: AO RLM GJP.

References


12. RLM GJP. Wrote the first draft of the article: AO RLM GJP. Contributed to the writing of the manuscript: AO RLM GJP.


