An original phylogenetic approach identified mitochondrial haplogroup T1a1 as inversely associated with breast cancer risk in BRCA2 mutation carriers


* Correspondence: david.cox@lyon.unicancer.fr
Deceased
1INSERM U1052, CNRS UMR5286, Université Lyon 1, Centre de Recherche en Cancérologie de Lyon, Lyon, France
2Université de Lyon, 69000 Lyon, France

Full list of author information is available at the end of the article

© 2015 Blein et al; licensee BioMed Central. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Introduction

Breast cancer is a multifactorial disease with genetic, lifestyle and environmental susceptibility factors. Approximately 15% to 20% of the familial aggregation of breast cancer is accounted for by mutations in high-penetrance susceptibility genes [1-3], such as BRCA1 and BRCA2. Pathogenic mutations in BRCA1 and BRCA2 confer lifetime breast cancer risk of 60% to 85% [4,5] and 40% to 85% [4,5], respectively. Other genomic variations (for example, in genes encoding proteins interacting with BRCA1 and BRCA2) have been identified as modifiers of breast cancer risk and increase or decrease the risk initially conferred by BRCA1 or BRCA2 mutation [6].

BRCA1 and BRCA2 are involved in DNA repair mechanisms, including double-strand break (DSB) repair by homologous recombination [7,8]. DSBs are considered to be among the most deleterious forms of DNA damage because the integrity of both DNA strands is compromised simultaneously. These breaks can lead to genomic instability resulting in translocations, deletions, duplications or mutations when not correctly repaired [9]. Reactive oxygen species (ROS) are one of the main causes of DSBs, along with exposure to ionizing radiation, various chemical agents and ultraviolet light [10].

ROS are naturally occurring chemical derivatives of metabolism. Elevated levels of ROS and downregulation of ROS scavengers and/or antioxidant enzymes can lead to oxidative stress, which is associated with a number of human diseases, including various cancers [11]. The electron transport chain process, which takes place in the mitochondria, generates the majority of ROS in human cells. Variations in the mitochondrial genome have been shown to be associated with metabolic phenotypes and
oxidative stress markers [12]. Mitochondrial dysfunction recently was shown to promote breast cancer cell migration and invasion through the accumulation of a transcription factor, hypoxia-inducible factor 1α, via increased production of ROS [13].

Human mitochondrial DNA (mtDNA) has undergone a large number of mutations that have segregated during evolution. Those changes are now used to define mitochondrial haplogroups. Some of these changes slightly modify metabolic performance and energy production; thus, not all haplogroups have identical metabolic capacities [14]. It has been hypothesized that the geographic distribution of mitochondrial haplogroups results from selection of metabolic capacities driven mainly by adaptation to climate and nutrition [15,16].

Mitochondrial haplogroups have been associated with diverse multifactorial diseases, such as Alzheimer's disease [17], hypertrophic cardiomyopathy [18], retinal diseases [19] or age-related macular degeneration [20]. Variations in mtDNA have also been linked to several types of cancer, such as gastric cancer [21] or renal cell carcinoma [22]. Interestingly, variations in mtDNA have been linked to several types of female cancers, including endometrial [23], ovarian [24] and breast cancer [25,26]. A recent study underlined the possibility that mtDNA might be involved in the pathogenic and molecular mechanisms of familial breast cancer [27].

The Collaborative Oncological Gene-environment Study (COGS) is a European project designed to improve understanding of genetic susceptibility to breast, ovarian and prostate cancer. This project involves several consortia: the Breast Cancer Association Consortium (BCAC) [29], the Ovarian Cancer Association Consortium [30], the Prostate Cancer Association Group to Investigate Cancer Association Alterations in the Genome (PRACTICAL) [31] and the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) [32]. CIMBA is a collaborative group of researchers working on genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers. As part of the COGS project, more than 200,000 single-nucleotide polymorphisms (SNPs) were genotyped for BRCA1 and BRCA2 female mutation carriers on the iCOGS chip, including 129 mitochondrial polymorphisms. The iCOGS chip is a custom Illumina™ Infinium genotyping array (Illumina, San Diego, CA, USA) designed to test, in a cost-effective manner, genetic variants related to breast, ovarian and prostate cancers.

In this study, we explored mitochondrial haplogroups as potential modifiers of breast cancer risk in women carrying pathogenic BRCA1 or BRCA2 mutations. Our study includes females diagnosed with breast cancer and unaffected carriers belonging to CIMBA. We used an original analytic phylogenetics-based approach implemented in a homemade algorithm and in the program ALTree [33,34] to infer haplogroups and to detect associations between haplogroups and breast cancer risk.

**Methods**

**Ethics statement**

A signed informed written consent form was obtained from all participants. All contributing studies involved in CIMBA received approvals from the institutional review committees at their host institutions. Ethical committees that approved access to the data analyzed in this study are listed in Additional file 1.

**BRCA1 and BRCA2 mutation carriers**

Final analyses included 7,432 breast cancer cases and 7,104 unaffected BRCA1 mutation carriers, as well as 3,989 invasive breast cancer and 3,689 unaffected BRCA2 mutation carriers, all belonging to CIMBA. Supplementary specifications regarding inclusion profiles and studies belonging to CIMBA are available in the reports by Couch et al. [35] and Gaudet et al. [36]. All analyses were conducted separately on CIMBA BRCA1 and BRCA2 mutation carriers (abbreviated pop1 and pop2, respectively). Eligible female carriers were aged 18 years or older and had a pathogenic mutation in BRCA1 and/or BRCA2. Women with both BRCA1 and BRCA2 mutations were included in downstream analyses. Data were available for year of birth, age at study recruitment, age at cancer diagnosis, BRCA1 and BRCA2 mutation description and self-reported ethnicity. Women with ovarian cancer history were not excluded from analyses, and they represented 15% and 7% of BRCA1 and BRCA2 mutation carriers, respectively. Information regarding mastectomy was incomplete and was therefore not used as an inclusion or exclusion parameter.

**Genotyping and quality filtering**

Genotyping was conducted using the iCOGS custom Illumina Infinium array. Data from this array are available to the scientific community upon request. Please see [37] for more information. Genotypes were called using Illumina's proprietary GenCall algorithm. Genotyping and quality filtering were described previously [35,36]. Initially, 129 mitochondrial SNPs were genotyped for both BRCA1 and BRCA2 mutation carriers. SNPs fulfilling the following criteria were excluded from downstream analyses: monoallelic SNPs (minor allele frequency = 0), SNPs with more than 5% data missing, annotated as triallelic, or having probes cross-matching with the nuclear genome. Heterozygous genotypes were removed from analyses, and we further filtered out SNPs having more than 5% of heterozygous calls to limit the potential for heteroplasmy affecting our results. We also did not retain SNPs representing private mutations. These mutations are rare, often restricted to a few families, and not sufficiently prevalent in the general population to be included in the reference...
mitochondrial evolutionary tree (see below). This last step of filtration yielded 93 and 92 SNPs for the pop1 and pop2 analyses, respectively (see Additional file 2). Only individuals with fully defined haplotypes (that is, non-missing genotypes for the 93 and 92 SNPs selected for pop1 and pop2, respectively) were included in downstream analyses (14,536 and 7,678 individuals, respectively).

Mitochondrial genome evolution and haplogroup definition
Analyses were based on the theoretical reconstructed phylogenetic tree of the mitochondrial genome (mtTree) known as PhyloTree [38] (v.15). The mtTree is rooted by the Reconstructed Sapiens Reference Sequence (RSRS). RSRS has been identified as the most likely candidate to root the mtTree by refining human mitochondrial phylogeny by parsimony [39]. Each haplogroup in mtTree is defined by the set of mtDNA SNPs that have segregated in RSRS until today in the mitochondrial genome. Each haplogroup is fully characterized by the 16,569-bp sequence resulting from the application of all the substitutions that are encoded by the corresponding SNPs in the RSRS sequence.

Haplogroups imputation
The phylogenetic approach used to infer haplogroups is described in Figure 1. Mitochondrial genome sequences can be reconstructed at each node of mtTree, given the

![Figure 1](image-url)
substitutions that have segregated in RSRS. Each haplogroup therefore has a corresponding full-length mitochondrial sequence. However, the full-length mitochondrial sequence is not available in the data, because the iCOGS platform captured only 93 and 92 SNPs for pop1 and pop2, respectively. Thus, for each of the 7,864 nodes of the phylogenetic tree, the corresponding short haplotype (that is, the full-length sequence restricted to available loci) was defined. Some of the short haplotypes are unique, and they can be matched with their corresponding haplogroup directly. However, most of the time, given the small number of SNPs analyzed, several haplogroups correspond to the same short haplotype. Consequently, a unique haplogroup cannot confidently be assigned to each short haplotype. Therefore, each short haplotype was assigned the most recent common ancestor of all the haplogroups that share the same short haplotype. Once this matching was done, short haplotypes were reconstructed in the same way for each individual in our dataset and were assigned the corresponding haplogroup. The accuracy of the method used was assessed by application to a set of 630 mtDNA sequences of known European and Caucasian haplogroups (see Additional file 3).

Association detection
This phylogenetic approach is based on the identification of subclades in the reference phylogenetic tree of the mitochondrial genome differentially enriched for cases and unaffected controls compared with neighboring subclades. We used ALTTree [33,34] to perform association testing. ALTTree—standing for Association detection and Localization of susceptibility sites using haplotype phylogenetic Trees—is an algorithm used to perform nested homogeneity tests to compare distributions of affected and unaffected individuals in the different clades of a given phylogenetic tree. The objective is to detect if some clades of a phylogenetic tree are more or less enriched in affected or unaffected individuals compared with the rest of the tree. There are as many tests performed as there are levels in the phylogenetic tree. The \( P \)-value at each level of the tree is obtained by a permutation procedure in which 1,000 permutations are performed. Individual labels (“affected” or “unaffected”) are permuted 1,000 times to see to what extent the observed distribution of affected or unaffected is different from a random distribution. A procedure to correct for multiple testing adapted to nested tests [40] is implemented in ALTTree. The objective of ALTTree is to detect an enrichment difference at the level of the whole tree. To conserve computational time and resources, only the most significant \( P \)-value obtained for all tests performed on one tree is corrected.

Handling genetic dependency
ALTTree is used to perform homogeneity tests to detect differences in enrichment or depletion of affected or unaffected individuals between clades in the phylogenetic tree. This kind of test can be performed only on independent data. However, because some individuals in the CIMBA dataset belong to the same family, we constructed datasets with genetically independent data by randomly selecting one individual from among all those belonging to the same family and sharing the same short haplotype. To take into account the full variability of our data, we resampled 1,000 times. The results of the analysis pipeline are obtained for each resampling independently and then averaged over the 1,000 resamplings to obtain final results.

Character reconstruction at ancestral nodes
Before the ALTTree localization algorithm was launched, ancestral sequences were reconstructed at each internal tree node; that is, short haplotypes were inferred with maximum likelihood at all nodes that were not leaves. We used the software PAML [41] to perform the reconstruction at ancestral nodes using a maximum likelihood method. The phylogeny model used was the general time-reversible model (either GTR or REV).

Localization of susceptibility sites
ALTTree also includes an algorithm used to identify which sites are the most likely ones to be involved in the association detected. For each short haplotype observed, the ALTTree add-on altree-add-S adds to the short haplotype sequence a supplementary character called \( S \), which represents the disease status associated with this short haplotype. Are individuals carrying this short haplotype more often affected or unaffected? \( S \) is calculated based on the affected and unaffected counts, the relative proportion of affected and unaffected in the whole dataset, and sensitivity parameter \( e \). \( e \) was set to its default value, which is 1. After \( S \) character computation, haplotypes including character \( S \) are reconstructed at ancestral nodes. Susceptibility site localization is achieved with ALTTree by computing a correlated evolution index calculated between each change of each site and the changes of the character \( S \) in the two possible directions of change. The sites whose evolution are the most correlated with the character \( S \) are the most likely susceptibility sites.

Selected subclades
The analyses were carried out on the full evolutionary tree. However, the more haplogroups there are at each level, the less statistical power homogeneity tests have. Therefore, analyses were also applied to subclades extracted from the tree. Subclades were defined using counts of individuals in each haplogroup of the clade to maximize statistical power. The chosen subclades
and corresponding affected and unaffected counts are presented in Table 2.

Statistical analysis
We quantified the effect associated with enrichment discovered by applying ALTree by building a weighted Cox regression in which the outcome variable is the status (affected or non-affected) and the explicative variable is the inferred haplogroup. Analyses were stratified by country. Data were restricted to the clades of interest. The uncertainty in haplogroup inference was not taken into account in the model. The weighting method used takes into account breast cancer incidence rate as a function of age [42] and the gene containing the observed pathogenic mutation (that is, \(BRCA1\) or \(BRCA2\)). Familial dependency was handled by using a robust sandwich estimate of variance (R package \texttt{survival}, \texttt{cluster()} function).

Results

Haplogroup imputation
In Additional file 4, absolute and relative frequencies are recapitulated for each haplogroup imputed in \(BRCA1\) and \(BRCA2\) mutation carriers. For \(BRCA1\) mutation carriers, we reconstructed 489 distinct short haplotypes of 93 loci from the genotypes data. Only 162 of those 489 short haplotypes matched theoretical haplotypes reconstructed in the reference mitochondrial evolutionary tree. These 162 haplotypes represented 13,315 of 14,536 individuals. Thus, 91.6% of \(BRCA1\) mutation carriers were successfully assigned a haplogroup. For \(BRCA2\) mutation carriers, we reconstructed 350 distinct short haplotypes of 92 loci from our genotype data. Only 139 of those 350 short haplotypes matched theoretical haplotypes reconstructed in the reference mitochondrial evolutionary tree. These 139 haplotypes represented 6,996 of 7,678 individuals. Thus, 91.1% of \(BRCA2\) mutation carriers were successfully assigned a haplogroup. Because more \(BRCA1\) than \(BRCA2\) mutation carriers were genotyped (14,536 vs. 7,678 individuals), we logically observed more distinct haplotypes in \(pop1\) than in \(pop2\) (489 vs. 350 haplotypes).

The accuracy of the main haplogroup inference method used was estimated at 82% and reached 100% for haplogroups I, J, K, T, U, W and X. Given the set of SNPs we disposed of, our method has difficulty differentiating between H and V haplogroups (see Additional file 3).

Association results
For both populations of \(BRCA1\) or \(BRCA2\) mutation carriers, as well as for the full tree as for all selected subclades (see Table 1), we extracted the mean corrected \(P\)-values for association testing over all resamplings performed (see Table 2). The only corrected \(P\)-value that remained significant was that obtained for subclade T (abbreviated T*) in the population of individuals of \(BRCA2\) mutation carriers (\(P = 0.04\)).

The phylogenetic tree of subclade T (see Figure 2a) contains only three levels; thus, only three tests were performed within this clade. Raw \(P\)-values were examined to determine at which level of the tree ALTree detects a difference of enrichment in affected or unaffected individuals (see Table 3). Only the \(P\)-value associated with the test performed at the first level of the tree is significant. We looked more closely at the mean frequencies of affected and unaffected individuals in the tree at this level (see Figure 2b). In the T1a1 subclade, the mean count of affected and unaffected are 32 and 47, respectively. In the T2* subclade, we observed, on average, 217 and 148 affected and unaffected individuals, respectively, whereas in the T subclade, we observed, on average, 13 and 11 affected and unaffected individuals, respectively. The ranges observed for each of these values over the 1,000 resamplings are represented in Figure 2b. On the basis of these observations, we conclude that subclade T1a1 is depleted in affected carriers compared with the neighboring subclades T and T2.

### Table 1 Counts of participants in selected subclades

<table>
<thead>
<tr>
<th>Subclade</th>
<th>(BRCA1) mutation carriers</th>
<th>(BRCA2) mutation carriers</th>
</tr>
</thead>
<tbody>
<tr>
<td>U8</td>
<td>1,458</td>
<td>863</td>
</tr>
<tr>
<td>T</td>
<td>1,243</td>
<td>651</td>
</tr>
<tr>
<td>J</td>
<td>1,270</td>
<td>630</td>
</tr>
<tr>
<td>J1</td>
<td>1,043</td>
<td>513</td>
</tr>
<tr>
<td>H</td>
<td>3,706</td>
<td>1,967</td>
</tr>
<tr>
<td>H1</td>
<td>582</td>
<td>337</td>
</tr>
<tr>
<td>U5</td>
<td>868</td>
<td>458</td>
</tr>
<tr>
<td>X1'2'3</td>
<td>221</td>
<td>103</td>
</tr>
<tr>
<td>K1a</td>
<td>608</td>
<td>364</td>
</tr>
</tbody>
</table>

*pop1, \(BRCA1\) mutation carrier; \(pop2, BRCA2\) mutation carrier. Bold indicates a significant \(P\)-value.

### Table 2 Mean corrected \(P\)-values for association testing with ALTree

<table>
<thead>
<tr>
<th>Subclade</th>
<th>(pop1) corrected (P)-value</th>
<th>(pop2) corrected (P)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>0.083</td>
<td>0.681</td>
</tr>
<tr>
<td>U8</td>
<td>0.146</td>
<td>0.626</td>
</tr>
<tr>
<td>T</td>
<td>0.285</td>
<td>0.040</td>
</tr>
<tr>
<td>J</td>
<td>0.718</td>
<td>0.112</td>
</tr>
<tr>
<td>J1</td>
<td>0.621</td>
<td>0.150</td>
</tr>
<tr>
<td>H</td>
<td>0.747</td>
<td>0.930</td>
</tr>
<tr>
<td>H1</td>
<td>0.268</td>
<td>0.804</td>
</tr>
<tr>
<td>U5</td>
<td>0.829</td>
<td>0.747</td>
</tr>
<tr>
<td>X1'2'3</td>
<td>0.416</td>
<td>0.629</td>
</tr>
<tr>
<td>K1a</td>
<td>0.170</td>
<td>0.162</td>
</tr>
</tbody>
</table>
Localization results
We performed a localization analysis with ALTTree. The correlated evolution index for all non-monomorphic sites observed in short haplotype sequences of subclade T are displayed in Additional file 5. The higher the correlated evolution index, the more likely it is that corresponding sites will be involved in the observed association. Three short haplotype sites numbered 44, 57 and 72 and corresponding to SNPs T988C, G11812A/rs4154217 and G13708A/rs28359178, respectively, clearly distinguish themselves, with correlation index values of 0.390, 0.324 and 0.318, respectively, whereas the correlation index values of all other sites ranged from −0.270 to −0.101. Table 4 shows the details for these three loci.

Effect quantification
The ALTTree method is able to detect an association, but cannot to quantify the associated effect. We estimated the risk of breast cancer for individuals with the T1a1 haplogroup compared with individuals with another T subclade haplogroup in the population of BRCA2 mutation carriers using a more classical statistical method, a weighted Cox regression. We found a breast cancer HR of 0.55 (95% CI, 0.34 to 0.88; \( P = 0.014 \)). We also tested haplogroup T1a1 and compared it with other T* haplogroups and the H haplogroup (the main haplogroup in the general population), and we found a breast cancer HR of 0.62 (95% CI, 0.40 to 0.95; \( P = 0.03 \)).

Discussion
We employed an original phylogenetic analytic method, coupled with more classical molecular epidemiologic analyses, to detect mitochondrial haplogroups differentially enriched for affected BRCA1/2 mutation carriers. We successfully inferred haplogroups for more than 90% of individuals in our dataset. After haplogroup imputation, the ALTTree method identified T1a1 in the T clade as differentially enriched in affected BRCA2 mutation carriers, whereas no enrichment difference was found for BRCA1 mutation carriers. The T subclade is present in 4% of African populations compared with 11% in Caucasian and Eastern European populations [43]. In our data, the T subclade represented 9.34% of BRCA1 mutation carriers and 9.30% of BRCA2 carriers. The ALTTree method also identified three potential breast cancer susceptibility loci in mtDNA. The main goals of using the phylogenetic method we used were to improve statistical power by regrouping subclades according to genetic considerations, to limit the number of tests performed and to precisely quantify this number. ALTTree identified three SNPs of interest. Whereas the association we observed could possibly be driven by a single SNP, no difference was observed between multivariate and univariate cox models including the three SNPs identified by ALTTree (data not shown).

In this study, we investigated to what extent mtDNA variability modified breast cancer risk in individuals.

<p>| Table 3 Non-corrected P-values by level of phylogenetic tree for subclade T in BRCA2 mutation carriers |
|--------------------|-----------------|----------------|</p>
<table>
<thead>
<tr>
<th>Level</th>
<th>Degrees of freedom</th>
<th>Mean of non-corrected P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>0.02141039</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>0.14355900</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>0.22249700</td>
</tr>
</tbody>
</table>

Figure 2 Phylogenetic tree of subclade T tested for association with ALTTree. (a) Phylogenetic tree of subclade T with all observed haplogroups. A homogeneity test is performed at each level of the tree. (b) First level of the phylogenetic tree of subclade T. Averaged counts, ranges and proportions of affected and unaffected observed in resamplings are indicated below each subclade. T2* represents the entire T2 subclade.
Table 4 Description of loci identified as potential susceptibility sites by ALTree\textsuperscript{a}

<table>
<thead>
<tr>
<th>Site</th>
<th>SNP name</th>
<th>Position</th>
<th>Direction of change</th>
<th>Correlated evolution index</th>
<th>Major allele</th>
<th>Minor allele</th>
<th>MAF in pop2</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>MitoT19900C</td>
<td>9,899</td>
<td>T → C</td>
<td>0.390</td>
<td>T</td>
<td>C</td>
<td>0.016</td>
</tr>
<tr>
<td>57</td>
<td>rs41544217</td>
<td>11,812</td>
<td>G → A</td>
<td>0.324</td>
<td>A</td>
<td>G</td>
<td>0.071</td>
</tr>
<tr>
<td>72</td>
<td>rs28359178</td>
<td>13,708</td>
<td>G → A</td>
<td>0.318</td>
<td>G</td>
<td>A</td>
<td>0.111</td>
</tr>
</tbody>
</table>

\textsuperscript{a}MAF, Mean allele frequency; pop2, BRCA2 mutation carrier.

carrying pathogenic mutations in BRCA1/2. A large proportion of breast cancer heritability still remains unexplained today [44]. Different methods exist to study genomic susceptibility to a disease, such as linkage analyses (which identified the BRCA1 and BRCA2 susceptibility genes) or genome-wide association studies (GWASs). However, classical linkage analysis cannot be applied to the haploid mitochondrial genome. Furthermore, commercial GWAS chips available do not adequately capture the majority of mtDNA SNPs. A non-genome-wide and mtDNA-focused approach was required to explore how mtDNA variability influences breast cancer risk. Here we have shown that BRCA2 mutation carriers with the subclade T1a1 have between 30% and 50% less risk of breast cancer than those with other clades, which, if validated, is a clinically meaningful risk reduction and may influence the choice of risk management strategies.

The association we observed among BRCA2, but not BRCA1, mutation carriers may reveal a functional alteration that would be specific to mechanisms involving BRCA2-related breast cancer. Today, it is established that BRCA1- and BRCA2-associated breast cancers are not phenotypically identical. These two types of tumors do not harbor the same gene expression profiles or copy number alterations [45]. Breast cancer risk modifiers in BRCA1/2 mutation carriers have already been identified [46]. However, most of them are specific from one or the other type of mutation carried [47]. It is therefore not surprising that this observation is observed in BRCA2 mutation carriers only.

Our inability to assign haplogroups to 9% of study participants could have three main explanations. First, given the high mutation rate in the mitochondrial genome, observed combinations of mtDNA SNPs might have appeared relatively recently in the general population, and the corresponding haplotypes might not yet be incorporated into PhyloTree. Second, only one genotyping error could lead to chimeric haplotypes that do not exist, although, given the quality of our genotyping data, this is unlikely. Third, the mitochondrial reference evolutionary tree PhyloTree is based on phylogeny reconstruction by parsimony, and, for some subclades, it might be suboptimal, especially for haplogroups relying on few mitochondrial sequences, as is the case for African haplogroups [48]. In cases of uncertainty, the choice we made to assign the most recent common ancestor to the studied haplotype enabled us to improve statistical power without introducing a bias in the detected association. For the association detected between T, T1* and T2* subclades, the haplogroup inference method used did not bias the counts of affected and unaffected individuals in these subclades. More details are presented in Additional file 6. Furthermore, on the basis of the haplogroup inference with our method of 630 European and Caucasian mtDNA sequences whose haplogroup is known, we successfully assigned the correct main haplogroup and subhaplogroup of 100% of sequences belonging to T, T2* and T1a1* haplogroups.

We quantified the effect corresponding to the detected association by using a more classical approach. We built a weighted Cox regression including inferred haplogroup as an explicative variable. However, the uncertainty in haplogroup inference was not taken into account in this model. Nevertheless, based on haplogroup assignment and regrouping performed in clade T, affected and unaffected counts of individuals in this clade were not biased.

With only 129 loci genotyped over the 16,569 nucleotides composing the mitochondrial genome, we certainly did not explore the full variability of mitochondrial haplotypes. A characterization of individual mitochondrial genomes would require more complete data acquisition methods to be used, such as next-generation sequencing. However, next-generation sequencing has its own limits and challenges, because some regions of the mitochondrial genome are not easily mappable, owing to a high homology with the nuclear genome, among other factors, and important bioinformatics treatment is necessary to overcome sequencing technology biases. Finally, even for a relatively short genome of “only” 16,569 bp, mtDNA sequencing of more than 20,000 individuals would represent a major increase in cost relative to genotyping 129 SNPs.

ALTTree identified T9899C, G11812A/rs41544217 and G13708A/rs28359178 as three potential susceptibility sites for the discovered association (see Additional file 7). These three SNPs are located in the coding part of genes MT-Co3, MT-Nd4 and MT-Nd5, respectively. When looking at PhyloTree, T9899C seems to be involved in T1 subclade definition, whereas G13708A and A11812G are involved in T2 subclade definition. Whereas T98899C and
G11821/rs41544217 are synonymous SNPs, G10398A leads to a change of amino acid in the final protein (from alanine to threonine). These two synonymous SNPs have never been described in a disease context in the literature. G13708A is also known for being a secondary mutation for Leber’s hereditary optic neuropathy (LHON) and multiple sclerosis [49]. Although the role of secondary mutations in LHON is still controversial, G13708A could be associated with impairment of the respiratory chain in this pathology. G13708A has also been described as a somatic mutation in a breast cancer tumor, whereas it was not present in adjacent normal tissue or in blood leukocytes [50]. A high proportion of mitochondrial somatic tumor-specific variants are also known mtDNA SNPs, which is consistent with the hypothesis that tumor cells are prone to acquire the same mutations that segregate into mtDNA by selective adaptation when humans migrated out of Africa and confronted new environments [51]. Interestingly, the germline variant G13708A has already been shown to be inversely associated with familial breast cancer risk (with the same direction of the association), with a breast cancer odds ratio of 0.47 (95% CI, 0.24 to 0.92) [52]. None of these SNPs have been described in the context of ovarian cancer.

The corrected P-value obtained using ALTREE in studying clade T is 0.02, which is not highly significant. A replication step should be performed to validate these results. However, it will be difficult to include enough women in this replication step, given the specific profile studied here. In fact, the estimations of BRCA2 pathogenic mutations in the general population range from 0.068% [5] to 0.69% [53]. T1a1 represents only a small percentage of European haplogroups (from 1% to 2%). The number of women who have this association is therefore low. However, women carrying such mutations are confronted with drastic choices regarding the prevention of breast cancer, notably prophylactic mastectomy or complete hysterectomy. If breast cancer risk is really reduced by a factor of 2 for women with T1a1, this could be an important fact to take into account for breast cancer prevention.

Conclusions

This study and our results suggest that mitochondrial haplogroup T1a1 may modify the individual breast cancer risk in BRCA2 mutation carriers. For now, this observation cannot be extended to the general population. Further investigation of the biological mechanism behind the associations we observed may further reinforce the hypothesis that the mitochondrial genome is influential in breast cancer risk, particularly among carriers of BRCA2 mutations, and, if validated, is of a level to influence cancer risk management choices.

Additional files

**Additional file 1:** List of ethical committees that approved the access to the data analyzed in this study.

**Additional file 2:** SNPs selected for downstream analyses.

**Additional file 3:** Description and results of the procedure used to estimate the accuracy of our haplogroup inference methodology.

**Additional file 4:** Absolute and relative frequencies of imputed haplogroups by population. Table containing absolute and relative frequencies of imputed haplogroups for BRCA1 and BRCA2 mutation carriers.

**Additional file 5:** Correlated evolution index for all non-monomorphic sites observed in short haplotype sequences of subclade T. Table containing correlated evolution index for all non-monomorphic sites observed in short haplotypes sequences of subclade T.

**Additional file 6:** Details of haplogroups inference results for subclade T.

**Additional file 7:** Methods used to compute coevolution index.

**Abbreviations**

BCAC: Breast Cancer Association Consortium; CI: Confidence interval; CIMBA: Consortium of Investigators of Modifiers of BRCA1/2; COGS: Collaborative Oncological Gene-environment Study; DSB: Double-strand break; GWAS: Genome-wide association study; HR: Hazard ratio; LHON: Leber’s hereditary optic neuropathy; MAF: Mean allele frequency; mtDNA: Mitochondrial DNA; mtTree: Phylogenetic tree of the mitochondrial genome; OCAC: Ovarian Cancer Association Consortium; pop1: BRCA1 mutation carrier; pop2: BRCA2 mutation carrier; PRACTICAL: Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome; ROS: Reactive oxygen species; RSRS: Reconstructed Sapiens Reference Sequence; SNP: Single-nucleotide polymorphism.

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

Conception and design: DGC and GT. Development of methodology: SB, CBa and VD. Acquisition of data: LM, Shie, DB, ALe, JD, KBK, PS, MRt, WKc, DRG, SSB, Rl, DT, NT, CMd, EJvr, SLN, YCD, AMG, BE, FCN, TVOht, A0, JBe, RA, ES, JNW, MtHe, PP, Pp, R0d, BB, BP, DZ, GSC, SmAn, LV, GLC, LP, LO, DY, IK, JGa, UH, AD, Brer, CB, DGe, DF, DFd, FDo, JCe, JMo, A, Jlr, WI, LI, LES, MIk, MTI, MTR, MElp, PM, RP, RE, Rda, SHo, TCo, EKg, OJ, K, KCl, DML, AM, AG, BW, CS, CE, DN, DS, HP, KK, KR, ND, NA, RV, RkS, SP, NB, SW, APd, CLe, CLaS, DL, ER, FDa, GSC, HD, LB, LG, NU, VB, YS, YB, JCa, UlL, MP, PAd, MdH, TCa, HN, KA, AjAg, MsNd, DCM, CMa, FEvL, FBH, HEM, JCo, KVR, MAR, PD, RBvdl, EO, OD, AT, CLaZ, iB, JDv, AKa, GsS, JG, JlU, KO, KJ, BAA, CM, AA, MM, MRT, ABS, WF, CO, NLI, VSP, CS, AlLnc, LJ, MC, MR, IV, ABe, AF, CFS, CR, DGK, GP, MTe, MhG, PLM, Gr, El, AMM, GG, ILA, ST, AET, ISP, MThO, TAK, U BJ, MAC, EF, JZ, YL, AlLnD, BmL, BA, NLo, RR, OId, RLN, SR, KlN, SMD, Trr, BKA, GM, BYK, JLe, So, DSl, GT, JS, FJc, KO, DFE, GC, ACA, SMAz, OMp and OMS. Analysis and interpretation of data: SB, DGC and ACA. Writing of the manuscript: SB, DGC, ACA, SH, ABS, GC, SLN, AET, ILA, JCO, KO, MThO and GM. All authors read and approved the final manuscript.

**Acknowledgements**

Collaborative Oncological Gene-environment Study (COGS): This study would not have been possible without the contributions of the following: Per Hall (CIMBA); Kyriaki Michailidou, Marjoke K Bell and Qin Wang (Breast Cancer Association Consortium (BCAC)); Rosalind A Eeles, Ali Amin Al Olama, Zsofia Kote-Jarai and Sara Benlloch (PRACTICAL); Alison M Dunng, Craig Luccarini, Michael Lush and the staff of the Centre for Cancer Genetic Epidemiology; Simard and Daniel C Tessier, Francois Bacot, Daniel Vincent, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre; and Julie M Cunningham, Sharon A Windebank, Christopher A Hillker, Jeffrey Meyer and the staff of Mayo Genotyping Core Facility.

**Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA):** Maggie Angelakos, Judi Maskiell, Gillian Dite and Helen Tsimiklis; members of and participants in the New York site of the Breast Cancer Family Registry.
members of and participants in the Ontario Familial Breast Cancer Registry for their contributions to the study; Vilus Rudatish, Laimonas Grinkevicius, Drs Janis Egils, Anna Kitlova and Aivars Stengrevics; the families who contribute to the BRCA1/2 mutations and breast cancer in South African women (BMBSA) study; Chun Ding and Linda Steele; Alicia Barroso, Rosario Alonso, Guillermo Pita, Alessandra Vie and Lara delia Puppa of the Centro di Riferimento Oncologico, IRCCS, Aviano (PN), Italy; Laura Papi of the University of Florence, Florence, Italy; Monica Barile of the Istituto di Eziologia di Oncologia, Milan, Italy; Liliana Varesco of the IRCCS AOI San Martino – IST Istituto Nazionale per la Ricerca sul Cancro, Genoa, Italy; Stefania Tommasi, Brunella Pilato and Rossana Lambro of the Istituto Nazionale Tumori "Giovanni Paolo II" – Bari, Italy; Aline Martayan of the Istituto Nazionale Tumori Regina Bina, Rome, Italy; Maria Grazia Tistelli of the Ospedale di Circolo-Università dell’Insuescia, Varese, Italy; and the personnel of the Cogentech Genetic Test Laboratory, Milan, Italy.

**Epidemiological Study of BRCA1 and BRCA2 Mutation Carriers (EMBRACE) Collaborating Centers:** coordinating center, Cambridge, UK: Debra Frost, Steve Ellis, Elena Fineberg and Radka Piatte; North of Scotland Regional Genetics Service, Aberdeen, UK: Zsuzsa Arvai and Maria Novak; Northern Ireland Regional Genetics Service, Belfast, UK: Patrick Morrison and Lisa Jeffers; West Midlands Regional Clinical Genetics Service, Birmingham, UK: Trevor Cole, Kai-ren Ong and Jonathan Hoffman; South West Regional Genetics Service, Bristol, UK: Alan Donaldson and Margaret James; East Anglian Regional Genetics Service, Cambridge, UK: Marc Tischkowitz, Joan Paterson and Amy Taylor; Medical Genetics Services for Wales, Cardiff, UK: Alexandra Murray, Mark T Rogers and Emma McCann; St James’s Hospital, Dublin, and National Centre for Medical Genetics, Dublin, Ireland: M. John Kennedy and David Barton; South East of Scotland Regional Genetics Service, Edinburgh, UK: Mary Porteous and Sarah Drummond; Peninsula Clinical Genetics Service, Exeter, UK: Carole Bewer, Emma Kivuva, Anne Searle, Selina Goodman and Kathryn Hill; West of Scotland Regional Genetics Service, Glasgow, UK: Rosemary Davidson, Victoria Murray, Nadia Brodthau, Lesley Snadden, Mark Longmuir, Catherine Watt, Sarah Gibson, Elisha Haque, Ed Tobias and Alexis Duncan; South East Thames Regional Genetics Service, Guy’s Hospital London: Louise Izatt, Chris Jacobs and Caroline Langman; North West Thames Regional Genetics Service, Harrow, UK: Huw Dorkins; Leicestershire Clinical Genetics Service, Leicester, UK: Julian Banwell; Yorkshire Regional Genetics Service, Leeds, UK: Julian Adlard and Gemma Serra-Fellu; Cheshire & Merseyside Clinical Genetics Service, Liverpool, UK: Ian Ellis and Catherine Houghton; Manchester Regional Genetics Service, Manchester, UK: D Gareth Evans, Fiona Laloo and Jane Taylor. North East Thames Regional Genetics Service, NE Thames, London: Lucy Side, Alison Male and Cheryl Berlin; Nottingham Centre for Medical Genetics, Nottingham, UK: Jacqueline Eason and Rebecca Collier; Northern Clinical Genetics Service, Newcastle, UK: Fiona Douglas, Oonagh Clabber and Irene Jobson; Oxford Regional Genetics Service, Oxford, UK: Lisa Walker, Diane McLeod, Dorothy Halliday, Sarah Drewell and Barbara Stayner; The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London: Ros Eeles, Susan Shanahan, Nasneen Rahman, Richard Houlston, Elizabeth Bancroft, Elizabeth Page, Audrey Ardem-Jones, Kelly Kohut, Jennifer Wiggins, Elena Castro, Emma Killick, Sue Martin, Gillian Rea and Anjana Kulkarni; North Trent Clinical Genetics Service, Sheffield, UK: Jackie Cook, Oliver Quarrer and Kathryn Bardsley; South West Thames Regional Genetics Service, London: Shirley Hodgson, Sheila Goff, Glen Brice, Lizze Winchester, Charlotte Eddy, Vishalika Tripathi, Virginia Attard and Anna Lehmann; Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, UK: Diana Ecrich, Amanda Low, Gillian Crawford, Donna Mbendi and Sarah Smallley; and JollEen Weaver and Dr Betsy Bove for their technical support.

**Genetic Modifiers of Cancer Risk in BRCA1/2 Mutation Carriers (GEMO) study:** National Cancer Genetics Network UNICANCER Genetic Group, France; GEMO Collaborating Centers: coordinating centers, Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Hospices Civils de Lyon – Centre Léon Bérard, and équipe Génétique du cancer du sein, Centre de Recherche en Cancérologie de Lyon; Olga Sinilnikova, Sylvie Mazoyer, Francesca Demagni, Laure Barjouhé, Carole Verno-Peyre, Alain Calender, Sophie Giraud and Mélanie Léone; and Service de Génétique Oncologique, Institut Curie, Paris: Dominique Stoppa-Lyonnet, Marion Gauthier-Villars, Bruno Buechler, Claude Houdayer, Virginie Moncoutier, Muriel Belotti, Carole Tirapo, Antoine de Pauw; Institut Gustave Roussy, Villejuif; France; Brigitte Bressac-de-Paillents and Olivier Caron; Centre Jean Perrin, Clermont-Ferrand, France; Yves-Jean Bignon and Nancy Uthman; Centre Léon Bérard, Lyon, France; Christine Lasset, Valérie Bonadona and Sandrine Handallou; Centre Français Bactéries, Caen, France; Agnès Hardouin and Pascaline Berthet; Institut Paoli Calmettes, Marseille, France; Hayag Sobol, Violaine Bourdon, Tetsuro Noguchi, Audrey Remenieres and François Esiriger; Centre Hospitalier RégionalUniversitaire Arnaud-de-Villeneuve, Montpellier, France; Isabelle Coupler and Pascal Puel; Centre Oscar Lambret, Lille, France: Jean-Philippe Peyrat, Joëlle Fournier, François Révillon, Philippe Vennin and Claude Adenis; Hospital Renè Huguenin, Institut Curie, Saint-Cloud, France; Etienne Rouleau, Rosette Lidereau, Liliane Demange and Catherine Nougès; Centre Paul Strauss, Strasbourg, France: Danièle Muller and Jean-Pierre Fricker; Institut Bergonie, Bordeaux, France: Emmanuelle Barouk-Simonet, François Bonnet, Virginie Bubien, Nicolas Sevenet and Michel Longy; Institut Claudius Regaud, Toulouse, France: Christinne Toulou, Rosine Guimbard, Laurent Gladieff and Viviane Feillet; Centre Hospitalier Universitaire de Grenoble, Grenoble, France; Dominique Leroux, Hélène Dreyfus, Christine Rebischung and Magalie Paysselon; Centre Hospitalier Universitaire de Dijon, Dijon, France: Fanny Coror and Laurence Faivre; Centre Hospitalier Universitaire de Saint-Etienne, Saint-Etienne, France: Fabienne Prieur, Marine Lebrun and Caroline Kientz. Hôtel Dieu Centre Hospitalier, Chambéry, France: Sandra Fert Ferrer, Centre Antoine Lacassagne, Nice, France: Marc Frémy; Centre Hospitalier Universitaire de Lille, Lille, France: Janis Eglitis, Anna Krilova and Aivars Stengrevics; the families who contribute to this resource; and the many clinicians and patients for their contributions to the study; Vilius Rudatish, Laimonas Grinkevicius, Drs Janis Egils, Anna Kitlova and Aivars Stengrevics; the families who contribute to the BRCA1/2 mutations and breast cancer in South African women (BMBSA) study; Chun Ding and Linda Steele; Alicia Barroso, Rosario Alonso, Guillermo Pita, Alessandra Vie and Lara delia Puppa of the Centro di Riferimento Oncologico, IRCCS, Aviano (PN), Italy; Laura Papi of the University of Florence, Florence, Italy; Monica Barile of the Istituto Nazionale Tumori Regina Bina, Rome, Italy; Maria Grazia Tistelli of the Ospedale di Circolo-Università dell’Insuescia, Varese, Italy; and the personnel of the Cogentech Genetic Test Laboratory, Milan, Italy.
Czech Republic); Michal Zikan, Petr Pohlreich and Zdenek Klebl (Oncogenetic Center and Department of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, Prague, Czech Republic); Anne Lincoln and Lauren Jacobs; the National Israeli Cancer Control Center (NICC) National Familial Cancer Consultation Service team led by Sara Dshon; the laboratory team led by Dr Flavio Lejbkowicz; and the research field operations team Ma Pinchev; members and participants in the Ontario Cancer Genetics Network for their contributions to the study. Leigha Senter, Kevin Sweet, Caroline Cavene and Michelle O’Connor were instrumental in accrual of study participants, ascertainment of medical records and database management; The Ohio State University Human Genetics Sample Bank; the Merav Breast Center team at the Sheila Medical Center, Åke Berg, Hilak Lilson, Helena Jerntoftin, Karin Herbstjat, Maria Soller and UK Kirstofferson; Sahlgrenska University Hospital, Gothenburg, Sweden: Anna Överholm, Margareta Nordling, Per Karlsson and Zakaria Einberg; Stockholm and Karolinska University Hospital, Stockholm: Anna von Wachenfeldt, Annelle Liljegren, Annika Lindblom, Brita Arver, Gisela Barbany Buxtina and Johanna Ranttia; Umeå University Hospital, Umeå, Sweden: Beatrice Melin, Christina Edwindsdotter Adorne and Monica Emanuelsen; Uppsala University, Uppsala, Sweden: Per Henricson, Anna Hallström Piggar and Yvonne Rosenquist; Linköping University Hospital, Linköping, Sweden: Marie Steenmark-Askmal and Sigurn Liedegren; Cecilia Zvocer, Qiu Niu, and physicians, genetic counselors, research nurses and staff of the Comprehensive Cancer Risk and Prevention Clinic of University of Chicago Medicine, Chicago, IL, USA, for their contributions to this resource; Joyce Seldon, MSGC, and Lorna Kwan, MPH, Dr Robert Nussbaum and the following genetic counselors: Beth Crawford, Kate Loranger, Julie Mak, Nicola Stewart, Robin Lee, Annie Blanco and Peggy Conrad; Salina Chan; Paul DP Harris, Simon Gayther, Susan Rams, Carole Pye, Patricia Harrington and Eva Wozniak for their contributions to the UK Familial Ovarian Cancer Registry (UFOOCR); Geoffrey Lidman, Marion Harris, Martin Delatycki of the Victorian Familial Cancer Trials Group; and Sarah Sawyer, Rebecca Dreissen and Ella Thompson.

**Funding**

**Higher-level funding:** The COGS project is funded through a European Commission Seventh Framework Program grant (agreement number 223175: HEALTH-F2-2009-223175). The OIMBA data management and data analysis were supported by Cancer Research UK grants C12/292/A11174 and C1287/A10118. SH is supported by a National Health and Medical Research Council (NHMRC) program grant (to GCT).

**Individual researcher support:** ACA is a Cancer Research UK Senior Cancer Research Fellow (C12292/A11174). DFE is a Principal Research Fellow of Cancer Research UK. GC, M/C and IC are supported by the National Health and Medical Research Council (NHMRC). BK holds an American Cancer Society Early Detection Professorship (SIP-06-258-01-COLUN). MHG and P/LM were supported by funding from the Intramural Research Program of the National Cancer Institute, National Institutes of Health. OIO is an American Cancer Society Clinical Research Professor. JS is Chairholder of the Canada Research Chair in Oncogenetics.

**Funding of constituent studies:** The Breast Cancer Family Registry (BCFR) was supported by grant U01 CA164920 from the National Cancer Institute, National Institutes of Health. The content of this article does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the BCRF, nor does mention of trade names, commercial products or organizations imply endorsement by the US Government or the BCRF. The Baltic Familial Breast Ovarian Cancer Consortium (BFBOCC) is partly supported by Lithuania (BFBOCC-LT): Research Council of Lithuania grant LG-07/2012; BFBOCC-LV (Latvia) is partly supported by LSC grant 10.0011.08 and in part by a grant from the European Social Fund number 2009/022/1DP/1.1.1.0/09/APIA/IAIA/016 and the Liepāja City Council, Liepāja, Latvia; Beth Israel Deacess Medical Center Cancer Center is supported by the Breast Cancer Research Foundation; BRCA2-gene mutations and breast cancer in South African women (BMBSA) was supported by grants from the Cancer Association of South Africa (Cansa) to Elizabeth J van Rensburg; SLN (Beckman Research Institute, City of Hope, Duarte, CA, USA) was partially supported by the Morris and Horowitz Families Professorship in Cancer Biology and Outcomes Research; the Copenhagen Breast Cancer Study (CBCS) was supported by the NEYE Foundation; the Spanish National Cancer Research Center (Centro Nacional de Investigaciones Oncológicas (CNIO)) was partially supported by the Spanish Association against Cancer (Asociación Española Contra el Cáncer AECC08), Thematic Network Cooperative Research in Cancer (Red Temática Investigación Cooperativa en Cáncer (RTICC), Centro de Investigación Cáncer, Salamanca, Spain) RTICC 06/0020/1060, Spanish Ministry of Science and Innovation grants FIS PI08 1120 (Fondo de Investigación Sanitaria (FIS)) and SAF2010-20493, and the Foundation Mutua Madrileña (FMMA); the City of Hope Clinical Cancer Genetics Community Network and the Hereditary Cancer Research Group (CHG-CGON), supported in part by award number RC4CA153828 (Principal Investigator: JNM) from the National Cancer Institute and the Office of the Director, National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**CONSOrdio Studi Italiani sui Tumori Ereditari Alla Mammella, Italy (CONSOrdio Team):** Funds from Italian citizens who allocated the "5 × 1,000" share of their tax payment in support of the Fondazione Ricerca Istituto Nazionale Tumori, according to Italian laws (Istituto Nazionale dei Tumori INT) institutional strategic project "5 × 1,000" (to SM); the Italian Association for Cancer Research (ARI [to LO]; National Centre for Scientific Research "Demokritos" has been cofunded by the European Union (European Social Fund (ESF)) and Greek national funds through the "Education and Lifelong Learning" operational program of the National Strategic Reference Framework (NSRF) – Research Funding Program of the General Secretariat for Research and Technology: ARISTEA; "Heracleitus II: Investing in knowledge society through the European Social Fund"; the DKFZ study was supported by the Deutsches Krebsforschungszentrum (DKFZ); Epidemiological Study of BRCA1 and BRCA2 Mutation Carriers (EMBRACE) is supported by grants: Cancer Research UK grants C1287/A10118 and C1287/A11190; DGE and FL are supported by a National Institute for Health Research (NIHR) grant to the Biomedical Research Centre, Manchester, UK; the investigators at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust are supported by an NIHR grant to the Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust, London; RE and EB are supported by Cancer Research UK grant CS047/AA385; Kansas University Medical Center: The authors acknowledge support from The University of Kansas Cancer Center (P30 CA168524) and the Kansas Bioscience Authority Eminent Scholar Program; AKG was funded by grants SU01 CA113916 and R01 CA140323 and by the Chancellors Distinguished Chair in Biomedical Sciences Professorship; The German Consortium of Hereditary Breast and Ovarian Cancer (GC-HBOC) is supported by German Cancer Aid (grant 109076) (to RKS) and by the Center for Molecular Medicine Cologne (CMMMC); GC-HBOC is deeply grateful to Dr Sabine Preisler-Adam for providing information and samples; the GEMO Study was supported by the Ligue Nationale Contre le Cancer; the Association "Le cancer du sein, parlons-en!" Award; and the Canadian Institutes of Health Research for the CIHR Team in Familial Risks of Breast Cancer program; G-FAST: KDL is supported by GOA grant BOF10/GOA/019 (Ghent University) and spearhead funding of Ghent University Hospital; the Gynecologic Oncology Group (GOG) was supported by National Cancer Institute grants to the GOG Administrative Office and Tissue Bank (grant CA 72469), the GOG Statistical and Data Center (grant CA 37517) and GOG’s Cancer Prevention and Control Committee (grant CA 101169); HSCC was supported by grants RD1/0036/00012 and 12/00339 from Instituto de Salud Carlos III (ISCIII), Madrid, Spain, partially supported by European Regional Development Fund (Fonds européen de développement régional (FEDER)) funds; the Helsinki Breast Cancer Study (HEBCS) was financially supported by the Helsinki University Central Hospital Research Fund, Academy of Finland (266528), the Finnish Cancer Society Study and the Sigrid Juselius Foundation; HEBON is supported by the Dutch Cancer Society grant NKI1998-1854, NK02004-3988 and NK02007-3756, the Netherlands Organization of Scientific Research grant NWO 11090924, the Pink Ribbon grant 110005 and Biobanking and Molecular Resource Infrastructure (BBMRI) grant NWO 184.021.007/CP46; HEBON thanks the registration teams of the Comprehensive Cancer Centre Netherlands and Comprehensive Centre South (together the Netherlands Cancer Registry) and PALGA (Dutch Pathology Registry) for part of the data collection; the High Risk Breast Cancer Program (HRBPC) is supported by the Hong Kong Hereditary Breast Cancer Family Registry and the Dr Ellen Li Charitable Foundation, Hong Kong; the Hungarian Breast and Ovarian Cancer Study (HUNBOCS) was supported by Hungarian Research and Technological Innovation Fund (KIA)/Hungarian Scientific Research Fund (Országos Tudományos Kutatási Alapprogramok (OTKA) research grants KIA-OTKA C5047/A8385; Instituto Català d’Oncologia (ICO): contract grant sponsor: Asociación Española Contra el Cáncer, Spanish Health Research Foundation; Ramón Areces Foundation; Instituto de Salud...
Carlos III (ISCIII); Catalan Health Institute; and Autonomous Government of Catalonia; contract grant numbers ISCIII RTI06/0020105, 10P9/02483, P11/01422, P11/00748, P11/00285, P11/00189 2009SGR290 and P11/00189 2009SGR283; the International Hereditary Cancer Center (Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland) was supported by grant P2Z_KBN_122/P07/2004; Landsprit – The National University Hospital of Iceland was supported by the Icelandic Association "Walking for Breast Cancer Research" and by the Landsprit University Hospital Research Fund; the Interdisciplinary Health Research Internal Team Breast Cancer Susceptibility Study (INHERT) was supported by the Canadian Institutes of Health Research (CIHR) for the "CIHR Team in Familial Risks of Breast Cancer" program, Canadian Breast Cancer Research Alliance grant O91515 and Ministry of Economic Development, Innovation and Export Trade grant PSR-SIRI-701; the Istituto Oncologico Veneto Hereditary Breast and Ovarian Cancer Study (IOV-HOBCS) is supported by the Ministero della Salute and a "> 1,000" Istituto Oncologico Veneto grant; the Portuguese Oncology Institute–Porto Breast Cancer Study (IPOBCS) was supported in part by Liga Portuguesa Contra o Cancro; KcnFab is supported by a grant from the Netherlands Organisation for Health Research and Development (NWO-VICI) (CA116201), a US Department of Defense Ovarian Cancer Idea award (W81XWH-10-1-0341), a grant from the Breast Cancer Research Foundation, a generous gift from the David F and Margaret T Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation; McGill University Jewish General Hospital Weekend to End Breast Cancer, Quebec Ministry of Economic Development, Innovation and Export Trade; Modifier Study of Quantitative Effects on Disease (ModSQaD) was supported by the Ministry of Health of the Czech Republic to Maxaryl Memorial Cancer Institute (MHCZ – DRO) (MMD 00209805) and by the European Regional Development Fund and the State Budget of the Czech Republic (RECAMO, CZ.1.05/ 2.1.00/03.0101) (to LF), and by Charles University in Prague project UNCE204024 (M2), Memorial Sloan Kettering Cancer Center (MSKCC) is supported by grants from the Breast Cancer Research Foundation and Robert and Kate Niehaus Clinical Cancer Genetics Initiative; National Cancer Institute, National Institutes of Health: The research of MHG and PLM was supported by the Intramural Research Program of the National Institutes of Health, and by support services contracts N02-CP-11019-50 and N02-CP-65504 with Westat, Inc, Rockville, MD, USA; the National Cancer Institute, National Institutes of Health is supported by the Intramural Research Program of the National Institutes of Health, and by support services contracts N02-CP-11019-50 and N02-CP-65504 with Westat, Inc, Rockville, MD, USA.

Author details

1INSERM U1052, CNRS UMR5286, Université Lyon 1, Centre de Recherche en Cancérologie de Lyon, Lyon, France. 2Université de Lyon, 69000 Lyon, France. 3Université Lyon 1, 69100 Villeurbanne, France. 4UMR CNRS 5558, Laboratoire de Biométrie et Biologie Evolutive (LBBE), "Biométrie et Biologie Evolutive", Université Claude Bernard Lyon 1, Bâtiment Géor Gendel, 46 boulevard du 11 novembre 1918, 69622 Villeurbanne, cedex, France. 5Université Grenoble Alpes, UMR 5217, Laboratoire d’Informatique de Grenoble (LIG), équipe-projet Multi-programmation et Ordonnancement sur ressources pour les Applications Interactives de Simulation (MOAIS), 38041 Grenoble, France. 6INRIA Rhône-Alpes, équipe-projet MOAIS, 38334 Saint Ismier, Cedex, France. 7New York for Cancer Genetics Research, Columbia University and Primary Care, University of Cambridge, Cambridge, UK. 8Department of Genetics and Computational Biology, QIMR Berghofer, Brisbane, Australia. 9Centre de recherche du Centre hospitalier universitaire de Québec, Laval University, Charlesbourg, PQ, Canada. 10Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA. 11Department of Pediatrics, Columbia University College of Physicians and Surgeons, New York, NY, USA. 12Department of Medicine, Columbia University College of Physicians and Surgeons, New York, NY, USA. 13Department of Dermatology, University of Utah School of Medicine, Salt Lake City, UT, USA. 14Department of Epidemiology, Cancer Prevention Institute of California, 2201 Walnut Avenue, Suite 300, Fremont, CA 94538, USA. 15Hematology, Oncology and Transfusion Medicine Center, Vilnius University Hospital Santariskiu Clinics, Vilnius, Lithuania. 16Department of Molecular and Regenerative Medicine, Centre for Innovative Medicine, State Research Institute, Vilnius, Lithuania. 17Latvian Biomedical Research and Study Centre, Ratsupelis iela 1, Riga LV-1067, Latvia. 18Division of Hematology Oncology, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Shapiro 9, Boston, MA 02215-5400, USA. 19Department of Genetics, University of Pretoria, Private Bag X20, Hatfield, 0028 Pretoria, South Africa. 20Department of Population Sciences, Beckman Research Institute, City of Hope, 1500 East Duarte Road, Duarte, CA 91010, USA. 21Department of Clinical Genetics, Righi-hostipal, Copenhagen University Hospital, Copenhagen, Denmark. 22Department of Oncology, Righi-hostipal, Copenhagen University Hospital, Copenhagen, Denmark. 23Center for Genomic Medicine, Righi-hostipal, Copenhagen University Hospital, Copenhagen, Denmark. 24Human Genetics Group, Spanish National Cancer Research Center (CNIO), Madrid, Spain. 25Center for Biomedical Network Research on Rare Diseases (CIBERER), Madrid, Spain. 26Medical Oncology Service, Hospital Clinico Universitario Lozano Blesa, Avenida San Juan Bosco, 15, 50009 Zaragoza, Spain. 27Holy Cross Hospital, Michael and Dianne Bienes Comprehensive Cancer Center, Fort Lauderdale, FL, USA. 28Division of Clinical Cancer Genetics, City of Hope (for the Clinical Cancer Genetics Community Research Network), City of Hope, 1500 East Duarte Road, Duarte, CA 91010, USA. 29Department of Preventive and Predictive Medicine, Istituto di Ricerca e Cur a a Carattere Scientifico (IRCCS), Istituto Nazionale dei Tumori (I.N.T), Via Venezia 1, 20133 Milan, Italy. 30Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Istituto di Ricerca e Cur a a Carattere Scientifico (IRCCS), Istituto Nazionale dei Tumori (I.N.T), Via Venezia 1, 20133 Milan, Italy. 31Gogentech Cancer Genetic Test Laboratory, Via Adamelio 16, 20139 Milan, Italy. 32University of Milan, Italy. 33Biomedical Research and Clinical Trial Unit, Centro di Riferimento Oncologico (CRON), Via Franco Gallini 2, 33081 Aviano, Italy. 34Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia, Via Ripamonti 435, 20141 Milan, Italy. 35Division of Medical Genetics, Department of Preventive and Predictive Medicine, Istituto di Ricerca e Cur a a Carattere Scientifico (IRCCS), Istituto Nazionale dei Tumori (I.N.T), Via Venezia 1, 20133 Milan, Italy. 36Division of Medical Genetics, Department of Preventive and Predictive Medicine, Istituto di Ricerca e Cur a a Carattere Scientifico (IRCCS), Istituto Nazionale dei Tumori (I.N.T), Via Venezia 1, 20133 Milan, Italy. 37Division of Medical Genetics, Department of Preventive and Predictive Medicine, Istituto di Ricerca e Cur a a Carattere Scientifico (IRCCS), Istituto Nazionale dei Tumori (I.N.T), Via Venezia 1, 20133 Milan, Italy.
Prevention and Special Functions, Istituto di Ricovero e Cura a Carattere Scientifico (IRCSS), Azienda Ospedaliera Universitaria “San Martino” di Genova, IST Istituto Nazionale per la Ricerca sul Cancro, Largo Rosanna Benzi, 10, 16132 Genova, Italy. 44FlorGen Foundation for Pharmacogenomics, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy. 45Unit of Medical Genetics, Department of Biomedical, Experimental and Clinical Sciences of Florence, Florence, Italy. 46Department of Molecular Medicine, Sapienza University, Rome, Italy. 47Department of Medical Oncology, Papageorgiou Hospital, Aristotle University of Thessaloniki School of Medicine, Thessaloniki, Greece. 48Molecular Diagnostics Laboratory, INRATES, National Centre for Scientific Research “Demokritos”, Aghia Paraskevi Attiki, Athens, Greece. 49Piana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215, USA. 50Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany. 51Clinical Genetics Department, St Michael’s Hospital, Southwell Street, Bristol BS2 8EG, UK. 52North West Thames Regional Genetics Service, Kennedy-Galton Centre, Harrow, UK. 53Department of Clinical Genetics, Royal Devon & Exeter Hospital, Barnack Road, Exeter EX2 SDW, UK. 54Menesseyl and Cheshire Clinical Genetics Service, Liverpool Women’s NHS Foundation Trust, 7th floor, Borough Wing, Guy’s Hospital, Southwell Street, Bristol BS2 8EG, UK. 55Academic Medicine, Manchester Academic Health Science Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK. 56Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, Strangeways Research Laboratory, University of Cambridge, Worts Causeway, Cambridge CB1 8RN, UK. 57Faculty of Medicine, University of Southampton, Southampton University Hospitals NHS Trust, Maplepoint SU1 8DY, UK. 58Inserm U823, Université Joseph Fourier, 38041 Grenoble, France. 59General Medical Science, Department of Clinical Genetics, University Hospitals of Leicester NHS Trust, Leicester LE1 5WW, UK. 60Oxford Regional Genetics Service, Churchill Hospital, Old Road, Headington, Oxford OX3 7JE, UK. 61Clinical Genetics Service, Guy’s and St Thomas’ NHS Foundation Trust, 7th floor, Borough Wing, Guy’s Hospital, Great Maze Pond, London SE1 9RT, UK. 62North East Thames Regional Genetics Service, Great Ormond Street Hospital for Children NHS Trust, Barlays House, 37, Queen Square, London WC1N 3BH, UK. 63Academic Unit of Clinical and Molecular Oncology, Trinity College Dublin, College Green, Dublin 2, Ireland. 64Medical Oncology Service, St James’s Hospital, James’s Street, Dublin 8, Ireland. 65Department of Clinical Genetics, East Anglian Regional Genetics Service, Addenbrooke’s Hospital, Level 6, Addenbrooke’s Treatment Centre, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK. 66All Wales Medical Genetics Services, University Hospital of Wales, Heath Park, Cardiff CF14 4XW, UK. 67South East Scotland Regional Genetic Service, Western General Hospital, David Brock Building, Crewe Road South, Edinburgh EH4 2XU, UK. 68Centre for Cancer Research & Cell Biology, School of Medicine, Dentistry and Biomedical Sciences, Queen’s University Belfast, 97 Lisburn Road, Belfast BT9 7AE, UK. 69Department of Medical Genetics, Belfast Health and Social Care Trust, Belfast City Hospital, Lisburn Road, Belfast BT9 7AB, UK. 70Oncogenetics Team, The Institute of Cancer Research and Royal Marden NHS Foundation Trust, 123 Old Brompton Road, London SW7 3RP, UK. 71Ferguson-Smith Centre for Clinical Genetics, Yorkhill Hospitals, Block 4, Glasgow G3 8SJ, UK. 72South West Thames Regional Genetics Service, Department of Medical Genetics, St George’s University of London, Cranmer Terrace, London SW17 0UT, UK. 73West Midlands Regional Genetics Service, Birmingham Women’s Hospital Healthcare NHS Trust, Mindeslohn Way, Edgbaston, Birmingham B15 2TG, UK. 74Department of Pathology & Laboratory Medicine, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, KS 66160, USA. 75Lombardi Comprehensive Cancer Center, MedStar Georgetown University Hospital, 3800 Reservoir Road NW, Washington, DC 20057, USA. 76Center for Medical Genetics, Ghent University Hospital, De Pintelaan 185, 9000 Ghent, Belgium. 77Division of Tumor Genetics, Department of Gynaecology and Obstetrics, University Hospital Klinikum Rechts der Isar, Technische Universität München, Ismaninger Strasse 22, 81675 Munich, Germany. 78Center of Familial Breast and Ovarian Cancer, Department of Medical Genetics, Institut für Humangenetik, Boztenum, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany. 79Center for Hereditary Breast and Ovarian Cancer, Medical Faculty, Center for Integrated Oncology (CIO) Cancer Center Cologne, University Hospital Cologne, Cologne, Germany. 80Center for Molecular Medicine Cologne (CMM), University of Cologne, Robert-Koch-Straße 21, 50931 Cologne, Germany. 81Department of Human Genetics, Institute of Human Genetics, University Hospital Heidelberg, Heidelberg, Germany. 82Institute for Medical Informatics, Statistics and Epidemiology, Medical Faculty, University of Leipzig, Leipzig, Germany. 83Department of Gynaecology and Obstetrics, University Hospital Düsseldorf, Heinrich-Heine Universität Düsseldorf, Moorenstrasse 5, 40225 Düsseldorf, Germany. 84Institute of Cell and Molecular Pathology, Centre for Pathology and Forensic and Genetic Medicine, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany. 85Institute of Human Genetics, University Medical Center Schleswig-Holstein, Campus Kiel, Arnold-Heller-Straße 3, 24105 Kiel, Germany. 86Department of Medical Genetics and Human Genetics, Campus Virchow-Klinikum, Charité Berlin – Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany. 87German Consortium of Hereditary Breast and Ovarian Cancer (G-BCOC), Cologne, Germany. 88Institute of Human Genetics, University Hospital Münster, Vessenstrasse 12-14, 48149 Münster, Germany. 89Department of Molecular Biology and Genetics, University Medical Center Schleswig-Holstein, Campus Kiel, Arnold-Heller-Straße 3, 24105 Kiel, Germany. 90Department of Medical Genetics and Human Genetics, Campus Virchow-Klinikum, Charité Berlin – Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany. 91Unité de Prévention et d’Épidémiologie Génétique, Centre Léon Bérard, 28 rue Laennec, 69008 Lyon, France. 92Génétique Clinique, Centre Hospitalier Universitaire de Grenoble, CS 10217, 38043, Grenoble cedex 09, France. 93Institut Albert Bonniou – Inserm U823, Université Joseph Fourier, Rond-point de la Chaufferie, 38076 La Tronche, France. 94Laboratoire d’OncoGénétique, Hôpital René Huguenin, Institut Curie, 35 rue Daily, 92210 Saint-Cloud, France. 95Department d’Endocrinogénétique, Centre Jean Perrin, Université de Clermont-Ferrand, 58 rue Montalet, BP 392, 63011 Clermont-Ferrand, France. 96Gynaecological Oncology, Sydney Cancer Centre, Royal Prince Alfred Hospital and University of Sydney, Missenden Road, Camperdown, NSW 2050, Australia. 97Gynecologic Oncology Group, Department of OB-GYN, University of North Carolina at Chapel Hill, 103B Physicians’ Office Building, CB 7572, Chapel Hill, NC 27599-7572, USA. 98Gynecological Oncology Group Statistical and Data Center, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263-0001, USA. 99Women & Infants Hospital, 1 Blackstone Place, Providence, RI 02905, USA. 100Molecular Oncology Laboratory, Health Research Institute of the San Carlos Clinical Hospital (IDISCI), 28040 Madrid, Spain. 101Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Biomedical Helsinki, PO Box 700, 00202 Helsinki, Finland. 102Department of Clinical Genetics, Helsinki University Central Hospital, Biomedicus Helsinki 1, Haartmaninkatu 8, 00290 Helsinki, Finland. 103Department of Medical Oncology, Family Cancer Clinic, Erasmus University Medical Center, PO Box 2040, 3000 CA Rotterdam, the Netherlands. 104Department of Clinical Genetics, Family Cancer Clinic, Erasmus University Medical Center, Rotterdam, the Netherlands. 105Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands. 106Department of Clinical Genetics, Academic Medical Center, Amsterdam, the Netherlands. 107Department of Epidemiology, Netherlands Cancer Institute, Amsterdam, the Netherlands. 108Family Cancer Clinic, Netherlands Cancer Institute, Amsterdam, the Netherlands. 109Department of Clinical Genetics, VU University Medical Center Amsterdam, De Boelelaan 1118, 1081 HV Amsterdam, the Netherlands. 110Department of Genetics, University Medical Center, Groningen, the Netherlands. 111Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, the Netherlands. 112Department of Human Genetics, Center for Human and Clinical Genetics, Leiden University Medical Center, S4-P PO Box 9600, 2300 RC Leiden, the Netherlands. 113Regional Genetics Service, University of North Carolina at Chapel Hill, NC 27599-7572, USA. 114Department of Human Genetics, Hospital Duran i Reynals, 3a planta - Gran Via de l’Hospitalet, 199, 08905, Barcelona, Spain. 115Institut d’Investigació Biomèdica de Bellvitge (IDIBELL)-Catalan Institute of Oncology, Hospital Duran i Reynals, 3a planta - Gran Via d’l’Hospitalet, 199, 08908 Hospital de Llobregat, Barcelona, Spain. 116Molecular Diagnostic Unit, Hereditary Cancer Program, Institut d’Investigació Biomèdica de Bellvitge (IDIBELL)-Catalan Institute of Oncology, Hospital Duran i Reynals, 3a planta - Gran Via d’l’Hospitalet, 199, 08908 Hospital de Llobregat, Barcelona, Spain. 117Department of Genetics and Pathomorphology, Faculty of Medicine and Dentistry, Pomeranian Medical
References


