ELSEVIER

Cancer Genetics ■■ (2015) ■■–■■

Cancer Genetics

ORIGINAL ARTICLE

Protein-truncating variants in moderate-risk breast cancer susceptibility genes: A meta-analysis of high-risk case-control screening studies

Fatima Aloraifi [a,b,*], Damian McCartan [c], Trudi McDevitt [d], Andrew J. Green [a,d], Adrian Bracken [b,1,*], James Geraghty [a,c,1]

[a] *School of Medicine and Medical Science, University College Dublin, Belfield, Dublin, Ireland;* [b] *Smurfit Institute of Genetics, Trinity College Dublin, College Green, Dublin, Ireland;* [c] *St. Vincent's University Hospital, Elm Park, Dublin, Ireland;* [d] *National Centre for Medical Genetics, Our Lady's Hospital, Crumlin, Dublin, Ireland*

Several "moderate-risk breast cancer susceptibility genes" have been conclusively identified. Pathogenic mutations in these genes are thought to cause a two to fivefold increased risk of breast cancer. In light of the current development and use of multigene panel testing, the authors wanted to systematically obtain robust estimates of the cancer risk associated with loss-of-function mutations within these genes. An electronic search was conducted to identify studies that sequenced the full coding regions of *ATM*, *CHEK2*, *BRIP1*, *PALB2*, *NBS1*, and *RAD50* in a general and gene-targeted approach. Inclusion was restricted to studies that sequenced the germline DNA in both high-risk cases and geographically matched controls. A meta-analysis was then performed on protein-truncating variants (PTVs) identified in the studies for an association with breast cancer risk. A total of 10,209 publications were identified, of which 64 studies comprising a total of 25,418 cases and 52,322 controls in the 6 interrogated genes were eligible under our selection criteria. The pooled odds ratios for PTVs in the susceptibility genes were at least >2.6. Additionally, mutations in these genes have shown geographic and ethnic variation. This comprehensive study emphasizes the fact that caution should be taken when identifying certain genes as moderate susceptibility with the lack of sufficient data, especially with regard to the *NBS1*, *RAD50*, and *BRIP1* genes. Further data from case-control sequencing studies, and especially family studies, are warranted.

Keywords ATM, PALB2, BRIP1, CHEK2, hereditary breast cancer

Undoubtedly, the impact of breast cancer has been substantial to women, and it is the most common cancer-related mortality among women ages 20 – 59 worldwide. Familial clustering has been observed since the Roman times, and 10 – 15% of cases are thought to be caused by a hereditary predisposition. In the past two decades, we have witnessed a plethora of genetic studies that successfully expanded our understanding of the inherited predisposition to breast cancer. The discovery of highly penetrant variants in the *BRCA1* and *BRCA2* genes has contributed to increased surveillance, tailored therapy, and preventative surgery among carriers (1–3).

Received December 5, 2014; received in revised form June 2, 2015; accepted June 3, 2015.

* Corresponding authors.
E-mail addresses: faloraifi@rcsi.ie, andrew.green@olchc.ie
[1] Joint senior authors.

However, the prevalence of protein-truncating variants (PTVs) in high-risk individuals varies within these genes. In a study conducted on 46,276 women with early-onset or familial breast cancer of different ethnicities, the frequency of truncating mutations in *BRCA1* and *BRCA2* were 7.2% and 5.3%, respectively (4). However, when combining several studies from large cohorts, variants in the *BRCA1* and *BRCA2* genes can vary from 12.5–31.0% (4,5), depending on the studied population (see review (6)). Other high-penetrance variants in genes, such as *TP53*, *STK11*, *CDH1*, have been identified but usually present as part of extremely rare cancer syndromes. However, the vast majority of hereditary breast cancer is largely unaccounted for (7–9).

In contrast to the highly penetrant rare variants successfully identified through linkage studies, efforts have been made to identify moderate- and low-risk variants through different case-control approaches. With the notion that breast cancer

2210-7762/$ - see front matter © 2015 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cancergen.2015.06.001

is a common complex disease, almost 1,000 genome-wide association studies have concentrated on common alleles shared among individuals with and without breast cancer. By definition, however, these confer low effects (see reviews (10,11)). On the other hand, with the use of candidate gene screening, pathogenic variants in *ATM*, *CHEK2*, *BRIP1*, *PALB2*, *NBS1*, and *RAD50* were found to confer a two- to fivefold increased risk of developing breast cancer (see review (12)). The term "moderate-risk gene" has been used to categorize these genes and has caused much controversy, because of the varying magnitudes of mutation prevalence within these genes. In addition, the frequency of pathogenic mutations in these genes reportedly varies significantly among different populations, as exemplified by the frequently studied founder mutation *CHEK2* 1100delC (13).

One of the most accepted models for the etiology of hereditary breast cancer involves a combination of a number of common low-effect alleles (i.e., "common disease, common alleles") or a combination of one or more rare large-effect alleles (i.e., "common disease, rare alleles") (14). Genes that are termed "moderate-risk susceptibility" are typically encountered in either a monogenic or polygenic manner; therefore, evidence supporting their pathogenicity remains obscure (12). The frequency of these mutations is relatively rare (<5%). Generally, candidate gene studies are comprised of smaller sample sizes than those used in genome-wide association studies. Elucidation of the true effects of mutations in these genes, with better statistical power, can be achieved by meta-analysis through the examination of published epidemiological evidence.

Although some meta-analysis studies have focused on single variants, or a combination of variants from the plethora of genome-wide associations that have low effects, we decided to concentrate on PTVs in genes that have been conclusively identified as the moderate susceptibility genes in breast cancer: *ATM* (15), *CHEK2* (13,16), *BRIP1* (17), *PALB2* (18), *NBS1* (19), and *RAD50* (20). Variants resulting in premature protein truncation (i.e., nonsense mutations, deletions/insertions that result in translational frameshift, and mutations that affect splice sites) are the most recognized disease-associated mutations (21). Studies involving sequence analysis of these genes have been published for different populations. By approaching the problem of unresolved heritability in breast cancer, we focused on studies with the highest likelihood of heritable genetic involvement. Thus, we selected studies where patient cohorts involved familial, bilateral, and early-onset breast cancer.

Our study elucidates population differences in those genes and summarizes their effects in terms of clinical utility with a view to development of multigene, next generation sequencing panels. Deep sequencing of whole genome and whole exome are still substantially costly and statistically challenging for large sample size, case-control studies; therefore, a reasonable approach for now would be to pool the data we have to date.

Materials and methods

Study strategy and approach

All methods for this meta-analysis follow the guidelines proposed by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (22). Data was extracted by one reviewer (F.O.) and a second independent reviewer (D.M.).

To identify all the relevant publications, we used a two-stage strategy (see Figure 1). The first stage was a general approach in which we queried terms (inherited OR familial OR hereditary OR family OR young OR "early onset" OR bilateral) AND (mutation OR variant OR truncating OR nonsense OR missense OR gene OR genetic) AND ("breast cancer" OR "breast neoplasm"). The second stage was a gene-targeted approach, where we queried ("breast cancer" OR "breast neoplasm") AND (inherited OR familial OR hereditary OR family OR young OR "early onset" OR bilateral) AND (*ATM* OR *BRIP1* OR *FANCJ* OR *BACH1* OR *CHEK2* OR *CHK2* OR *NBS1* OR *NBN* OR *PALB2* OR *FANCN* OR *RAD50* OR *NBSLD*) to include the gene name (OR alternative gene names) (see Figure 1). Finally, the references of all studies included were scanned, as were reference lists from relevant reviews and meta-analyses. This exhaustive search was conducted by computer-based searches on PubMed and PubMed Central (PMC) on or before June 1, 2014, without language restriction.

Interrogated genes and variants

We focused our search on PTVs because this class of mutations is strongly associated with disease, especially breast cancer (13,15,17,18). These include mutations that introduce a stop codon (nonsense), cause a frameshift in the open reading frame, and occur within splice sites. We also restricted our search to mutations that only occur in the germline (as identified in DNA extracted from peripheral blood leucocytes), because those mutations are the ones that can be tested with a simple blood test.

Although the high-risk genes *BRCA1* and *BRCA2* have been widely studied and implemented in clinical genetic testing, we aimed to investigate mutations only within the genes that are currently termed as moderate breast cancer–predisposing genes described previously (12). The five genes investigated have been convincingly identified as moderate-susceptibility genes (12). They include *ATM*, *BRIP1* (*FANCJ*/*BACH1*), *CHEK2*, *NBS1* (*NBN*), *PALB2* (*FANCN*), and *RAD50*.

Study population

It is widely accepted that in order to improve the power in selection of predisposing variants, analyzing mutations in high-risk individuals (instead of unselected individuals) is substantially more efficient (23). The relative risk of breast cancer can be greater than fourfold if a family history exists; a similar increased relative risk is also seen if cases include bilateral breast cancer (21). Another important factor that increases the likelihood of a predisposing genetic mutation is early age of onset. We reasoned that we could enrich for widely recognized parameters for genetic predisposition by specifically selecting studies that sampled cases with a family history, bilateral breast cancer, and/or early age of onset of the disease.

We restricted our summary statistics to exclude studies in which genetic screening was conducted in the absence of healthy individuals from the population screened.