PALB2, CHEK2 and ATM rare variants and cancer risk: data from COGS

Melissa C Southey,1 David E Goldgar,2 Robert Winqvist,3 Katri Pylkäs,3 Fergus Couch,4 Marc Tischkowitz,5 William D Foulkes,6 Joe Dennis,7 Kyriaki Michailidou,7 Elizabeth van Rensburg,8 Tuomas Heikkinen,9 Heli Nevanlinna,9 John L Hopper,10 Thilo Dörk,11 Kathleen BM Claes,12 Jorge Reis-Filho,13 Zhi Ling Teo,1 Paolo Radice,14 Irene Catucci,15 Paolo Peterlongo,15 Helen Tsimiklis,1 Fabrice A Odefrey,1 James G Downty,10 Marjanka K Schmidt,16 Annegien Broeks,16 Frans B Hogervorst,16 Senno Verhoef,16 Jane Carpenter,17 Christine Clarke,18 Rodney J Scott,19 Peter A Fasching,20,21 Thilo Dörk,11 Kathleen BM Claes,12 Fergus Couch,4 Marc Tischkowitz,5 William D Foulkes,6 Joe Dennis,7

For numbered affiliations see end of article.

Correspondence to Professor Melissa C. Southey, Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Melbourne, Victoria 3010, Australia; msouthey@unimelb.edu.au

Received 29 March 2016
Revised 1 June 2016
Accepted 21 June 2016

To cite: Southey MC, Goldgar DE, Winqvist R, et al. J Med Genet Published Online First, published on September 5, 2016 as 10.1136/jmedgenet-2016-103839

Copyright Article author (or their employer) 2016. Produced by BMJ Publishing Group Ltd under licence.
**ABSTRACT**

**Background** The rarity of mutations in **PALB2**, **CHEK2** and **ATM** makes it difficult to estimate precisely associated cancer risks. Population-based family studies have provided evidence that at least some of these mutations are associated with breast cancer risk as high as those associated with rare **BRCA2** mutations. We aimed to estimate the relative risks associated with specific rare variants in **PALB2**, **CHEK2** and **ATM** via a multicentre case-control study.

**Methods** We genotyped 10 rare mutations using the custom iCOGS array: **PALB2** c.1592delT, c.2816T>G and c.3113G>A, **CHEK2** c.1343T>G, c.349A>G, c.7271T>G, c.1312G>T, and c.1036C>T and c.13437>G in **ATM**. We assessed associations with breast cancer risk (42,671 cases and 42,164 controls), as well as prostate (22,301 cases and 22,320 controls) and ovarian (14,542 cases and 14,507 controls) cancer risk, for each variant.

**Results** For European women, strong evidence of association with breast cancer risk was observed for **PALB2** c.1592delT OR 3.44 (95% CI 1.84 to 9.60, p=6.9×10^{-12}) and **CHEK2** c.1343T>G OR 3.03 (95% CI 1.89 to 4.77, p=5.1×10^{-9}) for European men and **CHEK2** c.1343T>G OR 11.0 (95% CI 5.0 to 25.0, p=7.0×10^{-11}) for African men. Evidence for prostate cancer risk was observed for **PALB2** c.1592delT OR 1.84 (95% CI 1.39 to 2.43, p=7.1×10^{-5}) and **ATM** c.7271T>G OR 11.0 (95% CI 5.0 to 25.0, p=7.0×10^{-11}) and **ATM** c.7271T>G OR 11.0 (95% CI 5.0 to 25.0, p=7.0×10^{-11}) for African men. Evidence for ovarian cancer risk was observed for **PALB2** c.2816T>G OR 3.03 (95% CI 1.89 to 4.77, p=5.1×10^{-9}) for European women and **CHEK2** c.1343T>G OR 3.03 (95% CI 1.89 to 4.77, p=5.1×10^{-9}) for African men. Evidence for ovarian cancer risk was observed for **PALB2** c.2816T>G OR 3.03 (95% CI 1.89 to 4.77, p=5.1×10^{-9}) for European women and **CHEK2** c.1343T>G OR 3.03 (95% CI 1.89 to 4.77, p=5.1×10^{-9}) for African men.
men. No evidence of association with ovarian cancer was found for any of these variants.

**Conclusions** This report adds to accumulating evidence that at least some variants in these genes are associated with an increased risk of breast cancer that is clinically important.

**INTRODUCTION**

The rapid introduction of massive parallel sequencing (MPS) into clinical genetics services is enabling the screening of multiple breast cancer susceptibility genes in one assay at reduced cost for women who are at increased risk of breast (and other) cancer. These gene panels now typically include the so-called ‘moderate-risk’ breast cancer susceptibility genes, including *PALB2*, *CHEK2* and *ATM*. However, mutations in these genes are individually extremely rare and limited data are available with which to accurately estimate the risk of cancer associated with them.

Estimation of the age-specific cumulative risk (penetrance) of breast cancer associated with specific mutations in these three genes has been limited to those that have been observed more frequently, such as *PALB2* c.1592delT (a Finnish founder mutation), *PALB2* c.3113G>A and *ATM* c.7271T>G. These mutations have been estimated to be associated with a 40% (95% CI 17% to 77%), 91% (95% CI 44% to 100%) and 52% (95% CI 28% to 80%) cumulative risk of breast cancer to the age of 70 years, respectively. These findings, based on segregation analyses in families of population-based case series, indicate that at least some mutations in these ‘moderate-risk’ genes are associated with a breast cancer risk comparable to that of the average pathogenic mutation in *BRCA2*: 45% (95% CI 31% to 56%). However, such estimates are imprecise and, moreover, may be confounded by modifying genetic variants or other familial risk factors.

Case-control studies provide an alternative approach to estimating cancer risks associated with specific variants. This design can estimate the relative risk directly, without making assumptions about the modifying effects of other risk factors. However, because these variants are rare, such studies need to be extremely large to provide precise estimates.

The clearest evidence for association, and the most precise breast cancer risk estimates, for rare variants in *PALB2*, *CHEK2* and *ATM* relate to protein truncating and splice-junction variants. However, studies based on mutation screening in case-control studies, combined with stratification of variants by their evolutionary likelihood suggest that at least some evolutionarily unlikely missense substitutions are associated with a similar risk to those conferred by truncating mutations. For example, Tavtigan et al. estimated an OR of 2.85 (95% CI 0.83 to 8.61) for evolutionarily unlikely missense substitutions in the 3′ third of *ATM*, which is comparable to that for truncating variants. Specifically, *ATM* c.7271C>G has been associated with a high breast cancer risk in several studies. Le Calvez-Kelm et al. estimated that the ORs associated with rare mutations in *CHEK2* from similarly designed studies were 6.18 (95% CI 1.76 to 21.8) for rare protein-truncating and splice-junction variants and 8.75 (95% CI 1.06 to 72.2) for evolutionarily unlikely missense substitutions.

It is plausible that monoallelic mutations in *PALB2*, *CHEK2* and *ATM* could be associated with increased risks of cancers other than breast cancer, as has been observed for *BRCA1* and *BRCA2* and both ovarian and prostate cancers. However, with the exception of pancreatic cancer in *PALB2* carriers, there is little evidence to support or refute the existence of such associations, although a few individually striking pedigrees have been observed.

In this study we selected rare genetic variants on the basis that they had been observed in breast cancer candidate gene case-control screening projects involving *PALB2*, *CHEK2* or *ATM*. These included three rare variants in *PALB2*: the protein truncating variants c.1592delT (p.Leu531Cysfs) and c.3113 G>A (p.Trp1038*) and the missense variant c.2816T>G. (p. Leu939Trp), six rare missense variants in *CHEK2*: c.349A>G (p.Arg117Gly) and c.1036C>T (p.Arg346Cys) predicted to be deleterious on the basis of evolutionary conservation, c.538C>T (p.Arg180Cys), c.715G>A (p.Glu239Lys), c.1312G>T (p.Asp437Tyr) and c.1343T>G (p.Ile448Ser) and *ATM* c.7271T>G (p.Val2424Gly). We assessed the association of these variants with breast, ovarian and prostate risk by case-control analyses in three large consortia participating in the Collaborative Oncological Gene-environment Study.

**METHODS**

**Participants**

Participants were drawn from studies participating in three consortia as follows:

The Breast Cancer Association Consortium (BCAC), involving a total of 48 studies: 37 of women from populations with predominantly European ancestry (42,671 cases and 42,164 controls), 9 of Asian women (5795 cases and 6624 controls) and 2 of African-American women (1046 cases and 932 controls). All cases had invasive breast cancer. The majority of studies were population-based or hospital-based case-control studies, but some studies of European women oversampled cases with a family history or with bilateral disease (see online supplementary table S1). Overall, 79% of BCAC cases with known Estrogen Receptor (ER) status (23% missing) are ER-positive. The proportion of cases selected by family history that are ER-positive is 78% (38% missing).

The Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) involving a total of 26 studies: 25 included men with European ancestry (22,301 cases and 22,320 controls) and 3 included African-American men (623 cases and 569 controls). The majority of studies were population-based or hospital-based case-control studies (see online supplementary table S2).

The Ovarian Cancer Association Consortium (OCAC), involving a total of 46 studies. Some studies were case-only and their data were combined with case-control studies from the same geographical region (leaving 36 study groupings). Of these groupings, 33 included women from populations with predominantly European ancestry (16,287 cases (14,542 with invasive disease) and 23,491 controls), 25 included Asian women (813 cases (720 with invasive disease) and 1574 controls), 17 included African-American women (186 cases (150 with invasive disease) and 200 controls) and 29 included women of other ethnic origin (893 cases (709 with invasive disease) and 864 controls). The majority of studies were population-based or hospital-based case-control studies (see online supplementary table S3).

Details regarding sample quality control have been published previously. All study participants gave informed consent and all studies were approved by the corresponding local ethics committees (see online supplementary tables S1–S3).

**Variant selection**

We selected for genotyping 13 rare mutations that had been observed in population-based case-control mutation screening studies. These variants were *PALB2* (c.1592delT, p.

Genotyping
Three PALB2 variants c.2323C>T (p.Gln775*), c.3116delA (p.Asn1039Ilefs) and c.3549C>G (p.Tyr1183*) were unable to be designed for measurement on the custom Illumina iSelect genotyping array and were not considered further (table 1). Genotyping was conducted using a custom Illumina Infinium array (iCOGS) in four centres, as part of a multicentred collaboration—details provided elsewhere.22Genotypes were called using Illumina’s proprietary GenCall algorithm and then, for the data generated from the rare variant probes, manually confirmed with reference to the positive control sample. Two per cent of samples were provided in duplicate by all studies and 270 HapMap2 samples were genotyped in all four genotyping centres. Subjects with an overall call rate <95% were excluded. Plates with call rates <90% were excluded on a variant-by-variant basis. Cluster plots generated for all of the 10 rare variants were manually checked to confirm automated calls (see online supplementary figure S1).

Statistical methods
The association of each variant with breast, prostate and ovarian cancer risk was assessed using unconditional logistic regression, with reference to the positive control sample. Two per cent of samples were provided in duplicate by all studies and 270 HapMap2 samples were genotyped in all four genotyping centres. Subjects with an overall call rate <95% were excluded. Plates with call rates <90% were excluded on a variant-by-variant basis. Cluster plots generated for all of the 10 rare variants were manually checked to confirm automated calls (see online supplementary figure S1).

Table 1  Rare genetic variants included in the iCOGS array.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant*</th>
<th>Amino acid*</th>
<th>dbSNP rs</th>
<th>OR (95% CI)</th>
<th>Penetrance† (95% CI)</th>
<th>Align-GVGD</th>
<th>Reference(s)</th>
<th>Designed†</th>
<th>Genotyped</th>
</tr>
</thead>
<tbody>
<tr>
<td>PALB2</td>
<td>c.1592delT</td>
<td>p.Leu531Cysfs</td>
<td>rs180177110</td>
<td>3.94 (1.5–12.1)$</td>
<td>40% (17–77)</td>
<td>na</td>
<td>4, 5, 10</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>c.2323C&gt;T</td>
<td>p.Gln775*</td>
<td>rs18017711</td>
<td>na</td>
<td>25, 26, No</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.2816T&gt;G</td>
<td>p.Tryr1183*</td>
<td>rs45478192</td>
<td>C55</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.3113G&gt;A</td>
<td>p.Tryr1183*</td>
<td>rs180177132</td>
<td>na</td>
<td>2, 6, 20, Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.3116delA</td>
<td>p.Asn1039Ilefs</td>
<td>rs180177133</td>
<td>na</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.3549C&gt;G</td>
<td>p.Tyr1183*</td>
<td>rs11820399</td>
<td>na</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHEK2</td>
<td>c.349A&gt;G</td>
<td>p.Arg117Gly</td>
<td>rs28909982</td>
<td>8.75 (1.06–72.2)$</td>
<td>C65</td>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.538C&gt;T</td>
<td>p.Arg180Cys</td>
<td>rs77130927</td>
<td>2.47 (0.45–13.49)$</td>
<td>C5</td>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.715G&gt;A</td>
<td>p.Glu239Lys</td>
<td>rs12190870</td>
<td>1.82 (0.62–5.34)$</td>
<td>C15</td>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.1036C&gt;T</td>
<td>p.Arg346Cys</td>
<td>rs28909982</td>
<td>8.75 (1.06–72.2)$</td>
<td>C65</td>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.1312G&gt;T</td>
<td>p.Asp438Tyr</td>
<td>rs28909982</td>
<td>2.47 (0.45–13.49)$</td>
<td>C5</td>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.1343T&gt;G</td>
<td>p.Ilele448Ser</td>
<td>rs17886163</td>
<td>1.82 (0.62–5.34)$</td>
<td>C15</td>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ATM</td>
<td>c.7271T&gt;G</td>
<td>p.Val2424Gly</td>
<td>rs28909982</td>
<td>52% (28–80)</td>
<td>C65</td>
<td>7, 13, 23, 27</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

HGVSN Variation Society (HGVSN); reference sequences PALB2, NM_024675.3, NP_078951.2; CHEK2, NM_007194.3, NP_009125.1; ATM, NM_000051.3, NP_000042.3.
†Age-specific cumulative risk of breast cancer to age 70 years.5
‡Penetrance estimated in a combined group of 35 CHEK2 variants.11
§Breast cancer cases unselected for family history of breast cancer.2
$OR estimated in a combined group of 65 CHEK2 variants.2122
**OR estimated in a combined group of 25 CHEK2 variants.11
††OR estimated in a combined group of 15 CHEK2 variants.11
†††na, not available.

RESULTS
PALB2
In BCAC, PALB2 c.1592delT (Leu531Cysfs) was only observed in 35 cases and 6 controls, all from four studies from Sweden and Finland (Helsinki Breast Cancer Study (HEBCS), Kuopio Breast Cancer Project (KBPC), Oulu Breast Cancer Study (OBSC) and Karolinska Mammography Project for Risk Prediction Breast Cancer (pKARMA); see online supplementary figure S1).
Table 2  Summary results from Breast Cancer Association Consortium studies of white Europeans (42 671 invasive breast cancer cases and 42 164 controls)

<table>
<thead>
<tr>
<th>Variant</th>
<th>Frequency* Controls</th>
<th>Frequency* Cases</th>
<th>OR (95% CI)</th>
<th>LRT p Value</th>
<th>OR† (95% CI)</th>
<th>LRT p Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>PALB2‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1592deT (p.Leu531Cysfs)</td>
<td>0.00018</td>
<td>0.00031</td>
<td>2.06 (0.59 to 7.11)</td>
<td>0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.2816T&gt;G (p.Leu939Trp)</td>
<td>0.00354</td>
<td>0.00381</td>
<td>0.95 (0.69 to 1.29)</td>
<td>0.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.3113G&gt;A (p.Trp1038*)</td>
<td>0.00045</td>
<td>0.00027</td>
<td>0.49 (0.18 to 1.36)</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHEK2†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.349A&gt;G (p.Arg117Gly)</td>
<td>0.00063</td>
<td>0.00081</td>
<td>1.46 (0.71 to 3.02)</td>
<td>0.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.538C&gt;T (p.Arg180Cys)</td>
<td>0.00341</td>
<td>0.00296</td>
<td>1.02 (0.73 to 1.44)</td>
<td>0.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.715G&gt;A (p.Glu239Lys)</td>
<td>0.00021</td>
<td>0.00027</td>
<td>1.47 (0.41 to 5.35)</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1036C&gt;T (p.Arg348Cys)</td>
<td>0.00005</td>
<td>0.00022</td>
<td>1.07 (0.28 to 4.07)</td>
<td>0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1312G&gt;T (p.Asp438Tyr)</td>
<td>0.00078</td>
<td>0.00082</td>
<td>1.03 (0.62 to 1.71)</td>
<td>0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1343T&gt;G (p.Ile448Ser)‡</td>
<td>0.00002</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATM†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.7271T&gt;G (p.Val2424Gly)</td>
<td>0.00002</td>
<td>0.00028</td>
<td>11.6 (1.50 to 89.9)</td>
<td>0.0012</td>
<td>11.0 (1.42 to 85.7)</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

*Proportion of subjects carrying the variant.
†Excluding women from five studies that selected all cases based on family history or bilateral disease and the subset of selected cases from other studies (based on 34 488 unselected cases and 34 059 controls).
‡CHEK2 c.1343T>G (p.Ile448Ser) was only observed in one control and no cases of white European origin.
§PALB2 c.3113G>A (p.Trp1038*) only observed in Finland and Sweden.
LRT, likelihood ratio test; OR, OR for carriers of the variant versus common-allele homozygotes, adjusted for study and seven principal components.

Table 3  Summary results from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome studies for white European men* (22 301 prostate cancer cases and 22 320 controls)

<table>
<thead>
<tr>
<th>Variant</th>
<th>Frequency† Controls</th>
<th>Frequency† Cases</th>
<th>OR (95% CI)</th>
<th>LRT p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PALB2‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1592deT (p.Leu531Cysfs)</td>
<td>0.00018</td>
<td>0.00031</td>
<td>2.06 (0.59 to 7.11)</td>
<td>0.24</td>
</tr>
<tr>
<td>c.2816T&gt;G (p.Leu939Trp)</td>
<td>0.00354</td>
<td>0.00381</td>
<td>0.95 (0.69 to 1.29)</td>
<td>0.73</td>
</tr>
<tr>
<td>c.3113G&gt;A (p.Trp1038*)</td>
<td>0.00045</td>
<td>0.00027</td>
<td>0.49 (0.18 to 1.36)</td>
<td>0.16</td>
</tr>
<tr>
<td>CHEK2†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.349A&gt;G (p.Arg117Gly)</td>
<td>0.00063</td>
<td>0.00081</td>
<td>1.46 (0.71 to 3.02)</td>
<td>0.30</td>
</tr>
<tr>
<td>c.538C&gt;T (p.Arg180Cys)</td>
<td>0.00341</td>
<td>0.00296</td>
<td>1.02 (0.73 to 1.44)</td>
<td>0.90</td>
</tr>
<tr>
<td>c.715G&gt;A (p.Glu239Lys)</td>
<td>0.00021</td>
<td>0.00027</td>
<td>1.47 (0.41 to 5.35)</td>
<td>0.55</td>
</tr>
<tr>
<td>c.1036C&gt;T (p.Arg348Cys)</td>
<td>0.00005</td>
<td>0.00022</td>
<td>1.07 (0.28 to 4.07)</td>
<td>0.93</td>
</tr>
<tr>
<td>c.1312G&gt;T (p.Asp438Tyr)</td>
<td>0.00078</td>
<td>0.00082</td>
<td>1.03 (0.62 to 1.71)</td>
<td>0.91</td>
</tr>
<tr>
<td>c.1343T&gt;G (p.Ile448Ser)‡</td>
<td>0.00002</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATM†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.7271T&gt;G (p.Val2424Gly)</td>
<td>0.00002</td>
<td>0.00028</td>
<td>11.6 (1.50 to 89.9)</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

*For white European men, unless otherwise indicated.
†Proportion of subjects carrying the variant.
‡CHEK2 c.1343T>G (p.Ile448Ser) was the only CHEK2 variant observed in African men and was identified in two cases and no controls of white European origin.
§Based on data from 623 and 569 African-American cases and controls, respectively.
LRT, likelihood ratio test; OR, OR for carriers of the variant versus common-allele homozygotes, adjusted for study and seven principal components.

Table S1), giving strong evidence of association with breast cancer risk (p=7.1×10^{-8}); the OR estimate was 4.52 (95% CI 1.90 to 10.8) based on all studies and 3.44 (95% CI 1.39 to 8.52) based on unselected cases and controls (table 2). We also found evidence of heterogeneity by ER status (p=0.0023), the association being stronger for ER-negative disease (OR 6.49 (95% CI 2.17 to 19.4) versus 2.24 (95% CI 1.05 to 7.24) for ER-positive disease).

PALB2 c.3113G>A (p.Trp1038*) was identified in 44 cases and 8 controls from nine BCAC studies. Only one carrier of the variant was of non-European origin. Strong evidence of association with breast cancer risk was observed (p=6.9×10^{-8}), with an estimated OR of 5.93 (95% CI 2.77 to 12.7) based on all studies and 4.21 (95% CI 1.85 to 9.61) based on unselected cases and controls. There was no evidence of a differential association by ER status (p=0.15).

Based on unselected cases, the estimated OR associated with carrying either of these PALB2 variants (c.1592deT or c.3113G>A) was 3.85 (95% CI 2.09 to 7.09).

PALB2 c.2816T>G (p.Leu939Trp) was identified in 150 cases and 145 controls and there was no evidence of association with risk of breast cancer. There was no evidence of association with risk of prostate or ovarian cancer for any of the three PALB2 variants (see tables 3 and 4).
CHEK2

CHEK2 c.349A>G (p.Arg117Gly) was identified in 44 cases and 18 controls in studies participating in BCAC; all of these women were of European origin. We found evidence of association with breast cancer (p=0.003), with little change in the OR after excluding selected cases (OR 2.03 (95% CI 1.10 to 3.73)).

CHEK2 c.538C>T (p.Arg180Cys) was identified in 158 breast cancer cases and 142 controls in studies of white Europeans. Evidence of association with breast cancer risk (p=0.016) was observed, with an unbiased OR estimate of 1.34 (95% CI 1.06 to 1.70). A consistent OR estimate was observed for Asian women, based on 45 case and 45 control carriers (OR 1.16 (95% CI 0.75 to 1.76)).

CHEK2 c.715G>A (p.Glu239Lys) mutations were identified in 15 cases and 9 controls, all African women participating in BCAC and no evidence of association with risk of breast cancer was observed (p=0.21).

CHEK2 c.1036C>T (p.Arg346Cys) was identified in nine cases from seven studies and two controls from two different studies in BCAC (neither control carrier was from a study that had case carriers), all of European origin. We found evidence of association with breast cancer risk (p=0.017) with reduced OR estimate of 3.39 (95% CI 0.68 to 16.9) after excluding selected cases.

None of the above four CHEK2 variants (CHEK2 c.349A>G (p.Arg117Gly); c.538C>T (p.Arg180Cys); c.715G>A (p.Glu239Lys) and c.1036C>T (p.Arg346Cys)) were found to be associated with an increased risk of prostate or ovarian cancer (tables 3 and 4). CHEK2 variant c.1312G>T (p.Asp438Tyr) was not associated with risk of breast cancer for European women (p=0.91). Variant c.1343T>G (p.Ile448Ser) was not observed in any breast cancer cases of European or Asian origin. It was detected in 48 cases and 29 controls of African origin, giving weak evidence of association (OR 1.52 (95% CI 0.95 to 2.43, p=0.083)). CHEK2 c.1312G>T (p.Asp438Tyr) was identified in 23 cases and 11 controls from PRACTICAL, all European, providing evidence of association with prostate cancer risk (OR 2.21 (95% CI 1.06 to 4.63, p=0.030)). CHEK2 c.1343T>G (p.Ile448Ser) was observed in 35 cases and 11 controls, all African, participating in PRACTICAL and was also associated with an increased risk of prostate cancer (OR 3.03 (95% CI 1.53 to 6.03, p=0.00059)). There was no evidence that these CHEK2 variants were associated with risk of ovarian cancer (table 4).

ATM

ATM c.7271T>G (p.Val2424Gly) was identified in 12 cases and 1 control in studies participating in BCAC, all of European origin, giving evidence of association with breast cancer risk (p=0.0012). The OR estimate based on unselected studies was 11.0 (95% CI 1.42 to 85.7). There was no evidence of association of this variant with prostate or ovarian cancer risk (see tables 3 and 4).

DISCUSSION

The present report adds to an accumulating body of evidence that at least some rare variants in so-called ‘moderate-risk’ genes are associated with an increased risk of breast cancer that is of clinical relevance.

These findings are presented at a time when detailed information about variants in these genes is becoming more readily available via the translation of diagnostic genetic testing from Sanger sequencing-based testing platforms to MPS platforms that test panels of genes in single assays. The vast majority of information about PALB2, CHEK2 and ATM, variants generated from these new testing platforms is not being used in clinical genetics services due to lack of reliable estimates of the cancer risk associated with individual variants, or groups of variants, in each gene. Previous analyses have been largely based on selected families, relying on data on the segregation of the variant. The present study is by far the largest to take a case-control approach. Consistent with previous reports, PALB2 c.3113G>A (p.Trp1038*) and ATM c.7271T>G (p.Val2424Gly) were found to be associated with substantially increased risk of breast cancer all with associated relative risk estimates of 3.44 or greater.

The estimates for the two loss-of-function PALB2 variants (c.1592delT and c.3113G>A) were consistent with each other and with estimates based on segregation analysis. We found no evidence of association with breast cancer for PALB2 c.2816T>G (p.Leu939Trp), with an upper 95% confidence limit excluding an OR >1.5 which is notable given the

Table 4 Summary results from the Ovarian Cancer Association Consortium studies for white European women (14 542 invasive ovarian cancer cases and 23 491 controls)

<table>
<thead>
<tr>
<th>Variant</th>
<th>Frequency* Controls</th>
<th>Frequency* Cases</th>
<th>OR (95% CI)</th>
<th>LRT p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PALB2</td>
<td>c.1592delT (p.Leu531Cysfs)</td>
<td>0.00004</td>
<td>0.00012</td>
<td>2.50 (0.21 to 29.1)</td>
</tr>
<tr>
<td></td>
<td>c.2816T&gt;G (p.Leu939Trp)</td>
<td>0.00413</td>
<td>0.00399</td>
<td>0.96 (0.69 to 1.34)</td>
</tr>
<tr>
<td></td>
<td>c.3113G&gt;A (p.Trp1038*)</td>
<td>0.00034</td>
<td>0.00031</td>
<td>1.34 (0.36 to 4.97)</td>
</tr>
<tr>
<td>CHEK2</td>
<td>c.349A&gt;G (p.Arg117Gly)</td>
<td>0.00038</td>
<td>0.00031</td>
<td>1.07 (0.32 to 3.60)</td>
</tr>
<tr>
<td></td>
<td>c.538C&gt;T (p.Arg180Cys)</td>
<td>0.00128</td>
<td>0.00160</td>
<td>1.49 (0.83 to 2.67)</td>
</tr>
<tr>
<td></td>
<td>c.715G&gt;A (p.Glu239Lys)</td>
<td>0.00021</td>
<td>0.00037</td>
<td>1.47 (0.42 to 5.22)</td>
</tr>
<tr>
<td></td>
<td>c.1036C&gt;T (p.Arg346Cys)</td>
<td>0</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>c.1312G&gt;T (p.Asp438Tyr)</td>
<td>0.00081</td>
<td>0.00074</td>
<td>0.92 (0.42 to 1.99)</td>
</tr>
<tr>
<td></td>
<td>c.1343T&gt;G (p.Ile448Ser)</td>
<td>0.00009</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>ATM</td>
<td>c.7271T&gt;G (p.Val2424Gly)</td>
<td>0</td>
<td>0.00012</td>
<td>–</td>
</tr>
</tbody>
</table>

*Proportion of subjects carrying the variant.
†c.1036C>T (p.Arg346Cys) was not observed in any sample.
LRT, likelihood ratio test; OR, OR for carriers of the variant versus common-allele homozygotes, adjusted for study and seven principal components.
Align-Grantham Variation Granthan Deviation (Align-GVGD) score and the observed impact on protein function. The estimate for ATM c.7271T>G (p.Val2424Gly) was also consistent with that found by segregation analysis. The substantial increased risk of breast cancer associated with ATM c.7271T>G (p.Val2424Gly) could be due to the reduction in kinase activity (with near-normal protein levels) observed for ATM p.Val2424Gly, thus this variant is likely to be acting as a dominant negative mutation.

In contrast, we found no evidence of an association with risk of prostate or ovarian cancer with any of these three variants: however, the confidence limits were wide; based on the upper 95% confidence limit we could exclude an OR of >1.4 for prostate cancer for the loss-of-function PALB2 c.3113G>A and 1.9 for c.1592delT and c.3113G>A combined. We analysed six rare missense variants in CHEK2. Two of these (CHEK2 c.349A>G (p.Arg117Gly; rs28909982) and c.1036C>T (p.Arg346Cys)) had evidence of a significant impact on the protein based on in silico prediction. We proposed these variants for inclusion in the iCOGS design as they had been identified in 3/1242 cases and 1/1089 controls and 3/1242 cases and 0/1089 controls, respectively, in a population-based case-control mutation screening study of CHEK2. In that study, Le Calvez-Kelm et al., estimated an OR of 8.75 (95% CI 1.06 to 72.2) for variants with an Align-GVGD score C65 (based on nine cases and one control). The current analysis provides confirmatory evidence of this association in a much larger sample (OR 2.18 (95% CI 1.23 to 3.85)) including 40 unselected case and 18 control carriers. The evidence that CHEK2 is a breast cancer susceptibility gene is largely based on studies of protein truncating variants, in particular CHEK2 1100delC. Reports of the association of the missense variant I157T (C15) and breast cancer risk have been conflicting but a large meta-analysis involving 15,985 breast cancer cases and 16,699 controls estimated a modest OR of 1.58 (95% CI 1.42 to 1.75). We also found evidence (p=0.015) of an association for c.538C>T (Align-GVGD C25); OR 1.34 (95% CI 1.06 to 1.70), a risk comparable to I157T.

The p values reported above have not been adjusted for multiple testing. This was not considered appropriate for the associations with breast cancer risk of PALB2 c.1592delT, c.3113G>A and ATM c.7271T>G because these associations had previously been reported; our aim was to more precisely estimate the associated relative risks. All three associations with breast cancer risk reported for CHEK2 variants remained statistically significant after adjusting for the other tests conducted in relation to breast cancer risk, but not after correcting for all tests for all cancers. Nevertheless, the findings for CHEK2 c.349A>G and c.1036C>T confirmed those reported previously, although collectively the association observed with CHEK2 c.538C>T requires independent replication.

Do this approach and new data have an impact on clinical recommendations for women and families carrying these rare genetic variants? Although age-specific cumulative risks for cancer are more informative for genetic counselling and clinical management of carriers, our study provides information that is relevant to clinical recommendations. As discussed in Easton et al., a relative risk of 4 will place a woman in a ‘high-risk’ category (in the absence of any other risk factor) and a relative risk between 2 and 4 will place a woman in this category if other risk factors are present. Thus, several of the variants included in this report (PALB2 c.1592delT; c.3113G>A and ATM c.7271T>G) would place the carrier in a high-risk group, especially if other risk factors, such as a family history, are present. The high level of breast cancer risk associated with PALB2 c.1592delT and c.3113G>A reported here is consistent with the penetrance estimate reported for a group of loss-of-function mutations in PALB2 and has an advantage in terms of clinical utility that the estimates in this study have been made at a mutation-specific level. Therefore, this work provides important information for risk reduction recommendations (such as prophylactic mastectomy and potentially salpingo-oophorectomy) for carriers of these variants. However, further prospective research is required to characterise these risks and to understand the potential of other risk-reducing strategies such as salpingo-oophorectomy and chemoprevention.

The consistency of the relative risk estimates with those derived through family based studies supports the hypothesis that these variants combine multiplicatively with other genetic loci and familial risk factors; this information is critical for deriving comprehensive risk models. Even with very large sample sizes such as those studied here, however, it is still only possible to derive individual risk estimates for a limited set of variants, and even for these variants the estimates are still imprecise. This internationally collaborative approach also has limited capacity to improve risk estimates for rare variants that are only observed in specific populations. Inevitably, therefore, risk models will depend on combining data across multiple variants, using improved in silico predictions and potentially biochemical/functional evidence to synthesise these estimates efficiently. It will also be necessary develop counselling and patient management strategies that can accommodate a multifactorial approach to variant classification.

### Author affiliations

1. Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Melbourne, Melbourne, Australia
2. Huntsman Cancer Institute, Salt Lake City, UT, USA
3. Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit and Biocenter Oulu, University of Oulu, Nordlab Oulu, Oulu, Finland
4. Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA
5. Department of Medical Genetics and National Institute for Health Research Cambridge Biomedical Research Centre, University of Cambridge, and the Department of Clinical Genetics, East Anglian Regional Genetics Service, Addenbrooke’s Hospital
6. Program in Cancer Genetics, Department of Human Genetics and Oncology, Lady Davis Institute, and Research Institute, McGill University Health Centre, McGill University, Montreal, Canada
7. Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Laboratory, Worts Causeway, Cambridge, UK
8. Department of Genetics, University of Pretoria, South Africa
9. Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland
10. Centre for Epidemiology and Biostatistics, School of Population and Global Health, The University of Melbourne, Melbourne, Australia
11. Gynaecology Research Unit, Hannover Medical School, Hannover, Germany
12. Center for Medical Genetics, Ghent University Hospital, De Pinteelaan 185, 9000 Ghent, Belgium
13. Department of Pathology and Human Oncology and Pathogenesis Program, Sloan-Kettering Cancer Center, New York, New York, USA
14. Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Milan, Italy
15. FOM, the FIRC Institute of Molecular Oncology, Milan, Italy
16. Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands
17. Australian Breast Cancer Tissue Bank, University of Sydney at the Westmead Institute for Medical Research, NSW, Australia
18. Centre for Cancer Research, University of Sydney at the Westmead Institute for Medical Research, NSW, Australia
19. Division of Molecular Medicine, Pathology North, Newcastle and University of Newcastle, NSW, Australia
20. University Breast Center Franconia, Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany
Acknowledgements

The authors thank the following for their contributions to this study: Qin Wang (BCAC), Lesley McGuinness, and Ken Offit (CIMBA), Andrew Lee, and Ed Dicks and the staff of the Centre for Genetic Epidemiology Laboratory, staff of the CNIO genotyping unit, Sylvie Laboisse and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, staff of the Copenhagen DNA laboratory, and Sharon A Windenbak, Christopher A Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility. The authors also thank UM1 CA167552 (Willett) Cancer Epidemiology Cohort in Male Health Professionals, Duke University, Durham, NC, USA.

The authors thank the National Institute of Health for their support to the Institute of Cancer Research, Beatson Institute for Cancer Research, Glasgow, UK.

The Cancer Research UK Clinical Trials Unit, Beatson West of Scotland Cancer Centre, 1053 Great Western Road, Glasgow, G12 0YN

Department of Gynaecological Oncology, Glasgow Royal Infirmary

Department of Health Research and Policy - Epidemiology, Stanford University

School of Medicine, Stanford CA, USA

Epidemiology Center, College of Medicine, University of South Florida, Tampa, Florida, USA

Public Health Ontario, Toronto, Canada

Women’s College Research Institute, University of Toronto, Toronto, Ontario, Canada

Department of Epidemiology, Moffitt Cancer Center, Tampa, FL, USA

Department of Biostatistics and Bioinformatics, Moffitt Cancer Center, Tampa, FL, USA

Women’s Cancer, Institute for Women’s Health, UCL, London, United Kingdom

Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, California, USA

Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York, USA

Department of Pathology and Laboratory Diagnostics, The Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland

Department of Medicine, The University of Melbourne Health, Australia

The Royal Melbourne Hospital, Victoria 3050, Australia

Cancer Epidemiology Centre, Cancer Council Victoria, Victoria, Australia

A full list of the 1037 iCOGS and 368 BCAC participants and their genetic variants is available in the online supplemental material and therefore has made substantial contributions to the conception or design of this work. Many authors played multiple roles across these activities. Specifically, MCS conceived this study, worked to include the rare variants into the iCOGS and drafted the manuscript. RLM led the statistical analysis and drafted the paper. Members of the PALB2 interest group, MCS, DEG, RW, KP, FC, MT, WF, JD, KM, EJr, TH, HN, JLI, TD, KC, JRF, ZLT, PR, IC, PF, HT, FAO, JGD contributed to the inclusion of the PALB2 rare variants on iCOGS. DFE coordinated the BCAC project and contributed to statistical analysis along with DEG. SVT contributed to the selection of the CHEK2 rare variants. GC-T contributed to the selection of the rare variants in ATM. AD, CL and JD made significant contribution to the data quality related to the calling of the rare genetics variants on iCOGS. MKS, AB, FBH, SV, JC, CC, RJS, PAF, LH, ABE, MBW, JP, IDSP, OF, NJ, MKB, EJS, IT, MIK, NM, FM, BB, RG, PT, FF, TM, MS, SB, SFN, HF, JB, MPZ, IAP, PM, HAC, SN, AZ, CCD, HB, VA, CS, HB, TB, YDK, TAM, KA, CB, NHB, NNA, AL, SM, AM, VM, VK, N-MK, JMH, AS, EW, DS, BB, GF, JC, AR, PS, DJSF, JEO, CSV, CP, CM, CAH, BEH, FS, LLM, VK, GGA, WZ, DH, SL, SEH, PK, IA, JAK, GIS, AMM, AV, MG, SK, PD, RAEMT, CS, AH, MSG, IF, SJL, JCB, HD, ME, DME, SRT, WIT, SMG, MJH, IWM, JMCL, MTL, PH, JS, JB, KH, AC, MMWR, CL, CB, AD, UH, DT, HU, TR, AJ, KJ, KD, SS, AET, CBA, DY, AS, AA, NO, MJ, AGN, MPA, NA, DH, DCT, DF, VS, JD, MD, PS, RE, KM, FW, HG, JS, MW, BNG, RCT, DN, JLD, FCH, KT, JLS, WIB, ST, DIS, JLK, CM, ASK, CC, LCA, KB, JP, RK, JB, MRT, ZKQ, AAAAA, SB, SPR, AH, AH, MR, DL, EN, IV, NL, JAD, MAR, SN, UE, SWG, KO, LES, GF, JS, GL, JK, LRMG, MTG, NL, IV, RLR, RB, FP, RME, JPF, ASHG, AMG, HJF, FB, PH, HRA, HU, JNM, EB, EH, BP, BB, LB, ELG, BFL, RAV, JMC, MLC, ZCF, KRK, DL, KHL, MATH, XW, DAL, FD, MB, ABO, ESI, JRM, LA, DWC, KLT, ELP, MS, SST, EVB, IO, SHO, LB, HBS, AAVA, KKHK, LAK, LAFGM, TP, YB, ABW, LEC, LSD, NDL, BG, JG, JIM, CHL, LL, LNE, SAJE, ED, ET, J, CN, IP, NS, RG, AW, HJR, VG, WS, HC, XOS, RRT, RS, JRMN, SAN, CP, AMN, DF, HYL, JRM, TAS, TAC, YIT, ZCIM, AGA, SAG, SRM, UM, AHW, CLP, DVBH, MRIM, UH, JMS, JK, PDS, AEG, JC, GGG, SVT, DFE, RLM provided DNA samples and/or phenotypical data. All authors read and approved the final manuscript.

Funding

Funding for the iCOGS infrastructure came from: the European Community’s Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C5047/A8384, C5047/A15007, C5047/A10692, CRUK C8197/A10123), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (No. 1 U19 CA 148537—the GAME-ON initiative), the Department of Defense (WB1X1H-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, the Ovarian Cancer Research Fund and Susan G Komen (WF).

Further information about the financial report received is outlined in the supplemental file (WF).

Competing interests

None declared.

Provenance and peer review

Not commissioned; externally peer reviewed.

Data sharing statement

This would vary for each study—each study is listed in the supplemental material.

Open Access

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

REFERENCES


Cancer genetics


PALB2, CHEK2 and ATM rare variants and cancer risk: data from COGS