Author's response to reviews

Title: Comparing the frequency of common genetic variants and haplotypes between carriers and non-carriers of BRCA1 and BRCA2 deleterious mutations in Australian women diagnosed with breast cancer before 40 years of age

Authors:

Lidija Turkovic (lidija78@yahoo.com)
Lyle C Gurrin (lgurrin@unimelb.edu.au)
Melanie Bahlo (bahlo@wehi.EDU.AU)
Gillian S Dite (g.dite@unimelb.edu.au)
Melissa C Southey (msouthey@unimelb.edu.au)
John L Hopper (j.hopper@unimelb.edu.au)

Version: 3 Date: 22 April 2010

Author's response to reviews: see over
Dear Maria,

We have now addressed the comments of each of the 3 reviewers in point by point form and have made the relevant changes that are highlighted in this letter as well as in the manuscript. We have also made the formatting changes as requested. Details of both are in separate sections: “Formatting changes” and “Response to reviewers” and are included below.

Regards,

Lidija Turkovic

---

**Formatting Changes**

We have:

- Changed the title page to include e-mail addresses for all authors
- Changed the heading “Introduction” to “Background” in both, the abstract and the main text.
- Added a sentence on the Ethics in “Study methods” section of the manuscript: “The protocols for ABCFS have been approved by the Human Research Ethics Committee of University of Melbourne”.
- Added the Competing interests section to now state: “The authors declare that they have no competing interests”.
- Changed Author’s contributions section to now include author’s initials rather than their full names.
- Added the Acknowledgement section saying: “We are grateful to all the individuals who participated in this research, and thank the interviewing staff and co-ordinators, the database management team and directors and the laboratory staff and managers. This research was funded in part by the National Health and Medical Research Council (NHMRC) of Australia and the National Institutes of Health (NIH) in the USA. Lyle C Gurrin, Melissa J Southey and John L Hopper are supported by the NHMRC”.
Response to reviewers

Reviewer's report 1

Reviewer: Nancy Hamel

Reviewer's report:
This manuscript compares the frequency of common variants in the genes BRCA1 and BRCA2 between individuals affected with breast cancer who are either carrier or non-carrier of deleterious mutations in these same genes. A subset of 4 tag SNPs per gene are chosen for analysis and their frequency in the “deleterious” and “non-deleterious” subgroups of patients for each gene are compared. In addition, the various haplotypes in which these SNPs are found are inferred using statistical methods, and the relative frequencies of these haplotypes are again compared between the “deleterious” and ”non-deleterious” groups of breast cancer patients.

The study design is simple and the statistical approaches used to reconstruct the haplotypes and perform genotype and haplotype frequency comparisons are accepted in the field as appropriate to this type of study.

Major Compulsory Revisions:

1. The goal of the study is not immediately obvious from the abstract. Allele frequencies are compared, but what the authors were looking for and how the observations answered the questions is not clearly understood without reading the entire manuscript. I was unsure whether the authors were interested in the frequency of variants in cis with the deleterious mutation, or variants in trans (modifiers), or both.

RESPONSE: We have clarified in the abstract that our hypothesis is that some low prevalence haplotypes of common variants occur more frequently among chromosomes that carry rare, deleterious mutations than chromosomes that do not. We comment further on the plausibility of this hypothesis in the response to point 6 below. It was not our intention to infer statistically whether deleterious variants were in cis or trans phase with the common variants used to form haplotypes (since this would be subject to considerable uncertainty given the very low frequency of specific, individual deleterious mutations), but a reasonable explanation for the hypothesis proposed above is that one or more rare, deleterious variants are in cis phase with the minor alleles of a series of common variants on a single chromosome.

We have amended the introduction of the abstract to include the following phrase stating the hypothesis “Some low prevalence haplotypes of common variants occur more frequently among chromosomes that carry rare, deleterious mutations than chromosomes that do not.”
2. There is a lack of raw data in the paper. In the various tables the total number of variants observed for each gene is listed, but there is no sense of how many individuals were observed to be carriers of these variants. Was each deleterious mutation observed only once, with the total number of distinct mutations and the total number of mutation carriers being equal? This seems unlikely, as some of the deleterious mutations listed are known to be more common than others, and the number of mutation carriers would thus be very low. If some mutations were observed more than once, what was the total number of deleterious mutation carriers for each gene? For each mutation? Is it possible that the weak genotype/haplotype associations observed in this study were driven by one particular deleterious mutation rather than associated with the presence of several deleterious variants? With no data available for review, this is impossible to assess.

RESPONSE: With the exception of one deleterious mutation (2800 del AAG) each of these rare variants was observed only once in our sample.

We have included the following sentence at the end of the second paragraph of the results section “Each of the deleterious mutations identified in our sample appeared only once, with the exception of 2800 del AAG which occurred once in each of two study participants”.

In light of this it is not possible that the associations observed in our analysis were the result of one particular mutation that was over-represented in our sample. Given that the number of mutations known to exist in both \textit{BRCA1} and \textit{BRCA2} number almost one thousand in the case of each gene (a point highlighted by the reviewer and to which we have now updated our reference in the text) and we had sample sizes of breast cancer cases that were also in the hundreds, it is not particularly surprising that no single rare, deleterious mutation appeared more than once in our sample.

3. Throughout the article, the authors refer to mutation carrier patients as opposed to carrier chromosomes. Was phase between the deleterious mutation and the other variants established? Understandably, this would be difficult without several carriers for the same mutation available to identify the haplotype linked to the mutation. However, without phase information, a positive association between a given haplotype and presence of a deleterious mutation does not necessarily reflect a physical association between mutation and haplotype, as seems to be assumed by the authors throughout the article. It could reflect a functional association, such as modifying penetrance in trans and increasing the disease-causing potential of the deleterious variant. Would this affect the interpretation of any association observed?

RESPONSE: We completely agree with the reviewer that it is not possible to establish phase between the rare, deleterious mutations and other common variants without a much larger sample size, extended pedigrees or expensive biological methods. We wish to acknowledge, however, the reviewer’s point about the very limited conclusion that can be drawn about the physical association between mutation and haplotype, and that mutations in \textit{cis} and \textit{trans} would have different interpretations.
To accommodate this concern we have added the following sentence to the discussion: “The suggestive associations that we have observed do not imply a physical association on the same chromosome (as would be the case if the rare, deleterious mutation was in \textit{cis} phase with a haplotype consisting of, for example, the minor alleles of several common variants) or a functional association (as might be the case even if the rare, deleterious mutation was in \textit{trans} phase with a common variant haplotype, since it may still act to modify the penetrance of the disease causing variant). Establishing the phase of rare, deleterious mutations and the common variants we used to define haplotypes for both \textit{BRCA1} and \textit{BRCA2} would require either a much larger sample size than was available for this study, genetic data from extended pedigrees or expensive laboratory investigation.”

Minor Essential Revisions:

4. The number of variants listed in Table 2 is 22 (and not 23 as stated in the text). In addition, the first two mutations listed, 188delAA and 189 delAA are incorrect because 1) they would theoretically be the same mutation and should have the same name, yet they appear to be different since they have different minor allele frequencies and 2) there are no AAs in the vicinity of nucleotides 188 and 189 so the labeling must be wrong. Please verify the accuracy of these 2 listings.

RESPONSE: There are in fact only 22 variants as stated in Table 2 and it has now been corrected in the text.

After re-checking all variants in our data set, we have found that 188delAA and 189delAA were wrongly labelled and they are 188del11 and 189del11 respectively and this has been changed in Table 2. They are listed in BIC database (http://research.nhgri.nih.gov/bic/) as different mutations (even though in very close proximity to each other). They also correspond to different stop codons: 188del11 to STOP 39 and 189del11 to STOP 36. All these changes have been highlighted in Table 2.

5. Please indicate where the minor allele frequencies in Table 2 and 3 were obtained (public databases, or are these the MAF as observed from the data in this study)?

RESPONSE: The minor allele frequencies were calculated from our sample using HaploView and were read directly from the output of this program.

We have now indicated in the text using the following sentence to open the statistical methods section “Haploview was also used to calculate minor allele frequencies for \textit{BRCA1} and \textit{BRCA2} from the data.”

Discretionary Revisions:

6. There is no hypothesis offered by the authors as to why an association between deleterious mutations and one or a few particular haplotypes should be expected to exist in BRCA1 and BRCA2. According to the BIC database, 890
alterations in BRCA1 and 975 alterations in BRCA2 were reported to occur only once; novel variations in these genes arise frequently and are spread throughout the gene sequence. By chance alone, a rare haplotype may be in strong LD with a given mutation but by virtue of being rare would provide little useful predictive power. Common haplotypes, on the other hand, may harbor many deleterious variants, but would also be present often enough in non-carriers that most people who carry them would need to be sequenced regardless, especially since they would be likely to carry them on the 2nd allele as well. So why is this study worth doing? Certainly, identifying one or several haplotypes associated with the presence of deleterious mutations would be highly desirable (in the end it may not matter why it works, as long as it does), but are there any reasons to suggest that finding an association strong enough to be useful for mutation testing is even possible?

RESPONSE: We’ve addressed the reviewer’s question about the need to state the hypothesis explicitly in our response to reviewer’s comment # 1 above. We have also updated the figures from the BIC about the number of “one-off” rare variants in the text. We agree that it is unlikely any haplotype of common variants will be able to successfully “tag” any single rare variant (due to the low predictive power to which the reviewer refers above). We do think, however, that such haplotypes may harbour many different deleterious variants with non-trivial predictive power. Even if the predictive power is low, identifying these haplotypes may eventually prove useful in multi-stage screening and mutation testing and would come at potentially very low cost if data on the relevant common variants was available from existing GWAS platforms.

One can only speculate on the biological mechanism by which this happens, but if there are regions with BRCA1 and BRCA2 that are less stable and more susceptible to mutation then these regions might also be more likely to be populated by minor rather than major alleles of more common variants, which is essentially the hypothesis that we state above.

The reason this study is worth doing is that a great deal of attention has been focussed in the last 5-10 years on the “complex disease common variant” paradigm which proposes that low to moderate prevalence disease causing variants can be captured by combinations of common variants. BRCA1 and BRCA2 are something of an anomaly in this framework, since although the genes are large in terms of kilobases compared to other genes, only a small number of tagging SNPs are required to capture common variation due to strong linkage disequilibrium and yet there are hundreds of very rare, potentially deleterious variants that most likely won’t be captured individually by common haplotypes. We thought it was important to put this to the test, and to see whether groups of rare, deleterious variants might clusters within haplotypes even if single rare variants do not.
Reviewer's report #2

Reviewer: Tom Scholl

Reviewer's report:

The authors evaluate an Australian Breast Cancer cohort to understand relationships between common genetic variants and common haplotypes and the incidence of clinically significant mutations. This work is similar to studies produced elsewhere where sometimes conflicting results can be understood in the context of small data sets, some with potentially restricted ethnic backgrounds. This field overall can be difficult to follow due to the use of different SNPs, and other markers, to define the haplotypes and a lack of common designations for haplotypes. BRCA1 can be an interesting model to study haplotypes since it is thought reside at a locus that is somewhat protected from recombination, which reduces complexity at the locus. Also, a series of about 8-14 common polymorphisms spread across the coding regions that appear in clinical sequencing tests can explain virtually all specimens from diverse populations as 9 or 10 prevalent haplotypes.

Major Compulsory Revisions:

None.

Minor Compulsory Revisions:

The authors provide little information regarding the cohort. The potential for bias exists if subsets of the cohort have differential haplotype composition and differential access to the BRCA test. More specifically, a concern could be that specimens in the cohort with Australian Aboriginal ancestry could differ in haplotype composition from those descended from Western Europeans. If selection for testing was also skewed where one group had greater a priori risk of a mutation (for example, where stronger family history is required for one group to receive testing than another), mutation prevalence by haplotype would be affected. Since this type of bias is not unheard of, the authors should address it in the text and if possible, elaborate on the cohort, ethnicity of specimens, and testing selection criteria.

RESPONSE: The vast majority of proband participants in the Australian Breast Cancer Family Study are white women, many of whom have either northern or southern European ancestry. Approximately 2% of the Australian population are Indigenous and only 30% of these live in major Australian cities, so the proportion of the population eligible for enrolment that are Indigenous would be less than 1%. Although there are a small number of Indigenous people participating in the ABCFS, none were represented in the sample from which the data were drawn for the analyses presented in this paper.

We have added a sentence in the “study methods” section stating that “The majority of proband participants in the ABCFS are white women with northern or southern European ancestry.”
**Discretionary Revisions:**

Papers in this field attribute value to this research (as on page 3) as a possible way to preclude expensive comprehensive mutation scanning by stratifying patient risk or subdividing the mutations to be tested based on patients’ BRCA haplotypes. This goal seems very unlikely based on previously published data and the results presented here due to the weak predictive power of haplotypes to explain the presence of mutations. Family history will likely remain the preferred clinical selection criteria for genetic testing in hereditary breast cancer. It would be interesting to address haplotype composition of different ethnic groups. It could also be interesting to describe the genetic variant prevalence and composition that reside on different haplotypes. There are also very practical, though less glamorous, uses of haplotypes that are not addressed by the authors. For example, haplotypes in the BRCA genes have been used to identify patients carrying unbeknownst deletion mutations, and to define the clinical significance of genetic variants. The haplotype composition and localization of genetic variants to particular haplotypes can also be useful quality controls in labs performing clinical testing. The authors might choose to expand their discussion to include any of these topics.

**RESPONSE:** As we pointed out in our response to the discretionary revisions suggested by reviewer #1, it is not out of the question that combinations of common variants could generate haplotypes that have similar frequencies to rare, deleterious variants, and that groups of these rare variants will cluster within haplotypes. As suggested in our response to reviewer 1, we think that common haplotypes might still have potentially in a multi-stage approach to assessing the risk of an individual carrying *any* deleterious mutation. Unfortunately we are unable to address the issue of BRCA1 and BRCA2 haplotype composition in different ethnic groups due to the relative homogeneity of the ethnic background of participants in the ABCFS.
Reviewer's report #3

Reviewer: Patricia Tonin

Reviewer's report:

This study investigates the possibility that haplotypes revealed by the analysis of SNPs are associated with deleterious (or pathogenic) BRCA1 or BRCA2 mutations. These breast-ovarian cancer susceptibility genes are large and complex. Numerous pathogenic mutations have been identified that vary in their nature which confound mutation screening and thus screening remains costly and laborious. Identifying haplotypes associated with rare pathogenic mutations may facilitate the identification of carriers. As mentioned by the authors prior studies have shown that this approach has some merit.

The investigators used data from 680 participants for BRCA1 genetic analysis and 245 participants for BRCA2 genetic analyses. However the haplotype analyses was restricted to 392 (for BRCA1) and 179 (for BRCA2) which comprise the cohort diagnosed with breast cancer under 40 years of age. They identified tagSNPs using Haploview suitable for genetic analyses and assayed the frequency of these variants with respect to pathogenic mutation status. They observed differences in the frequency of the common genetic variants of BRCA1 and BRCA2 and their haplotypes inferring phase between early-onset breast cancer cases who did not carry pathogenic mutations in these genes. However the analysis was limited by sample size and thus evidence presented was suggestive (or weak) and would have to be substantiated using a larger pool of data.

The study appears sound, but could be improved by addressing the following points that should be considered as major compulsory revisions:

REVIEWER: It is not clear why the study was restricted to investigating haplotypes from cases diagnosed under 40 years of age? The first paragraph of the Results section states that “some sequencing was performed for cases in the other groups but sample sizes were small and not sufficient to warrant a separate analysis.” Does this imply that only a portion of cases were sequenced and/or investigated for SNP analyses? This does not correspond to the information indicated in Table 1 referring to “total DNA sequenced”. Given that only a ‘weak’ association was observed have authors investigated the possibility of using all available data – particularly as BRCA1-positive cases also occur in the later-age-of –onset group? While it may not be possible to analyze each age group a combined analysis may improve outcome.

RESPONSE: It is true that only a portion of cases in the 40-49 and 50-59 age groups were sequenced for mutations in BRCA1 and BRCA2, whereas all ABCFS participants aged 40 or under were sequenced for these two genes. This is one reason to restrict the analysis to this group of participants. The second reason is that genetic causes of breast cancer are far more prevalent in cases aged under 40 than over 40, maximising the statistical power of our analysis and ensuring that any findings can be generalised to a specific age bracket.

Analysing additional data from our sample is unlikely to alter the conclusion
since the number of rare, deleterious mutation carriers was very small in the older age brackets and in fact it was zero for BRCA2 gene.

REVIEWER: The population studied is from the Australian Breast Cancer Study from participants from metropolitan areas of Melbourne and Sydney. It is evident from the genotyping results presented in Tables 2 and 3 which contain the BRCA1 and BRCA2 pathogenic mutations that there is little evidence of recurrent (or founder) mutations at least within the population studied. Given the finding that study would have to be extended to include a larger sample size is it feasible to do so in the Australian population? A comment in the Discussion section addressing this issue could provide a useful discussion point for researchers investigating this avenue of research and the applicability of haplotype analyses for genetic testing of rare essentially monogenic disorders.

RESPONSE: It is certainly feasible to expand the sample size in the Australian population. The ACBFS is on-going and continues to be funded by the NHMRC and NIH. We plan to submit an application to the University of Melbourne Human Research Ethics Committee (the IRB at the institution housing the study) to engage in community recruitment of breast cancer patients and their family members, whereas previous recruitment was via the Victoria and NSW Cancer Registries.