Author’s response to reviews

Title: Genetic Determinants of Sporadic Breast Cancer in Sri Lankan Women

Authors:

Nirmala Sirisena (nirmalasirisena@yahoo.com)
Adebowale Adeyemo (adeyemoa@mail.nih.gov)
Anchala Kuruppu (anchalak@gmail.com)
Nilaksha Neththikumara (nilakshafreezon@gmail.com)
Nilakshi Samaranayake (nilakshis35@gmail.com)
Vajira Dissanayake (vajirahwd@hotmail.com)

Version: 1 Date: 21 Aug 2017

Author’s response to reviews:

21st August 2017

The Editor
BMC Cancer

Dear Dr. Solera,

Response to reviewers’ comments on the manuscript titled “Genetic Determinants of Sporadic Breast Cancer in Sri Lankan Women” – BCAN-D-17-00449R1

Thank you for giving us the opportunity to submit a revised version of the manuscript. A detailed description is annexed to this letter explaining how we have dealt with each of the reviewers’ concerns. The revised manuscript is also attached.

We look forward to hearing from you in due course.

Sincerely,

Dr. Nirmala Sirisena
Reviewer 1 Comments:

This is a well-written, congruent, and succinct manuscript about a genetic epidemiology study of breast cancer patients in a Sri Lankan population. Their approach overall is consistent with small cohorts and identifies a priori SNPs to test in the population. Their approach resulted in the identification of SNPs associated with sporadic breast cancer in the population.

We thank the expert reviewer for this comment.

The only concern I have is the use of the Guajarati Indian data for haplotype tagging. How do we know that the population is the best for the identification of tagging SNPs?

The Guajarati Indian (HapMap GIH) genotype data was used for haplotype tagging because at the time of the study design, this was the only South Asian population group for whom dense genotype data was available. Therefore, at the time, this group had the closest resemblance to the Sri Lankan population compared to other populations in the International HapMap Project. This statement has been included in the Methods section on page 6, lines 5-8. We note that if the study were to be designed today, it would be able to take advantage of the additional South Asian population groups genotype data that are now available through the 1000 Genomes Project.

There was little innovation in the study other than the fact that very little research has been done in the Sri Lankan population. Since the population was being considered, it would have been appropriate to mention other studies performed in Indian or other Asian populations. Have there been SNPs ID'ed in other Asian populations that may have been worth considering?

We now include studies performed in other Asian populations in the Introduction section of the revised manuscript on page 3, lines 17-25 and page 4, lines 1-11. However, it should be noted that the SNPs investigated in most of these studies were in intronic regions in contrast to our study which focused mainly on non-synonymous coding SNPs, and SNPs in the regulatory regions.

Reviewer 2 Comments:

The paper by Sirisena et al is well written. The study is logical and the experiments performed well planned, described well and the conclusions drawn are appropriate.

We thank the expert reviewer for these comments.

There are some minor typographical and grammatical errors throughout, but these can easily be corrected with another proof read.

Typographical and grammatical errors have been corrected in the revised manuscript.

I would like to have seen more structure in the discussion section, potentially some sub headings and some more information presented about the LD data in Table 2 and any interactions of the
three genes identified in the Sri Lankan population. As an example, it is postulated that XRCC2 may interact with other biological pathways, was a cursory look preformed? Is there any evidence of this?

The discussion section has been structured with the addition of subheadings. While LD was taken into account in selecting tag SNPs and in constructing haplotypes, we did not present LD data in Table 2. We did not formally test for interactions in the present study because of the limited sample size (350 cases, 350 controls) which while adequate for single variant analysis is too underpowered to detect interactions. Indeed, it should be noted the studies we cited on interactions each had a sample size of over 35,000 (Yu et al, Breast Cancer Res Treat 2010; Kong et al, Int J Clin Exp Med 2015). For these reasons, we prefer not to do formal tests for interaction between the SNPs in this study as this would essentially be a “fishing expedition” whose results are likely to be difficult to interpret (i.e. a positive finding may be a false positive because of the sample size, a negative finding would not necessarily mean that it would be so in an adequately powered sample).

Our discussion of how XRCC2 acts in various pathways that may be involved in breast cancer is based on the literature and not our dataset. In view of the limited nature of the present study, we do not have sufficient data to test for evidence of such interactions with biological pathways in our study.

The legends for the Tables presented do not clearly identify the contents, this should be improved.

The legends for the Tables have been revised to make them more informative.

An inclusion of the expected MAF is important for interpretation in the tables.

We now included MAF in the appropriate places in all relevant tables.

The discussion section is a little long, some structuring of this and potentially the inclusion of a summary table/figure would help reduce the content and better convey the overall findings and interactions of the identified SNPs and genes.

The discussion section has been structured and shortened, with subheadings added in the revised manuscript. Please see the note about interactions above. Thank you.