Author's response to reviews

Title: Genetic and functional evaluation of the role of CXCR1 and CXCR2 in susceptibility to visceral leishmaniasis in India

Authors:

Sanjana Mehrotra (anniebhu19@gmail.com)
Michaela Fakiola (mf300@cam.ac.uk)
Joyce Oommen (joommen@ichr.uwa.edu.au)
Sarra E Jamieson (sjamieson@ichr.uwa.edu.au)
Anshuman Mishra (anshuindia@gmail.com)
Medhavi Sudharshan (medhavisudarshan@gmail.com)
Puja Tiwary (tiwarypuja@gmail.com)
Deepa S Rani (deepa@ccmb.res.in)
Kumarasamy Thangara (thangs@ccmb.res.in)
Madhukar Rai (upicon2007@gmail.com)
Shyam Sundar (drshyamsundar@hotmail.com)
Jenefer M Blackwell (jmb37@cam.ac.uk)

Version: 2 Date: 21 October 2011

Author's response to reviews: see over
20 October 2011

For attention: Guian Paolo Declaro
Journal Editorial Office
BioMed Central

Dear Editors

**MS: 2399221345798089** - Genetic and functional evaluation of the role of CXCR1 and CXCR2 in susceptibility to visceral leishmaniasis in India.

Thank you for your email outlining compulsory changes required for our manuscript to be reconsidered for publication in *BMC Medical Genetics*. We thank the reviewers for their very helpful comments and, in particular, for picking up on one silly error (mine in final edit) in transcribing haplotypes from numbered alleles to nucleotides called on the incorrect strand, which we have corrected.

We have submitted a revised manuscript which we hope deals with all the issues of both reviewers, as outlined in the detailed response to reviewers appended below.

We hope that our revised manuscript will now be suitable for publication in *BMC Medical Genetics*.

With thanks,

Yours faithfully,

Jenefer M. Blackwell
For The Authors
Detailed response to reviewers – Mehrotra et al.

Relevant reviewer comment in italics; our response non-italicised – bold summarises changes made to manuscript; changes highlighted on the manuscript in yellow for reviewer #1 and lime for reviewer #2.

Reviewer #1

This reviewer had no major concerns about the manuscript. Minor points included:

1. *If any heritability studies have been done for VL, what is the % range attributed to the host’s genetic make-up?*

Studies from Brazil demonstrate that resistance to VL, as determined by a positive antigen-specific DTH response without clinical symptoms, is ~80% heritable in family-based studies. Sample sizes for VL in family studies have not been large enough to obtain accurate measures of heritability to date. **We have included a sentence about heritability in the introduction to the paper (Page 4).**

2. *Given the small odds ratios detected, do the authors consider these SNPs to be major contributors to VL?*

The odds ratios obtained are in line with those expected for complex traits like infectious or autoimmune disease, in which many loci contribute to overall susceptibility and larger sample sizes such as those used here are able to identify. Together, they provide a picture of the genes and pathways that contribute to disease pathogenesis. **We have commented on the size of our OR in relation to those expected for complex disease in the manuscript and provided an appropriate reference (Page 9).**

3. *Is there any reason to believe that population stratification should be corrected for?*

We do not have GWAS data for this sample, but in preliminary analysis of the GWAS data on a discovery sample from India we have shown that caste forms a very good surrogate for population substructure in this region of India. **We have now included results of analyses using caste (and religion, cf. reviewer #2) as a covariates (Pages 6 and 9).**

4. *The samples described in this study could in the future be used in a genome-wide approach.*

A GWAS is being undertaken on the basis of discovery and replication samples in India and will be published in due course. This work arose directly as a result of our hypothesis-driven candidate gene interest in this region of Chromosome 2q35, for which we have functional as well as genetic data.

Reviewer #2

Major essential revision:
1. The selection of such a small number of loci for such big sample size is surprising. Robust justification is needed. How and why were only these rsIDs selected for analysis? Give specific reasons.

With respect, just because one has a large sample size doesn’t mean that we need to analyse more SNPs. The whole point of the HapMap project was to provide information on SNPs that tagged LD blocks, so that less money needs to be spent genotyping redundant SNPs. As outlined in the methods (Pages 5/6), we selected SNPs that tagged the two major LD blocks that occur in CXCR1 and CXCR2 based on HapMap CHB/JPT data, plus an additional potentially functional non-synonymous SNP. As it happened, this non-synonymous SNP was not associated in the replication data, and is therefore unlikely to be the etiological variant. 

We have provided a supplementary figure 1 to show how the SNPs selected tagged these major blocks. Since we obtained associations using the tag-SNPs, and showed that haplotype associations extended across both genes, the issue then was to determine which gene was more likely to be genetically regulated and functionally involved in disease pathogenesis. To do that we turned to analysis of gene expression in splenic aspirates to see if either gene changed in expression as part of the change in disease pathogenesis associated with response to treatment.

2. In the title: Only population from BIHAR is considered and that cannot serve as representative of all of INDIA. It should read something like north India or north-east India as per IGVDB. Analyses must be done keeping the a/m database as reference.

We have altered the title accordingly.

3. In methods: although the selection of tag SNPs is described selection of only three loci for such a sample size is surprising. Given the fact that Sequenom platform was used inclusion of more loci in the study could have generated some more interesting unexplored data. This is not acceptable and needs justification.

See response to point 1 above.

4. It should be made clear whether strand conversion was done before haplotype analysis as the SNPs lie on different strands. If no, then these haplotypes are not possible. Needs clarification.

We thank the reviewer for pointing this out. The haplotype analyses were carried out based on the original allele calling before author JB made the decision to assign nucleotides according to the strand on which each gene was encoded – so indeed the haplotypes as presented were incorrect. We have now corrected the haplotypes in Table 5 and text (page 10) so that all alleles are called on the same strand.

5. The major issue of population stratification that can be a serious confounding factor in genetic studies has not been addressed which is a serious drawback as there are different religious groups involved in the population included in the study. Kindly perform this analysis.
See response to query 3 reviewer #1 and query 11 for reviewer #2 below

6. The role of CXCR1 and 2 are very confusing. association found in CXCR1 and downregulation expression of CXCR2 does not necessarily imply a role of CXCR2 and not of CXCR1. Moreover only a tiny fraction 19 samples used for expression analysis as against considerably larger samples used for genetic study. Therefore to discredit genetic study saying that the association could reflect some LD effect and the expression analysis has binding is not acceptable.

With respect, we disagree. The sample size of N=19 is perfectly adequate to purpose (see response to minor question 4 below). The discussion of how the expression is possibly related to genetic regulation is our interpretation of the data (e.g. favour CXCR2 as the etiological gene regulating susceptibility to disease) based on current results. We have not over-interpreted our data. Future work will determine whether this interpretation is correct.

Minor essential revision:

1. The research question are not well formulated and findings are not suitably explained.

With respect, we believe our hypotheses are outlined clearly in the introduction to the paper: “Here we use genetic and functional approaches to evaluate the role of PMN in VL caused by L. donovani in humans through analysis of the receptors CXCR1, which is a specific receptor for IL-8 (=CXCL8), and CXCR2, which is promiscuous in binding a variety of CXC chemokines (CXCL-1, 2, 3, 5, 6, 7) in addition to CXCL8”.

2. The study is concluded in a haphazard manner and needs to be concluded better, specifically and clearly.

With respect, we believe that we have made a clear conclusion regarding the findings of the paper: “Here we examined CXCR1 and CXCR2 as candidate genes for susceptibility to VL in India. Whilst SNPs at both loci were associated with VL, functional analysis of expression in splenic aspirates together with a common risk haplotype favour CXCR2 as the etiological gene regulating susceptibility to disease. Our data contribute to increasing evidence for an important role for PMN in directing the outcome of leishmanial infections in humans.”

3. Define nuclear families? Because a good part of analysis is depends on the relevant definition.

A nuclear family is the parents plus offspring (see Wikipedia: A nuclear family is defined as a family group consisting of a father and mother and their children). This term is used routinely in genetic studies and should not require further explanation in a genetics journal.

4. Why was qRT PCR done on such a small number of samples? This number is too small given the fact that the inference it draws and the number of cases the authors had access to.
With respect, N=19 is a well-powered sample for this kind of functional immunological study. Power calculations confirm that 19 paired samples has 98.2% power to detect a difference in the mean values between pre- and post-treatment samples of 0.9 with a standard deviation of 0.9 at alpha level = 0.05; 92.2% power at alpha 0.01.

Therefore, the sample was adequately powered for the purposes of the experiment being performed, which was to compare CXCR1 or CXCR2 expression levels between pre- and post-treatment samples.

**We have added a sentence about power calculations in the methods (page 7/8).**

5. *In abstract results include “using FBAT”. Details must be given in the methodology.*

With respect, FBAT is a standard genetic statistical tool for family-based analysis that was fully referenced in the methods section of the paper. **We have expanded the description of what FBAT does in the methods (Page 6).**

6. *Introduction: “Genetically regulated variability in the immune system”; Modify the sentence.*

With respect, there is no sentence that has this precise wording. **We have modified the wording to the sentence we believe the reviewer referred to (Page 4): Genetic variability in ability to mount an innate immune response, such as…..**

7.xi. *While mentioning the power of study authors say that the “the 313 VL trios had #95% power to detect an odds ratio #2 at P=0.01 for markers with MAF#0.1, and 49% power for an odds ratio of 1.5” since most of the detected odds ratio is in this range of 1.5 and none greater than 2; justify only 49% power of study.*

This information is provided to give the reader a feel for the relative level of power in the two samples. **We have now simplified this to indicate the power of the two samples under the sample conditions, OR=1.5, MAF>0.1, P=0.01 (Page 6).** Just because you only have 50% doesn’t mean that you won’t find an association. The important thing then is to verify the association by replication, which we have done in this study.

8. *The name of the country should be mentioned along with the products.*

**Corrected throughout methods (Pages 6 and 7).**

9. *In methods: while mentioning the softwares write it fully as :GraphPad Prism (version 5.00 for Windows, Graph Pad Software, San Diego California USA, www.graphpad.com)*

**Corrected (Page 7).**

10. *p values are one tailed or two tailed? Write clearly.*

**We have indicated in the methods (Page 7) that two-tailed tests are used throughout.**
11. The major issue of population stratification that can be a serious confounding factor in genetic studies has not been addressed which is a serious drawback as there are different religious groups involved. Kindly perform this analysis.

See also response to query 3, referee #1. We have added information on using both religion and caste as covariates (Pages 6 and 9).

12. In discussion the authors mention that same blood marriages are very common in this population but that is true only for Muslim families not the majority of Hindus. Such statement should be removed or detailed analysis must be included to justify the statement.

With respect, we have found evidence for the some level of consanguinity in both Hindu and Muslim families from this region of India (Fakiola et al., reference #19).