Author's response to reviews

Title: Association of ADPRT1, AKR1B1, RAGE, GFPT2 and PAI1 gene polymorphisms with chronic renal insufficiency among Asian Indians with type-2 diabetes

Authors:

Pushplata Prasad (pushplata_prasad114@rediffmail.com)
Arun K Tiwari (arun_genet@yahoo.com)
Prasanna Kumar - KM (kmprasanna@vsnl.com)
AC Ammini (aca433@yahoo.com)
Arvind Gupta (arvindneelum@hotmail.com)
Ashok Sharma (draksharma@satyam.net.in)
Rajeev Gupta (rajeevg@satyam.net.in)
BK Thelma (humgen@vsnl.com)

Version: 2 Date: 3 August 2009

Author's response to reviews: see over
To

Danielle Burgess, PhD
Assistant Scientific Editor
BioMed Central

Sub: Point by point reply to the reviewers’ comments

Dear Dr. Burgess,

Thank you very much for having our Ms. titled “Association of ADPRT1, AKR1B1, RAGE, GFPT2 and PAI1 gene polymorphisms with chronic renal insufficiency among Asian Indians with type-2 diabetes” reviewed by the experts in this area of work and for the very valuable comments and suggestions made by the reviewers. We have been able to address all the issues raised therein and this has undoubtedly enhanced the quality of our Ms. Corrections/Modifications made according to the reviewers’ comments are marked in red in the revised Ms.

We hope that you find our Ms. worthy of publication in your esteemed journal.

Thank you

Sincerely

Thelma B.K.
Professor
Dept. of genetics
University of Delhi South Campus
New Delhi, India
Major Compulsory Revisions:

Query1:
Prasad et al. tested nine SNPs in five genes as well as interactions for association with chronic renal insufficiency. They did not clarify how and why these nine SNPs were selected. In particular they say in the introduction that 'V1762Ala and Leu54Phe of the ADPRT1 gene encoding PARP have been tested for association ...', but from table 2 it became apparent that they did not analyze the Leu54Phe but a Arg...Lys polymorphism. Why? Furthermore, in the discussion they state that 'association analysis of functionally significant variants in these genes becomes imperative' (p9), but it is not clear whether the SNPs they tested are indeed functional.

Reply 1: The SNPs were selected based on the minor allele frequency (MAF) reported in NCBI database for different populations (preferably Asians), heterozygosity content, linkage disequilibrium (LD) structure, prior published association reports, validation evidence, and functional importance.

Among various Asian populations, the reported (NCBI database) MAF and heterozygosity of SNP Leu54Phe (rs3738708) is lower (MAF = 0.00 to 0.04, heterozygosity = 0.023) as compared to the other two polymorphisms [Val762Ala (rs1136410): MAF = 0.45 to 0.47, heterozygosity = 0.35; Arg940Lys (rs3219145): MAF = 0.03 to 0.08, heterozygosity = 0.42] of ADPRT1 gene that we selected for analysis in our population. Therefore, we did not test association of Leu54Phe in our study.

In this study, SNPs were selected based on their functional relevance: SNPs leading to missense mutations, promoter polymorphisms, and those present in the regulatory regions (5' UTR and 3' UTR) were chosen in the same order of priority, and information content (MAF).

A statement, explaining the selection criteria for SNPs that were tested in the study has now been added in the revised Ms. (Please see page No. 8, line No. 156-159)

Query 2:
Despite the multiple SNPs and interactions tested they adopted a significance level of 0.05 indicating that no multiple testing corrections was applied. If this would be applied, none of the associations would remain significant.

Reply 2: We agree with the reviewer. Out of a total of nine polymorphisms that were initially chosen in the study, two were monomorphic in our population. Therefore, only seven SNPs were tested for their association with diabetic CRI. Bonferroni correction value for the univariate association analysis in this study is ($\beta$) = 0.05/7 = 0.007. None of the association observed in the study stands Bonferroni correction. We have now added a statement in this regard in the revised Ms. under results section (Page No. 9, Line No. 189-
Therefore, the associations that were observed in the study would be better regarded as “marginal associations” than “significant associations”.

Query 3:
No explanation is given why the interactions they analyzed were selected (should be in introduction, not in results section) and why others weren't. For this interaction analysis genes from their previous studies were included in the interaction analysis. It needs clarification why they looked into interaction with these genes (and as it seems not between either of the five genes of this study). Is there any biological explanation why there could be an interaction effect? If so, mention this in the introduction; if not, it is not justified why only two interactions are picked out and it should be omitted. It is also advised to briefly describe the results for these genes from the previous studies.

Reply 3: Thank you very much for drawing our attention to this oversight. We have corrected this mistake in the revised Ms. (Please see Page No. 5-7, Line No. 95-131).

Query 4:
Also no information is given how these interactions were analyzed: what kind of inheritance model did they use and were main effects and covariates included?

Reply 4: Multiple logistic regression (Backward) analyses was carried out to correlate various clinical parameters with genotypes; and to study gene-gene interactions between SNPs of different pathways. Additive inheritance model in multiple regression analysis was employed, and candidate variables with a P value < 0.05 were selected, for gene-gene interaction analysis. Correction in this regard has been incorporated in the revised Ms. (Please see Page No. 9, Line No. 171-174).

Query 5:
In the discussion they state that they have moderate power. However for this calculation they used a significance threshold of 0.1 and not of 0.05, which was the significance threshold used in the analyses. Furthermore, power depends on the allele frequency, so for the PAI polymorphism the study has more power than for the other SNPs.

Reply 5: In the revised Ms., in Table 3, we have reported power (G) for each of the SNPs at 5% significance level.

Query 6:
They used a 2 df genotypic chi-square test. However, for seven of the nine polymorphisms the homozygotes of the wild type allele are rare and a chi-square test is not valid. They should have used a Fisher's exact test.

Reply 6:
We agree with the reviewer. In the Table 3 of the revised Ms. we have provided Fisher’s exact P value along with the P value for the Pearson’s chi square test.

Query 7: 
Furthermore, it was not specified what kind of inheritance model was used in the logistic regression model. It was clearly not the general model in contrast to the genotypic chi-square test, since only 1 odds ratio was given in the results section, but was it dominant, recessive or multiplicative (log additive)?

Reply 7: Candidate variables with a P value < 0.05 were selected and assessment of cumulative contribution and epistatic interaction between different candidate pathways was done using an additive inheritance model in multiple regression analysis. Correction in this regard has been incorporated in the revised Ms. (Please see Page No. 9, Line No. 171-174).

Query 8: 
In the discussion on p10 they say that ‘a heterozygote excess of this marker in our control group (DM) may be masking the allelic association ...’. As explanation they state that there is a 'very low frequency of CC homozygotes in the control group' and argue that this 'may suggest that it may confer major susceptibility to severe kidney impairments ...'. However, 5% is NOT very low and hence this SNP is not likely to be having a selective effect. In my opinion this finding is more likely the result of a chance finding than a true positive result. Furthermore, it is not clear to me what they mean by their conclusion that these 'severe kidney impairments' can lead to 'consequent early representation among the group'.

Reply 8: We are sorry for this mistake. By mistake we reported the genotypic frequencies of all seven SNPs for CRI subjects under control (DM) column and vice versa. This mistake has been rectified in the table 3 of the revised Ms. Therefore, in the discussion section of our Ms, we stated that there is a 'very low frequency of CC homozygotes in the control group' which may be suggestive of its potential nature to confer major susceptibility to severe kidney impairments.

However, considering your suggestion that this “may be the result of a chance finding than a true positive result”, we have added the following sentence for caution.

“Further, since the power to detect association of this SNP is very low (G=7%) in this study, we should be cautious in interpreting association of this SNP”. Changes in this regard has been made in the revised Ms. (Page No. 13, Line No. 270-273)

Minor revisions: 
Query: Abstract p2: in the objectives 'eight' polymorphisms must be 'nine'.
Reply: Sorry for the mistake. It has been corrected in the revised Ms. (Page No. 2, Line No. 32)

Abstract p2: two SNPs were found significantly associated but only one OR is given.
Reply: Sorry for the mistake. Considering the word limit (350) for the abstract, we have omitted the OR and CI values from the abstract in the revised Ms.

Query: Abstract p2: it is unclear where TGF-b1 comes from. This gene was not part of this study and hence needs further clarification.

Reply: In the objective of the abstract, we have mentioned that in addition to determine association of nine single nucleotide polymorphisms (SNPs) in ADP ribosyltransferase-1 (ADPRT1), aldo-keto reductase family 1 member B1 (AKR1B1), receptor for advanced glycation end-products (RAGE), glutamine:fructose-6-phosphate amidotransferase-2 (GFPT2), and plasminogen activator inhibitor-1 (PAI-1) genes with chronic renal insufficiency (CRI) among Asian Indians with type 2 diabetes we also attempted to identify epistatic interactions between genes from the present study and those from renin-angiotensin-aldosterone system (RAAS), and chemokine-cytokine, dopaminergic, and oxidative stress pathways (previously analysed using the same sample set).

TGF-β 1 is a cytokine, and the association of SNPs in TGF-β 1 was tested in the same sample set in a previously reported analysis (Reference 4). MLR analysis involving test of epistatic interactions between genes from different pathways identified SNP -800 G>A in TGF-β 1 and rs7725, C>T in GFPT2 to be significantly interacting SNP pairs.

Clarification in this regard has been added in the abstract of revised Ms. (Please see Page No. 3, line No. 42)

Query: Introduction p4: a reference is needed for the results of clinical trials (line 12-14).
Reply: Added references 7 to 9.

Query: Introduction p3: 'has' must be 'have' in line 10.
Reply: Corrected

Query: Methods p6: insert 'was' in line 7 ('informed consent was obtained').
Reply: Corrected

Query: Methods p7: was HWE calculated among controls or on the entire sample? It should be done on controls only since SNPs can deviate from HWE in cases as a result of disease.
Reply: Thank you for the information. We understand this completely. In this study, HWE was calculated using control samples only.

Query: Methods p7: they say that they used both linear and logistic regression. However, results are given only for CRI versus non-CRI so no linear regression is performed for this paper.

Reply: Sorry for this mistake. However, linear regression analysis was carried out using GFR (glomerular filtration rate) as a dependent variable. GFR is a linearly distributed clinical parameter crucial in determining the extent of kidney damage, and therefore may act as a shadow marker for kidney disease. Therefore, a linear regression analysis was
carried out using GFR values as a dependent variable and various genetic markers and clinical parameters (age, gender, BMI, duration of diabetes) as independent variables. We did not observe association of any of the independent variables with GFR.

We have described the result of linear regression analysis in the last paragraph under the result section of the revised Ms. (Please see Page No. 11, Line No. 230-235)

Query: Methods p7: in the discussion they mention that they used backward logistic regression. That should be added in the methods. It is also not clear in the results that they used this procedure. In addition, it should be specified what was included in this model.
Reply: Done. (Page No. 10, Line No. 206-210)

Query: Methods, statistical analysis: they should describe and elaborate on the Mirbase tool that they used for miRNA target identification.
Reply: Thank you for drawing our attention to this over-sight. We have added the following text in the Methods section. Of the revised Ms. (Please see Page No. 9, Line No. 176-181)

“The miRBase Sequence Database is a searchable database of published miRNA sequences and annotation [miRBase::Sequences (http://microrna.sanger.ac.uk/sequences)]. Each entry in the miRBase Sequence database represents a predicted hairpin portion of a miRNA transcript (termed mir in the database), with information on the location and sequence of the mature miRNA sequence (termed miR).”

Query: p8. last line states (Ref) instead of the reference.
Reply: Correction has been made in the revised Ms.

Query: p8: D allele of ACE needs further clarification.
Reply: Added (Please see Page No. 6-7, Line No. 122-125).
Query: p9: confidence interval (0.04-0.14) is not correct
Reply: The mistake has been corrected in the revised Ms. (Please see Page No. 13, Line No. 266-267).

Query: p10: the second paragraph starts with 'Two functional variants ... in this gene'. What gene is referred to?
Reply: The words “in this gene” are replaced by “RAGE”. (Please see Page No. 13, Line No. 262).

Reply: Thank you for drawing our attention to this over-sight. Reference (29) has been inserted at the appropriate place in the revised Ms. (Page No. 13, Line No. 280)

Query: p11: it is unclear what 'Gene bearing the 'T' (minor allele)' means.
Reply: The sentence has been modified to read as “T allele of rs7725, C>T in GFPT2 has been shown to be approximately 2-fold over-expressed resulting in increased mRNA levels
with resultant increased hexosamine flux” in the modified Ms. (Page No. 13-14 and Line No. 282-284).

Query: p11: abbreviation ECM needs to be explained
Reply: Done

Query: p11: 'predisposing' in line 18 should be removed and 'which' should be added in the same sentence after 'TGF beta1 gene'
Reply: Done

Query: p11: 'protective towards' must be 'protective against'.
Reply: Done

Query: p12: in the final paragraph they state that the interaction is 'confirmed by pair-wise as well as multiple logistic regression analysis'. Analyzing the data in two ways is not a confirmation. I suggest to remove this.
Reply: Done

Query: Table 1: abbreviations UAER and GFR need clarification
Reply: Done

Query: Table 2: Heading of first column should change and specify all the information in this column (i.e. gene, OMIM number, location of gene, polymorphism)
Reply: Done

Query: Table 2: not only give the substitution but the position as well if that is known. Eg. change Val>Ala to Val762Ala
Reply: Done

Query: Table 3: Heading is incorrect (different genes than analyzed).
Reply: Sorry for the mistake. The heading has been corrected in the revised Ms.

Query: Table 3: change PARP to ADPRT1
Reply: Done

Query: Table 3: give rs-numbers or other polymorphism coding instead of only the substitution
Reply: Done
Query: Table 3: The genotypic association of the RAGE 63bp ins/del polymorphism is not a 2 df test, since no del/del homozygotes are present.
Reply: We agree. Corrected P value (df=1) has now been incorporated in the table.

Level of interest: An article whose findings are important to those with closely related research interests
Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests: I declare that I have no competing interests
**Point by point reply to Reviewer II: Goisa Trynka**

**Major Compulsory Revisions**

1) Authors should specify if the calculated p-values are calculated as one or two Tailed.

Reply: We have used a 2 df genotypic chi-square test and calculated two tailed P value. However, as also pointed by the first reviewer, for seven of the nine polymorphisms the homozygotes of the minor allele are rare and thus a Fisher’s exact test (instead of a chi-square test) may be more appropriate. Therefore, in the Table 3 of the revised Ms. we have provided Fisher’s exact P value along with the P value for the Pearson’s chi square test.

2) Authors should take into account and state in the manuscript the p-values after correction for the number of tests performed e.g. the Bonferroni correction for the seven SNPs tested in this study.

Reply: Sorry for this oversight. We have now added a statement in this regard in the revised Ms. under results section (Page No. 9, Line No. 189-191).

3) P-values in the table 3 should be followed by ORs and confidence intervals.

Reply: OR and 95% CI values for the associated SNPs have been provided in the Table 3 of the revised Ms.

**Minor Essential Revisions**

1) page 5, introduction, 2nd paragraph. Authors mentioned that the SNPs investigated for the association in this study were chosen based on their significant functional role. I think more information should be added in the introduction or methods section, describing more precisely the criteria for selecting the particular SNPs for this study (e.g. Previously reported associations, bioinformatics predicted strength of the deleterious effect of the SNP on the protein product etc.)

Reply: In this study, SNPs were selected based on their functional relevance: SNPs leading to missense mutations, promoter polymorphisms, and those present in the regulatory regions (5' UTR and 3' UTR) were chosen in the same order of priority; and information content (minor allele frequency).

A statement, explaining the selection criteria for SNPs that were tested in the study has now been added in the revised Ms. (Please see page No. 8, line No. 157-159)
2) Is the polymorphism nomenclature used throughout the manuscript correct (according to the Human Genome Variation Society)? All except two of the tested polymorphisms have the rs ID numbers, it would be more convenient for the reader to use this nomenclature instead.

Reply: rs ID for all the polymorphisms investigated in this study has been provided in the Table 2 of the revised Ms.

3) Page 8, paragraph describing association to the GFPT2 gene. Authors state that the significant allelic and genotypic association of the 3'UTR polymorphism was detected. However the p-value for the genotype association, provided in the table 3 is p=0.11, which is not significant!

Reply: We are sorry for this mistake. The sentence has been corrected to in the revised Ms. (Please see page No. 13, Line No. 281-282).

4) The same paragraph. The authors state that according to the "Mirbase tool" the miRNAs (has-miR-378 and has-miR-4229) are targeting the sequence that captures variant (rs7725) tested in this study. Given that there are several bioinformatics tools available on-line that can be used for predicting miRNA-target site interaction I would require Authors to provide the reference to the Mirbase tool they have used.

Reply: Thank you for drawing our attention to this over-sight. In view of your query, we have now added the reference for the miRBase tool and the following text in the Methods section of the revised Ms. (Please see Page No. 9, Line No. 176-181)

“The miRBase Sequence Database is a searchable database of published miRNA sequences and annotation [miRBase::Sequences (http://microrna.sanger.ac.uk/sequences)]. Each entry in the miRBase Sequence database represents a predicted hairpin portion of a miRNA transcript (termed mir in the database), with information on the location and sequence of the mature miRNA sequence (termed miR).”

5) Page 10, 5th line of the 2nd paragraph; 'On the other hand CC genotype of -429 T>C SNP is predisposing (...)’ I would be more careful stating any strong effects of the SNPs tested in the study. Although the SNP that authors are referring to has an OR = 0.11 its confidence are very wide, CI 0.01-0.91, which of course can partly be explained by the small sample size.

Reply: We agree with the reviewer. Therefore, considering the concern raised by the reviewer, we have added a sentence for caution in the revised Ms. (Page No. 13, Line No. 270-273)

Discretionary Revisions
1) Regarding the discussion of the miRNA interactions with the sequence harbouring the rs7725 variation in the GFPT2 gene I was wondering if there is anything known about the genes targeted by has-miR-378 and has-miR-4229, eg. If there are genes known to be regulated by these miRNAs and they would fall into the same pathway as GFPT2 (also relevant for the CRI type 2 diabetes) it could significantly improve the discussion of the current manuscript.

Reply: Taking this brilliant suggestion of the reviewer a search was made in this direction. However, none of the genes which are implicated in diabetic chronic renal insufficiency/nephropathy were found to be regulated by has-miR-378 and has-miR-4229.

2) Page 11, first paragraph, the authors mention study in the African Americans where a significant association to the rs7725, T allele has been reported. It would be interesting to perform the meta-analysis of that study and the study described in this manuscript (and more if published) and see if the association holds.

Reply: True. However, to date, we have not received any response from the corresponding author of the above mentioned research article and thus a meta-analysis could not be performed.

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests