Reviewer’s report

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

**Version:** 1  **Date:** 25 June 2009

**Reviewer:** Ilja Nolte

**Reviewer’s report:**

Major Compulsory Revisions:

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 'VI762Ala 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.

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

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.

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?

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.

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. 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)?

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'.

Minor revisions:

Abstract p2: in the objectives 'eight' polymorphisms must be 'nine'.
Abstract p2: two SNPs were found significantly associated but only one OR is given.
Abstract p2: it is unclear where TGF-b1 comes from. This gene was not part of this study and hence needs further clarification.
Introduction p4: a reference is needed for the results of clinical trials (line 12-14).
Introduction p3: 'has' must be 'have' in line 10.
Methods p6: insert 'was' in line 7 ('informed consent was obtained')
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.
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.
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.
Methods, statistical analysis: they should describe and elaborate on the Mirbase tool that they used for miRNA target identification
p8. last line states (Ref) instead of the reference.
p8: D allele of ACE needs further clarification.
p9: confidence interval (0.04-0.14) is not correct
p10: the second paragraph starts with 'Two functional variants ... in this gene'. What gene is referred to?
p11: 'A study involving African-American subjects reported ...' requires a
reference.

p11: it is unclear what 'Gene bearing the 'T' (minor allele)' means.
p11: abbreviation ECM needs to be explained
p11: 'predisposing' in line 18 should be removed and 'which' should be added in the same sentence after 'TGF beta1 gene'
p11: 'protective towards' must be 'protective against'.
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.

Table 1: abbreviations UAER and GFR need clarification
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)
Table 2: not only give the substitution but the position as well if that is known. Eg. change Val>Ala to Val762Ala
Table 3: Heading is incorrect (different genes than analyzed).
Table 3: change PARP to ADPRT1
Table 3: give rs-numbers or other polymorphism coding instead of only the substitution
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.

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