Author's response to reviews

Title: Inflammation gene variants and susceptibility to albuminuria in the U.S. population: analysis in the Third National Health and Nutrition Examination Survey (NHANES III), 1991-1994

Authors:

Renee M Ned (RNed@cdc.gov)
Ajay Yesupriya (AYesupriya@cdc.gov)
Giuseppina Imperatore (GImperatore@cdc.gov)
Diane T Smelser (dtsmelser@geisinger.edu)
Ramal Moonesinghe (RMoonesinghe@cdc.gov)
Man-huei Chang (MChang@cdc.gov)
Nicole F Dowling (NDowling@cdc.gov)

Version: 2 Date: 21 September 2010

Author's response to reviews: see over
September 21, 2010

Melissa Norton, MD
Editor-in-Chief
BMC Medical Genetics

Dear Dr. Norton and editors of BMC Medical Genetics,

Thank you for consideration of our manuscript entitled “Inflammation gene variants and susceptibility to albuminuria in the U.S. population: analysis in the Third National Health and Nutrition Examination Survey (NHANES III), 1991-1994” for publication as a Research Article. We have responded to all reviewer comments/suggestions and have revised the manuscript accordingly. Please find our point-by-point responses on the pages that follow.

Once again, thank you for providing us the opportunity to publish our manuscript in BMC Medical Genetics. We anxiously await your response. Please contact me for any further information.

Sincerely,

Renée M. Ned, M.M.Sc., Ph.D.
Health Scientist
Centers for Disease Control and Prevention
OSELS/Office of Public Health Genomics
1600 Clifton Rd, NE
MS E-61
Atlanta, GA 30333

(404) 498-0125 (Office)
(404) 498-0160 (Fax)
Email: RNed@cdc.gov
Title: Inflammation gene variants and susceptibility to albuminuria in the U.S. population: analysis in the Third National Health and Nutrition Examination Survey (NHANES III), 1991-1994

Reviewer: Frederic Fumeron

Minor essential revisions

1) Page 9: usually, the codominant mode of inheritance means that each allele has an impact on phenotype so that heterozygotes have an intermediate phenotype, i.e. it is synonymous of “additive”. The assessment of risk by logistic regression for each genotype by comparison to a reference genotype does not assume any mode of inheritance. The most frequent genotype is not always the “wild type” genotype.
- The codominant mode of inheritance assesses individual genotypes, while the additive mode examines the per-allele effect. While heterozygotes may have an intermediate phenotype, this is not always the case. Consequently, the codominant mode is not exactly synonymous with the additive mode. We do agree that the most frequent genotype is not always “wild type”, and the text on pgs. 9-10 has been changed to reflect this.

2) Only few significant results were found and the associated variants in one group usually do not replicate in other groups. Provide power calculations taking into account the size of the different ethnic groups.
- Power calculations are now provided in the Discussion section, pg.17-18.

3) It has been shown for the “best” genes discovered by GWAS (TCF7L2 in T2D, FTO in obesity) that they replicate in many ethnic backgrounds. Are the results not replicating here due to a lack of power, or to interactions? (As stated by the authors themselves, other genes which may be of more importance need to be tested.) Although (or since) the results are significant for only one ethnic group, the authors may perform a kind of meta analysis which could evidence the effect of some SNP.
- The lack of replication may be due to differences in underlying linkage disequilibrium patterns between race/ethnic groups, untested interactions, and/or due to lack of power. We have added text that addresses this issue (Discussion section, pg. 21).
- We appreciate the suggestion of performing additional analyses to increase power to detect SNP effects since many variants were found significant in only one race/ethnic group. We performed analyses combining all participants that included an adjustment for race/ethnicity. These results have been added to the manuscript (pg. 14).

4) In the pdf, the last lines of table 1 are missing.
- Thank you. We will make sure that all Table rows and legends are visible in the submitted documents.
Reviewer: Lingyi Lu

Major Compulsory Revisions

1- Please discuss power in light of the allele frequency and provide some sort of post hoc power calculation.
   - Power calculations are now provided in the Discussion section, pg.17-18.

2- Please give a brief statement why choose additive model and co-dominant model instead of additive, dominant and recessive model.
   - We have added text to pg. 20 to outline some of the strengths and weaknesses of additive and codominant genetic models.

3- Please include a brief statement of the strategy in choosing snps from inflammation candidate genes for association test at the genes and polymorphisms section. How and why partial polymorphisms are selected from 27 genes? Do LD, allele frequency, or genotyping quality be considered as selection criteria? The authors indicate selecting snps rs1260326 and rs890945 from two published papers [ref35, ref36] however there are more associated snps in the two papers. Please give a brief description about why only two of them are selected.
   - We were able to include only variants that were genotyped in NHANES III samples that passed quality control/quality assurance guidelines from the National Center for Health Statistics (NCHS). We chose all inflammation genes and all available SNPs within these genes. There was no consideration for LD or allele frequency, but all variants passed QC/QA criteria as defined by NCHS. The two additional variants from refs 35 and 36 were added as they were the only ones from those studies that had been genotyped in NHANES III. Other variants in those studies were not available in NHANES III. We have added text to pg. 7 to clarify gene and SNP selection.

4- Recent admixture population of different ethnic groups with different alleles and disease frequencies may cause spurious association results between markers and complex traits. Non-Hispanic blacks and Mexican American can be considered as recent admixture population. The authors indicate that there are no ancestry-informative markers available to estimate the admixture proportion. So author should be more careful to present the association results in Non-Hispanic blacks and Mexican American population.
   - We agree that there needs to be added caution in interpreting the results in non-Hispanic blacks and Mexican Americans since we could not control for population substructure with ancestry-informative markers. We have added this language in the Discussion to reflect this (pg. 21).

Minor Essential Revisions

1- Risk alleles are nowhere presented throughout the report. Please include modeled allele as well as the frequency of the modeled alleles and include them into table2 to table 5.
   - Though prevalence estimates for the majority of the variants have been included in a previous report (Chang, et al. 2009), we concur that it would be informative to add them to this manuscript. Instead of adding prevalence estimates in multiple locations, we have
added them all to a new Table 2, which presents allele frequencies within each of the three included race/ethnic groups.

2- Please indicate each population (non-Hispanic whites, non-Hispanic blacks, Mexican American) sample size used in the analysis.
   • This information has been added to the Methods section (pg. 6). These sample sizes are already included in Table 1 and are now also in the revised Table 2.

Reviewer: Alexander Teumer

Major Compulsory Revisions:

1. The results of the association analyses of the SNPs and the outcome are not on line with the conclusions the authors made. Looking at tables 3 to 5, SNPs were assumed to be significantly associated with the outcome if their uncorrected p-value was below 5%. Due to the number of independent SNPs, independent population subgroups, and different genetic models that were used for association testing, correction for multiple testing has to be taken into account, even if not all SNPs were most probably completely independent from each other (what the authors correctly mentioned). The authors reported FDR corrected p-values, which were not significant at a 5% level for most SNPs, nevertheless, also those SNPs were reported as significantly associated with ACR or albuminuria, respectively. Corrected p-values should be the minimal basis for reporting an association in this case. So no SNPs were significantly associated for ACR after FDR.
   • We have revised the Abstract, Results, and Discussion to reflect the difference between marginally associated SNPs (with uncorrected P values <0.05) and those still significant after FDR adjustment.

2. The SNP rs1800750 was the only SNP that was significantly associated with albuminuria after FDR using the crude model. It was not associated in the adjusted model anymore. The authors should discuss this behavior. Furthermore, it was noted by the authors that this SNP was not in Hardy-Weinberg-Equilibrium for the tested ethnical subgroup. This is an important issue that could have resulted due to genotyping errors and should also be taken into account when reporting this SNP as associated.
   • We agree that results for this variant should be interpreted with caution due to its failure of Hardy-Weinbery proportions in non-Hispanic whites. We have revised the language in the Abstract and in the Discussion (pg. 16) to reflect this.

3. For other SNPs (rs1143623, rs1800947) there was no significant association for albuminuria after FDR in the crude model but they had an extreme low p-value in the full adjusted model. Could this effect be more likely by limitations of the logistic regression model regarding the type and number of covariates used instead of being a true association?
   • In age-(sex) adjusted models as well, the rs1143623 variant is significantly associated with each albuminuria outcome in Mexican Americans. This variant also has a high minor allele
frequency (~44%), so the regression models appear stable. For CRP rs1800947, however, we agree that such findings could be an effect of the limitations of the regression models, particularly due to its low minor allele frequency (2%). We have added language on pg. 17 regarding the CRP variant.

4. The remaining significantly associated SNP after FDR, rs2070744, had an association p-value that still needs confirmation in other independent samples, especially because it was not associated in the other ethnical subgroups.
   • We agree that all of our findings need confirmation in independent samples. However, we think it is useful to report significant findings within each race/ethnic group.

5. The authors should discuss why most of the SNPs were potentially associated only in one of the tested ethnical subgroups. Are there differences of the allele frequencies or haplotype structure of these loci among the subgroups that could explain this result? At least the allele frequencies of the SNPs and haplotype frequencies of the haplotype blocks in the ethnical subgroups used for association testing would be of interest and should be reported in a table. Does a meta-analysis or a combined analysis of the three ethnical subgroups reveal stronger associations?
   • We have added prevalence estimates for all variants to a revised Table 2. Haplotype frequencies have been added to Supplemental Table 7.
   • We have added text that addresses the lack of replication of findings across race/ethnic groups to the Discussion section (pg.21).
   • We also appreciate the suggestion of performing some sort of additional analyses to increase power to detect SNP effects since many variants were found significant in only one race/ethnic group. We performed analyses combining all participants that included an adjustment for race/ethnicity. These results have been added to the manuscript (pg. 14).

6. Due to quality control issues, for all SNPs the result of a test for Hardy-Weinberg-Equilibrium should be reported.
   • This information is now included in the revised Table 2.

7. Association results of an age and sex (where appropriate) only adjusted model would be informative and could help to find spurious associations due to limitations of the regression model and the type of covariates used.
   • Age and sex (where appropriate)-adjusted models have been examined for all three outcomes in each race/ethnic group. These data have now been added to the Supplemental Tables for each outcome.

8. It should be given a reference or discussed in more detail why hypertension and diabetes were not considered as covariates due to their involvement in causal pathways of chronic kidney disease, but e.g. waist-to-hip ratio and education were included as covariates.
   • We have added some discussion on this issue to pg.19-20.

Minor Essential Revisions:
Please describe briefly the function of the SAS/SUDAAN commands used to make the methods easier understandable for non SAS users.
• The Statistical Methods section has been revised to clarify the analyses performed, with the deletion of the majority of the SAS/SUDAAN commands.

The Satterthwaite-adjusted F-statistics used does not seem to be a commonly used F-statistics. Please explain the method briefly, esp. the advantages and possible limitations in this context or provide a reference.

• Text and a reference for use of Satterthwaite-adjusted F-statistics has been added (pg.10-11).