Author's response to reviews

Title: Chromosome 7p linkage and association study for diabetes related traits and type 2 diabetes in an African-American population enriched for nephropathy

Authors:

Tennille S Leak (leakt@edc.pitt.edu)
Carl D Langefeld (clangefe@wfubmc.edu)
Keith L Keene (kik4a@virginia.edu)
Carla J Gallagher (cgallagher@hmc.psu.edu)
Lingyi Lu (llu@wfubmc.edu)
Josyf C Mychaleckyj (jmychale@virginia.edu)
Stephen S Rich (ssr4n@virginia.edu)
Barry I Freedman (bfreedma@wfubmc.edu)
Donald W Bowden (dbowden@wfubmc.edu)
Michele M Sale (msale@virginia.edu)

Version: 2 Date: 18 December 2009

Author's response to reviews: see over
Reviewer: Nathalie Vionnet

- Major Compulsory Revisions:
1) It seems there has been a confusion between linkage and physical maps. In the discussion, the authors state that the linkage region is 38 Mb large whereas it is 38 cM large. This makes a difference when considering the selection of potential positional candidate genes. According to figure 1, the linkage region roughly spans the distance between D7S2514 and D7S2846 which are located at 11 cM and 60 cM respectively on the linkage map but at 7.7 Mb and 38 Mb respectively on the physical map. Three out of the 4 so-called positional candidate genes lie outside this interval (GCK at 44.2 Mb; IGFBP1 and IGFBP3 at ~45.9 Mb). Therefore it seems that in the context of positional candidate genes, the study of those 3 genes was not relevant. Please, provide the region of interest on a physical map scale, provide the number of genes lying in this interval and select positional candidate genes within this interval.

Response:
We selected plausible functional candidate genes that lie under the intersection of the optimal linkage peaks for early age of T2D, lower BMI and longer duration to ESRD. This regions roughly spans the distance between D7S513 and D7S1818 (11,517,762-49,459,633 bp), which corresponds to 38 Mb. The investigated candidate genes GCK1, IL6, IGFBP1 and IGFBP3 lie within this 38Mb region.

We have now included the following statement under Methods (p.7):
“Plausible functional candidate genes (IL6, GCK1, IGFBP1 and IGFBP3) that lie under the intersection of the optimal linkage peaks for early age of T2D, lower BMI and longer duration to ESRD was selected (Figure 1). This region roughly spans the distance between D7S513 and D7S1818.”

2) Linkage was confirmed in a subset of 21% of the families with early onset T2D. How do you explain that the mean age of diagnosis in this subset of families is 29 +/- 3 years despite the fact that T2D was diagnosed in patients developing diabetes after the age of 35 years. Could those families segregate MODY phenotype? The study of the GCK gene in this context, including sequencing, could be interesting.

Response:
While the proband was required to have diabetes diagnosed > 35 years, their sibling(s) recruited into the study could have been diagnosed at a younger age. We have clarified this in the Methods (p.5). Although we agree with the reviewer’s comment that sequencing of GCK may be interesting, unfortunately the majority of the 52 families that contributed to the linkage signal consist of small sibships without parental data, thus it would be challenging to distinguish genetically or clinically between MODY and T2D.

We also included the following statement in the Discussion (p.16): “In addition, several reports have shown that MODY genes segregate in late onset T2D cases and families[34-
In an effort to reduce the likelihood of including potential MODY2 patients, only individuals with T2D onset over the age of 35 were ascertained as probands.”

- Minor Essential Revisions

References 1 and 16 are the same.

In the article by Bonnycastle LL et coll (ref 32), the association between T2D and GCK markers was not significant with age at diagnosis.

Response:

We apologize for these errors that persisted from earlier drafts. Corrections have been made throughout the manuscript and adjustments have been made to the Discussion (p.16) as follows:

“T2D risk (recessive model: OR ranging from 1.36-1.87) in a case-control population from Finland [39]. Bonnycastle et al. [39] observed an 87% T2D risk with marker rs882020. Marker rs2908296 is approximately 14.7 kb from rs882020; it is plausible that these variants may tag or be in LD with the causal variant(s).”

The association with mean age of ESRD onset and GCK markers should be adjusted with age at diagnosis of T2D. The association with shorter duration of diabetes before onset of ESRD is marginally significant and would be contrary to the initial linkage finding with longer duration to ESRD.

Response:

We have made the following adjustments and results are presented in Table 2 and have also adjusted the Abstract (p.2) as follows:

“The most significant association was observed between ESRD age of diagnosis and SNP rs730497, located in intron 1 of the GCK1 gene (recessive T2D age-adjusted P=0.0006). Nominal associations were observed with GCK1 SNPs and T2D age of diagnosis (BMI-adjusted P=0.014-P=0.032). Also, one IGFBP1 and four IGFBP3 SNPs showed nominal genotypic association with T2D-ESRD. After correcting for multiple tests, only rs730497 remained significant.”

and Methods (p.10): “SNPs that showed nominal evidence for association with age of T2D onset were adjusted by BMI. Additionally, ESRD age, duration of T2D to ESRD onset and BMI were adjusted for age of T2D onset. All analyses were adjusted using linear regression on a priori genotypic models, and conducted using Stata 10 (College Station, TX)”
Reviewer: Katerina kankova

Reviewer's report:
Major Compulsory Revisions:
1. I think the level of significance in GWAS or multilocus studies is critical since many false associations are likely to be identified. Although authors did perform correction for multiple tests - which eliminated any significance at all - they conduct most of the MS without to get to this conclusion only at the end. The criteria for significance should be clearly stated beforehand and all text revised accordingly.

Response:
We agree that correction for multiple comparisons is key, especially in large scale studies such as GWAS. In the present study, we have performed a focused analysis of functional candidates with *a priori* evidence for a role in diabetes and/or nephropathy. Rather than apply the Bonferroni correction factor to all P-values, we have included the following statement (previously noted in the Abstract) under Methods (p.10) to clarify the significance level: “We report nominally significant associations at the P < 0.05 level. To correct for multiple tests at the gene level the conservative Bonferroni method was used, with a $P$-value ≤0.0007 considered significant evidence for association when assuming independence based on LD.” We feel that this approach allows the reader to evaluate the relative importance of the uncorrected P-values for associations in these genes, while making clear our significance threshold.

Minor Essential Revisions:
1. Result section should be make more clear - many analyses were performed and reader might have difficulties to follow the results.

Response:
We have clarified these points to make the results section more clearer.

2. I am confused by the numbers of SNPs analyzed - 68+70? There are some 11 SNPs mentioned in the Abstract?

Response:
To clarify the number of polymorphisms analyzed in the present study: 11 microsatellite markers were used to fine map the linkage peak on chromosome 7p, and a total of 68 SNPs were selected to test association with *GCK1* (n=24 SNPs), *IL6* (n=17), *IGFBP1* (n=16) and *IGFBP3* (n=11). In addition, 70 AIMs were used to estimate individual ancestral proportions and adjust for population substructure.

The following statement has been added to the Abstract (p.2): “…a total of 68 single nucleotide polymorphisms (SNPs) in…”