Author's response to reviews

Title: Genome wide association for substance dependence: convergent results from epidemiologic and research volunteer samples

Authors:

  Catherine Johnson (johnsoncat@intra.nida.nih.gov)
  Tomas Drgon (tdragon@intra.nida.nih.gov)
  Qing-Rong Liu (qrliu@intra.nida.nih.gov)
  Ping-Wu Zhang (pwzhang@intra.nida.nih.gov)
  Donna Walther (dwalther@intra.nida.nih.gov)
  Chuan-Yun Li (lichua@nida.nih.gov)
  Yulan Ding (yding@jhsph.edu)
  James Anthony (ianthony@msu.edu)
  William Eaton (weaton@jhsph.edu)
  George R Uhl (guhl@intra.nida.nih.gov)

Version: 2 Date: 29 October 2008

Author's response to reviews: see over
Dear Ms Whitaker,

Thank you for your recent communication indicating that our paper “Genome wide association for substance dependence: convergent results from epidemiologic and research volunteer samples” (MS: 9574760992124779) should be acceptable with attention to referees comments. We have addressed each of the referees’ comments, and your additional queries, as detailed below.

**Reviewer 1 (Lachman)**

1) *Many of the overlapping genes shown in the supplementary table do not show statistically significant differences between cases and controls. Their presence on the table is not clear.*

Response: We revise Table I. We now display only genes with $p < 0.05$ in the body of the table, while noting that large genes with localized association signals may be among those in the Table I legend that do not achieve nominal gene wise significance.

2) *… how do these findings relate to the group’s previously published results? A quick view shows that many, if not most, do not overlap. This reviewer realizes that the small sample size may affect results and that the primary aim was to show overlap in the different cohorts.*

   However, genetic data is being presented and it would be of interest to other investigators to explore the findings with a little more depth.
Response: We attempt to reduce confusion on this point by: 1) more clearly emphasizing to overall significance of the comparisons with prior GWA datasets in the body of the paper
2) emphasizing that the data in Table I comes from only part of the MNB samples (eg just European American subset) that has been previously reported.

3) …(the) method.. used to show that the overwhelming majority of targeted genes were expressed in the brain (primarily hippocampus) (was not clear)..

Response: We add detail to the description of the dbEST-based method for identifying brain expression of the genes identified by these GWA data in the Methods, (par 10), and also emphasize that a fuller discussion of this approach is currently in preparation.

Reviewer 2: (Everall):

1. (better describe the ECA sample to indicate how the numbers differ between) “ECA sample in the introduction.. 3,481…DNA sampled in 1,071 subjects… 662 European-American respondents (of) whom 85%, which equates to a sample of 550 subject provided DNA”. (better describe the MNB sample to indicate how the numbers differ between)” recruiting 3,800 volunteers… 34 pools of 400 “abusers” and 280 “controls”

Response: We eliminate a number of the numbers that obviously confused this reviewer, just focusing on the samples (eg only European American substance dependent and control that provided DNA) that were studied here.
2. “what are the authors examining as the manuscript states dependence but the MNB is comprised of abusers”.

Response: We reduce use of the shorthand “abusers” and replace it with “dependent individuals” or the equivalent wherever possible to reduce this ambiguity.

3. In the methods it is stated that DNA is prepared from blood or cell lines but in the description of the cohorts it is stated that DNA samples were obtained. Were the DNA samples obtained from blood or other tissue specimens?

“In addition there is a complete lack of description of the cell lines that the authors had access to. Were these immortalized lymphoblastoid cells? If so this need to be explained. I could not find description in the references either.

Response: We now clarify that MNB samples were DNA extracted from blood and that most ECA sample DNAs were extracted from lymphoblastoid cell lines (Methods, par 1).

4. In the methods it is stated that the authors utilized 100,000 Monte Carol simulation trials for SNPs of interest. However in the results it is stated that 25,000 or 10,000 Monte Carlo simulations were performed. Again the reviewer is confused by the lack of consistency.
Response: We add emphasis to our use of 100,000 Monte Carlo trials for clustering, 10,000 trials for convergence of the gene set and also for individual genes, and eliminate the typographical error that mistakenly referred to 25,000 trials.

5. The rationale for the comparison to brain libraries is inadequate... Were the authors examining particular brain regions or was this global brain cDNA libraries?....”In the results the authors state that they preferential brain and brain regional expression and refers to Table 1. However there is no information in this table on regional brain expression for the listed genes, again the reviewer is confused as to what the comparison exactly was with the brain cDNA libraries.

Response: As noted for referee 1, we add detail to the description of the dbEST-based method for identifying brain expression of the genes identified by these GWA data in the Methods (par 10), and also emphasize that a fuller discussion of this approach is currently in preparation.

6. The statistical p values listed in the results are inadequately explained and do not tally with the table.

Response: We add discussion to the Table I legend to emphasize the differences between the “overall convergence” and “individual gene” Monte Carlo results (Table I, legend).
7. There was no attempt at discerning any potential biological meaning of the 172 genes of interest. Do they populate particular GO categories that may implicate certain functional processes in substance dependence?

Response: We add to the discussion (par 3) the idea that, while the discussion of all classes of identified genes is beyond the scope of this article, cell adhesion related genes appear to be overrepresented, as we have previously observed.

Ethics concerns from the Editor: - Please clearly explain the ethical approval that covered your research.

Response: Approval of appropriate IRBs (and notation that pooled genotyping is ruled “exempt” by the NIH Office of Human Research Subjects Protection) are now included in Methods, pars 2 and 3.

We believe that these changes strengthen the manuscript, and hope that it is now acceptable for publication in BMC.

Regards

George Uhl MD PhD