Author's response to reviews

Title: Correlations of Gene Expression with Ratings of Inattention and Hyperactivity/Impulsivity in Tourette syndrome: A Pilot Study

Authors:

Yingfang Tian (yftian@ucdavis.edu)
Joan Gunther (joan.gunther@ucdmc.ucdavis.edu)
Boryana Stamova (boryana.stamova@ucdmc.ucdavis.edu)
Bradley Ander (bradley.ander@ucdmc.ucdavis.edu)
Glen Jickling (glen.jickling@ucdmc.ucdavis.edu)
Blythe Corbett (blythe.corbett@ucdmc.ucdavis.edu)
Netty Bos-Veneman (netty@bos-veneman.nl)
Pieter Hoekstra (p.hoekstra@accura.nl)
Julie Schweitzer (julie.schweitzer@ucdmc.ucdavis.edu)
Frank Sharp (frank.sharp@ucdmc.ucdavis.edu)

Version: 2 Date: 14 August 2012

Author's response to reviews: see over
August 14, 2012

RE: Major Revisions- MS: 5148098517252286

Dear Editor:

We thank you and the reviewers for the opportunity to revise our manuscript entitled “Gene expression correlated with Ratings of inattention and hyperactivity/impulsivity in Tourette syndrome: A pilot study”, which we previously submitted to BMC Medical Genomics. We have made the revisions suggested by the reviewers. These are included in the attached manuscript. Our responses to the reviewer comments are listed below.

Reviewer comments are in Italics.
Our Response is in black text.
Changes in the text in the manuscript are <<“ ”>>

Reviewer's report-1
Title: Correlations of Gene Expression with the Conners’ inattention and hyperactivity/impulsivity rating scales in Tourette syndrome: A Pilot Study
Version: 1 Date: 18 June 2012
Reviewer's report:

Study summary: The authors explored the possible relationship between gene expression in peripheral blood and the traits of impulsivity and hyperactivity in 21 children with Tourette Syndrome who were not formally diagnosed with ADHD. RNA was isolated from the blood of 21 TS subjects, and gene expression measured on Affymetrix human U133 Plus 2.0 arrays. An analysis of covariance found 1201 gene probesets that correlated with IA scales, 1625 that correlated with HI scales, and 262 that correlated with both IA and HI scales (p<0.05. Twenty-seven of the identified genes have been previously reported in ADHD genetic studies.

Minor Essential Revisions
1. The authors discuss the hypothetical mechanisms supportive of the use of gene expression in peripheral blood, rather than brain tissue. They should mention that these are largely hypothetical. Although receptors are present in CNS and periphery it does not necessarily follow that these are expressed in similar manner as the genetic environment may be quite different.

Response: We agree with the reviewer. The following sentence was added to the second paragraph of the discussion.
These mechanisms are hypothetical since the current studies cannot gauge what the relationship between blood and brain gene expression might be, particularly given the different genetic influences in blood compared to brain.”

2. Age effect is not mentioned and yet it is also worthy of discussion. Expression of various genes especially in the CNS would be quite different at age 7 vs. 15 What is the evidence that the age effects are similar in the CNS and the periphery?

Response: We agree with the reviewer that the gene expression may be affected by age since TS symptoms in childhood vary with age. Our previous studies identified the age-related gene expression in TS (Lit L, 2009). Thus age was considered in the ANCOVA model as described in the methods in current study. None of the genes related to inattention and hyperactivity in this study were shown to be related to age either in control or TS subjects in our previous studies.


3. Reference No. 8 is incomplete
Response: Reference No. 8 has been corrected as suggested.

4. It may be helpful if the authors offer guidance what would constitute a statistically acceptable sample size with sufficient power.

Response: We have not found any reference about how to calculate sample size on correlations between numerical factors with gene expression based on microarray data. Here we provide the guidance how to perform power analysis for effect sizes - when we were reporting a list of differentially expressed genes between groups.

The idea underlying power analysis is that we will eventually detect differentially expressed genes by performing separate gene-by-gene t-tests. The correct significance level, alpha, for these gene-by-gene tests depends both on the number of genes we are looking at on the microarray and the number of false positives we are willing to accept. The sample size also depends on how large a difference we want to be able to detect (typically measured as the fold difference between the two kinds of samples). It also depends on the power of the test, which we can think as the percentage of the differentially expressed genes that are likely to be detected by the experiment. The sample size can be calculated from the following website:

http://bioinformatics.mdanderson.org/MicroarraySampleSize/

Owing to the costly nature of microarray experiments, however, we chose a small sample size to show feasibility of the approach, and thus have titled this a pilot study. If the findings from a small study look promising, a large-scale study may be developed to confirm the findings in the future. Thus we point out in the title, abstract and limitations that our study is a pilot study, and needs to be confirmed in future studies with a large sample size.
Reviewer's report-2

Title: Correlations of Gene Expression with the Conners’ inattention and hyperactivity/impulsivity rating scales in Tourette syndrome: A Pilot Study

Version: 1 Date: 6 July 2012

Reviewer's report:

Major Compulsory Revisions

In this manuscript, Tian and colleagues aimed to identify peripheral gene expression associated to inattentive and/or hyperactive/impulsive rating scales. The authors’ succeeded in identifying several probesets significantly correlated to the scales and these probesets included genes and pathways previously identified as being associated to ADHD. None of the probesets passed FDR, nor were there additional validations or confirmations of these results. Further, the significance of the results given that the subjects are primarily diagnosed or ascertained for Tourettes is not clear.

Response: The largest limitation of this study, as the reviewer points out, is that none of the probesets passed FDR, nor were there additional validations or confirmations of these results. On the other hand, even with the small sample size this study was able to confirm possible involvement of 27 genes in ADHD that were the subject of over 20 previous studies (Table 2). Thus, though power was severely limited here, the data remarkably confirm many other previous studies. We suggest that in this case, the failure to pass FDR does not impact on the current results. In addition, given the many different platforms used in the other 20 previous studies, the current results suggest that the confirmation is likely platform independent. In addition, the data support the idea that many genes are involved in ADHD symptoms, and a given gene may only contribute a small percent to the symptoms (Comings et al., 2001, Franke et al., 2009, Faraone et al., 2010). This could explain the modest association between a single gene and ADHD symptoms. Thus, we also focused on pathways that are more likely to be confirmed in future studies, and confirmed a role for the integrin pathway proposed in a recent ADHD meta analysis. Importantly, a gene co-expression analysis did validate these pathway-related ADHD genes in our study (Supplementary Figures 1-3). Moreover, our gene-gene correlation results demonstrate that the multiple probesets targeting a specific gene on the Affymetrix human U133 plus 2.0 arrays were highly correlated each other (r>0.9) (Supplementary Table 3). This is now explained in the Limitations section.

We agree with the reviewer these microarray data should be confirmed using an independent method such as RT-PCR. This is now stated in the Limitations section. The validation was not performed because of limited amounts of RNA. However our own data (Gregg et al., 2008,) and reviews from the literature show a >90% concurrence for microarray and PCR data (Shi et al., 2008,). Lastly, as noted above, the current study confirms involvement of 27 genes in over 20 previous studies many of which used different platforms for the most part. Nevertheless, we added the following statement in the discussion in accord with the reviewers’ comments.

<<“The largest limitation of the study is that, in spite of that many genes correlated HI and/or IA symptoms, no gene passed multiple comparison correction testing using the Benjamini-Hochberg False Discovery rate (FDR<5%), and none of them confirmed using an independent method such as RT-PCR. Thus, a future confirmatory study with a larger sample size is needed to validate the candidate genes reported here.” >>

We chose the participants which were primarily diagnosed for TS in current study for several reasons. (1) Patients with TS often display comorbid symptoms of ADHD. Of subjects with TS who visit a physician, as many as 50-80% have comorbid ADHD, a rate that is 10 to 20 times that of the general population. (2) Examining these behaviors in participants with TS might provide more homogeneous
phenotypes since TS is highly heritable, is readily and objectively identifiable. (3) We have a large cohort of TS subjects with ADHD symptoms that made the study possible.

However, to address this issue raised by the reviewer, we added the following statement to the first paragraph in the Limitation section: “It is not known if the genes associated with IA and HI symptoms in the TS subjects could be replicated in general populations of children with ADHD. Given that many genes overlapped between IA and HI symptoms in participants with TS, some of these might also overlap in subjects with ADHD without TS.”


Additional specific points raised in the review are:
• Were blood cell counts significantly correlated to gene expression, and the rating scales used? Response: We did not address this question in this manuscript since blood cell count data were never obtained. This will be addressed in future studies. Virtually all gene expression studies of blood in psychiatric and neurological disorders to date have been performed on whole blood, and for the most part the results do not appear to mirror counts of an individual cell type. However, this has not been studied well in neuropsychiatric conditions and requires further investigation. We added a short caveat in the discussion to address this. “Thus, a future confirmatory study likely including RT-PCR and possible corrections for blood cell types in a much a larger sample size will be needed to validate the genes reported here”.

• The authors do not mention carrying out outlier detection for their subjects, which could be easily carried out by performing a principal components analysis. For the subject who stopped taking atomoxetine 40 hours before participation, what is the half-life of the medication? It would be interesting to see if this subject was an outlier with respect to the others because if by chance he/she was a poor metabolizer there may still have been small amounts of the medication in the bloodstream.

Response: We agree with reviewer and have performed the Principal Component analysis (PCA). This is now mentioned in the methods and in the discussion. The PCA did not show any outliers among the TS subjects studied here. Our previous study indicated that these two participants with medication were not outliers in behavioural fMRI analyses as well (Baym CL, 2008). Thus, these two subjects were
included in the current study. We acknowledge that the medication might affect blood gene expression, and this should be addressed in the future research. We added the following sentences to the Limitations section to address these questions:

>> “Two participants who had been previously prescribed medication were included in the current study. Our Principal Components Analysis (not shown) revealed that there were no outliers in gene expression data, suggesting these two individuals did not significantly bias the correlations observed. Moreover, our previous studies including these individuals did not show them to be outliers with regard to fMRI findings or alternative splicing [11,15]. Nevertheless, the fact that prior medications might have affected blood gene expression should be addressed in future research.”


• Although the authors did not identify predictors of TS or ADHD, it would have been interesting to confirm whether their findings could lead to diagnoses of ADHD subclasses in their Tourette’s subjects since this was not carried out as stated in the Methods section.
  Response: We agree entirely, and this might be one of the interesting directions for this research to go in the future. This was not performed because of the very small sample size, and because this would represent a cross-validation of sorts on the same sample rather than true prediction of ADHD subclasses. A future much larger study will need to be performed to address this question.

• In the Discussion, to make things a bit clearer, the authors should re-write the “How gene expression…” section, on page 10. Specifically, the authors’ need to make sure the sentences follow each other since there is some disconnect and the ideas are not conveyed as clearly as they could be.
  Response: We have revised the paragraph as requested by the reviewer. We are hopeful the ideas are more clear and logical in the revised section.

• The discussion could benefit from the addition of references of the similarity between brain and blood gene expression in general, as well as those pertaining to specific neuropsychiatric disorders.
  Response: We added a reference of the similarity between brain and blood gene expression in the manuscript as reference 11 as below:


• The authors’ could have considered carrying out some filtering prior to performing any analyses, such as filtering based on present or absence calls, this would have reduced the number of probesets and might have led to some passing multiple testing correction.
  Response: We did carry out filtering which reduced the numbers of probesets on Affymetrix U133 plus 2.0 from about 54,000 to 36,000, but still did not lead to major change of the FDR-corrected P-values.
Reviewer's report-3

Title: Correlations of Gene Expression with the Conners' inattention and hyperactivity/impulsivity rating scales in Tourette syndrome: A Pilot Study

Version: 1  Date: 4 July 2012

Reviewer's report:

This is an interesting manuscript on the field of molecular genomics of neuropsychiatric disorders. It is a well written manuscript, reporting novel findings. The impact of the manuscript will benefit from some improvements:

Major Compulsory Revisions
1. Authors should submit raw microarray data to a public online repository (NCBI GEO or EBI ArrayExpress, for example), following international recommendations in the field.

Response: We have submitted our data to GEO. This is described in the Data analysis section as below:
<<"We deposited the raw data at GEO under accession number GSE30470 and can confirm all details are MIAME (Minimum Information About a Microarray Experiment) compliant.”>>

Minor Essential Revisions
2. Authors should discuss about the possible confounding effects of inclusion of two patients taking medicaments.

Response: The two subjects who had been treated prior to entering the study are now discussed as above (Question 3 of reviewer 2).

3. Authors should expand on the description of the genetic basis of TS (mutations in SLITRK1 in a subset of patients, among others).

Response: We agree with the reviewer, and now briefly mention some of the genetic bases of TS:

<<“Cytogenetic and linkage analyses have uncovered a number of loci and several genetic mutations that are associated with Tourette syndrome. For example, mutation in SLIT and NTRK-like 1 (SLITRK1 ) can cause TS, and though there are other examples, each only accounts for a small fraction of cases[9, 10]”>>


4. Authors should provide a more detailed explanation of the contents and advantages of the Conners’ scales.

Response: We agree. The following statement was added to the Introduction:
<<“Thus, the current study quantified IA and HI behaviors using the Conners’ Parent Rating Scales-Revised (CPRS) in a group of subjects with TS. The well-validated Conners’ scale is widely used in research and clinical practice to diagnose ADHD and evaluate treatment effects in the disorder [3].”

The following was added to the Methods section <<“The CPRS-R: S was used to assess ADHD symptoms using continuous, standardized age and gender adjusted CPRS-T scores. The parent ratings are useful and valid as they have the opportunity to observe their children over extended periods of time and in a variety of situations. The scale contains 27 items and is composed of 4 subscales including: Cognitive Problems/Inattention, Hyperactivity, Oppositional and the ADHD Index [3]. A major advantage of the CPRS is that it uses a very large normative database (8,000+ children) to support the validity and reliability of it. Furthermore, the standardized data from the CPRS were derived from the means and standard deviations for children with and without ADHD.”>>

5. Authors should discuss other papers that explore the correlations of gene expression between brain tissues and blood (PMID: 16526044, for example).

Response: We added this reference (number 11).

6. Exploration of genomic data would benefit from other kinds of data mining approaches: Focus on genes known to be brain-expressed, integration with large datasets from GWAS for ADHD, exploration of KEGG derived pathways, among others. They could use freely available bioinformatic tools, such as DAVID or Babelomics.

Response: We agree. Table 2 shows the summary of regulated genes in our study and how they confirm those in previous studies – including those genes known to be brain expressed, those found in large GWAS studies, and those implicated in various pathways related to ADHD. IPA was used for pathway analysis simply because of its availability at our center. Those wishing to use DAVID, Babelomics, Egan and many other pathway programs should be able to do so since our entire data set is included in GEO, and specific data is included in Supplementary Table 1.

7. Supplementary files could be grouped in two files, to facilitate readers their downloading and visualization.

Response: Supplementary files have been grouped into three files. Supplementary Table 1 is ADHD symptom-related gene lists (Supplementary Table 1-3 in previous version). Supplementary Table 2 is over-represented pathways and chromosomes (Table 1and Supplementary Table 4 in previous version), and Supplementary Table 3 is co-expression results (Supplementary Table 5 in previous version).

8. Table 1 could be modified, focusing on significant pathways, leaving the extended lists of associated genes for the supplementary files.

Response: We have revised the Table 1 as the reviewer suggested. The revised table focused on significant pathways, and associated genes were put in Supplementary Table 2-1
9. Authors should revise the following statement “Due to the large numbers of probes on the microarrays no gene passed multiple comparison correction testing using the Benjamini-Hochberg False Discovery rate (FDR<5%) approach” on the Limitations section.
Response: We have revised the sentence in the Limitations section as below:

<< “The largest limitation of the study is that, in spite of that many genes correlated with HI and/or IA symptoms, no gene passed multiple comparison correction testing using the Benjamini-Hochberg False Discovery rate (FDR<5%), and none of them were confirmed using an independent method such as RT-PCR. Thus, a future confirmatory study with a larger sample size is needed to validate the genes reported here. “>>

Reviewer's report-4
Title: Correlations of Gene Expression with the Conners' inattention and hyperactivity/impulsivity rating scales in Tourette syndrome: A Pilot Study Version: 1 Date: 21 June 2012
Reviewer's report:

1. The article may be written more concisely.
Response: We have revised the manuscript to make it more concise.

2. All the statistical symbol “p” in the paper should be changed to “P”.
Response: All of the statistical symbols have been changed to “P”.

3. The gender and ethnic origin of the subjects should have been described in a more specific way.
Response: The sentence describing gender and ethnicity was revised: << “There were 17 males (81.9%) and 4 females (18.1%), including 15 persons identifying themselves as Caucasian (71.5%), 2 Hispanic (9.5%), and 4 as Other ethnic category (19.0%).”>>

4. There are no normal controls or controls with ADHD only in the experiment design.
Response: This study was specifically designed to determine if gene expression correlated with IA and HI (ADHD symptoms) in subjects with TS in order to focus the study on a clearly identified clinical neuropsychiatric condition where as many as 50-80% of the TS subjects have ADHD symptoms. We did not include normal controls and subjects with ADHD only because of the pilot nature of the study. In addition, true normal controls exhibit few IA and HI symptoms. The controls would have low scores on the Conners’ rating scale which would not display a normal distribution, with insufficient variability in their scores, and thus, they are not an ideal control group to address the current question. ADHD alone was not included in part because the genes that correlate with IA and HI might or might not overlap with those that occur with ADHD symptoms in TS subjects. In addition, this study avoided making the “clinical diagnosis of ADHD” so that subjects were not placed into some arbitrary clinical subgroup. We agree that including all three groups (TS, ADHD alone, controls) would be interesting in order to compare the results of all three groups. We submit that this is a pilot study, it shows the feasibility of performing these types of studies, and would form part of the rationale for performing much larger studies including all of the groups in the future.
5. The scale of the sample is too small to draw the conclusions. 
Response: We agree with the reviewer that the sample size is small and as stated in the Abstract and Limitations section, it is to be regarded as a pilot study. Even through the pathway and co-expression analysis was performed to decrease the false positive results, the results need to be confirmed with a much larger sample size in the future (Question 1 of reviewer 2). However, given these limitations, the study did confirm the possible involvement of 27 genes previously reported in the literature to be associated with ADHD (Table 2). We submit that this ameliorates the small size of the current sample, it demonstrates the potential usefulness of a gene expression approach in TS and ADHD, and helps support this as a stand alone pilot study.

6. The format of Table 1 is not accord with the requirement of BMC medical genomics. 
Response: The format of Table 1 was revised to focus on significant pathways, leaving the extended lists of associated genes as supplementary Table2-1.

Sincerely yours,

Yingfang Tian, Ph.D.
Department of Neurology and the M.I.N.D. Institute
University of California at Davis
2805 50th Street, Room 2434
Sacramento, CA 95817 Tel: 916-703-0384