Author's response to reviews

Title: Sequence variants in oxytocin pathway genes and preterm birth: a candidate gene association study

Authors:

Jinsil Kim (jinsil-kim@uiowa.edu)
Kara J. Stirling (jursak01@gmail.com)
Margaret E. Cooper (mec9@pitt.edu)
Mario Ascoli (mario-ascoli@uiowa.edu)
Allison M. Momany (allison-momany@uiowa.edu)
Erin L. McDonald (erin-l-mcdonald@uiowa.edu)
Kelli K. Ryckman (kelli-ryckman@uiowa.edu)
Lindsey Rhea (lindsey-rhea@uiowa.edu)
Kendra L. Schaa (kendraschaa@gmail.com)
Viviana Cosentino (vivicosentino@hotmail.com)
Enrique Gadow (egadow@cemic.edu.ar)
Cesar Saleme (csaleme@colemed.com.ar)
Min Shi (shi2@niehs.nih.gov)
Mikko Hallman (mikko.hallman@oulu.fi)
Jevon Plunkett (plunkett@stanford.edu)
Kari A. Teramo (kari.teramo@hus.fi)
Louis J. Muglia (louis.muglia@cchmc.org)
Bjarke Feenstra (fee@ssi.dk)
Frank Geller (fge@ssi.dk)
Heather A. Boyd (hoy@ssi.dk)
Mads Melbye (mme@ssi.dk)
Mary L. Marazita (marazita@pitt.edu)
John M. Dagle (john-dagle@uiowa.edu)
Jeffrey C. Murray (jeff-murray@uiowa.edu)

Version: 3 Date: 9 March 2013

Author's response to reviews: see over
February 28, 2013

The Editor
BMC Medical Genetics

Dear Editor,

We appreciate the constructive comments on our manuscript entitled “Sequence variants in oxytocin pathway genes and preterm birth: a candidate gene association study”. We have carefully considered all the suggestions we received to improve the quality of our manuscript. Below we provide our point-by-point response to each comment.

Response to Reviewer 1

1. Comment #1: The results section should be reduced in length by no less than 50%. The focus of the manuscript should be on the resequencing efforts, rather than the initial 16 SNP genotyping effort (which should be eliminated from the manuscript altogether).

Comment #2: The results should begin with the sequencing efforts, describing the identified variants, highlighting those that are novel and potentially pathogenic.

Response to Comments #1 and #2: It is true that most of the data presented in the Results section were obtained through resequencing efforts, which may make the SNP genotype data appear relatively less important. However, as stated in the text, the SNP genotyping study (which was performed on three selected candidate genes, one (LNPEP) of which was identified as being associated with preterm birth), served as a basis for expanding our efforts to find rare variants in the other two candidate genes (OXT, OXTR) by sequencing.

Not only did the SNP analysis set the stage for resequencing efforts, but it also revealed a potential contributing gene, LNPEP. As discussed in the text, there is already some biological evidence supporting the relevance of the gene to the disease, which is strengthened by our genetic data. Therefore, we think that sharing our findings would benefit the scientific community, allowing researchers in relevant fields to conduct follow-up studies of the LNPEP gene and its variants in the context of preterm birth.

For these reasons, we did not eliminate the results of the common variant (SNP) analysis from our manuscript, and those results, as in the original manuscript, were placed in the first part of the Results section in order to make a logical transition to the rare variant analysis data (that is, the resequencing data).

We, however, tried to present the resequencing data as concisely as possible and reduced the length of the Results section by about 20-25%. We summarized the variants identified in a table (Table 3) and highlighted those that are novel, potentially pathogenic, and/or noteworthy in other aspects. Please note that we have updated the Results section to
reflect new information on genetic variation made available to public databases (such as dbSNP) since our submission of this manuscript and as a result, we made a few changes in our description of some of the variants.

2. **Comment #3**: Discuss the findings considering all publicly available data (dbSNP and 1000 genomes) simultaneously. There is no need to have multiple discussions of each source of data

**Response**: At the time of our submission of this manuscript, the data from the 1000 Genomes (1000GP) project for the region of our interest were largely not available in dbSNP. However, while this manuscript was being reviewed, the dbSNP Build 137 (the latest Build) containing new submissions from the 1000GP phase I release has become available, now making it unnecessary to differentiate the two sources of data.

However, there was and is a big difference in the list of OXTR coding variants between the 1000GP and NHLBI Exome Sequencing Project (NHLBI_ESP) data with the NHLBI_ESP list containing a much larger number of coding variants, compared to the 1000GP. Due to this fact, we, for the purpose of clarity, mentioned a specific source of data (either the 1000GP or the NHLBI_ESP) just in one part of the Results section where we comparatively discussed the variants from our study and those from the public databases. Wherever possible in the manuscript, however, we discussed our findings without distinguishing between different sources of data.

3. **Comment #4**: Formal statistical comparisons of table 2 findings (The core of the paper) should be clearly outlined, with careful attention not to mix stratified populations (i.e. a combined analysis of Argentina, US, Fins altogether is not appropriate and should not be done).

**Comment #5**: Formal case control comparisons should be performed using methods developed for rare variant analysis.

**Response to Comments #4 and 5**: We appreciate this valuable comment. In our original manuscript, we did not perform formal statistical analyses on the findings shown in Table 2 (Table 3 in the revised manuscript) because of the small samples sizes of the populations examined. However, in a revised manuscript, as per your suggestion, we performed a formal association analysis on each population using a rare variant analysis method. We summarized the results from these analyses in Table 5.

4. **Comment #6**: Definitive conclusions should be added to both the abstract and the discussion. What do the genetics tell us that we did not already know. Does the genetic data strengthen the importance of the gene or not?

**Response**: We are grateful for this helpful suggestion. We added definitive conclusions to the Abstract as well as the Conclusions section.
Response to Reviewer 2

Major Compulsory Revisions

1. Comment: Formatting the paper in a standardized fashion: Abstract, Background, Methods, Results, Discussion and Conclusion to optimize readability.

Response: We formatted our manuscript to make it conform to the journal’s guidelines, and it is now organized in the order of Abstract, Background, Methods, Results, Discussion, and Conclusions.

2. Comment: Portions of the last paragraph of the Background section contain information that are better suited for the methods section (study details and statistical analyses).

Response: We revised the last paragraph of the Background section (on pages 6-7) so that it only briefly describes what is being reported in the current manuscript without going into too much methodological detail.

3. Comment: The last paragraph of page 14 also contains information that is best suited in the Background section to give the reader sufficient information regarding the defining characteristics and rationale for inclusion of LNPEP as a gene candidate.

Response: Part of the information on the LNPEP gene in the last paragraph on page 14 was integrated into the Background section (the first paragraph on page 6).

4. Comment: Clearly describing the phases of the study and the patient population (and demographics) in both the Methods and Results sections. For example defining the phases of the study (phase 1-discovery; phase 2-replication) and the sample sizes, demographics, and exclusion and inclusion criteria for each. In fact, some of this information may be best accomplished with a figure/table.

Response to Comments #4 and 5: We appreciate the reviewer’s comment. We clearly defined the phases (phase 1 and phase 2) of our study in the revised manuscript. Also, we summarized information about the study populations (sample sizes and demographics) in a table (Table 1) and stated exclusion and inclusion criteria for each population (including the U.S. population) in the Methods section (under “Study population” on pages 7-8).

Minor Essential Revisions
1. **Comment #1**: The last paragraph on page 12 is better suited in the Background and/or Discussion as appropriate.

**Response**: The information in the last paragraph on page 12 was incorporated into the Background section (the first paragraph on page 6).

2. **Comment #2**: Adding a statement in the Abstract and Results regarding the potential contribution that these variants may effect on the incidence of preterm birth (i.e. how much of the variance in preterm birth is explained by the identified sequence variations).

**Response**: Wherever appropriate and possible, we discussed about the potential contribution of the variants examined to the incidence of preterm birth. For example, on page 18 in the Results section, in reference to Table 6, we stated: “The results indicate that 2 SNPs significantly explain 5.5% of the variation in GA and 7.3% of the variation in BW with rs61740241 in a slightly more supportive role to the other SNP. The SNPs that best predict case-control status are identical to the SNPs involved as the main predictors of GA in weeks and BW in grams”.

3. **Comment #3**: In the conclusion, consideration should be given towards placing a statement regarding:
   a. The potential for future study of these genetic variations in women who deliver at term (when the oxytocin pathway is more likely to have a causative effect).
   b. The observed effect being primarily within the maternal unit.

**Response**: We appreciate this thoughtful comment of the reviewer. We added a statement regarding the potential importance of maternal effects in preterm birth to the Conclusions section (on pages 23-24).

We deeply appreciate the opportunity to revise and resubmit our manuscript. We hope that we have satisfactorily addressed the reviewers' comments and that the revised manuscript is now suitable for publication in BMC Medical Genetics.