Author's response to reviews

Title: 118 SNPs of Folate-Related Genes and Risks of Selected Congenital Anomalies

Authors:

Gary M Shaw (GShaw@marchofdimes.com)
Wei Lu (wlu@ibt.tamhsc.edu)
Huiping Zhu (hzhu@ibt.tamhsch.edu)
Wei Yang (WY@marchofdimes.com)
Farren BS Briggs (fbriggs@genepi.berkeley.edu)
Suzan L Carmichael (SCarmichael@marchofdimes.com)
Lisa F Barcellos (barcello@genepi.berkeley.edu)
Edward J Lammer (elammer@chori.org)
Richard H Finnell (rfinnell@ibt.tamhs.edu)

Version: 3 Date: 9 April 2009

Author's response to reviews: see over
April 9, 2009

BMC Genomics
Publications Department

Dear Editor in Chief;

Please consider our revised manuscript, “118 SNPs of Folate-Related Genes and Risks of Spina Bifida and Conotruncal Heart Defects” and also attached is the point-by-point response to the reviewers’ concerns for publication as an original article in BMC Genomics.

There are no prior publications or submissions with information that overlaps with this submitted manuscript.

This manuscript, or any of its contents, is not and will not be submitted to any other journal while it is under review by BMC Genomics

There is no potential conflict of interest related to all or parts of this manuscript.

Each author listed on the manuscript has seen and approved the submission of this version of the manuscript and takes full responsibility for the content of this manuscript. All authors participated in study design, laboratory analytic plan, epidemiologic analytic plan and writing of results.

Thank you for the continued consideration of our manuscript for publication in BMC Genomics.

If you should have any questions or concerns, please contact me at 510-597-7043 or you may email me at GShaw@marchofdimes.com. Thank you for your continued consideration of this paper.

Sincerely,

[Signature]

Gary M. Shaw, DrPh
Epidemiologist
First Reviewer’s report  (Jean-Louis Gueant)

Comment
The authors present a retrospective study evaluating case control association between 118 SNPs of genes related with folate pathway and the risk of spina bifida and conotruncal heart defects, using a SNPlax assay. The sample size is rather big with respectively 259 cases with spina bifida, 214 cases with conotruncal defects and 359 controls, recruited before folate fortification in US. The choice of genes is relevant. It includes genes with already described associations, such as MTHFR, MTRR, MTR, BHMT and CBS, with a rather exhaustive analysis of new SNPs. The haplotype analysis underlines the importance of gene variants involved in the re-methylation pathway of homocysteine. In conclusion, an important study, which could be improved by a minor revision

Major Compulsory Revisions
1) A calculation of population size could be included in method section, in function of study power, allele frequency and expected increased frequency in cases. In the “limitations” in last paragraph of page 9, the sample size effect should be considered according to this calculation.

Reply
We have added a phrase to the second to last paragraph of the Discussion that reads, “For example, our study had 80% power to detect risks of 2.5 or more associated with genotypes that were observed in at least 4% of controls.”

Comment
2) The race/ethnicity is not matched between Spina bifida and controls (P<0.0001). This limitation concerns particularly MTHFR (dramatic differences of frequency among Hispanics, Caucasians and Afro-American. For example, see Guéant-Rodriquez, Am J Clin Nutr, 2006). This comparison between controls and SB cases should be revised using a subgroup of matched controls.

Reply
The reviewer is correct that frequencies of genotypes vary by race/ethnicity. This is what motivated us to analyze our data with statistical adjustment for race/ethnicity. As we note in the Results section we did not observe evidence to indicate risk patterns were confounded by race/ethnicity. Moreover, analyses involving haplotype blocks were carried out specific to race/ethnic groups. Therefore, potential effects of population structure have been minimized in our analyses.

Comment
3) Correct the analysis for multiple comparisons.

Reply
We do not believe that correcting the analysis for multiple comparisons is necessary. Our goal was to describe the risk patterns for a large number of genotypes in a
population-based set of cases and controls. We do not make strong inferences about any particular finding – thus, correcting the observed confidence intervals for the number of tests performed adds very little to the overall message that we are trying to convey in our manuscript.

Comment
4) The citation of the literature should be revised. For example, concerning the evaluation of MTRR, MTR, MTRR with Spina bifida, two case control studies have observed a significant association with MTRR in Italy (Guéant Rodriguez, Neuroscience Letters, 2003) and in France (Candito M, American Journal of Medical Genetic, 2008).

Reply
We appreciate the reviewer pointing out these citations – they have been incorporated into our revised manuscript.

Comment
Specific minor comments:
1) The title is not informative (“Spina bifida and heart defect” instead of “selected anomalies”)

Reply
We have followed the reviewer’s suggestion to revise the title.

Comment
2) Please, avoid “Wild type” (throughout the ms). It refers to a mutation rather than to a polymorphism.

Reply
The use of the term “wildtype” has been replaced throughout the manuscript with the phrase “reference genotype.”

Comment
3) Abbreviations of FOLR1 and FOLR2 are not given in footnote of table 1

Reply
Thank you for pointing out this oversight, abbreviations are now clarified.

Comment
4) Table 2 should be divided into two tables for Spina Bifida and conotruncal defects, respectively, giving the frequencies not only in controls but also in cases. Some of the SNPs have a minor allele frequency lower than 0.1.

Reply
The reviewer appears to be referring to Table 3 and not Table 2. We have attempted to limit the potentially excessive number of tables by combining the two outcomes into one table and providing the genotype frequency among the controls. We do not believe that adding more information (genotype frequencies among cases) to an already very large data table would provide substantial benefit to the reader.
Comment
5) Please, indicate significant results of table 3 in bold, to help the reader! Tables 4 and 5 could be limited to significant haplotypes. Please, provide P-values in the tables and indications for adjustment in footnote.

Reply
We have noted the odds ratios in Table 3 whose confidence intervals do not overlap 1.0 in bold font.

Second Reviewer's report (Andrew E. Czeizel)

Comment
This paper presents an important and very work-consuming study with a bit disappointed results. All possible candidate gene polymorphisms in the origin of NDT and conotruncal cardiac defects were checked without the detection and/or confirmation strong association. Nevertheless this study is an important, though mainly negative step in the route of understanding better the etiology of two most important structural birth defects in the human being. I addition the detailed presentation of their laboratory analyses may help others to use their data for meta-analysis or to look for new directions.
I think the major problem is that this approach was not appropriate for the evaluation of maternal-fetus SNPs together. We know data regarding the role of hyperhomocyttenemia in the origin a placental malfunction and the placenta has a double: fetal-embryonic and maternal origin. Thus we can imagine that the maternal factors are important in the gene-environmental interaction inducing finally these defects.

Reply
We agree with the reviewer that it would be advantageous to also have maternal genotype. Unfortunately, we did not have that information as we note in the Discussion.

Comment
The international literature is very abundant in this topic, authors did their best to select more important previous publications, but I know some others as well. It would be better to cite the more relevant papers of the Hungarian group in this topic (Czeizel AE.: Reduction of urinary tract and cardiovascular defects by periconceptional multivitamin supplementation. Am J Med Genet 1996. 62: 179-183, Czeizel AE, Dobo M, Vargha P. Hungarian cohort-controlled trial of periconceptional multivitamin supplementation shows a reduction in certain congenital abnormalities. Birth Defects Res (Part A) 2004. 70: 853-861.)

Reply
We have incorporated both suggested citations into our revised manuscript.

Third Reviewer's report (James L Mills)

Comment
In this paper Shaw and colleagues search for associations between 118 SNPs in folate pathways and neural tube defects (NTDs) or cono-truncal heart defects. They use population based data on births in California.
This is a worthwhile topic for investigation and the researchers provide interesting data. The strengths of the study are the large number of variants examined, the representative sample selected for investigation and the moderately large number of subjects.

The analysis is reasonable with the exception of a few issues that will be discussed below. The conclusions are appropriately cautious. The paper is clearly written. Dr. Shaw’s inimitable style is evident in sentences that begin, “Eligible were live born infants…” and “Included were 259 infants…”.

The authors identify almost all the limitations in the Discussion section.

Suggestions:

Note that these fall into the category of Minor Essential Revisions because I believe that the authors will respond appropriately.

1. Most readers will know these variants by the polymorphisms, e.g. MTHFR 677C->T. Those should be used in the text. The rs numbers are helpful, but many people will not be able to relate them to the polymorphisms.

Reply

While some readers will recognize a few SNPs by name, we believe the majority of readers will benefit from using the standard approach of reflected by rs numbers. In the Discussion we use both SNP names and rs numbers when discussing previous literature to help the reader.

Comment

2. In the abstract and in the results section, confidence intervals should be included with the odds ratios.

Reply

We have included confidence intervals with the odds ratios presented in the Abstract and Results sections.

Comment

3. In the abstract conclusions the authors should note that the haplotype findings could be due to chance.

Reply

We have rewritten the last sentence of the Abstract to temper the interpretation of our results.

Comment

4. My major concern is that the NTD case population is quite different from the control population in race/ethnicity. The authors are aware of this problem. They say that they did not observe evidence that risk patterns were confounded by this difference. These data should be presented. In fact, the regression analysis including race/ethnicity should probably be the major analysis presented in the paper (results and tables). If the logistic regression results are presented in Table 3, that should be noted in the table. If population stratification can be excluded, it will make the findings much stronger.

Reply

While we agree that the adjusted analyses may be important to some readers, we believe the unadjusted analyses may have more general utility for future comparisons.
such as meta-analyses. Thus, we present the unadjusted analyses in table 3 and have indicated in the text that the adjusted results can be obtained from the authors upon request. Moreover, results from adjusted analyses were not substantially different from the results associated with the unadjusted analyses.

Comment
5. The authors note briefly in the limitations paragraph that multiple comparisons were made. They should expand this to provide some estimate as to what findings, if any, would be statistically significant if correction for multiple comparisons were performed. This does not have to be done in a way that dismisses all the findings with a p value below 0.05. Correcting by standard methods may be too severe. It does, however, have to indicate that p values less than 0.05 cannot be taken at face value when hundreds of comparisons are made.

Reply
We have added a phrase to the Discussion that reiterates a point we make in the Results. This point is that we conducted 472 analytic comparisons and would have expected to observe even more statistically significant findings to arise by chance alone.

Comment
6. At the end of the results section, the authors report on haplotype analyses stratified by race. Did the other positive results in the previous paragraph become non-significant in the stratified analyses? This needs to be clarified.

Reply
We see the reviewer’s concern and have revised the last paragraph of the Results section to clarify the presentation of results.

Comment
7. I have two comments regarding MTHFD1, one of several genes that we have studied. The first is that we (ref.39) actually performed two studies. The second, referenced in the paper, is the confirmatory study. So it would be helpful to indicate that this gene has been confirmed to be a maternal risk factor in the Irish population. The second is that this study could have found a modest effect in cases when there is a true effect in mothers. That point should be mentioned.

Reply
We have added that the study findings in the Irish population were confirmatory to the text in the 7th paragraph of the Discussion. In that same paragraph, we note that our results did show an elevated risk of 1.6 for this SNP in infants for spina bifida.

Comment
8. On table 4, I was not able to discern why some rows were highlighted.

Reply
We have removed the “highlighted” rows in Table 4.
Fourth Reviewer’s report (Liborio Stuppia)

Comment
The manuscript of Shaw et al. “118 SNPs of Folate-Related Genes and Risks of Selected Congenital Anomalies” describes a study aimed to investigate the influence played by 118 SNPs associated with the complex folate pathway on the risk of spina bifida or conotruncal heart defects. Obtained results showed few odds ratios revealing sizable departures from 1.0 with respect to spina bifida, and no odds ratios with confidence intervals that did not include 1.0 for any of the studied SNPs with conotruncal heart defects. Authors conclude that these results do not implicate a particular folate transport or metabolism gene to be strongly associated with risks for spina bifida or conotruncal defects.

The study appears well designed, and the manuscript well written. Authors demonstrate to have clear also the limits of the study, the main of which being represented by the lack of data about the mothers’ genotypes, which likely play a crucial role in the folate metabolism.

Minor essential revisions
1) My only major criticism to this study is related to the conclusions. Authors seem to suggest that their study “failed” to evidence a gene-only effect on risk of spina bifida and conotruncal heart defects. However, hundred of studies in the last years have already demonstrated that the contribute of the variants in the genes involved in folate metabolism to the risk of congenital diseases must be considered as a typical example of “poligenic effect”, with several genes playing a limited role. In this view, the results obtained by the authors are exactly those expected, while the detection of the presence of a single gene with major effects should have been unexpected. Therefore, I strongly suggest to the authors to concentrate their discussion on this point.

Reply
We do not share the reviewer’s opinion that there have been hundreds of studies in the past demonstrating effects between gene variants involved in folate metabolism and risk of congenital abnormalities. We capture the most salient and rigorous studies on this topic in the Discussion section – its certainly not hundreds. The unique contribution of our study is that it contained the largest number of SNPs in folate-related genes interrogated as risk factors for human spina bifida or conotruncal heart defects. We note this point in the Discussion (2nd paragraph).

Comment
2) In the Abstract session, the Conclusion is represented by the sentence “We did not observe any ORs with confidence intervals that did not include 1.0 for any of the studied SNPs with conotruncal heart defects. Haplotype reconstruction showed associations with TYMS, MTHFR, BHMT and MTR for spina bifida.” In my opinion, this is not a conclusion of the study, but just a portion of the results. The conclusion should be represented by a short comment on the obtained results.

Reply
The reviewer is correct. The two sentences are indeed a further summary of the observed results. It was not our intention to provide a conclusion in the Abstract. It was our intention to provide a summary of results.

**Fifth Reviewer’s report (Anne Parle-McDermott)**

**Comment**
This manuscript by Shaw et al., sought to test a number of polymorphisms for risk of spina bifida and conotruncal heart defects in a Californian population. Their candidate gene, case-control approach consisted of 13 genes involved in the metabolism or transport of folate. The list of candidate genes is not novel; many if not all have been examined previously in relation to spina bifida risk (some in relation to conotruncal heart defects) with often conflicting results. Although, some of the variants showed an association with spina bifida, the overall conclusions indicate lack of a strong association of any of the genes with risk of either malformation. However, the genetic risk associated with these malformations has not been addressed extensively and thus, appropriately designed association studies, yielding positive or negative results should be published. I have a number of issues with the manuscript as follows:

**Major compulsory revisions**

1. Population stratification: the breakdown of the ethnic groups as described in Table 2 is a major cause for concern. There is a significant difference in the percentage of individuals of White Hispanic versus non-Hispanic in the case group compared to the control group. Allele frequencies that simply vary between ethnic groups and are unrelated to these malformations may produce a false positive association in such a study design. Although the authors do indicate that they adjusted for this in their logistic regression analysis more detail in relation to how they adjusted for this should be provided.

**Reply**
The well accepted approach of logistic regression was used to compute risk estimates adjusted for maternal race/ethnicity based on 3 groups, white Hispanic; white nonHispanic, and Other. This is noted in the last paragraph of the Methods section.

**Comment**

2. Abstract & Main Text: Results- a number of SNPs from the same gene are listed as having significant associations: the level of linkage disequilibrium between markers from the same gene should be included i.e, D’ and r2 values. Are they separate risks from the same gene or are they simply acting as markers of each other? The positive results should be put into the context of LD patterns of the gene itself- this is most important.

**Reply**
We agree with the reviewer that this is important information to add. We have added the linkage disequilibrium information to the Results section as well as to the Abstract.

**Comment**

3. The genomic DNA samples were extracted from dried blood spots on filter paper and subsequently amplified by whole genome amplification (WGA). The authors should provide some details into relation to the WGA procedure in terms of ensuring both copies of each gene are efficiently amplified.
Reply
We used the GenomiPhi DNA amplification kit, which is based on Phi29 DNA polymerase to replicate linear genomic DNA by multiple displacement amplification (MDA) method. MDA, which relies on isothermal amplification using the DNA polymerase of the bacteriophage phi29, is a recently developed technique for high performance WGA. In comparison to other WGA methods, MDA appears to be most reliable for genotyping, with the most favorable call rates, best genomic coverage, and lowest amplification bias (Lovmar L, Syvänen AC. Multiple displacement amplification to create a long-lasting source of DNA for genetic studies. Hum Mutat. 2006 Jul;27(7):603-14). Studies indicate no discernable difference between WGA samples with GenomiPhi kit and the original DNA templates (Holbrook JF, Stabley D, Sol-Church K. Exploring whole genome amplification as a DNA recovery tool for molecular genetic studies. J Biomol Tech. 2005:16:125-33. and Bergen AW, Qi Y, Haque KA, Welch RA, Chanock SJ. Effects of DNA mass on multiple displacement whole genome amplification and genotyping performance. BMC Biotechnol. 2005:16:5:24. We have added some of this salient information to the Methods section, 3rd paragraph.

Comment
4. Statistical correction: given the large number of comparisons performed some form of correction should be applied. The authors should address this or justify why they feel it unnecessary to correct the P-values. In addition, P-values are not included anywhere in the manuscript even though significance can be inferred from the OR. P-values for all significant associations should be included.

Reply
As we noted above to Reviewer 3’s concern on this topic - we have added a phrase to the Discussion that reiterates a point we make in the Results. This point is that we conducted 472 analytic comparisons and would have expected to observe even more statistically significant findings to arise by chance alone. We do not believe that correcting the analysis for multiple comparisons in this study is necessary. Our goal was to describe the risk patterns for a large number of genotypes in a population-based set of cases and controls. We do not make strong inferences about any particular finding – thus, correcting the observed confidence intervals for the number of tests performed adds very little to the overall message that we are trying to convey in our manuscript. We have estimated risks using odds ratios. The statistical precision of the odds ratio is captured by confidence intervals. We have not included p-values because they are redundant to the confidence intervals and provide far less information.

Comment
5. Table 1: the information contained within Table 1 would be easier for the reader to access if presented in the form of a diagram with each gene represented separately (probably more appropriate in a supplementary document). Each diagram should include the location of each polymorphism from the 5’ to 3’ end of the gene. Also Table 1 is not actually referred to in the text.

Reply
Table 1 provides a wealth of standardized information to the reader about the SNPs investigated in this manuscript. Table 1 is indeed referred to in the text - in the Methods section at an appropriate place (5th paragraph).
Comment
6. Table 4: also very long. This information could be contained within a supplementary section or simply 'data not shown', except for significant associations. Also details of how exactly the haplotype associations were computed should be included. ‘Identified blocks were assessed with odds ratios’ is too vague. A permutation analysis using Phase 2.1.1. software is more commonly used to test for haplotype risk (Stephens & Donnelly, Am J Hum Genet, 73(5): 1162-1169, 2003.)

Reply
As we note a few replies above our goal was to describe the risk patterns for a large number of genotypes in a population-based set of cases and controls. Reducing the presentation of the results to a set of results that meets an arbitrary criterion of statistical significance would undermine the value of the current data to future applications such as meta-analysis. In terms of haplotype analyses, we note in the Methods section that these analyses were performed using Haploview version 3.32.

Comment
7. Discussion: acknowledgement that full gene coverage for each gene was not achieved with this study design i.e., enough markers genotyped across the entire gene to ensure all variants are captured.

Reply
We have added two sentences to the end of the second-to-last-paragraph of the Discussion that acknowledges that full gene coverage was not achieved.

Minor Essential Revisions

Comment
1. The manuscript would benefit from a review of the structure of some of the sentences and the overall ‘flow’ of the information. I’ve highlighted a couple of sentences:
   • Abstract background: should probably read ‘Folic acid taken in early pregnancy….’
   • ‘Eligible were live born infants only….’ ??
   • ‘Few odds ratios (Ors) revealed sizable departures from 1.0.’

Reply
Thank you we have corrected a few of these sentences.