Author's response to reviews

Title: Expression profiling of clonal lymphocyte cell cultures from Rett syndrome patients

Authors:

Ivan J Delgado (idelgado@identigene.com)
Dong S Kim (doskim@knu.ac.kr)
Karen N Thatcher (knthatcher@ucdavis.edu)
Janine M LaSalle (mlasalle@ucdavis.edu)
Ignatia B Van den Veyver (iveyver@bcm.tmc.edu)

Version: 4 Date: 7 May 2006

Author's response to reviews: see over
Editor in Chief,
BMC Medical Genetics

May 5, 2006:

RE: manuscript: Expression profiling of clonal lymphocyte cell cultures from Rett syndrome patients

Dear Editor:

We appreciate the reviewers’ thoughtful comments and suggestions to improve our manuscript. We have addressed all comments systematically and have included revisions in appropriate sections of the text and figures, as indicated below, following each of the reviewers’ comments:

RESPONSE TO REVIEWER 1 (A Renieri):
Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. Authors state repeatedly in the manuscript that they compared their list of deregulated genes to “expression profiles of independent microarray experiments in cells and tissues of RTT patients and mouse models with Mecp2 mutations”. Apparently, they refer both to experiments performed from the authors themselves and still to be published and to results of previously published experiments (P6, L7-8; P12, L1-7). However, authors report only results of the comparison to the first group of genes. It is not at all clear which are the results of the comparison to the previously published de-regulated genes. Authors must clearly state in the manuscript whether the genes they found down- or up-regulated in their cells have been previously identified by other gene expression profiling experiments. If none of their gene is present in the previously published lists of de-regulated genes, authors should discuss the reasons of this result.

REPLY: We have indicated more clearly throughout the manuscript which comparisons were made and what the results of all these comparisons were. New supplementary tables 1 and 2 outlining the comparisons were included: In summary: there was little overlap between our data and analyzed published datasets, as well as between the different published datasets. We focused QPCR analysis on SPOCK1, which was identified by comparing our own (unpublished) data on gene expression in in vitro differentiated wild-type and mutant Mecp2<sup>308y</sup> mouse ES cells. We have provided more discussion regarding discrepancy between our data and published literature (page 16).

2. Authors dedicate just one sentence to discuss the possible reasons of the conflicting results of microarray and quantitative RT-PCR experiments for SPOCK gene. However, this discrepancy represents the most significant finding of the manuscript. Authors should extensively discuss the reasons of such difference in the “Discussion” section. Authors conclude their discussion saying that one reason for such discrepancy might be that “the changes in SPOCK expression in the presence of MECP2 mutations are subtle and not detectable with the current methods”; however, this is not consistent with the fact that they detected a statistically significant difference with microarray experiments.

REPLY: Repeated and improved analysis of the microarray data and qPCR does not show the same discrepancy (i.e. previously found opposite expression results between microarray and qPCR). The gene is downregulated in mutant in both of our microarray experiments. qPCR analysis, however only showed a non-significant trend for downregulation in the lymphocyte RNAs (see figure 2A). In addition, we want to highlight that with FDR analysis (see below in response to reviewer 2, none of the genes showed significantly altered expression. For qPCR, the analysis in mice used CNS tissues, while the human data were obtained from lymphocytes, a non-neuronal tissue. This may have contributed to the
expression discrepancies between various tissues.. Results and discussion have been modified to include revised information on Spock1 qPCR data.

3. An important data presented in the manuscript is the fact that, in spite of a relatively large number of de-regulated genes identified with each of two different normalization algorithms (684 and 643 respectively), only a small subset of these genes were consistently identified with both algorithms (95). This outlines the limitations of the algorithms presently employed for data analysis and the dimension of the risk of false positive and false negative results. Authors should discuss this issue.

REPLY: In response to comments from the second reviewer, we have now reanalyzed the data using GC-RMA and have included results of false discovery rate (FDR) analysis. (please see reply to question 1 of the second reviewer). The resulting new gene list contains 121 genes (77 upregulated, with being 13 being >2fold, and 44 downregulated, with 12 being >2 fold). The list is partially overlapping with prior analysis, and still contains SPOCK as a 2-fold downregulated gene.

4. Authors performed data validation by quantitative RT-PCR with SYBR Green as the detection agent. Since the fold change observed for SPOCK gene in their microarray experiments seems quite subtle (-0.935) it could be more appropriate to perform the analysis with a specific primer/probe set which should allow a greater sensitivity.

REPLY: Our current resources do not allow us to change the method from SYBR green to use of specific primers and probes. We have included all the recommended controls and methods of data validation and analysis for use of this method, and we therefore believe that our quantitative RT-PCR data are valid. Furthermore, this method is commonly used for this type of analysis.

5. Methods, section "Microarray data analysis": Understanding of the statistical elaboration of data is quite difficult; authors should describe data analysis more clearly. In particular, they should clearly state which criteria they followed to define statistically significant expression differences. Moreover, they should define how exactly the “Fold Change Mut>WT” value was calculated. Is it just a ratio of the two values?

REPLY: We have provided more clarification on the data analysis and statistics, which is now entirely focused on GC-RMA. The fold change Mut > Wt is calculated using Genespring software as the absolute fold difference in expression of the combined mutant over wild-type GCRMA-normalized intensities of genes in the p<0.05 list from the one-way ANOVA analysis.

6. P6, L9-10: Authors state that by quantitative RT-PCT SPOCK “expression was up- regulated in T lymphocytes”. It is not clear if they are speaking of lymphocytes isolated from their RTT patients and, in this case, if they refer to clonal cells expressing mutant or wild type MECP2. In addition, lymphocytes RT-PCR analysis is not described in results’ section of the manuscript, where they just report results of analysis on mouse brain (P12, last lines). SPOCK expression in patients lymphocytes is an essential point to consider and authors must clearly state in the results’ section whether or not they analysed its expression and which results they obtained. If the analysis has not been performed, authors should perform it and include results in the manuscript.

REPLY: As outlined above, the data on lymphocyte RNAs have changed after reanalysis. The experiment was done using the clonal lymphocyte RNA used for the microarray experiment. This information was included in the paper. The qPCR analysis is described in more detail in methods and results sections and a panel showing the data on Lymphocyte RNAs is now included in figure 2. We did not analyze lymphocytes of additional patients, because they are female and heterozygous for a wild-type and mutant Mecp2 gene and we do not have lymphocyte clones from them.

7. P16, L2: The sentence “No statistically significant differences in expression were found” is not correct. At least 1 gene with a statistically significant difference in expression resulting from microarray data was found: SPOCK. Even if subsequent experiments did not confirm the result, the
down-regulation found by microarray experiments was statistically significant both for the experiments presented in this manuscript and, I suppose, for the preliminary unpublished experiments performed by authors on MECP2 knock-out ES cell lines.

REPLY: We revised this statement to read "most candidate genes" (in the abstract conclusion) or "upon stringent analysis" (in the manuscript conclusion, page 17). We clarified in the results that based on p-values of <0.05, there were up and down-regulated genes, but that upon FDR analysis, none of these remained.

8. Figure 3: Authors must report the standard deviation also in the histograms representing SPOCK expression in controls even if they set them as one-fold baseline.

REPLY: This was corrected. The referenced figure 3 is figure 2 in this version.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. Abstract, L. 6-7: Authors state that “MECP2 mutations result in genome-wide transcriptional repression”. They must change “repression” in “de-repression” or “deregulation”.

REPLY: This was changed as suggested to "deregulation"

2. Background, P3, L14: Authors say that MeCP2-e2 “contains exons 2, 3 and 4”. This is not correct. Since authors are speaking of the protein they should say that it is “encoded” by exons 2, 3 and 3. Otherwise, if they refer to mRNA, it contains exons 1, 2, 3 and 4.

REPLY: This was changed as suggested.

3. Methods, P9, L14: The sentence “An equal amount of total RNA brain regions from 5-6-week-old mice or…….” should be corrected in “an equal amount of total RNA isolated from brain regions of 5-6-week-old mice or…….”. In addition, authors should specify from which brain regions they isolated RNA.

REPLY: This was changed as suggested.

4. P12, L8-9: The title of the paragraph should be changed to better reflect its content: The title anticipates the analysis of SPOCK without having Preventively explained why authors are interested in analyzing SPOCK. If authors do not want to change the title of the paragraph, they should move the sentence reporting the identification of SPOCK gene (L.10-12) to the previous paragraph (microarray analysis of matched lymphocyte clones) as a conclusion to the last sentence of this paragraph.

REPLY: The title was changed as suggested.

5. P12, L13-14: If the function of the functional domains present in SPOCK is known, authors should briefly state which function each domain exert.

REPLY: This was changed as suggested.

6. P12, L16-17: The sentence should be corrected in “Mouse Spock1 localizes predominantly to the postsynaptic region of a subpopulation of pyramidal neurons in the CA3 region of the hippocampus”.

REPLY: This was changed as suggested.

7. P12, L22-23: Did authors perform expression analysis on 5-6-week-old mice and E16 embryos (as indicated in methods, P9, L14) or in 5-week-old mice and E16.5 embryos?

REPLY: This was changed for consistency as suggested (5 week-old mice and E16.5 embryos).

8. P14, L20: The sentence “….and compared the overlap between the gene lists obtained from each
patient…..” should be changed in “….and compared the gene lists obtained from each patient…..”.

**REPLY:** This was changed as suggested

9. P.15: Authors should include more references related to SPOCK role in CNS and its expression in astrocytes and after injury.

**REPLY:** This was changed as suggested. Additional references were included

10. Figure 1: The figure would be clearer if authors indicate patient code, in addition to mutation both in the figure and in its legend.

**REPLY:** This was changed as suggested

11. Figure 2: There seems to be a mistake in the number of genes reported: for RMA-normalized genes they have 588588+95= 683 (and not 684 as it should be) and for MAS5-normalized genes they have547+95= 642 (and not 643 as it should be). Did authors exclude SPOCK from the count? Why? Or it is just a tiping error?

**REPLY:** This was a typing error. However, since we have refocused on GC-RMA analysis, the figure was removed from the current version of the manuscript.

12. Figure 3: There is a mistake in figure description (P22, L16): the dark bars should represent wild-type samples and the light bars the mutant samples, and not vice-versa as said in figure legend. In addition, in the right diagram of figure 3A the histogram referring to wild-type cortex do not have the one-fold baseline value. Finally, a legend inserted in the figure reminding to the reader what the different colours represent would greatly improve figure clarity.

**REPLY:** this figure, now figure 2 and its legend, has been modified for additional data and clarity, taking the reviewer's comments into account.

**RESPONSE TO REVIEWER 2 (Carlo Colantuoni):**

**Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)**

1. It is my strong recommendation that the authors decide on a single optimal normalization method and use the resulting list for cross referencing with other lists. RMA, even GC-RMA, is the method I would recommend. Among other problems, MAS 5.0 introduces a lot of noise into the low intensity range. The lack of success in most of the authors’ cross references as well as the lack of confirmation in the PCR in the mouse CNS model may be a result of the original screening down to the list of 95 genes (which draws heavily on MAS 5.0 data). Short version - run RMA or GC-RMA, toss the MAS 5.0 data, and repeat all the cross references and PCR (with a new gene from the new cross references).

**REPLY:** Our goal was to perform a more stringent data analysis to reduce the number of altered genes to a small workable list for further confirmation. Hence, we elected to perform MAS5 as well as RMA analysis, and compare results. MAS5 has indeed high noise, but also detects genes not found by RMA. Nevertheless, according to the suggestion, we have now performed GC-RMA analysis followed by one-way ABNOVA. This yielded a shorter gene list (121) with partial overlap to the one in the previous version of this manuscript. Importantly, the limited overlap with published gene lists as well as with the data of our ES cell experiment remained nearly unchanged.

2. Additionally, the authors may consider completing a statistical analysis involving FDR (false discovery rate) analysis, or other more rigorous approach to the two gene lists, rather than simply
reporting hundreds of p-values from t-tests that are not correct for multiple comparisons.

**REPLY:** the p-values supplied are now from one way ANOVA analysis. We also performed a multiple testing correction, using the Benjamini-Hochberg False Discovery Rate (FDR) analysis. This did not result in any genes with significantly altered expression. This information was added to the manuscript. Although very valuable, this method could be too stringent for some situations, hence we elected to use the uncorrected gene lists for further analysis.

-----------------------------------------------------------------------------------

**Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)**

1. on page 8, the authors need an affymetrix reference for the MAS 5.0 algorithm even if they se R to implement the algorithm. and the use of RMA requires the bioconductor reference.

**REPLY:** we corrected the references for the various methods of analysis. We did no longer use MAS5 for the repeat analysis.

Thank you for reviewing this manuscript and allowing us to resubmit the revised version, on behalf of my co-authors, I am looking forward to a positive review of the resubmitted revised paper.

With Sincere Regards,

Ignatia B. Van den Veyver, M.D.
Associate Professor
Department of Obstetrics and Gynecology
Department of Molecular and Human Genetics
Baylor College of Medicine