Author’s response to reviews

Title: High resolution chromosomal microarray analysis in paediatric obsessive-compulsive disorder

Authors:

Edna Grünblatt (edna.gruenblatt@kjpd.uzh.ch)
Beatrice Oneda (oneda@medgen.uzh.ch)
Arif Ekici (arif.ekici@uk-erlangen.de)
Juliane Ball (Juliane.ball@puk.zh.ch)
Julia Geissler (Geissler_J@ukw.de)
Steffen Uebe (stenf.uebe@uk-erlangen.de)
Marcel Romanos (romanos_m@ukw.de)
Anita Rauch (anita.rauch@medgen.uzh.ch)
Susanne Walitza (susanne.walitza@puk.zh.ch)

Version: 1 Date: 24 Aug 2017

Author’s response to reviews:

We thank the reviewers for their constructive comments and helpful suggestions for improving the manuscript. Accordingly, we followed, point by point, all reviewers’ suggestions and made the necessary changes, as well as answered each of the points (see below). We wish to thank all reviewers and the editor for the opportunity to improve the paper and to resubmit it for your consideration.

- Juan Rodriguez-Flores (Reviewer 1)

Grunblatt and colleagues conduct an interesting study of CNVs 50 kb or larger in OCD, and while none of the CNVs studied are enriched in the cases, the pathway-based distribution of the CNVs was different, with an excess of genes related to brain and synapse pathways. The results are very interesting and potentially of relevance, however a few minor concerns about the pathway analysis remain:

Reply: We thank the positive review of our manuscript and encouragement.
1. The p-values for enrichment are marginal (between 0.05 and 0.01). Were multiple tests conducted, and if so is a multiple testing correction applied?

Reply: The p-values for enrichment analysis according to PLINK was corrected for multiple testing as indicated in Suppl. Table S4. For pathway studio enrichment analysis cutoff was set to a stringent p-value of 0.005 as indicated in Supplementary Table S3a, while for DAVID analysis the cutoff was p<0.05 however Bonferroni, Benjamini and FDR are given also in Suppl. Table S3b. In order to make this clearer we added this information also in the methods (page 9).

2. Additional information about the details of the tests should be provided, in terms of model and test. Possibly a simpler and easier to conduct test result such as GSEA (as implemented in R) can be used. Tests based on proprietary tools and online tools are difficult to confirm and replicate.

Reply: Pathway studio analysis is based on the GSEA model, while DAVID and the PLINK are described in details in the cited papers (Huang et al. 2009; Raychaudhuri et al. 2010). We added the GSEA model into the text for clarification (page 8-9).

3. In terms of gene expression, can the authors support their claims based on public gene expression data? Are these genes high/low expressed in brain tissue?

Reply: Gene expression omnibus (GEO) profiles were scanned for each of the genes filtering with the key words brain and Homo sapiens checking for expression levels in the brain. The main inclusion criteria was that these genes are at all expressed in the brain in order to avoid a bias due to an arbitrary cut-of level of relative expression. Nevertheless, we added into Suppl. Table S2 the GTEx transcriptomic database link for the brain related gene expression profiles, where expression levels are compared between brain tissue and peripheral tissues.

- Ryan Yuen (Reviewer 2)

Grunblatt et al reported a study of copy number variation in early onset obsessive compulsive disorder (OCD). Using a high resolution microarray, the authors found a significantly higher number of rare CNVs (>50kb) affecting genes related to brain and synapse functions in OCD patients. Similar to previous studies, they did not find an enrichment of overall number of rare CNVs or difference in CNV's size in cases compared to the controls. Some of the associated genes reported here have also been found by previous studies on neurodevelopmental disorders.
Analyses done here were mostly adequate and appropriate. The findings were consistent with previous studies.

Reply: We thank the reviewer for the positive evaluation.

1. They mentioned that there is only one CNV found from the 74 samples that were included in the previous study, but is there any CNV found in 74 samples from the previous study that was not identified in the present study?

Reply: No, there was none that was identified in the previous study, and was not identified in the current study. We added this in methods page 8.

2. Please provide more information on the control samples: Are the control samples run on the same microarray platform? Where were they recruited from? Were they blood samples?

Reply: The control samples, recruited at the Institute of Medical Genetics, University of Zurich, were processed on the same platform as the patient samples and the DNA was extracted from native tissue. We added this in methods page 7.

3. The authors should indicate whether the reported CNVs have been validated by independent experiments (e.g. qPCR). If so, please provide details in methods.

Reply: The detected CNVs were not validated by independent experiments, because they fulfilled our established criteria for reliable calls. The reliability of a CNV call and the necessity of confirmation of detected CNVs with alternative methods depends on the platform used and the experience acquired by users for the specific technique. For this purpose, we would like to mention that we started using molecular karyotyping in 2009 and have been using this technique for regular prenatal and postnatal routine diagnoses, running more than 1000 samples a year. During the first couple of years we conducted extensive confirmatory testing with alternative techniques, such as MLPA and FISH. We therefore trust that we can tell with adequate certainty, which CNV calls are reliable and which are not. We also published our experience and the quality control of our setting in 2014 (Asadollahi, Oueda et al. Journal of Medical Genetics 2014), as reported in the main manuscript, lines 10-12, page 8. In addition, as the technique in our lab is accredited, we regularly participate in external and internal quality controls, with excellent results. All the CNVs reported in the actual manuscript fulfil our proven criteria for real aberrations and thus do not need any additional confirmatory testing. Moreover, the finding that the number and size of CNVs do not differ between cases and controls indicates that there is no technical bias.
4. They said there is no significant difference in the number of all detected rare CNVs between patients and controls. What if they compare the numbers by deletion and duplication separately? How about restricting to CNVs that overlap with genes or coding regions of the genes?

Reply: Since there was no surplus in adding the results of the separate analysis for deletions or duplication as they resulted also in insignificance, we did not add it into the manuscript. Nevertheless, as mentioned in the results page 9-10, in the subgroup of CNVs carrying brain/synaptic genes, deletions seemed to have even more significance rather than when taking all rare CNVs together.

5. Is there a sex bias on the number of rare CNVs found?

Reply: No sex bias could be found for carriers of rare CNVs versus non-carriers; not between OCD and controls, nor within the groups. However, we did not include all these tests into the manuscript because of the relative small sample size for the subgroups.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Gender</th>
<th>Count</th>
<th>Gender</th>
<th>Count</th>
<th>Count</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNV carrier</td>
<td>male</td>
<td>female</td>
<td>male</td>
<td>female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>55</td>
<td>30</td>
<td>52</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>21</td>
<td>18</td>
<td>23</td>
<td>18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. This study may be more interesting if they have utilized the phenotypic information available for more analyses. For example, are the samples with rare CNVs tend to have more comorbidities? Are they tend to be more severe based on different measures? Is there any difference in IQ?

Reply: As reviewer #1 already requested we added comorbidities information into Suppl. Table S2. For IQ and severity (measured by CY-BOCS), as there was no significant results showing influence of CNVs on these parameters, we added this into results page 9.

7. The two cases that the authors discussed, 9025082001 and 9025043001, have additional duplications involved. In particular for 9025043001, the additional duplication was de novo. They may want to discuss the potential effect of the additional rare CNVs.
Reply: The duplications in the two participants were discussed shortly in the manuscript as requested on page 13 and 14.

8. In Table 1, the authors should include other comorbidities investigated from Supp Table 1. Also gene names should be italicized.

Reply: In order not to overload table 1 we added the comorbidities into Suppl. Table S2, where already the comorbid tics were included. We have also italicized all gene names accordingly.