Author’s response to reviews

Title: Alternative Cleavage and Polyadenylation of genes associated with protein turnover and mitochondrial function are deregulated in Parkinson’s, Alzheimer’s and ALS disease.

Authors:
Radhika Patel (radhika.patel168@gmail.com)
Cillian Brophy (cillian.brophy@pmb.ox.ac.uk)
Jonathan Neve (jonathanneve89@googlemail.com)
Mark Hickling (mark.hickling@queens.ox.ac.uk)
Andre Furger (andre.furger@bioch.ox.ac.uk)

Version: 1 Date: 29 Oct 2018

Author’s response to reviews:

Technical Comments:
Editor Comments:

This article was reviewed by two field experts and associate editors. Article uses publicly available limited amount of RNA-seq data to identify APA events in neurodegenerative diseases. Authors used appropriate methodology to analyze RNA-seq data to identify APA events. However, the analyses carried out in this work provide only an initial step in identifying interesting leads/candidates for subsequent exploration/validation, but on their own are not of sufficient breadth or interest. For example, this study compared four LOAD patients with four control samples (not matched for age and gender) to identify genes with APA in Alzheimer's, and is of very limited statistical power. Without any independent validation of the reported APA events confidence of the findings are highly questionable. Additional validation is necessary for this work.

Response to editors comments:

Thank you for considering our paper and we are pleased that the editors found that we used appropriate methodology to analyse the RNA-seq data. The editors find that we identify a number of genes that undergo APA in the respective diseases but that there are shortcomings in regards to matched control samples and independent validation of the identified events. As we stated in the paper, we used publically available RNaseq data to assess global APA in neurodegenerative disease and are thus reliant on the controls that were used in the original transcriptome analysis’. We agree that age and gender matched samples would be ideal but as the
Alzheimer’s samples stem from diseased individuals, this is extremely difficult to achieve and in our case, where we rely on published data sets, simply impossible. We are therefore surprised about this comment. Furthermore, none of the referees highlighted these issues.

BMC Medical Genomics operates a policy of open peer review, which means that you will be able to see the names of the reviewers who provided the reports via the online peer review system. We encourage you to also view the reports there, via the action links on the left-hand side of the page, to see the names of the reviewers.

Reviewer reports:

Zheng Xia (Reviewer 1): Alternative Polyadenylation is a widespread post-transcriptional regulation mechanism. But the APA usages in neurodegenerative diseases have not been studied thoroughly. Here, Patel et al. profiled APA and their functions in Parkinson's, Alzheimer's and ALS disease. Overall, this is a very interesting work and provides knowledge of the APA usages in neurodegenerative diseases.

We would like to thank Professor Xia for his time to assess our manuscript and we are delighted that he finds the manuscript is very interesting and advances our knowledge of APA in neurodegenerative diseases.

Response to minor points:

1. Is it possible to investigate the global miRNA bindings of the differential APA genes?

The miRNAs targeting regions in the aUTR that were of note are given on the figures (miR-26 and miR-128-3p for UBR1, miR-133b for NEFLA). These miRNAs had links to altered expression in neurodegeneration as highlighted in other studies. We did search for miRNA target sites in the aUTRs of the other genes but there were either no target sites, no unique target sites to the aUTRs, or no target sites of miRNAs that had a clear link to AD or PD.

2. In Table, the number of samples in each dataset will be helpful.

We thank Professor Xia for highlighting this important issue and we have now included a table that provides this information in the supplementary Figure 2C.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Background to Dataset</th>
<th>Number of control</th>
<th>Number of diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig 1</td>
<td>Late Onset AD</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Fig 2</td>
<td>Temporal Lobe AD</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
3. The heatmap plots of APA usages of genes in Figure 1C and Figure 5C will be more informative.

We agree and have now included a heat map for the genes referred to in Figure 1C. However, we feel for Figure 5C this to be less appropriate as there are 80 different comparisons.

Bin Tian (Reviewer 2): In the paper, Patel et al. analyzed public RNA-seq data sets for AD, PD, and ALS for APA changes with the computational tool DaPars. The authors reported limited, variable APA changes in data sets related to these neurodegenerative diseases. Overall, the work was well carried out, and it represents the most comprehensive APA analysis of neurodegenerative diseases to date. The results will be important for the neurodegenerative field and RNA processing field alike. I only have some minor suggestions.

We thank Professor Tian for his time to assess this manuscript and we are pleased that he judges it as important for both the neurodegenerative and RNA processing field.

Response to minor points:

1. It would much strengthen the paper if the authors could also examine whether the APA profiles can distinguish diseased and normal samples and compare the result with that based on gene expression levels.

This is a valid suggestion however, we did not look at gene expression changes of the APA affected genes in the same samples because we believe it would be difficult to control for noise between the different samples as they “lack spike” in controls and hence they would be difficult to normalise.

2. The authors mentioned that the DaPar program depends on annotated polyA sites. Perhaps the authors were referring to the last polyA site, as the proximal site is predicted. This needs to be clarified. In addition, it would be helpful to show the annotated polyA sites in figures, using the latest polyA site databases, such as polyA_DB 3, to indicate accuracy of proximal polyA site prediction by DaPars.
We clarified this point in the manuscript. The proposed proximal polyA site was confirmed to be annotated from the polyA database, see below:

ODGHL: pred: 50943074 / pADb: 50942687
UBR1: pred: 43237467 / pADb: 43237375
BIN1: pred: 127805770 / pADb: 127805639
VAMP2: pred: 8063745 / pADb: F:8063994 / M:8063734
LONP1: pred: 5692018 / pADb: 5692012
NELFA: pred: 1984894 / pADb: 1984850
Churc1: pred: 65399057/ not in pADb

We are not sure how useful these numbers are in the figures as the original and subsequent paper have shown that DaPars poly(A) site prediction is accurate.

3. Figure 1B. Since there are 16 comparison, how did authors define lengthened and shortened genes?

All genes that showed APA, shortening or lengthening between any of the patient and control comparisons were used for the GO analysis. There was no elimination/selection of genes from the data for this analysis.

4. Figure 2. Authors stated that "Unlike in the previous data set, a clear trend toward 3'UTR lengthening was observed in both the temporal and frontal lobes (Figure 2A)." Since there are only 30-50 genes with significant APA, this statement seems a bit too strong.

We agree and have changed this accordingly in the manuscript.

5. The authors may also want to discuss other approaches to study APA using RNA-seq data, in addition to DaPars.

We agree and refer now to other approaches that were recently developed, see second paragraph in the discussion.