STAU2 binds a complex RNA cargo that changes temporally with production of diverse intermediate progenitor cells during mouse corticogenesis

Rebecca Chowdhury, Yue Wang, Melissa Campbell, Susan K. Goderie, Francis Doyle, Scott A. Tenenbaum, Gretchen Kusek, Thomas R. Kiehl, Suraiya A. Ansari, Nathan C. Boles and Sally Temple
DOI: 10.1242/dev.199376

Editor: Paola Arlotta

Review timeline

Original submission:	20 December 2020
Editorial decision:	12 February 2021
First revision received:	14 May 2021
Editorial decision:	16 June 2021
Second revision received:	28 June 2021
Accepted:	5 July 2021

Original submission

First decision letter

MS ID#: DEVELOP/2020/199376

MS TITLE: STAU2 binds a Complex RNA Cargo that Changes Temporally With Production of Diverse Intermediate Progenitors During Mouse Corticogenesis

AUTHORS: Rebecca Chowdhury, Jenny Wang, Melissa Campbell, Susan K Goderie, Francis Doyle, Scott A. Tenenbaum, Gretchen Kusek, Thomas R. Kiehl, Suraiya Ansari, Nathan C. Boles, and Sally Temple

I have now received all the referees' reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

As you will see, the referees express considerable interest in your work, but have some significant criticisms and recommend a substantial revision of your manuscript before we can consider publication. In particular, there were concerns about the analysis pipeline and the lack of controls. If you are able to revise the manuscript along the lines suggested, which will involve further experiments and new analysis, I will be happy receive a revised version of the manuscript. Your revised paper will be re-reviewed by one or more of the original referees, and acceptance of your manuscript will depend on your addressing satisfactorily the reviewers' major concerns. Please also note that Development will normally permit only one round of major revision.

We are aware that you may be experiencing disruption to the normal running of your lab that make experimental revisions challenging. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Please attend to all of the reviewers' comments and ensure that you clearly highlight all changes made in the revised manuscript. Please avoid using 'Tracked changes' in Word files as these are lost in PDF conversion. I should be grateful if you would also provide a point-by-point response detailing how you have dealt with the points raised by the reviewers in the 'Response to Reviewers' box. If you do not agree with any of their criticisms or suggestions please explain clearly why this is so.

Reviewer 1

Advance summary and potential significance to field

This paper follows previous work by the Temple lab showing that the double stranded RNA-binding protein called STAU2, which is differentially present in one daughter cell following progenitor cell division (Kusek et al. 2012), may regulate cell fate on a developmental schedule. As this molecule has been hypothesised to play a role as a cargo molecule, the authors previous work suggested that STAU2 may be responsible for moving proteins into the differentiating daughter cell, and thus play a role in determining cell fates across cortical development. In their previous study the lab used RIP-ChIP to identify STAU2 targets. In the present study they used RIP-seq as this would allow them to identify further cargo-bound RNA transcripts in an unbiased manner across different developmental stages. They hypothesise that STAU2 targets would be dynamic across different developmental stages. They conclude that STAU2 cargo changes across development in a manner consistent with known schedules of the production of different cell types across cortical development.

The manuscript is based largely on bioinformatic analyses which unfortunately include significant methodological issues with the analysis that decrease enthusiasm for this paper.
The results include long lists of molecule types that are the putative identified cargo of STAU2 and it was difficult to get a sense of exactly how important STAU2 is in progenitor cell fate determination from this analysis (as opposed to being just a general shuttling molecule). This prompted us to perform a more in-depth review of the authors results and methodology which revealed a number of serious concerns detailed below.

In order to provide some experimental test for their hypothesis, the authors select one identified cargo molecule, Taf13 and knock it down in E13.5 cortical progenitor cells in vitro. They find a fairly modest (although significant) decrease in Olig2 following this manipulation and no decrease in the neuronal marker BTUB. Overall, this paper provides long lists of putative RNA cargo bound by STAU2 at different stages of development but serious issues exist with the analysis, decreasing confidence in the results. Moreover, the analysis of only one of the putative targets, Taf13, falls short of providing compelling evidence for the hypothesis that STAU2 is driving, or even influencing, progenitor cell fate determination.

Comments for the author

Major Concerns:
Page 6 Results section: RIP-Seq Defines Global STAU2 Cargo at Different Stages of Corticogenesis
"After the RNA-seq count data was mapped to the UCSC mm10 reference genome HOMER (Heinz et al. 2010) was used to identify peaks statistically enriched across replicates and thus determine differential peak expression."

RIP-seq data was analysed with HOMER, which was designed for analyses of sequencing data arising from genomic DNA. It is not clear how HOMER was utilized for peak calling and identifying differential peaks based on the description of the methods in the main manuscript and supplementary material. A link to access the code through github was made available in the manuscript but did not work.

The authors need to justify their use of HOMER instead of a tool specifically designed for RIP-seq or RNA-seq data. Moreover, further information on how HOMER was utilized to call peaks should be included in the manuscript.
A limitation of HOMER that is evident in Table S1 (column F) is the loss of strand-specific information, which should be readily available considering the kit used for library preparation.

Consequently, sequence reads that could correspond to either sense or antisense transcripts are lumped together within a peak and considered as one transcript. Another concern with this analysis is that sequencing reads that could correspond to other exons of a transcript not considered as peaks, would have been discounted from further analysis. Could the Bayes factor analysis yield different results if these reads were also included in the analysis?
“This allowed us to focus on the bound cargo at this level, consisting of 16071 transcripts, corresponding to 9412 genes, for further analysis (Table S1).”

This is a very large gene list and a concern is that some proportion of this could be non-specific binding. The authors should comment on this.
“Of the bound transcripts, most (78.6%) were classified as protein coding, 6.84% as pseudogenes, 2.48% as long intergenic non-coding RNA (lincRNA), 2.08% as microRNA (miRNA), and less than 1% were miscellaneous RNA (miscRNA), ribosomal RNA (rRNA), small Cajal body-specific RNA (scaRNA), small nucleolar RNA (snoRNA), and small nuclear RNA (snRNA) (Fig. 2B).”
It is unclear how peaks were annotated in this instance. Looking at Table S1 however, it seems that each individual peak was assigned to the nearest TSS (columns L to T in Table S1). If this is the case, this method is flawed. A peak that is far away from a gene’s TSS could be erroneously assigned to a different gene with a nearer TSS.

Our analysis of two putative targets Nfib and DCC identified such errors were apparent. For example, the peak ‘Merged-chr4-82350385-1’ overlapping exon 6 of Nfib is annotated as overlapping intron 6 of Nfib in columns I and J in Table S1. However, when assigned to the nearest TSS, this peak is incorrectly assigned to the rRNA n-R5s188 despite being 100kb from this rRNA. Similarly, the peak ‘Merged-chr18-71299375-1’ overlapping exon 27 of Dcc is annotated as overlapping this exon of Dcc in columns I and J of Table S1 but incorrectly assigned to the nearest TSS of the pseudogene AC103658.1. Results in Figures 2D, E and Figure 3 may also need to be re-analysed because of this.

Page 7 Results section: A Subset of STAU2 Target Transcripts is Temporally Regulated
“Using R, a Bayes factor (BF) was assigned to each gene representing the degree of differential binding between two consecutive timepoints, which was then summed across the time course (summed ln(BF))(Table S2).”

Was this analysis performed using read counts within peaks? A description should be included in the Methods or supplementary material of how this was performed including how data between the different time points were normalized to account for sequencing depth.
"In contrast, the top biological process and pathways for the 'dynamic' genes included ion transport, regulation of neuron differentiation, synapse organization, ER to Golgi transport, axon guidance, and extracellular matrix (ECM) organization (Fig. 3E,F), suggesting that STAU2 participates in developmentally regulated processes by altering subsets of genes bound at different time-points."
As a control, could the Bayes factor analysis also be performed using RNA-seq data from identical or similar tissue at these ages? For example, using ENCODE forebrain RNA-seq data at E11.5, E13.5, E15.5 and E18.5 (or other similar datasets). How would the result of such an analysis differ from the results with RIP-seq? Is the STAU2 cargo a reflection of differential gene expression due to differences in gene expression and tissue composition between the ages, or is STAU2 specifically binding to some highly expressed transcripts but not others?

Additional comments:
The abstract is rambling and raises the interesting motivation for the study to investigate the hypothesis “that STAU2 cargo would change as IPCs become specified to produce first neurons and then glia” but never comes back to this.

“..of a total of 29285 bound transcripts..”
The term ‘transcripts’ used from this sentence onwards is confusing and the term ‘peaks’ may be better suited. This is because multiple peaks could arise from separate exons of a single gene (all corresponding to one transcript).

Reviewer 2

Advance summary and potential significance to field

Previous work from the Temple group and others has shown that, as in Drosophila Staufen is asymmetrically inherited by one daughter cell during neural progenitor cell division. In this manuscript, the authors report data on analysis of the RNAs bound to the Staufen homolog, STAU2, during mouse cerebral cortex development. Based on a timecourse analysis, where the analysis of STAU2-associated RNAs was carried out at four developmental stages, the authors conclude that STAU2 changes its RNA cargoes in cortical radial glial cells during development, and those changes are functionally important for regulating the fate of the intermediate progenitor cells (IPCs) which inherit STAU2 and its RNA cargoes.

Comments for the author

Unfortunately, the data as reported lack several important controls and replicates, which make the conclusions made by the authors challenging to substantiate. This may be in part due to how the paper is presented, rather than the data themselves, as the manuscript skimps on many details, the methods are very short, and the figure legends are not sufficiently detailed. For example the authors have carried out IPs for STAU2 from lysates of developing cortex following by RNA-seq of RNAs harvested from that material, as well as the input lysate. Enrichment of STAU2-bound RNAs is then calculated by comparison with the input whole tissue lysate. Although it does not appear to be stated in the text or methods, each RIP seems to have been carried out once. The lack of replicates given the very large number of RNAs found in each RIP is a concern. This could be mitigated by comparing the E15 and E17 data, which are likely to be the most similar, in terms of cell composition and developmental stage.

In any case, identifying almost 10,000 genes as STAU2-bound, using relatively conservative calling methods from an original total of 29,285 transcripts, is puzzling and raises many questions. Previous studies identified around 1,000 RNAs bound by STAU2 in rat E17 brain (Heraud-Farlow et al., Cell Reports 2013), estimated as around 8.5% of mRNAs present in the input RNA. Subsequent research using the more accurate CLIP-seq technique identified 356 mRNAs bound by STAU2 in E18 mouse brain (Sharangdhar et al., EMBO Reports, 2017). The authors' own previous work in 2012 (Kusek et al., 2012), identified 1566 STAU2-bound RNAs.

Thus the 10,000 RNAs/genes identified in this paper as bound by STAU2 is 6-10-fold higher than that found in previous research, and is likely to include more than half of the mRNAs expressed in the developing cortex in progenitor cells neurons or both.

The very high number of RNAs identified here may be due to the almost-ubiquitous expression of STAU2 in the developing cortex, in different classes of progenitor cells and in neurons. However, the lack of technical controls in the paper, which might address this, is a weakness. The authors state that a control antibody brought down very little RNA, so was not used here. This is unusual, as RNA is notoriously "sticky" and typically a good deal of total RNA is recovered from non-specific binding to the beads, protein A/G etc. Combined with what appears to be a lack of replicates, it raises questions about whether the RNAs identified are essentially a sample of the whole lysate, either bound non-specifically to the IP material, or associating with STAU2 in solution. The authors' reference in the methods to the approach taken (Baroni et al., 2008) is not included in the reference list, but it would appear that the method does not involve crosslinking RNA to protein, so novel associations in solution between STAU2 and RNA during the IP process may be a confounding issue. Alternatively, the very high number of identified STAU2-bound RNAs may reflect the analysis method used to call peak enrichment, which is not well described in the methods, and which seems to be enrichment relative to the input RNA.

This very high number of STAU2-bound RNAs makes analyses of functional associations difficult to assess, as such a high fraction of all RNAs appear to be bound by the protein. The change in RNAs identified over time, in this context, may be itself artefactual, reflecting the change in RNA composition in the lysate – this is hard to judge, without appropriate negative controls. As such, it makes detailed consideration of the subsequent claims from the paper redundant, without addressing the question of the very high number of STAU2-bound RNAs and the various technical points raised above.

Reviewer 3

Advance summary and potential significance to field

The manuscript entitled "STAU2 binds a complex RNA cargo that changes temporally with production of diverse intermediate progenitors during mouse corticogenesis" comes from the lab of one of the pioneers of neural stem cell biology and makes a very important contribution to developmental neurobiology. In a nutshell, the authors have identified the various RNA molecules that bind to the RNA-binding protein STAU2 at four different stages of mouse cortical development (E11.5-E17.5). Interestingly, they show that there are two subsets of these RNA molecules, (i) those present at all developmental stages analysed, and (ii) those that changed with a given developmental stage. In addition, the authors performed functional studies on one STAU2 target, Taf13, and demonstrate that its knockdown in vitro results in reduced oligodendrogenesis.

This study not only provides key information on the behaviour of this important RNA-binding protein during corticogenesis, but also constitutes an enormously useful resource for future studies. The study is therefore highly appropriate for publication in Development. I only have one major point and a few minor points that should be addressed before publication.

Comments for the author

Major point

The authors use lysed cortical tissue to pull down and isolate RNA molecules bound to STAU2. How can the authors be certain that there is no post-lysis binding of RNA molecules to STAU2 that are not bound to STAU2 in vivo? The control data shown in Fig. 4 are nice, but this reviewer wonders whether there are additional controls to exclude potential post-lysis binding? For example, on the assumption that STAU2 is not expressed in endothelial cells of the embryonic mouse cortex, if the authors could show that RNA molecules specific to endothelial cells were absent from the sets of STAU2-bound RNA molecules, that would be a strong piece of evidence against post-lysis binding. If such control experiments are not feasible, the authors should at least include a disussion of this issue in the revised manuscript. In this context, the authors should comment on categories related to inner-mitochondrial processes, such as mitochondrial translation or TCA cycle.

Minor points
1. Summary Statement: "...contribute to IPC..." I guess the word "diversity" is missing?

2. On page 10, the authors state that they used "...transcriptomes obtained at mid-gestation in the mouse (Fietz et al. 2012).". This reviewer is curious if the authors also used other transcriptomes, for example those reported by Florio et al. 2015?

First revision

Author response to reviewers' comments

We are grateful to the reviewers for their thoughtful comments and helpful suggestions for improving our manuscript. In our response, we have redone our analysis to identify peaks of bound genes on positive and negative strands, and then defined a cutoff based on the elbow of the peak trajectory. This identified 7,968 genes above the cutoff as the total STAU2 cargo across four timepoints. Following, we redid all the analyses and remade all figures but for Figures 1 (and S1), 6 (and S6). While this re-analysis didn't change the main points and conclusions of the manuscript, some of the genes and pathways included in dynamic and stable categories were altered. Relevant changes in the manuscript such as edits to address reviewers' comments or details in genes or pathways changed as a result of re-analysis are highlighted in red.
Below we have responded in detail (purple text) to the reviewer's comments (in italics).

Reviewer 1 Advance Summary and Potential Significance to Field:

This paper follows previous work by the Temple lab showing that the double stranded RNA-binding protein called STAU2, which is differentially present in one daughter cell following progenitor cell division (Kusek et al. 2012), may regulate cell fate on a developmental schedule. As this molecule has been hypothesised to play a role as a cargo molecule, the authors previous work suggested that STAU2 may be responsible for moving proteins into the differentiating daughter cell, and thus play a role in determining cell fates across cortical development. In their previous study the lab used RIP-ChIP to identify STAU2 targets. In the present study they used RIP-seq as this would allow them to identify further cargo-bound RNA transcripts in an unbiased manner across different developmental stages. They hypothesise that STAU2 targets would be dynamic across different developmental stages. They conclude that STAU2 cargo changes across development in a manner consistent with known schedules of the production of different cell types across cortical development.

The manuscript is based largely on bioinformatic analyses which unfortunately include significant methodological issues with the analysis that decrease enthusiasm for this paper. The results include long lists of molecule types that are the putative identified cargo of STAU2 and it was difficult to get a sense of exactly how important STAU2 is in progenitor cell fate determination from this analysis (as opposed to being just a general shuttling molecule). This prompted us to perform a more in-depth review of the authors results and methodology which revealed a number of serious concerns detailed below.

In order to provide some experimental test for their hypothesis, the authors select one identified cargo molecule, Taf13 and knock it down in E13.5 cortical progenitor cells in vitro. They find a fairly modest (although significant) decrease in Olig2 following this manipulation and no decrease in the neuronal marker BTUB. Overall, this paper provides long lists of putative RNA cargo bound by STAU2 at different stages of development but serious issues exist with the analysis, decreasing confidence in the results. Moreover, the analysis of only one of the putative targets, Taf13, falls short of providing compelling evidence for the hypothesis that STAU2 is driving, or even influencing, progenitor cell fate determination.

Reviewer 1 Comments for the Author: Major Concerns:

Page 6 Results section: RIP-Seq Defines Global STAU2 Cargo at Different Stages of Corticogenesis

"After the RNA-seq count data was mapped to the UCSC mm10 reference genome, HOMER (Heinz et al. 2010) was used to identify peaks statistically enriched across replicates and thus determine differential peak expression." RIP-seq data was analysed with HOMER, which was designed for analyses of sequencing data arising from genomic DNA. It is not clear how HOMER was utilized for peak calling and identifying differential peaks based on the description of the methods in the main manuscript and supplementary material. A link to access the code through github was made available in the manuscript but did not work. The authors need to justify their use of HOMER instead of a tool specifically designed for RIP-seq or RNA-seq data. Moreover, further information on how HOMER was utilized to call peaks should be included in the manuscript.

We apologize for the lack of detail regarding our utilization of HOMER for peak calling. In the revised manuscript, we now provide a more in-depth description of the methods used to analyze the RIP-seq data in the materials and methods (p.21) and an overview of our RIP-seq pipeline in Fig. 2A. We are sorry the GitHub code was not made public by our institution; we have corrected this, and reviewers will have access to the code (https://github.com/neural-stem-cell-institute/Stau2).

While it is true that HOMER was originally designed to analyze CHIP-sequencing data, the developers of the software actively maintain the software and have expanded its use to analyze a variety of sequencing data types including csRNA-seq, RNA-seq, GRO-seq, Hi-C, DNase- Seq, and BS-seq among others. A pipeline for utilizing HOMER to analyze CLIP-seq (which is similar to RIP-seq) has also been published (Liu et al., 2015). Liu et al. demonstrated that normalizing CLIP-data to input mRNA sequences using HOMER as we did can improve specificity of RBP-bound mRNA targets.

We considered several tools for analyzing the RIP-seq data such as RIP-Seeker (Li et al., 2013), and AsPeaks (Kucukural et al., 2013) designed specifically for RIP-seq or CLIP-seq data. However, it appears that due to lack of maintenance, these software packages are no longer fully operational. For example, RIP-seeker was last updated 6 years ago and the associated R packages are now outdated and therefore difficult to use. Piranha (Uren et al., 2012) is another package that has not been updated, although it is available through the GALAXY webserver in a limited way.

Based on the status of these alternative software packages and the fact that HOMER has been shown to be an effective tool for analyzing CLIP-seq as well as a large variety of RNA base sequencing technologies, we feel it is justified to use HOMER in our RIP-seq analysis pipeline.

A limitation of HOMER that is evident in Table S1 (column F) is the loss of strand-specific information, which should be readily available considering the kit used for library preparation. Consequently, sequence reads that could correspond to either sense or antisense transcripts are lumped together within a peak and considered as one transcript.

We agree with the reviewer that it is important to preserve strand specific information to accurately identify binding sites on specific genes. HOMER does not inherently lose strand- specific information, although we did not use this information in our prior analysis. Hence, we redid our analysis preserving strand-specific information. First, we identified the peaks using HOMER, then using the GenomicRanges and GenomicAlignments packages in R, the reads associated with either the positive or negative strand at the identified peak locations were counted to generate a counts matrix. This was then normalized and annotated in R using the gencode annotation for mouse (M21).

Another concern with this analysis is that sequencing reads that could correspond to other exons of a transcript not considered as peaks, would have been discounted from further analysis. Could the Bayes factor analysis yield different results if these reads were also included in the analysis?

We agree that transcripts not considered peaks would have been discounted from the analysis. Our goal was to measure peaks enriched above input in order to identify transcripts specifically bound to STAU2 at levels higher than background. Including exons of a transcript not enriched above input will increase false positives, therefore we did not do this. However, including non-enriched exons would not affect the Bayes factor analysis since such sequences or transcripts would have a zero weight within the analysis.

"This allowed us to focus on the bound cargo at this level, consisting of 16071 transcripts, corresponding to 9412 genes, for further analysis (Table S1)." This is a very large gene list and a concern is that some proportion of this could be non-specific binding. The authors should comment on this.

After the reanalysis performed in response to reviewers' comments, we identified 18,975 peaks corresponding to 7,968 genes. We acknowledge that our study identifies a larger STAU2 RNA cargo than previous studies. This could be due to several factors:

i) We utilized RNA-sequencing while most other studies used microarrays (Heraud-Farlow et al., 2013; Kusek et al., 2012); RNA-seq generally reveals a higher number of transcripts than microarrays. As indicated in the discussion, several previous studies using RNA-seq for other RBPs identified >10,000 peaks associated with a single RBP using CLIP-seq (Ascano Jr et al., 2012; Maurin et al., 2018). For example, Maurin et al., 2018 found 14,732 peaks bound to the RBP FMRP across 3 brain regions using HITS CLIP-seq.

ii) Maurin et al., 2018 trimmed their large data set to exclude those having less than 50 counts or reads. The authors mention that *'taking in consideration only those having at least 50 fragments (counts) co-immunoprecipitated with FMRP, we found a total of 1,065 targets.'* However, we chose not to use an arbitrary cutoff e.g. of 50 counts or reads, since such pruning of the dataset could lead to a loss of bona fide targets. Rather, we assessed the peak trajectory (Fig. S2B) and used the point of maximum curvature (the elbow of the curve) as a cutoff, corresponding to a peak score of 12.5. By removing peaks in the more incremental regime (below the cutoff) we

sought to reduce the likelihood of including false positives in our analysis. As the scores increase more rapidly we expect them to more accurately reflect meaningful features of our system. Changes to the text explaining this are made (p.6, p.21).

iii) We would like to also point out that the 18,975 peaks/7,968 genes were identified over four time points when the tissue is changing gene expression. Given that approximately half of the total cargo was dynamically bound, demonstrating variation between timepoints, we expect to see an increase in the total number of genes identified vs a single timepoint experiment.
With regards to nonspecific binding, at the suggestion of Reviewer 3, we performed an analysis of endothelial cell markers in the STAU2 cargo and found only low levels of binding, and we note that some of these may be real captured targets as there is a low level of STAU2 expression in endothelial cells (see p.10 of this document).

"Of the bound transcripts, most (78.6%) were classified as protein coding, 6.84% as pseudogenes, 2.48% as long intergenic non-coding RNA (lincRNA), 2.08% as microRNA (miRNA), and less than 1% were miscellaneous RNA (miscRNA), ribosomal RNA (rRNA), small Cajal body-specific RNA (scaRNA), small nucleolar RNA (snoRNA), and small nuclear RNA (snRNA) (Fig. 2B)." It is unclear how peaks were annotated in this instance. Looking at Table S1 however, it seems that each individual peak was assigned to the nearest TSS (columns L to T in Table S1). If this is the case, this method is flawed. A peak that is far away from a gene's TSS could be erroneously assigned to a different gene with a nearer TSS.

Our analysis of two putative targets Nfib and DCC identified such errors were apparent. For example, the peak 'Merged-chr4-82350385-1' overlapping exon 6 of Nfib is annotated as overlapping intron 6 of Nfib in columns I and J in Table S1. However, when assigned to the nearest TSS, this peak is incorrectly assigned to the rRNA n-R5s188 despite being 100kb from this rRNA. Similarly, the peak 'Merged-chr18-71299375-1' overlapping exon 27 of Dcc is annotated as overlapping this exon of Dcc in columns I and J of Table S1 but incorrectly assigned to the nearest TSS of the pseudogene AC103658.1. Results in Figures 2D, E and Figure 3 may also need to be re-analysed because of this.

We thank the reviewer for bringing up this point. We realize that the raw HOMER output generated a column with the predicted nearest TSS. This information was not relevant since we focused on peaks within genes (Column I of Table S1 in prior submission). Please note that the erroneous column was not used in the analyses shown in the prior version of the paper.
Nevertheless, after reanalysis, we have regenerated the table based on the GENCODE annotation (Table S1) and made sure that no inappropriate columns appear in the table.

"Using R, a Bayes factor (BF) was assigned to each gene representing the degree of differential binding between two consecutive timepoints, which was then summed across the time course (summed ln(BF))(Table S2)." Was this analysis performed using read counts within peaks? A description should be included in the Methods or supplementary material of how this was performed, including how data between the different time points were normalized to account for sequencing depth.

We apologize that the original methods provided were not sufficiently detailed. As mentioned above, we now include an overview of the RIP seq pipeline in Fig. 2A and we have expanded the RIP-seq analysis methods section in Methods (p.21) and supplemental methods.

We confirm that the Bayes factor analysis was based on read counts within peaks. We normalized the counts by sampling 10 million reads so that the counts for each sample all sum to approximately 10 million.

"In contrast, the top biological process and pathways for the 'dynamic' genes included ion transport, regulation of neuron differentiation, synapse organization, ER to Golgi transport, axon guidance, and extracellular matrix (ECM) organization (Fig. 3E,F), suggesting that STAU2 participates in developmentally regulated processes by altering subsets of genes bound at different time-points."

As a control, could the Bayes factor analysis also be performed using RNA-seq data from identical

or similar tissue at these ages? For example, using ENCODE forebrain RNA-seq data at E11.5, E13.5, E15.5 and E18.5 (or other similar datasets). How would the result of such an analysis differ from the results with RIP-seq? Is the STAU2 cargo a reflection of differential gene expression due to differences in gene expression and tissue composition between the ages, or is STAU2 specifically binding to some highly expressed transcripts but not others?

We would like to thank Reviewer 1 for raising these important issues. Since our peak calling method was designed to identify STAU2-bound transcripts significantly enriched above input, we do not expect the cargo to be simply a reflection of the input. Note that we used a limit of 4- fold above input seen across the three replicates to define a peak, which is quite stringent. Taking into account the normalization to 10 million reads along with the requirement for a 4-fold enrichment, the Bayes factor analysis will self-correct for increases in gene expression and the rate of STAU2 binding: if the gene expression goes up and STAU2 binds the same proportion, then no change occurs in the enrichment; conversely, if more or less enrichment occurs then the proportion of STAU2 binding must be changing relative to input.

The Bayes factor analysis works by utilizing peaks enriched above background (input in this case) to define dynamic and stable genes, and so it is not directly applicable to standard RNA- seq analysis such as the ENCODE forebrain RNA-seq data.

Additional comments:

The abstract is rambling and raises the interesting motivation for the study to investigate the hypothesis "that STAU2 cargo would change as IPCs become specified to produce first neurons and then glia" but never comes back to this.

Thank you for pointing this out, we have tightened up the abstract to make it flow better (p.2). Regarding IPC specification over time, we have made the following points more clearly in the manuscript that support the model that STAU2 cargo changes as IPCs become specified to produce the normal order of cortical cell types:

- We mined existing databases for known IPC specification genes and demonstrated enrichment of subsets of these in the STAU2 cargo at different stages of development. This underscores that STAU2 cargo includes known IPC determinants and that the combination of determinants changes over time.
- Our analysis of a database of layer-specific genes has revealed differential association in the cargo at different times, with deep layer genes declining earlier than upper layer genes, mirroring the normal order of cortical neuron subtype production. Several known gliogenesis genes were present at later stages. Hence this temporal analysis reveals deep then upper and gliogenic genes in the cargo, mirroring the timing in vivo.

"..of a total of 29285 bound transcripts.." The term 'transcripts' used from this sentence onwards is confusing and the term 'peaks' may be better suited. This is because multiple peaks could arise from separate exons of a single gene (all corresponding to one transcript).-

We agree and have changed the term 'transcripts' to 'peaks' where appropriate.

Reviewer 2 Advance Summary and Potential Significance to Field: *Previous work from the Temple group and others has shown that, as in Drosophila, Staufen is asymmetrically inherited by one daughter cell during neural progenitor cell division. In this manuscript, the authors report data on analysis of the RNAs bound to the Staufen homolog, STAU2, during mouse cerebral cortex development. Based on a timecourse analysis, where the analysis of STAU2-associated RNAs was carried out at four developmental stages, the authors conclude that STAU2 changes its RNA cargoes in cortical radial glial cells during development, and those changes are functionally important for regulating the fate of the intermediate progenitor cells (IPCs) which inherit STAU2 and its RNA cargoes.*

Reviewer 2 Comments for the Author:

Unfortunately, the data as reported lack several important controls and replicates, which make the conclusions made by the authors challenging to substantiate. This may be in part due to how

the paper is presented, rather than the data themselves, as the manuscript skimps on many details, the methods are very short, and the figure legends are not sufficiently detailed.

We appreciate the reviewer's concerns, and we have expanded the methods and figure legends.

For example, the authors have carried out IPs for STAU2 from lysates of developing cortex, following by RNA-seq of RNAs harvested from that material, as well as the input lysate.

Enrichment of STAU2-bound RNAs is then calculated by comparison with the input whole tissue lysate. Although it does not appear to be stated in the text or methods, each RIP seems to have been carried out once. The lack of replicates, given the very large number of RNAs found in each RIP is a concern. This could be mitigated by comparing the E15 and E17 data, which are likely to be the most similar, in terms of cell composition and developmental stage.

We apologize for the lack of clarity in the methods. We did in fact sequence 3 replicates of STAU2 IP samples and 3 replicates of the input RNA at each time point so we did not need to perform the comparison suggested between E15 and E17. We have made this point clear in the text (p.6), methods (p.21) and in Fig. 2A in the revised manuscript.

In any case, identifying almost 10,000 genes as STAU2-bound, using relatively conservative calling methods from an original total of 29,285 transcripts, is puzzling and raises many questions. Previous studies identified around 1,000 RNAs bound by STAU2 in rat E17 brain (Heraud-Farlow et al., Cell Reports 2013), estimated as around 8.5% of mRNAs present in the input RNA. Subsequent research using the more accurate CLIP-seq technique identified 356 mRNAs bound by STAU2 in E18 mouse brain (Sharangdhar et al., EMBO Reports, 2017). The authors' own previous work in 2012 (Kusek et al., 2012), identified 1566 STAU2-bound RNAs. Thus the 10,000 RNAs/genes identified in this paper as bound by STAU2 is 6-10- fold higher than that found in previous research, and is likely to include more than half of the mRNAs expressed in the developing cortex in progenitor cells, neurons or both.

Reviewer 1 had the same concern and we addressed it, as repeated below:

> After the reanalysis performed in response to reviewers' comments, we identified 18,975 peaks corresponding to 7,968 genes. We acknowledge that our study identifies a larger STAU2 RNA cargo than previous studies. This could be due to several factors:
>
> i) We utilized RNA-sequencing while most others studying STAU2 cargo used microarrays (Heraud-Farlow et al., 2013; Kusek et al., 2012); RNA-seq generally reveals a higher number of transcripts than microarrays. As indicated in the discussion, several previous studies using RNA- seq for other RBPs identified >10,000 peaks associated with a single RBP using CLIP-seq (Ascano Jr et al., 2012; Maurin et al., 2018). For example, Maurin et al., 2018 found 14,732 peaks bound to the RBP FMRP across 3 brain regions using HITS CLIP-seq.
>
> ii) Maurin et al., 2018 trimmed their large data set to exclude those having less than 50 counts or reads. The authors mention that *'taking in consideration only those having at least 50 fragments (counts) co-immunoprecipitated with FMRP, we found a total of 1,065 targets.'* However, we chose not to use an arbitrary cutoff e.g. of 50 counts or reads, since such pruning of the dataset could lead to a loss of bona fide targets. Rather, we assessed the peak trajectory (Fig. S2B) and used the point of maximum curvature (the elbow of the curve) as a cutoff, corresponding to a peak score of 12.5. By removing peaks in the more incremental regime (below the cutoff) we sought to reduce the likelihood of including false positives in our analysis. As the scores increase more rapidly we expect them to more accurately reflect meaningful features of our system.
>
> iii) We would like to also point out that the 18,975 peaks/7,968 genes were identified over four time points when the tissue is changing gene expression. Given that approximately half of the total cargo was dynamically bound, demonstrating variation between timepoints, we expect to see an increase in the total number of genes

identified vs a single timepoint experiment.

With regards to nonspecific binding, at the suggestion of Reviewer 3, we performed an analysis of endothelial cell markers in the STAU2 cargo and found only low levels of binding, and we note that some of these may be real captured targets as there is a low level of STAU2 expression in endothelial cells (see p.10 of this document).

As mentioned by the reviewer, CLIP followed by RNA seq (Sharangdhar et al., 2017) identified 356 mRNAs bound by STAU2 in E18 mouse brain. However, only targets bound to STAU2 isoform >60 KDa, the least abundant isoform, were assessed. Therefore, RNAs bound to the remaining isoforms (52, 56, 59 KDa, (Duchaîne et al., 2002)) were excluded, potentially accounting for the low number of transcripts identified compared to our studies using an antibody that also detected the more abundant isoforms.

We intentionally used the non-cross-linked RIP method to enable the capture and identification of both direct and indirectly associated mRNA cargo because our goal was to identify intact mRNP complexes rather than STAU2 binding sites, which cross-linked RIP protocols like CLIP are better suited for. Furthermore, double stranded (ds) RBPs such as STAU2 bind the deep and major groove of A-form RNA helical structures with little or no sequence specificity (Liu et al., 1996; Wheeler et al., 2018), and due to this property, may crosslink relatively inefficiently using traditional UV-crosslinking methods (Wheeler et al., 2018).

The very high number of RNAs identified here may be due to the almost-ubiquitous expression of STAU2 in the developing cortex, in different classes of progenitor cells and in neurons. However, the lack of technical controls in the paper, which might address this, is a weakness. The authors state that a control antibody brought down very little RNA, so was not used here. This is unusual, as RNA is notoriously "sticky" and typically a good deal of total RNA is recovered from non-specific binding to the beads, protein A/G etc. Combined with what appears to be a lack of replicates, it raises questions about whether the RNAs identified are essentially a sample of the whole lysate, either bound non-specifically to the IP material, or associating with STAU2 in solution.

Please note as explained above that we did have 3 replicates. Regarding the fact that we chose not to use a negative control antibody, our previous experience and that of our collaborator Dr. Tenenbaum who pioneered RIP studies, has been that we recovered too little RNA to perform bulk sequencing. Moreover, we want to stress that the Liu et al, 2015 study found CLIP-data normalized to input mRNA sequences was best able to identify specific targets, and this was our approach. Finally, we addressed this empirically by adding a new QC figure. We assessed the expression of 16 candidate genes selected randomly from the E15.5 STAU2 RIP-seq dataset for binding to STAU2 or a negative control IgG antibody by qRT-PCR (Fig. S2C). In each case, we found a many-fold (ranging between 4-290 fold) higher gene expression in the STAU2 IP samples compared to the negative control IgG antibody. This supports our experience to date that using the negative control antibody is not necessary for accurate target identification.

The authors' reference in the methods to the approach taken (Baroni et al., 2008) is not included in the reference list, but it would appear that the method does not involve crosslinking RNA to protein, so novel associations in solution between STAU2 and RNA during the IP process may be a confounding issue. Alternatively, the very high number of identified STAU2- bound RNAs may reflect the analysis method used to call peak enrichment, which is not well described in the methods, and which seems to be enrichment relative to the input RNA. This very high number of STAU2-bound RNAs makes analyses of functional associations difficult to assess, as such a high fraction of all RNAs appear to be bound by the protein. The change in RNAs identified over time, in this context, may be itself artefactual, reflecting the change in RNA composition in the lysate - this is hard to judge, without appropriate negative controls. As such, it makes detailed consideration of the subsequent claims from the paper redundant, without addressing the question of the very high number of STAU2-bound RNAs and the various technical points raised above.

We have ensured that the Baroni et al., 2008 reference is listed in the References section. We hope that the specific responses to each of the points detailed above, especially concerning

replicates and non-specific binding will be satisfactory. In addition, the issue of STAU2-RNA association in solution or post-lysis binding is addressed in comments to Reviewer 3.

Reviewer 3 Advance Summary and Potential Significance to Field

The manuscript entitled "STAU2 binds a complex RNA cargo that changes temporally with production of diverse intermediate progenitors during mouse corticogenesis" comes from the lab of one of the pioneers of neural stem cell biology and makes a very important contribution to developmental neurobiology. In a nutshell, the authors have identified the various RNA molecules that bind to the RNA-binding protein STAU2 at four different stages of mouse cortical development (E11.5-E17.5). Interestingly, they show that there are two subsets of these RNA molecules, (i) those present at all developmental stages analysed, and (ii) those that changed with a given developmental stage. In addition, the authors performed functional studies on one STAU2 target, Taf13, and demonstrate that its knockdown in vitro results in reduced oligodendrogenesis.

This study not only provides key information on the behaviour of this important RNA-binding protein during corticogenesis, but also constitutes an enormously useful resource for future studies. The study is therefore highly appropriate for publication in Development. I only have one major point and a few minor points that should be addressed before publication.

Reviewer 3 Comments for the Author:

Major point

The authors use lysed cortical tissue to pull down and isolate RNA molecules bound to STAU2. How can the authors be certain that there is no post-lysis binding of RNA molecules to STAU2 that are not bound to STAU2 in vivo?

We agree with the reviewer that post-lysis binding of RNA to RBPs is a potential confounding factor. However, extensive and rigorous tests (Penalva et al., 2004) using RIP protocols similar to ours (Tenenbaum et al., 2002) did not detect any post-lysis RNA reassortment (discussed in (Nicholson et al., 2017)). Nicholson et al. further pointed out that studies reporting post-lysis RNA binding (Mili and Steitz, 2004; Riley et al., 2012) used harsh lysis procedures such as sonication, which is likely to have caused fragmentation of mRNAs and ribonucleoproteins (mRNPs), possibly leading to post-lysis RNA reassortment. Our RIP protocols involved mild lysis conditions and gentle washing and no sonication, and were designed to minimize this issue (Kusek et al., 2012; Tenenbaum et al., 2002); explanatory text was added to p. 6-7.

The control data shown in Fig. 4 are nice, but this reviewer wonders whether there are additional controls to exclude potential post-lysis binding? For example, on the assumption that STAU2 is not expressed in endothelial cells of the embryonic mouse cortex, if the authors could show that RNA molecules specific to endothelial cells were absent from the sets of STAU2- bound RNA molecules, that would be a strong piece of evidence against post-lysis binding. If such control experiments are not feasible, the authors should at least include a discussion of this issue in the revised manuscript.

Thank you for this extremely useful suggestion. Using the single cell dataset (Loo et al., 2019), we found that STAU2 expression in endothelial cells (ECs) is indeed much lower than other cell types at E14.5. However, other studies indicate STAU2 is expressed in ECs in single cells derived from developing mouse brains and spinal cords (P2 and P11) (Rosenberg et al., 2018) as well the adult mouse brain ((Zhang et al., 2014)- https://www.brainrnaseq.org/, (Zeisel et al., 2015)- http://linnarssonlab.org/cortex/) therefore we were unable to rule out expression of STAU2 in ECs of the developing mouse brain. Nevertheless, we thought it would be valuable to pursue this concept because the level of STAU2 in ECs is significantly lower than that in neurons. Therefore, if we found high levels of EC transcript binding, this could be indicative of post-lysis binding.
The Daneman lab has reported mouse EC markers enriched over whole brain (Daneman et al., 2010). We evaluated the top 100 genes most highly enriched in ECs over whole brain and found only 7 present in the cargo in our reanalyzed data. We also assessed EC markers identified in

other publications and found ~10% of top EC markers in the STAU2 cargo (Daneman et al., 2010; Hupe et al., 2017). While most common EC markers such as *Pecam1*, *Pglyrp1*, *Tie1*, *Epas1*, *Flt1* and *Fli1* were absent from our dataset, we found 2 specific EC markers, *Tek* and *Kdr* in the STAU2 cargo. However, based on studies mentioned above (Rosenberg et al., 2018; Zeisel et al., 2015; Zhang et al., 2014), STAU2 maybe be expressed at a low level in ECs, and therefore we cannot rule out the possibility that STAU2 is genuinely binding these RNAs in ECs. A description of this analysis is added to supplementary information, p.3.

While using ECs or other non-neural cell type markers for elimination of false positives is not entirely secure without clear negative STAU2 expression, we would like to point out that other measures we have taken in this revision, including improving peak annotation, applying a slightly higher cutoff based on the point of maximum curvature or elbow of the peak trajectory (Fig. S2B), and qRT-PCR validation (Fig. S2C) provide further confidence that we have generated an authentic and more comprehensive list of STAU2 cargo genes.

In this context, the authors should comment on categories related to inner-mitochondrial processes, such as mitochondrial translation or TCA cycle.

As this reviewer correctly pointed out, pathways related to inner-mitochondrial processes in our previous analysis were associated with stable (Fig. 3D) and deep layer STAU2 cargo genes (Fig. 5E). Within the reanalyzed dataset, we found pathways related to mitochondrial gene expression, translation (all STAU2 cargo-Table S3, stable STAU2 cargo- Table S6) and TCA cycle (layer V and DL cargo-Table 3, S3, Fig. 5E).

It is our understanding of the reviewer's comment, that since STAU2 is not known to be expressed in mitochondria, the reason that mitochondrial genes are present in the cargo is not clear. If we have understood correctly, we theorize the following: mitochondrial DNA codes for 13 polypeptides (constituents of the enzyme complexes of the oxidative phosphorylation system), two mitochondrial ribosomal RNAs (mt-rRNAs) and 22 mt-tRNAs, whereas all other protein components are encoded by nuclear genes and imported into mitochondria from the cytosol (Pearce et al., 2017). Because mitochondrial DNA products were not found in the STAU2 cargo, we suggest STAU2 may be involved in the transport or transcription of the nuclear genes involved in mitochondrial gene expression or TCA cycle that are then imported into the mitochondria. Moreover, STAU2-bound RNAs have previously been shown to be enriched for mitochondrial genes in E17 ER depleted rat brain extracts (Heraud-Farlow et al., 2013).

Minor points

1. *Summary Statement: "...contribute to IPC..." I guess the word "diversity" is missing?*

Thank you, 'diversity' is now included in the summary statement.

2. *On page 10, the authors state that they used "...transcriptomes obtained at mid-gestation in the mouse (Fietz et al. 2012).". This reviewer is curious if the authors also used other transcriptomes, for example those reported by Florio et al. 2015?*

We did use two other transcriptome databases, Aprea et al., 2013 (described in Fig. 4C and 4D) and Ayoub et al., 2011 and made similar findings as we did using the Fietz at al., 2012 dataset, although we did not show this analysis. (Aprea et al., 2013; Ayoub et al., 2011; Fietz et al., 2012). Thank you for bringing to our attention this pioneering work on expansion of the human neocortex (Florio et al., 2015). Since our data is derived purely from the mouse brain, we had difficulty comparing the mouse/human filtered data from Florio et al., 2015 against ours, and due to this, the analysis seems too preliminary to be included within the current manuscript, although it will be helpful in future studies examining the role of STAU2 targets in human neurons.

We greatly appreciate the reviewers' input and hope we have addressed each of their concerns satisfactorily.

References

Aprea, J., Prenninger, S., Dori, M., Ghosh, T., Monasor, L. S., Wessendorf, E., Zocher, S., Massalini, S., Alexopoulou, D., Lesche, M., et al. (2013). Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment. *EMBO J.* **32**, 3145–3160.

Ascano Jr, M., Mukherjee, N., Bandaru, P., Miller, J. B., Nusbaum, J. D., Corcoran, D. L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M., et al. (2012). FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* **492**, 382–386.

Ayoub, A. E., Oh, S., Xie, Y., Leng, J., Cotney, J., Dominguez, M. H., Noonan, J. P. and Rakic, P. (2011). Transcriptional programs in transient embryonic zones of the cerebral cortex defined by high-resolution mRNA sequencing. *Proc. Natl. Acad. Sci.* **108**, 14950 LP- 14955.

Daneman, R., Zhou, L., Agalliu, D., Cahoy, J. D., Kaushal, A. and Barres, B. A. (2010). The Mouse Blood-Brain Barrier Transcriptome: A New Resource for Understanding the Development and Function of Brain Endothelial Cells. *PLoS One* **5**, e13741.

Duchaîne, T. F., Hemraj, I., Furic, L., Deitinghoff, A., Kiebler, M. A. and DesGroseillers, L. (2002). Staufen2 isoforms localize to the somatodendritic domain of neurons and interact with different organelles. *J. Cell Sci.* **115**, 3285–3295.

Fietz, S. A., Lachmann, R., Brandl, H., Kircher, M., Samusik, N., Schröder, R., Lakshmanaperumal, N., Henry, I., Vogt, J., Riehn, A., et al. (2012). Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal. *Proc. Natl. Acad. Sci.* **109**, 11836 LP-11841.

Florio, M., Albert, M., Taverna, E., Namba, T., Brandl, H., Lewitus, E., Haffner, C., Sykes, A., Wong, F. K., Peters, J., et al. (2015). Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science (80-.).* **347**, 1465 LP-1470.

Heraud-Farlow, J. E., Sharangdhar, T., Li, X., Pfeifer, P., Tauber, S., Orozco, D., Hörmann, A., Thomas, S., Bakosova, A., Farlow, A. R., et al. (2013). Staufen2 regulates neuronal target RNAs. *Cell Rep.* **5**, 1511–1518.

Hupe, M., Li, M. X., Kneitz, S., Davydova, D., Yokota, C., Kele, J., Hot, B., Stenman, J. M. and Gessler, M. (2017). Gene expression profiles of brain endothelial cells during embryonic development at bulk and single-cell levels. *Sci. Signal.* **10**, eaag2476.

Kucukural, A., Özadam, H., Singh, G., Moore, M. J. and Cenik, C. (2013). ASPeak: an abundance sensitive peak detection algorithm for RIP-Seq. *Bioinformatics* **29**, 2485–2486.

Kusek, G., Campbell, M., Doyle, F., Tenenbaum, S. A., Kiebler, M. and Temple, S. (2012). Asymmetric segregation of the double-stranded RNA binding protein Staufen2 during mammalian neural stem cell divisions promotes lineage progression. *Cell Stem Cell* **11**, 505–516.

Li, Y., Zhao, D. Y., Greenblatt, J. F. and Zhang, Z. (2013). RIPSeeker: a statistical package for identifying protein-associated transcripts from RIP-seq experiments. *Nucleic Acids Res.* **41**, e94–e94.

Liu, Z. R., Wilkie, A. M., Clemens, M. J. and Smith, C. W. (1996). Detection of double- stranded RNA-protein interactions by methylene blue-mediated photo-crosslinking. *RNA* **2**, 611–621.

Liu, Q., Zhong, X., Madison, B. B., Rustgi, A. K. and Shyr, Y. (2015). Assessing Computational Steps for CLIP-Seq Data Analysis. *Biomed Res. Int.* **2015**, 196082.

Loo, L., Simon, J. M., Xing, L., McCoy, E. S., Niehaus, J. K., Guo, J., Anton, E. S. and Zylka, M. J. (2019). Single-cell transcriptomic analysis of mouse neocortical development. *Nat. Commun.* **10**, 134.

Maurin, T., Lebrigand, K., Castagnola, S., Paquet, A., Jarjat, M., Popa, A., Grossi, M., Rage, F. and Bardoni, B. (2018). HITS-CLIP in various brain areas reveals new targets and new modalities of RNA binding by fragile X mental retardation protein. *Nucleic Acids Res.* **46**, 6344–6355.

Mili, S. and Steitz, J. A. (2004). Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* **10**, 1692–1694.

Nicholson, C. O., Friedersdorf, M. and Keene, J. D. (2017). Quantifying RNA binding sites transcriptome-wide using DO-RIP-seq. *RNA* **23**, 32–46.

Pearce, S. F., Rebelo-Guiomar, P., D'Souza, A. R., Powell, C. A., Van Haute, L. and Minczuk, M. (2017). Regulation of Mammalian Mitochondrial Gene Expression: Recent Advances. *Trends Biochem. Sci.* **42**, 625–639.

Penalva, L. O. F., Burdick, M. D., Lin, S. M., Sutterluety, H. and Keene, J. D. (2004). RNA-

binding proteins to assess gene expression states of co-cultivated cells in response to tumor cells. *Mol. Cancer* **3**, 24.

Riley, K. J., Yario, T. A. and Steitz, J. A. (2012). Association of Argonaute proteins and microRNAs can occur after cell lysis. *RNA* **18**, 1581–1585.

Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L. T., Peeler, D. J., Mukherjee, S., Chen, W., et al. **(2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding.** ***Science (80-.)***. 360**, 176 LP-182.**

Sharangdhar, T., Sugimoto, Y., Heraud- Farlow, J., Fernández- Moya, S. M., Ehses, J., Ruiz de los Mozos, I., Ule, J. and Kiebler, M. A. (2017). A retained intron in the 3′- UTR of *Calm3* mRNA mediates its Staufen2- and activity- dependent localization to neuronal dendrites. *EMBO Rep*. e201744334.

Tenenbaum, S. A., Lager, P. J., Carson, C. C. and Keene, J. D. (2002). Ribonomics: identifying mRNA subsets in mRNP complexes using antibodies to RNA-binding proteins and genomic arrays. *Methods* **26**, 191–198.

Uren, P. J., Bahrami-Samani, E., Burns, S. C., Qiao, M., Karginov, F. V, Hodges, E., Hannon, G. J., Sanford, J. R., Penalva, L. O. F. and Smith, A. D. (2012). Site identification in high-throughput RNA–protein interaction data. *Bioinformatics* **28**, 3013– 3020.

Wheeler, E. C., Van Nostrand, E. L. and Yeo, G. W. (2018). Advances and challenges in the detection of transcriptome-wide protein–RNA interactions. *WIREs RNA* **9**, e1436.

Zeisel, A., Muñoz-Manchado, A. B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science (80-.)*. **347**, 1138 LP-1142.

Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., Phatnani, H. P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. (2014). An RNA- Sequencing Transcriptome and Splicing Database of Glia, Neurons, and Vascular Cells of the Cerebral Cortex. *J. Neurosci.* **34**, 11929 LP-11947.

Second decision letter

MS ID#: DEVELOP/2020/199376

MS TITLE: STAU2 binds a Complex RNA Cargo that Changes Temporally With Production of Diverse Intermediate Progenitors During Mouse Corticogenesis

AUTHORS: Rebecca Chowdhury, Yue Wang, Melissa Campbell, Susan K Goderie, Francis Doyle, Scott A. Tenenbaum, Gretchen Kusek, Thomas R. Kiehl, Suraiya Ansari, Nathan C. Boles, and Sally Temple

I have now received all the referees reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

The overall evaluation is positive and we would like to publish a revised manuscript in Development, provided that the referees' comments can be satisfactorily addressed. Please attend to all of the reviewers' comments in your revised manuscript and detail them in your point-by-point response. If you do not agree with any of their criticisms or suggestions explain clearly why this is so.

We are aware that you may currently be unable to access the lab to undertake experimental revisions. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Reviewer 1

Advance summary and potential significance to field

The paper provides a resource of identified RNA cargo molecules of STAU2 across different developmental stages during cortical development. it is potentially significant for understanding mechanisms regulating neural development.

Comments for the author

Overall, the authors have greatly improved their analysis which strengthens the authors' claims that the STAU2 cargo changes temporally. This basic analysis is important but the authors do not address how or why the "dynamic" cargo changes.
This could be due to changes in gene expression and/or RNA modifications as cells undergo different stages of development, but how the cargo is changed for STAU2 is unclear. The authors may wish to comment on how this occurs in the Discussion.

Additional minor comments:
1. In the analysis of RIP-seq data, why was strand specificity considered after and not before peak calling with HOMER?

2. "Of all bound genes, 4508 were bound at a single peak, 2,797 had two or three peaks per gene and the remainder (661) had four or more with a maximum of 14 peaks (Fig. 2C)."
The authors should not refer to peaks as 'bound' regions. The whole RNA transcript should be IP'ed during the experiment unless there is significant RNA degradation.

It is not clear how read pairs that map to separate exons of a gene are considered when calling peaks using factor mode in HOMER (originally designed for ChIP-seq analysis). Could these be discounted from peak calling if they are too far apart?
Two implications need to be considered if this is possible. 1. Are longer exons more likely to be called as peaks as compared to shorter exons? 2. As a consequence of point (1), transcripts that are bound by STAU2 but with intrinsically short exons may not have been detected in the analysis. Another possibility as to how HOMER might have considered these reads is to call a peak within the intron that separates these two exons. It is suggested that the authors check this, if they haven't already, by looking at how reads map to intronic peaks in a genome browser.

3. The y-axis for Figure 5B is not properly scaled.

4. Page 3, Introduction, should the word maroglia be followed by "such as astrocytes and oligodendrocytes", or do the authors mean "microglia" and if so these are not derived from the neuroepithelium.

Reviewer 3

Advance summary and potential significance to field

In the revised version of the manuscript, the authors have satisfactorily addressed the points I had raised previously.

Comments for the author

In the revised version of the manuscript, the authors have satisfactorily addressed the points I had raised previously.

Second revision

Author response to reviewers' comments

We have also uploaded a correctly formatted PDF titled 'Response to Reviewers Revision 2'.

We are very grateful to the reviewers for their comments and suggestions. Relevant changes in the manuscript are highlighted in red text. Below we have responded in detail (purple text) to the reviewer's comments (in italics).

Reviewer 1 Advance Summary and Potential Significance to Field:
The paper provides a resource of identified RNA cargo molecules of STAU2 across different developmental stages during cortical development. it is potentially significant for understanding mechanisms regulating neural development.

Reviewer 1 Comments for the Author:
Overall, the authors have greatly improved their analysis which strengthens the authors' claims that the STAU2 cargo changes temporally. This basic analysis is important but the authors do not address how or why the "dynamic" cargo changes. This could be due to changes in gene expression and/or RNA modifications as cells undergo different stages of development, but how the cargo is changed for STAU2 is unclear. The authors may wish to comment on how this occurs in the Discussion.

We thank the reviewer for the positive feedback on the revised manuscript. Regarding the cargo dynamism, this is an important point to consider, and we added it to the discussion (p.17):

"This dynamicity in the cargo may be due to several factors such as altered gene expression over time, altered modification of RNAs that changes their binding affinity to STAU2, or alterations in STAU2 protein expression, modification, protein-protein interactions or localization."

Additional minor comments:
1. In the analysis of RIP-seq data, why was strand specificity considered after and not before peak calling with HOMER?

We used HOMER to determine the regions where STAU2 showed increased binding over the background, and then used those regions to calculate the counts on each strand from the data for further analysis. We chose to do it this way given that strand-specificity can be accounted for before or after.

2. "Of all bound genes, 4508 were bound at a single peak, 2,797 had two or three peaks per gene and the remainder (661) had four or more with a maximum of 14 peaks (Fig. 2C)." The authors should not refer to peaks as 'bound' regions. The whole RNA transcript should be IP'ed during the experiment unless there is significant RNA degradation.

We agree with the reviewer that entire transcripts may be immunoprecipitated with the antibody and therefore 'bound regions' is not the correct wording. We have edited the MS to clarify (p.7), and in addition, we searched the manuscript to make sure this error was not elsewhere.

New wording p. 7:
Distribution of the types of peaks identified and the genomic features associated with the peaks were analyzed using R. The majority of captured sequences were categorized as 3'UTR (6,750), followed by exons (5,983), intergenic (4,513), 5'UTRs (1,140) and introns (589) (Fig. 2B). A single peak was identified for 4508 genes, two or three peaks for 2,797 genes and four or more for the remainder (661), with a maximum of 14 peaks per gene (Fig. 2C). As an example, we examined 3 genes belonging to the TATA-box-binding-protein-associated factor (TAF) family enriched in the cargo, *Taf10, Taf11 and Taf13*, and found different patterns. *Taf10* and *Taf13* transcripts each had peaks at 2 sites within exonic and 3'UTR regions, while *Taf11* peaks were observed in the 5'UTR and 3'UTR regions (Fig. S2D). Note that the location of a peak was not used to identify binding location, rather we interpret a peak as indicative of the presence of the transcript in the cargo.

It is not clear how read pairs that map to separate exons of a gene are considered when calling peaks using factor mode in HOMER (originally designed for ChIP-seq analysis). Could these be discounted from peak calling if they are too far apart? Two implications need to be considered if this is possible. 1. Are longer exons more likely to be called as peaks as compared to shorter exons? 2. As a consequence of point (1), transcripts that are bound by STAU2 but with intrinsically short exons may not have been detected in the analysis. Another possibility as to how HOMER might have considered these reads is to call a peak within the intron that separates these two exons. It is suggested that the authors check this, if they haven't already, by looking at how reads map to intronic peaks in a genome browser.

The RIP-seq data is from single-end sequencing and not paired-end sequencing. We apologize for not making this clear, we previously included it in the GEO submission but now it is also stated in the methods section (p.21). Since our approach does not have read pairs, the concerns noted by the reviewer are not applicable to our data. In the case of single end sequencing, any count for a peak can be traced back to specific reads. Because the reads are mapped directly, we identified the peak location (exon/intron) and do not have to be concerned about interpolating sequences between read pairs that might have led to the errors noted by the reviewer. Furthermore, to be certain, we checked 20 intronic peaks and found they could all be matched to specific intronic reads. Nevertheless, as mentioned above and included in the text (p.7), we agree with the reviewer that the location of a peak does not reflect binding location, rather we interpret a peak as indicative of the presence of the transcript in the cargo. We hope that this explanation answers the reviewer's concern.

3. The y-axis for Figure 5B is not properly scaled.

Thank you for pointing this out, we have now scaled the y-axis text appropriately.

4. Page 3, Introduction, should the word maroglia be followed by "such as astrocytes and oligodendrocytes", or do the authors mean "microglia" and if so these are not derived from the neuroepithelium.

To avoid confusion, we have removed the word macroglia (meaning astrocytes and oligodendrocytes) from the introduction (p.3).

Reviewer 3 Advance Summary and Potential Significance to Field:
In the revised version of the manuscript, the authors have satisfactorily addressed the points I had raised previously.

We thank the reviewer for this positive response.

We hope that these final changes are acceptable to the reviewers and thank them all again for their valuable input throughout this process which significantly improved our manuscript.

Third decision letter

MS ID#: DEVELOP/2020/199376

MS TITLE: STAU2 binds a Complex RNA Cargo that Changes Temporally With Production of Diverse Intermediate Progenitors During Mouse Corticogenesis

AUTHORS: Rebecca Chowdhury, Yue Wang, Melissa Campbell, Susan K Goderie, Francis Doyle, Scott A. Tenenbaum, Gretchen Kusek, Thomas R. Kiehl, Suraiya Ansari, Nathan C. Boles, and Sally Temple

ARTICLE TYPE: Research Article

I am happy to tell you that your manuscript has been accepted for publication in Development, pending our standard ethics checks.