Sox2-*Evf2* lncRNA-mediated mechanisms of chromosome topological control in developing forebrain

Ivelisse Cajigas, Abhijit Chakraborty, Madison Lynam, Kelsey R. Swyter, Monique Bastidas, Linden Collens, Hao Luo, Ferhat Ay and jhumku D. Kohtz
DOI: 10.1242/dev.197202

Editor: Haruhiko Koseki

Review timeline
Original submission: 22 September 2020
Editorial decision: 4 November 2020
First revision received: 8 January 2021
Editorial decision: 29 January 2021
Second revision received: 4 February 2021
Accepted: 7 February 2021

Original submission

First decision letter

MS ID#: DEVELOP/2020/197202

MS TITLE: Sox2-RNA mechanisms of chromosome topological control in developing forebrain

AUTHORS: Ivelisse Cajigas, Abhijit Chakraborty, Madison Lynam, Kelsey Swyter, Monique Bastidas, Linden Collens, Hao Luo, Ferhat Ay, and jhumku D kohtz

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. If you are able to revise the manuscript along the lines suggested, which may involve further experiments, 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

The authors continue their dissection of the molecular functions of the Evf2 lncRNA by exploring its functional relationship with SOX2 and other components of the Evf2-containing RNP that they previously characterized. First they show that SOX2 directly binds to Evf2 and that this interaction can be inhibited by pGEM RNA, suggesitng a sequence-independent binding mode. Then they show that loss of Evf2 affects binidng of Sox2 at the Dlx5/6 locus and loss of Sox2 affects Evf2-regulated genes. By performing 4C the authors also analyze chromosome conformation in a large region adjacent fo the Dlx5/6 locus and ifnd tat Sox2 and Evf2 KO have complex effects on the various interactions. Next they perform FISH and immunofluorescence to show that Sox2 forms “protein pools” along with Evf2 at target loci that change in locailzation and dimension in Evf2 mutants. Finally, they show that other RNP components also also affeted in their localization by Evf2 mutation.

The mechanism of action of lncRNAs remain largely mysterious. This study represents a very detailed analysis of the spatial and functional interplay between a developmentally important lncRNA, Evf2, and a key transcription factor, SOX2, in a developmental context. As such, I believe the study would be of interest to the readers of Development. However, I find the current version exceedingly difficult to follow and lacking in a well defined, clear model for the function of Evf2-SOX2 interactions. I suggest a few points below that might help clarify.

Comments for the author

Major points

1. Model clarification. I feel that in an attempt to capture all the complexity of this process the figures and the text end up being very difficult to follow thus detracting from the conclusions. From reading this version I am not sure what is the mechanism actually proposed.
1.1. The model in Fig. 4F did not help me very much. Perhaps the authors could expand the model figure and have separate panels showing different steps of the proposed process?
1.2. One problem is that the different components of the Evf2-RNP seem to respond very differently to the Evf2 mutants. Perhaps the authors could consider re-focusing this study on SOX2, for which they have promising in vitro results and potential mutants to explore, and de-emphasize or keep for a future study the other proteins.

2. RNA binding by SOX2. It is unclear to me how SOX2 binds to Evf2 and where the specificity comes from.
2.1. In Fig. 1E the authors state that the binding is “promiscuous”. Is that different from sequence-independent? If not, I would maybe use this expression which is more commonly employed.
2.2. If indeed SOX2 is binding to RNA in sequence independent manner what drives its specific colocalization with Evf2? Some of the protein-interactions? If so can this be demonstrated?
2.3. How promiscuous is the binding? Would DNA compete with Evf2 just as efficiently as the pGEM RNA?

3. RNA binding mutant of SOX2. I think the authors could further leverage the knowledge that the HMG domain is required for the RNA interaction and use these mutants to sharpen the focus on SOX2.
3.1. In Fig. 3K he authors show that residues 40-206 are sufficient for colocalization of SOX2 and Evf2 but they should also be able to show they are necessary otherwise it is difficult to interpret these results. In is particularly surprising that the Δ68–97 deletion increases rather than decreases colocalization. The authors conclude that these internal aminoacids inhibits the protein–RNA interaction but this is hard to envision. Why not test the mutant that cannot bind Evf2 in vitro (Fig. 1E)?
3.2. Perhaps an effective RNA-binding mutant would help interpret some of the discrepancies in the Evf2 and Sox2 dependent effects, for example as shown in Fig. 2.

Minor points
- Fig. 1H: how were the genes classified in green and red here? If from previous work please add a citation. - Fig. 1J: I would not refer to CUT&RUN as a “native ChIPseq method”. I think most people by now are familiar with CUTRUN without explanations. Just cite the first Henikoff paper at first mention.
- FIg. 1J: I think the authors is reading too much into the 120 bp vs 150 bp profiles. If they wish to establish direct vs. indirect interactions additional data are required. In alternative I would tone down these conclusions which anyway do not add much to the main points of the study.
- I am not familiar with the expression “protein pools”. What are the authors refdrring to? Are these accumulation of the protein by homomultimeric interactions or just a reflection of multivalent binding of the protein to a large RNA? Is there a reference for this concept of “protein pools”?

Reviewer 2

Advance summary and potential significance to field

In this study, Cajigas et al. investigate molecular mechanisms of Evf2 ncRNA-mediated gene regulation through Dlx5/6 UCE. Using CUT&RUN, 4C and imaging, they show that Evf2 RNA antagonizes SOX2-dependent Dlx5/6 UCE activation through a direct binding to SOX2, suggesting an implication of SOX2 protein in the gene regulation by Evf2 RNA. This study could give a novel concept for RNA-mediated gene regulation.

Comments for the author

In this study, Cajigas et al. investigate molecular mechanisms of Evf2 ncRNA-mediated gene regulation through Dlx5/6 UCE. Using CUT&RUN, 4C and imaging, they show that Evf2 RNA antagonizes SOX2-dependent Dlx5/6 UCE activation through a direct binding to SOX2, suggesting an implication of SOX2 protein in the gene regulation by Evf2 RNA. They present a large amount of data and reach the conclusion. However, some of the data presented in this study are preliminary and not convincing for me to reach their conclusion. For instance, they present charts showing that SOX2 foci become smaller upon Evf1 KO (Figure 3I), but do not show any cellular images that are the source of these charts. The similar tendency is seen in the 4Cseq data. They show the processed results (e.g. Figure 2A) but do not show a total view of 4C results. These data are not convincing for me. I therefore suggest that the authors should present the data more carefully to concrete their conclusion.

Major points
(1) Figure 2A
In this figure, the authors present 4Cseq data as a simplified and processed form. However, it is not clear how 4Cseq data is processed and presented here. I think it is important to present this data in a more convincing way. For instance, 4C interactomes can also be shown like Figure S2B-D for both Evf2(+) and Evf2(-).

(2) Figure 3
In figure 3, most of fluorescent signals look over-saturated. Especially, gene loci visualized with BAC probes such as Akr1b8 and Rbm28 should show smaller pinpoint signals, not as big as the ones shown in Figure 3A and G. I think this makes difficult to assess a colocalization between these foci. Therefore, these data should be correctly displayed. Alternatively, I recommend the authors to use a super-resolution microscopy instead of a confocal microscopy to assess the colocalization between these molecules.

(3) The title should be modified. The current "Sox2-RNA" could imply that Sox2 binds many RNA species. I think this may be misleading. "Sox2-Evf2 RNA" should be used instead of "Sox2-RNA".

Minor points
(1) Page 5, line 21, “Evf2 RNA binding to Sox is promiscuous”
I think this may be misleading because they do not actually show the promiscuous RNA binding. Does Sox2 bind any part of Evf2 RNA? Otherwise, the author should remove these words.

(2) Page 7, line 6, “Evf2TS/TS (expressing only Evf2-3’)”
These words are confusing. In the Figure 3C, it is shown that Evf2TS/TS expresses Evf1-3’ RNA. The authors have to clarify these discrepancies.

(3) Figure 1H The Sox2 expression level in Evf2TS/+;Sox2fl/fl ;Dlx5/6cre and Evf2+/+;Sox2fl/fl ;Dlx5/6cre is lower than Evf2TS/+;Sox2fl/fl , but still certain amount of Sox2 is expressed compared to Evf2+/+;Sox2fl/fl. Does this mean that the ablation of Sox2 is not efficient with Dlx5/6cre?

(4) Figure 2C-D legend, “I indicates Dlx5/6UCEins that are independent Evf2.”
More explanation in the legend is necessary to understand this sentence.
The authors should clearly describe how this was calculated.

(5) Figure 3E It is not described clearly how “colocalization” is defined. It should be explained in the M&M section.

(6) Figure 3E, F, H, I and J Representative FISH images used for these charts should be shown as supplementary data.

(7) Figure 3 It is not clear that these colocalization between Evf2 RNA, Sox2 PPs, and Evf2-target loci (Akr1b8 and Rbm28)
occurs monoallelically or biallelically. This point should be described in the main text.

(8) Figures 3 and 4 All of colocalization analyses were done in 2D? It should be clearly described in the main text or in the M&M section.

(9) Figure S3 It is difficult to understand this figure because the letters are too small. There is the same tendency for all of figures.
Bigger size letters should be used. The only important part of this figure should be presented.

First revision

Author response to reviewers' comments

We thank the reviewers for the time they have taken to provide their insightful and detailed comments. In response, we have substantially revised the manuscript, focusing on Sox2 regulatory effects, and supplying previously missing details. A specific point by point is included below, as well as excerpts from some of the revised manuscript figures, and additional supporting information. An important point was raised regarding the Sox2-*Evf2* co-localization domain in vivo. Experiments to precisely define the functional amino acids are complicated by overlap with nuclear localization sig- nals (NLS). We believe that the ability of ectopic Sox2 to associate with endogenous *Evf2* RNA and Dlx5/6 DNA, and demonstration of the sufficient domains are important for beginning to understand RNP formation in vivo, as well as providing future directions. However, if reviewers believe that the ectopic mutational analysis is incomplete in its present form, we can remove the transfection experi- ments, and publish these results in a more complete form in a manuscript focused on localization. While we agree that more mutations are potentially valuable, we also believe that removal will not de- tract from the significance or innovation of this work.

With respect to the rest of the critiques, we hope that these revisions have addressed reviewers’ concerns, and significantly improved the manuscript.

Reviewer #1:
Major points
1. Model clarification. I feel that in an attempt to capture all the complexity of this process the figures and the text end up being very difficult to follow thus detracting from the conclusions. From reading this version I am not sure what is the mechanism actually proposed.

1.1. The model in Fig. 4F did not help me very much. Perhaps the authors could expand the model figure and have sep- arate panels showing different steps of the proposed process?

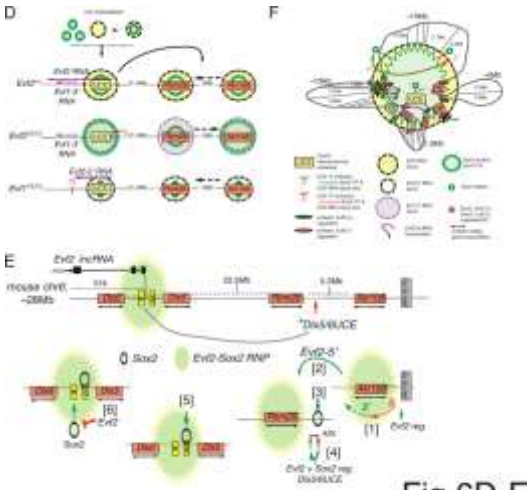


Fig 6D-F

-We now include a three part model (Fig 6D, E, F). In 6D, we show a linear representation of Evf2 RNA cloud -Sox2PP re- lationships with respect to the Dlx5/6UCE and repressed tar- get genes Rbm28 and Akr1b8. These relationships are altered, in two Evf2 mutants (Evf2 loss -Evf2$^{TS/TS}$ and trun- cation, Evf1$^{TS/TS}$) compared to wildtype. In 6E, we have sep- arated the multistep model into 6 steps, as requested by the reviewer. Each step is discussed in greater detail in the dis- cussion. In 6F, a potential arrangement of the Evf2 RNP- Sox2PP with respect to chromatin loops is shown.

1.2. One problem is that the different components of the Evf2-RNP seem to respond very differently to the Evf2 mutants. Perhaps the authors could consider re-focusing this study on SOX2, for which they have promising in vitro results and potential mutants to explore, and de-emphasize or keep for a future study the other proteins.

-We agree with the reviewer and have focused the revision on Sox-Dlx-Smarca4 data, and removed the Nono and Smarcc2 RNP data.

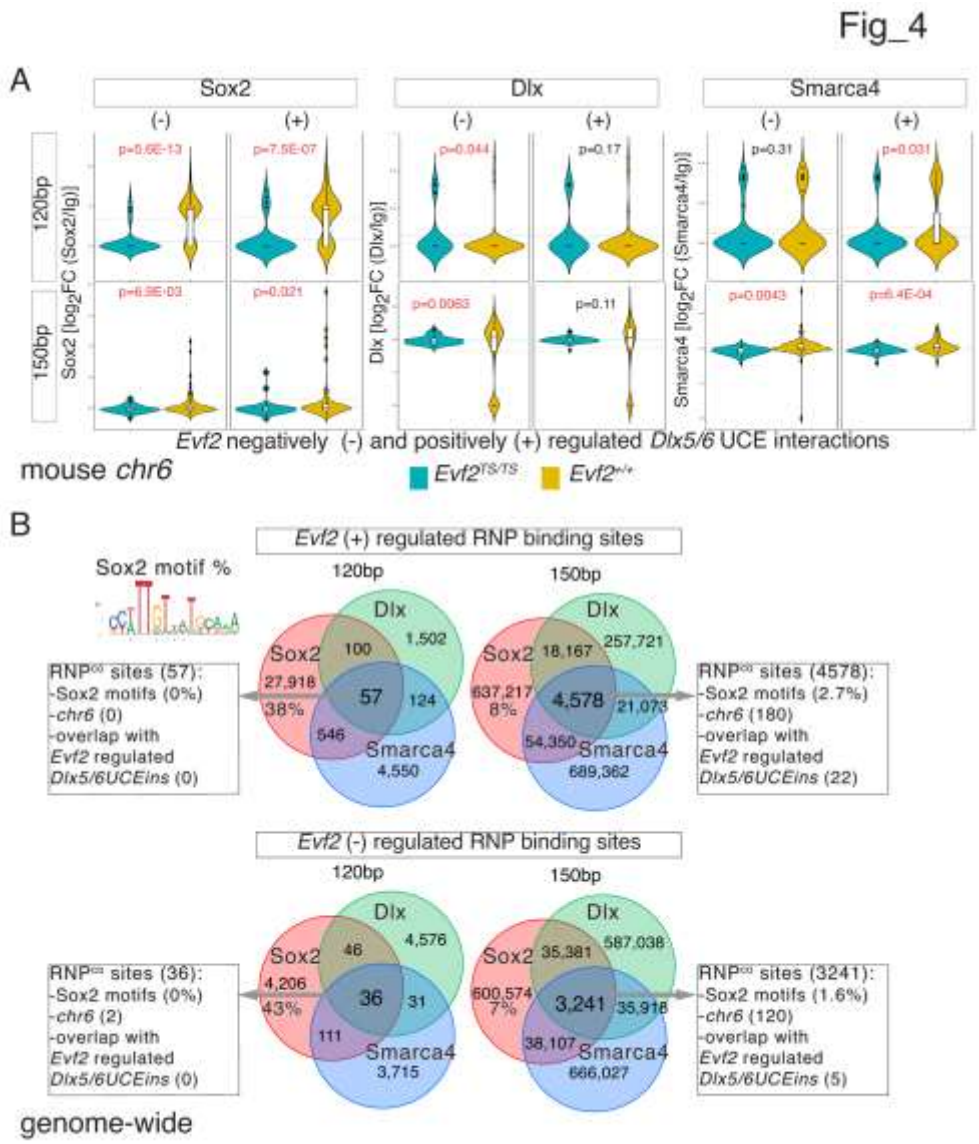


***Figure 4B**. Venn diagrams of the number of Evf2 pos- itively (+) and negatively (-) regulated Sox2, Dlx, and Smarca4 peaks identified by CUT&RUN in E13.5GEs. The percentage of Sox2 peaks that contain Sox2-DNA motifs is indicated (Sox2 DNA motif defined by JASPAR, using an FDR <0.01 cut-off). Sox2 peaks from 120bp fragment sequencing are enriched for Sox2 DNA motifs compared to 150bp (Evf2 (+) regulated: 120bp-38% vs 150bp-8%, and Evf2 (-) regulated: 120bp-43% vs 150bp-7%). Analysis of 150bp fragments shows that the percentage of co-regulated sites (arrow, bound by all three RNPs, Sox2, Dlx and Smarca4, RNPco) with Sox2 DNA motifs decreases compared to Sox2 singly bound sites (Evf2 (+) regulated: RNPco -2.7% vs Sox2-8%, Evf2 (-) regulated: RNPco -1.6% vs.Sox2-7%). RNPco sites identified in the 120bp analysis do not contain known Sox2 DNA motifs (0%) compared to singly bound Sox2 Evf2 (+) regulated (38%) or Evf2 (-) regulated (43%). RNPco sites are associated with 27 Dlx5/6UCEins, 22 at Evf2 (+) regulated RNPco sites, and 5 at Evf2 (-) RNPco regulated sites.*

We now include an analysis of *Evf2* co-regulated Sox2- Dlx-Smarca4 binding sites (Venn diagrams in Fig 4B), and Sox2 DNA binding motifs. These data support the work of Meers at al. 2019b showing that 120bp CUT&RUN fragments are enriched for directly bound TFs. The percentage of Sox2 motifs in Sox2 peaks at Evf2 positively regulated sites is 38% for 120bp and 8% for 150bp, and at Evf2 negatively regulated sites is 43% for 120bp and 7% for 150bp.

Importantly, *Evf2* co-regulated RNP sites are de- pleted for Sox2 motifs, supporting that Dlx and

Smarca4 influence Sox2 DNA binding genome-wide. These data highlight the importance of defining Sox2- nucleic acid interactions in the context of the RNP, as the reviewer recommends, but also emphasizes the re- sults of in vivo analysis.

2. RNA binding by SOX2. It is unclear to me how SOX2 binds to Evf2 and where the specificity comes from.
2.1. *In Fig. 1E the authors state that the binding is "promis-cuous". Is that different from sequence-independent? If not, I would maybe use this expression, which is more commonly employed.*
-Done
We use the term "promiscuous binding" first established by Davidovich et al (2013) for HOTAIR and PRC2. We showed similar promiscuous properties for Evf2-Smarca4 interactions (Cajigas et a. 2015, length- dependent binding rather than sequence specific binding). Most recently, Holmes et al (2020) show that Sox2 binds RNA with high affinity/low specificity. Since the investigation by Holmes is extensive with respect to Sox2-RNA binding, we now adopt their terminology (high affinity/low specificity).

2.2. If indeed SOX2 is binding to RNA in sequence independent manner what drives its specific colo- calization with Evf2? Some of the protein-interactions? If so can this be demonstrated?

-We now provide an analysis of *Evf2* co-regulated sites (Sox2-Dlx-Smarca4 co-bound sites, RNP^{co}), where Sox2 DNA motifs are depleted compared to singly bound Sox2 sites.These data support the idea that Dlx and Smarca4 influence Sox2 -DNA motif recognition in vivo.

For the ectopic analysis - we believe that since Sox2 co-localization domains overlap with nuclear lo- calization (discussed in detail below), it will be necessary to perform single amino acid substitution ex- periments across both NLS's to distinguish between Evf2 RNA cloud/Dlx5/6UCE colocalization and nuclear localization. If possible, this would need to be followed by determining whether any colocal- ization amino acids disrupt interactions with RNPs. We agree that these studies are potentially valu- able regarding the mechanism. However, interpretation will be complicated by the overlap with the NLS's, and may not feasible.

2.3. How promiscuous is the binding? Would DNA compete with Evf2 just as efficiently as the pGEM RNA?

-Holmes et al 2020 show that RNA and DNA compete for binding to the Sox2 HMG in vitro; they state that
"DNA and RNA binding to Sox2-HMG is mutually exclusive". We discuss their work in this manu- script, as it is important with respect to the models proposed, specifically the direct inhibition of Sox2- enhancer activation and Sox2 recruitment.

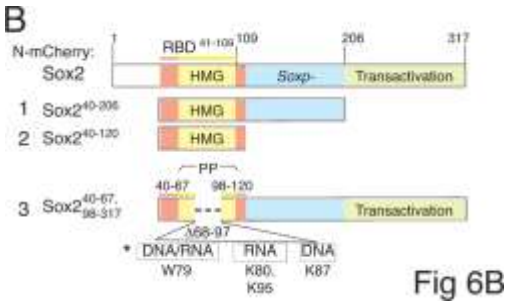


Fig 6B

3. RNA binding mutant of SOX2. I think the authors could further lev- erage the knowledge that the HMG domain is required for the RNA interaction and use these mutants to sharpen the focus on SOX2.
3.1. In Fig. 3K he authors show that residues 40-206 are sufficient for colocalization of SOX2 and Evf2 but they should also be able to show they are necessary otherwise it is difficult to interpret these re- sults. In is particularly surprising that the Δ68-97 deletion increases rather than decreases colocalization. The authors conclude that these internal aminoacids inhibits the protein-RNA interaction but this is hard to envision. Why not test the mutant that cannot bind

Evf2 in vitro (Fig. 1E)?

- We agree with the reviewer's assessment that increases in colocalization when RNA/DNA binding is removed are difficult to explain if the HMG domain is responsible for colocalization with Evf2 RNA clouds.. We also agree that a necessary domain would help in defining the mechanism. We have re- vised the interpretation of the mutant analysis, to include the alanine substitution data reported in Holmes et al. 2020 that identifies specific amino acids necessary for RNA and/or DNA binding. Since Holmes has defined a region that is removed in our mutant 3 as RNA/DNA binding, this supports that the NLS regions rather than DNA/RNA binding are involved in co-localization. We discuss the impor- tance of future experiments to determine whether amino acids involved in nuclear localization, Sox2 PP formation and Dlx5/6 localization can be distinguished, and how these amino acids contribute to Evf2-RNP assembly and protein-protein interactions with the RNP.

We now state the following:

We next used FISH analysis of N-terminally tagged, mcherry-Sox2 (mch-Sox2) transfected into E13.5GE's to determine whether ectopically expressed Sox2 associates with endogenous Evf2 RNA clouds and/or Dlx5/6. Transfected mch-Sox2 forms PPs that colocalize with endogenous Evf2 RNA clouds and/or Dlx5/6 (Fig 6A), through a minimal region spanning the HMG nucleic acid binding do- mains and adjacent NLS's (orange) (Sox240-120, Fig 6A-C). In Sox2 mutant 3 (Sox240-67, 98-317), critical RNA/DNA binding amino acids within the HMG domain are deleted: ▫66-97 deletes W79, K80, K87, and K95, defined as RNA and/or DNA binding (Holmes et al., 2020). However, Sox2 mut3 colo- calization with endogenous Evf2 RNA clouds and/or Dlx5/6 increases, supporting that nucleic acid binding is dispensable, while amino acids in the NLS's are critical for PP formation and RNA/DNA lo- calization. Future experiments that distinguish between Sox2 nuclear localization and RNA cloud/Dlx5/6 localization, and determining the role of individual Sox2-RNP interactions in localization will be important in understanding Evf2 RNP assembly and targeting in vivo.

3.2. *Perhaps an effective RNA-binding mutant would help interpret some of the discrepancies in the Evf2 and Sox2 dependent effects, for example as shown in Fig. 2.*

-Given that deletion of DNA/RNA binding amino acids (Mutant #3 in Fig 6B) does not eliminate colo- calization, we believe that future experiments aimed at dissecting NLS from RNP colocalization and RNP interactions will be required.

-While single amino acid deletions across aa40-67 and aa98-120 will be required to determine aa necessary for colocalization, this may be very difficult (or may even not be possible) to define, given the overlap of the sufficient region with NLS's. However, the transfection data is a small part of this manuscript. While we prefer to include the transfection data, we also believe that these experiments are not a critical part of the proposed mechanism. If the reviewer thinks that the colocalization/muta- tional analysis in its present form still requires single amino acid substitution analysis, then we can remove the transfection data from this manuscript. Transfection analysis is performed in primary ganglionic eminences and nuclear FISH colocalization, and for the number of mutations required, would likely take more than a year to complete, with an uncertain outcome.

Minor points
- Fig. 1H: how were the genes classified in green and red here? If from previous work please add a citation.
-added (Cajigas et al. 2018)

- Fig. 1J: I would not refer to CUT&RUN as a "native ChIPseq method". I think most people by now are famil- iar with CUTRUN without explanations. Just cite the first Henikoff paper at first mention.
- done

- Flg. 1J: I think the authors is reading too much into the 120 bp vs 150 bp profiles. If they wish to establish di- rect vs. indirect interactions additional data are required. In alternative I would tone down these conclusions which anyway do not add much to the main points of the study.

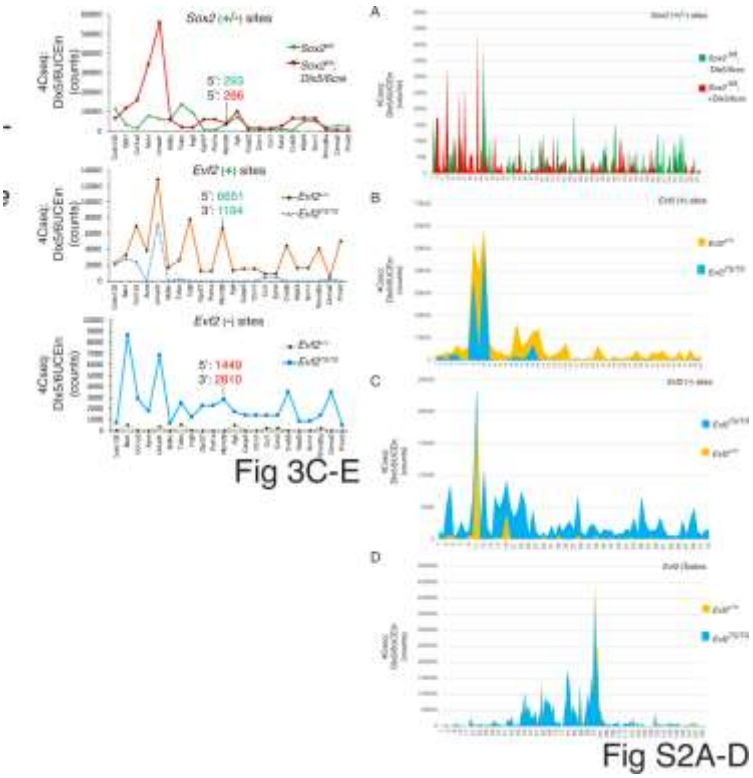
We now add Sox2 motif analysis at Evf2 regulated Sox2 singly bound and Sox/Dlx/Smarca4 (RNP^{CO}) co-regulated sites (described in response to reviewer's comment to refocus the paper on Sox2 see 1.2 above, Venn diagrams in Fig 4B), and Sox2 DNAbinding motifs comparing CUT&RUN 120bp and 150bp fragments. These data support the work of Meers at al. 2019b showing that 120bp fragments are enriched for Sox2 motifs.

- I am not familiar with the expression "protein pools". What are the authors referring to? -We are using pro- tein pools to describe proteins that associate with RNA clouds distinguishing these from "RNA clouds" themselves. RNA clouds were first termed by Redrup et al.2009 in reference to spherical formations by the lncRNA Kcnq1ot1 (Development 136, 525-530). Here, we name spherical for- mations of Evf2-RNP proteins within the nucleus as protein pools.

Are these accumulation of the protein by homomultimeric interactions or just a reflection of multivalent binding of the protein to a large RNA? We do not provide evidence for homomultimeric formation. However, the Evf2-RNP contains 87 proteins, many of which associate with the RNA cloud by FISH analysis, a subset associating with Sox2, with direct binding between Sox2-Dlx1, Sox2-Smarca4, Smarca4- Dlx1. Taken together with genome-wide co-binding/co-regulation of Sox2-Dlx-Smarca4, and low sequence specific binding of Sox2 and Smarca4 to RNA, we believe that the Sox2 protein pools re- flect multivalent binding to the RNA and other Evf2-RNPs, as shown in the Evf2-RNP Venn diagram (Fig1A).

Is there a reference for this concept of "protein pools"? -We are not aware of a reference using this term, but this could be similar to subnuclear compartments such as paraspeckles, that we refer to in the discussion about possible Evf2 RNPs that may regulate RNP assembly, such as Nono.

Reviewer #2:
...*They present a large amount of data and reach the conclusion. However, some of the data pre- sented in this study are preliminary and not convinc- ing for me to reach their conclusion. For instance, they present charts showing that SOX2 foci become smaller upon Evf1 KO (Figure 3I), but do not show any cellular images that are the source of these charts.*



The similar tendency is seen in the 4Cseq data. They show the processed results (e.g. Figure 2A)

but do not show a total view of 4C results. These data are not convincing for me. I therefore suggest that the authors should present the data more carefully to concrete their conclusion.
Major points
1) Figure 2A. In this figure, the authors present 4Cseq data as a simplified andprocessed form. However, it is not clear how 4Cseq data is processed and presented here. I think it is important to present this data in a more convincing way. For instance, 4C interactomes can also be shown like Figure S2B-D for both Evf2(+) and Evf2(-).

-We now include 4C interactome data in graph forms for specific co-regulated sites in Fig 3C-E. In FigS2 A-D, we now include the graphs with counts across chr6. The raw counts are available in an excel file in supplementary data. 4C (Dlx5/6UCE as bait) sequencing data for *Evf2*$^{+/+}$ and *Evf2*$^{TS/TS}$ is available on NCBI, associated with previous work (Cajigas et al. 2018). *Sox2*$^{fl/fl}$*; Dlx5/6cre+/cre-* sequencing data will also be made available on NCBI, as well as CUT&RUN data.

2) Figure 3 In figure 3, most of fluorescent signals look over-saturated. Especially, gene loci visualized with BAC probes such as Akr1b8 and Rbm28 should show smaller pinpoint signals, not as big as the ones shown in Figure 3A and G. I think this makes difficult to assess a colocalization between these foci. Therefore, these data should be correctly displayed. Alternatively, I recommend the authors to use a super-resolution micros- copy instead of a confocal microscopy to assess the colocalization between these molecules.

- The pictures appeared oversaturated likely because we showed an image that was too zoomed in - in order to match the adjacent 3-D reconstruction. In the new panels, we show a less zoomed image to match the first panel, containing a wider view of the nucleus (new panels in Fig 5A, G). We also include ad- ditional images in Supplement Figs S5, S6. Image files will be accessible on Mendeley.

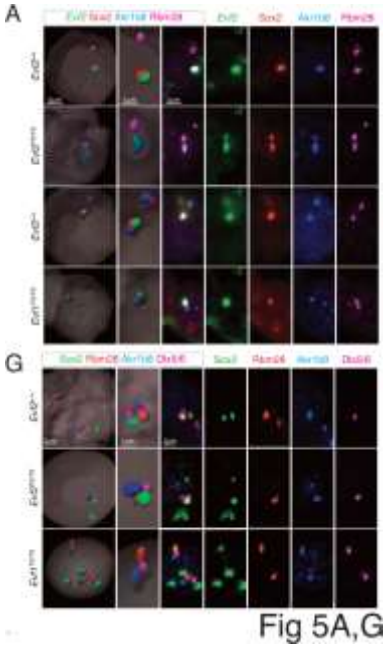


Fig 5A,G

3) The title should be modified. The current "Sox2-RNA" could imply that Sox2 binds many RNA species. I think this may be misleading. "Sox2-Evf2 RNA" should be used instead of "Sox2-RNA".
- done

Minor points
1) Page 5, line 21, "Evf2 RNA binding to Sox is promiscuous" I think thismay be misleading because they do not actually show the promiscuous RNAbind- ing. Does Sox2 bind any part of Evf2 RNA? Otherwise, the author should re- move these words.
-We now state that Sox2 binds RNA with low sequence specificity (based on pGEM RNA competition with Evf2 binding, and Holmes et. al. 2020 who reported SOX2-RNA high affinity low specificity binding), also raised by reviewer 1.

2) Page 7, line 6, "Evf2TS/TS (expressing only Evf2-3')"
These words are confusing. In the Figure 3C, it is shown that Evf2TS/TS expresses Evf1-3' RNA. The authors have to clarify these discrepancies.
-*Evf1-3'* RNA and *Evf2-3'* are transcribed from exon4, and overlapping. This is now clarified, and made consistent throughout.

3) Figure 1H
The Sox2 expression level in Evf2TS/+;Sox2fl/fl ;Dlx5/6cre and Evf2+/+;Sox2fl/fl ;Dlx5/6cre is lower than Evf2TS/+;Sox2fl/fl , but still certain amount of Sox2 is expressed compared to Evf2+/+;Sox2fl/fl. Does this mean that the ablation of Sox2 is not efficient with Dlx5/6cre? -The cre recombinase is under the control of Dlx5/6 regulatory sequences. Dlx5/6 is not expressed in 100% of E13.5GE cells. Therefore, we be- live that the likely source for the remaining levels of Sox2 expression are cells where Dlx5/6 is not ex- pressed.

4) Figure 2C-D legend, "I indicates Dlx5/6UCEins that are independentEvf2." More explanation in the legend is necessary to understand thissentence.
The authors should clearly describe how this was calculated. - Done, also reported in Cajigas et al. 2018, and is now referenced.

5) Figure 3E
It is not described clearly how "colocalization" is defined. It should be explained in the M&M section.
-now included

(6) Figure 3E, F, H, I and J
Representative FISH images used for these charts should be shown as supplementary data.
- Colocalized images for Evf2-5' and 3' are shown in Fig 3A, and colocalized images for Sox2 PPs in Fig 3G. Now also included (new images in Fig S5, Evf2-5', 3' RNA clouds, Fig S6, Sox2 PPs).

(7) Figure 3
It is not clear that these colocalization between Evf2 RNA, Sox2 PPs, and Evf2-target loci (Akr1b8 and Rbm28) occurs monoallelically or biallelically. This point should be described in the main text.
- We found examples of both monoallelic and biallelic colocalization. This is now stated, as well as in- dicated in the examples in Fig S5.

(8) Figures 3 and 4
All of colocalization analyses were done in 2D? It should be clearly described in the main text or in the M&M section.
-We now include a description in the M&M that nuclei were visualized using a Zeiss Laser Scanning Microscope 880 and the Zen 2.1 software, (2-D colocalization for Sox2 PP transfections, and 3-D co- localization for volume analysis). Co-localization of transfected mcherry-fused SoxPPs with Evf2 RNA clouds and Dlx5/6UCE was determined by Zen2.1 software and manual inspection of Z-stacks through each nucleus. For endogenous analysis, z-stacks generated from Zen2.1 software imaging were imported into Imaris software and used for 3-D reconstruction, co-localization analysis and size measurements. Colocalization was defined as any regional overlap in one of 4 channels. The numbers of overlapping clouds, and or Sox2PPs with DNA target genes were determined by IMARIS software following 3D reconstruction, and in conjunction with size determination. Threshold settings were determined in pilot experiments performed in Evf2+/+ nuclei, and remained constant throughout the analysis between genotypes.

(9) Figure S3

It is difficult to understand this figure because the letters are too small. There is the same tendency for all of figures. Bigger size letters should be used. The only important part of this figure should be presented.
-We have changed the sizes for many of the figures, and removed extra information as requested. In addition, we have now have 6 figures instead of 4, increasing lettering sizes.

Second decision letter

MS ID#: DEVELOP/2020/197202

MS TITLE: Sox2-Evf2 RNA mechanisms of chromosome topological control in developing forebrain

AUTHORS: Ivelisse Cajigas, Abhijit Chakraborty, Madison Lynam, Kelsey Swyter, Monique Bastidas, Linden Collens, Hao Luo, Ferhat Ay, and jhumku D kohtz

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 are basically appreciated your efforts to add more data. However, the reviewer2 still found the story a bit too complicated and some data were not sufficiently explained in the text. I basically support opinions of the reviewer 2 and, thus, like you to modify the text according to her/his suggestions. If you are able to revise the manuscript along the lines suggested, 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

The mechanism of action of lncRNAs remain largely mysterious. This study represents a very detailed analysis of the spatial and functional interplay between a developmentally important lncRNA, Evf2, and a key transcription factor, SOX2, in a developmental context.

Comments for the author

I am overall pleased with the changes made to this manuscript, in particular the efforts made by the authors to clarify their model of this very complex process.
I have no objections to its publication.

Reviewer 2

Advance summary and potential significance to field

In this study, Cajigas et al. propose a new working model for Evf2 ncRNA-mediated gene regulation. They found SOX2 protein is involved in this Evf2-mediated gene regulation through direct interaction with Evf2 ncRNA. I think their findings give a new insight to the field of non-coding RNA and genome topology. However, I think the manuscript needs to be more simplified because the whole story is still quite complicated.

Comments for the author

In the revised manuscript of Cajigas et al., the authors well responded to the points raised by the reviewers.
However, I still think that the overall story is complicated and contains several unclear points. Perhaps, the authors should ask someone outside of the field to read the manuscript carefully, especially about the configuration of the text and figures. In addition, the revised manuscript still needs to add at least several modifications listed below. For example, mouse mutants and CUT&RUN 120/150 bp difference need to be clearly described in the text. I believe that these modifications will improve the quality of this study and facilitate a comprehensive understanding by potential readers.

Major points
(1) Figure 2C
It is surprising for me that the authors do not explain what <120 bp and >150 bp peaks mean in the results section (they did explain it in the original version though). In the revised manuscript, they actually explain it in the M&M and discussion. However, this should be clearly described in the results section with proper references. Otherwise, I think readers cannot understand what these data mean. I think that most of readers are familiar with CUT&RUN by now but would not know much about the <120/>150 bp differences.

(2) Figure 5B-D
Mutant mice (Evf2TS/TS and Evf1TS/TS) and their output transcripts (Evf2-5', Evf2-3' and Evf1-3') are still complicated. Although Figures 5B-D and 6D explain them, I think this should be clearly explained earlier because these mutant mice first appeared in Supplementary Figure S1. Therefore, I think details of these genotypes and output transcripts need to be explained in Figure 1 or Supplementary Figure S1.

(3) Figure 3
In this figure, they demonstrate Sox2- and Evf2 RNA-dependent interactions with Dlx5/6UCEin. They just pick up certain genomic loci for the analysis. Are these changes of interaction also seen in other genomic regions? I think it is also important to show the whole chr 6 interaction map in WT and mutants. Does Supplementary Figure S2 correspond to it? What the X axis in this figure means? The locations of Evf2 and Dlx5/6UCE need to be indicated.

Minor points
(1) Figure 3A Presence of two * marks, one for Dlx5/6UCE and the other for Rbm28-5', in this panel is confusing. Actually, only the latter is explained in the figure legend. I think they should be presented with different markings and explained in the figure legend.

(2) Figure 3B-D The location of the 4Cseq bait, Dlx5/6UCEin, is not clearly indicated. Is it the broken line at Rbm28 locus? If so, it should be clearly indicated and described in the figure and figure legend.

(3) Figure 4
(+) and (-) need to be correctly placed.

Second revision

Author response to reviewers' comments

*In the revised manuscript of Cajigas et al., the authors well responded to the points raised by the reviewers. However, I still think that the overall story is complicated and contains several unclear points. *Perhaps, the authors should ask someone outside of the field to read the manuscript carefully, especially about the configuration of the text and figures. In addition, the revised manuscript still needs to add at least several modifications listed below. For example, mouse*

mutants and CUT&RUN 120/150 bp difference need to be clearly described in the text. I believe that these modifications will improve the quality of this study and facilitate a comprehensive understanding by potential readers.
Major points
(1)Figure 2C
It is surprising for me that the authors do not explain what <120 bp and >150 bp peaks mean in the results section (they did explain it in the original version though). In the revised manuscript, they actually explain it in the M&M and discussion. However, this should be clearly described in the results section with proper references. Otherwise, I think readers cannot understand what these data mean. I think that most of readers are familiar with CUT&RUN by now, but would not know much about the <120>150 bp differences.

On page 10 in the results section, we state the following: *This data is consistent with CUT&RUN analysis of Sox2 binding profiles where Sox2 120bp fragments are motif enriched, while 150bp fragments are motif depleted, and reflect nucleosomal binding (Meers et al., 2019b).*

4. Now also added at the bottom of page 6:
-
5. In the CUT&RUN method, sequencing of <120bp and >150bp fragments distinguish between proteins directly bound to DNA (less than 120bp), and indirect binding through protein-protein interactions (more than 150bp) (Meers et al., 2019a; Meers et al., 2019b; Meers et al., 2019c). Analysis of >150bp CUT&RUN peaks has the potential to detect proteins associated with the large *Evf2-RNP*, not previously possible using crosslinked ChIPseq (X-ChIP) methods.
-
- Figure 5B-D
Mutant mice (Evf2TS/TS and Evf1TS/TS) and their output transcripts (Evf2-5', Evf2-3' and Evf1- 3') are still complicated. Although Figures 5B-D and 6D explain them, I think this should be clearly explained earlier because these mutant mice first appeared in Supplementary Figure S1. Therefore, I think details of these genotypes and output transcripts need to be explained in Figure 1 or Supplementary Figure S1.

- Schematic added in S1A showing the relationships between Evf2-5' and Evf1-3' and TS insertions in Evf2TS and Evf1TS..
-
- Figure 3
In this figure, they demonstrate Sox2- and Evf2 RNA-dependent interactions with Dlx5/6UCEin. They just pick up certain genomic loci for the analysis. Are these changes of interaction also seen in other genomic regions?
-
Yes – We have made the following changes to clarify Fig 3:

4) 1. Fig 3B-D: added chr6 genes as a label on X-axis, 2. Fig 3B-D: labelled where Evf2/Dlx5/6UCE is located on the X-axis with a dotted purple line(between Col1a2 and Asns genes)
5) In the Figure legend, we include the following change:
6) Page 21: 4Cseq *Dlx5/6UCEin* counts (Y-axis) at genes on mouse chromosome 6 (X-axis)where *Evf2* and Sox2 co-regulate *Dlx5/6UCEin*. Dotted purple lines indicate the location of *Evf2* and Dlx5/6UCE bait.
changes on other regions of chr6 are indicated in Figs 3B-E, and Fig S2:
-In Fig 3B, C, D, counts from 4Cseq using Dlx5/6UCE as bait are indicated on the Y-axis, at chr 6 genes on the X-axis where co-regulation by Sox2 and Evf2 are detected.
-In 3B, counts at chr6 genes where Sox2 positively and/or negatively regulates Dlx5/6UCEin are indicated.
-In 3C, counts at chr6 genes where Evf2 positively regulates Dlx5/6UCEin are indicated .
-In 3D, counts at chr6 genes where Evf2 negatively regulates Dlx5/6UCEin are indicated. Only genes where co-regulation occurs are listed on the X-axis. In Figs 3E and 3F, the positions (but not counts) of co-regulated or oppositely regulated genes are indicated across chr6.
7) I think it
is also important to show the whole chr 6 interaction map in WT and mutants. Does Supplementary Figure S2 correspond to it? What the X axis in this figure means? The locations of Evf2 and Dlx5/6UCE need to be indicated.

-
Yes- in FigS2, we show the Dlx5/6UCE counts across chr6 for WTs and mutants. We include the following on page 8:
The complete Sox2 and *Evf2* regulated 4Cseq-*Dlx5/6UCEin* counts across *chr6* are shown in Fig S2, where the X-axis numbers correspond to the location sites on chr6 listed in the Excel sheet in the associated data). Complete CUT&RUN profiles for Rbm28-5 are shown in Fig S3A.
We also now clarify this in the legend, and indicate where Evf2/Dlx5.6UCE bait is located along the X-axis.
8) - We added the following to the S2 figure legend:
9) -Y-axis: count numbers from 4Cseq using Dlx5/6UCE as bait
10) -X-axis: numbers correspond to chr6 location sites listed for each genotype (Excel).
11) -Dlx5/6UCE bait location indicated by purple dotted lines
-
Minor points
(1)Figure 3A
*Presence of two * marks, one for Dlx5/6UCE and the other for Rbm28-5', in this panel is confusing. Actually, only the latter is explained in the figure legend. I think they should be presented with different markings and explained in the figure legend.*—**done (removed star from Dlx5/6UCE, w only one remaining star at Rbm28-5')**

(2) Figure 3B-D
The location of the 4Cseq bait, Dlx5/6UCEin, is not clearly indicated. Is it the broken line at Rbm28 locus? If so, it should be clearly indicated and described in the figure and figure legend.
-

12) **-we now include where the Dlx5/6UCE is indicated by purple dotted lines**

(3) Figure 4
(+) and (-) need to be correctly placed.- **corrected**

**Perhaps, the authors should ask someone outside of the field to read the manuscript carefully, especially about the configuration of the text and figures*

Comments from colleagues outside of this field are now addressed, including changes to the text or figure legends (in yellow below and in the revised manuscript):

5) Page 1 text modification: Removal of a select few UCEs
6) Page 1 text modification (split one sentence into two): Transcription of UCE sequences and enhancer-regulating activity of UCE transcripts (Calin et al., 2007; Feng et al., 2006), was followed by the identification of genome-wide scale enhancer transcripts with enhancer-like activities (Orom et al., 2010; Orom and Shiekhattar, 2011). Together, these data support mechanistic and functional diversity of RNA regulatory roles (Rinn and Chang, 2020)
7) Include description of E13.5GE: now included on page 3:
specifically at sites of GABAergic interneuron birth in E13.5 mouse ganglionic eminences (E13.5 GE)
-page 6: Sox2 contributes to *Evf2*-dependent gene activation or repression, we analyzed gene expression in E13.5GE (mouse embryonic GABAergic interneuron progenitors)
8) Sentence change, page 3:
Mechanisms of *Evf2* gene activation and repression are distinguished by different regional requirements of the RNA

9) Include background information on RNP87. Now included on page 4:

RNP^{87} was previously identified by comparing proteomic profiles from anti-DLX, affinity purified *Evf2*$^{+/+}$ and *Evf2*$^{TS/TS}$ E13.5GE complexes, showing that the number of Dlx associated proteins is 87 in the presence of *Evf2*, and 15 in the absence of *Evf2*.

10) Page 4 text modification:
-*Evf2* assembles a ribonucleoprotein complex (RNP^{87}) containing at least 87 functionally diverse proteins
- *Evf2*-dependent gene regulation across a 27MB region of mouse chromosome 6 (*chr6*) is

characterized by recruitment of individual RNPs and regulation of histone modifications

11) Clarify definition of Sox2 PPs on page 5: Sox2 colocalizes with *Evf2* RNA clouds in subnuclear domains that we have termed protein pools (PPs), detectable both in the presence and absence of Evf2 RNA.
12) Include better description of Evf2TS and Evf1TS mouse models. Addressed in (2) above.
-Figure 2D -schematic, legend to Fig 2D:
D. On mouse chr6, the sites of Evf1TS and Evf2TS transcription stop insertions (TS, blue) are shown with respect to repressed genes Dlx5 and Dlx6 (red boxes), and Evf1, Evf2, and Evf2-5' transcripts (exons 1-4 in black). Dlx5/6UCE (yellow/star).

-page 4: A description of site of transcription stop (TS) insertion that generate mice lacking Evf2 and Evf1 (Evf2TS/TS, Evf1TS/TS) is shown in Fig S1A.

-page 7: In *Evf1*$^{TS/TS}$, the transcription stop sequence is inserted into exon 3, preventing expression of *Evf1* (and also the *Evf2-3'* region), but producing a truncated *Evf2-5'* transcript (Fig S1A).
13) page 7: Loss of Dlx and Smarca4 binding to *Dlx5/6UCE* in *Evf2*$^{TS/TS}$ (expressing only *Evf1*-3', overlapping with the *Evf2-3'* region) is not rescued in *Evf1*$^{TS/TS}$ (expressing only *Evf2*-5'); transcripts resulting from *Evf1*^{TS} and *Evf2*^{TS} insertion are schematized in Fig 2D and S1A.

13) Page 5 sentence changes:
- At the genome-wide level, *Evf2* co-recruitment of Sox2 with RNPs Smarca4 and Dlx affects Sox2-DNA recognition. We propose that the *Evf2* lncRNA functions as a Sox2 subnuclear domain organizer, controlling *Dlx5/6UCE* targeting and activity by distributing Sox2 and associated RNPs Smarca4/Dlx to key DNA regulatory sites on *chr6*, with genome-wide effects.

-Previous work showed that Smarca4 bridges the *Evf2* RNA with the protein Dlx1, and other RNA binding proteins within the RNP
-moved this sentence to beginning of Results section- to begin with the question being asked: *Evf2* activates and represses genes across a 27Mb region on mouse chromosome 6, raising questions regarding the mechanistic basis for *Evf2*-dependent differential gene regulation (Cajigas et al., 2018).
14) What is SoxOT? Now included in FigS1 legend:
Sox2ot (a ~118kb overlapping transcript)
15) What is MACS2? Now included in S1 legend:
MACS2 (peak calling algorithm used for ChIPseq, Zhang et al. 2008)
16) Describe CUT&RUN. Answered in (1) above, also now include:
we used the native ChIPseq method CUT&RUN,
17) Page 7:
with the *Evf2-3'* region
18) Modify Venn diagrams in Fig 4B to explain results. Now we include numbers and categories in four boxed areas that list attributes of RNPco sites with respect to Sox2 motifs, chr6 location and Evf2-regulated Dlx5/6UCEins. The text contains the following change (page 10):
at 27 (22+5) out of 123 *Evf2* regulated

19) Page 8 clarifications:
-Previous work showed that *Evf2* regulated *Dlx5/6UCE* targeting…
- We performed ChIPseq using crosslinked E13.5GE chromatin to analyze *Evf2*- regulated Sox2 and Smc3 binding across *chr6* (Fig S1E). Analysis of overlapping regulatory sites identifies antagonistic sites of *Evf2* positively (+) regulated Sox2 binding and *Evf2* negatively (-) regulated Smc3 binding (Fig S1E).
20) Figure 1 legend, clarify difference between RNP87 and 79: now include:
Of the 87 proteins in the Evf2-Dlx RNP complex,
21) Define MACS2 in Figure 2 legend:
MACS2 a peak calling method for ChIPseq (Zhang et al., 2008),

Third decision letter

MS ID#: DEVELOP/2020/197202

MS TITLE: Sox2-Evf2 RNA mechanisms of chromosome topological control in developing forebrain

AUTHORS: Ivelisse Cajigas, Abhijit Chakraborty, Madison Lynam, Kelsey Swyter, Monique Bastidas, Linden Collens, Hao Luo, Ferhat Ay, and jhumku D kohtz
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.