EMBO reports

# Sequentially Acting SOX Proteins Orchestrate Astrocyte and Oligodendrocyte Specific Gene Expression

Susanne Klum, Cécile Zaouter, Zhanna Alekseenko, Åsa K Björklund, Daniel W Hagey, Johan Ericson, Jonas Muhr and Maria Bergsland

**Review timeline:**

| | |
|---|---|
| Submission to The EMBO Journal | 25 September 2017 |
| Referee Reports (EJ) | 5 December 2017 |
| Revision received (EJ) | 30 May 2018 |
| Editorial Decision (EJ) | 22 June 2018 |
| Submission to EMBO Reports: | 25 June 2018 |
| Editorial Decision: | 11 July 2018 |
| Revision received: | 24 July 2018 |
| Accepted: | 7 August 2018 |

Editor: Esther Schnapp

**Transaction Report: This manuscript was transferred to *EMBO reports* following peer-review at *The EMBO Journal***



The EMBO Journal – Referee Reports 5 December 2017

REFEREE REPORTS

**Referee #1:**

The study from Jonas Muhr and colleagues adds significant information about how Sox proteins determine astrocyte vs. oligodendrocyte lineages. While there is nothing seriously wrong with the findings presented, the nature of the data, however, renders this manuscript more appropriate as a resource paper than as a research article.

**Referee #2:**

In this manuscript, Klum et al. reported how SOX proteins sequentially act on neurogenesis and gliogenesis. Using RNA-Seq and ChIP-Seq, the author demonstrate that Sox9 promote astrocyte development together with NFIA, and activate oligodendrocyte maturing together with Sox10. The requirement of Sox9 in astrocytogenesis and Sox10 in oligodendrocytogenesis has been well characterized. Limited novel information is reported in the manuscript. The involvement of Sox3 in gliogenesis is novel but the authors only provided in vitro over expression data. Therefore, the in vivo requirement of Sox3 in gliogenesis has not been tested. Most of the manuscript contains descriptive scRNaseq and ChipSeq data.
Main comments:
1. The orange cell group in Figure 1B was omitted from subsequent analysis. The genes differentially expressed by this group are microglia markers. It's better to include this group in the analysis and label them as microglia.
2. In Fig2 A, it is hard to tell whether there is Sox9 expression in the day 4 image. Showing single

channel images would be helpful.
3. The authors draw a model (Figure 6 middle) in which Sox3 binds both Sox9 and NFIA without showing experimental evidence for direct interaction between sox3 and sox9 or between sox3 and NFIA. If they do not know whether Sox3 and Sox9 binds co-bind DNA at the same time at the same location. They should acknowledge this point in the text.
4. The authors should specify the sequence or specific motifs they used for the fgfbp3 and slc1a3 genes in the luciferase assay in Figure 5.

**Referee #3:**

Klum and colleagues dissect the roles of Sox3, Sox9 and Sox10 in specifying mouse glial lineages during embryogenesis. The study is largely based on ChIP-seq analysis of these transcription factors in different ES cell-derived lineages and leverages on single cell transcriptome datasets from the mouse spinal cord to link ChIP-seq data to differentially expressed genes. Overall the study is well-written and the data supportive of the conclusions.

Major comments:

1. While ChIP-seq was performed from in vitro differentiated ES cell populations, transcriptome datasets were obtained from single cells isolated from mouse embryos. While the latter is arguably a valuable resource, how can the authors be sure that these are actually comparable ? To which extent could the ES cell-derived populations be pre-patterned and thus diverge from the single cell datasets they obtained ? The authors should perform RNA-seq on the same populations they used for ChIP-seq to ensure that these populations are comparable.

2. The authors claim "SOX9 activates glial gene expression together with NFIA in maturing astrocytes". While the authors show that the NFIA motif is enriched close to genes regulated by SOX9 in astrocytes, there did not provide direct evidence that the NFIA protein is expressed and that it binds at these loci. The authors should consider performing ChIP-seq or ChIP-QPCR analysis on NFIA to confirm its binding to sites targeted by SOX9.

Minor comments

1. Figure 4: why are the tracks for Sox10 OL not shown ? The figure legend of Fig.4D refers only to the left panel ?

2. The expression patterns described in Fig.3C-D would benefit from a more extensive description.

3. Even though the method was previously published, the authors should briefly describe how ES cells are differentiated

The EMBO Journal - Revision received 30 May 2018

**Referee #1:**

*The referee points out that our manuscript is more appropriate as a resource paper than as a research article.*
Indeed, apart from providing novel information on how glial gene expression is controlled at the transcriptional level, our manuscript presents several genome wide data sets that will be of general interest and serve as valuable resources. More specifically, our study includes single cell RNA-seq data on mouse E11.5 and E15.5 spinal cord cells, as well as, SOX3 ChIP-seq (two replicates), SOX9 ChIP-seq (three replicates) and NFIA ChIP-seq data sets (two replicates) in GPCs. We also provide RNA-seq data of NPC, GPC and their differentiated progeny.

**Referee #2:**

 

*1) The referee points out that we omitted the orange cell group in Figure 1B from subsequent analysis. The referee argues that we should label them as microglia and include these genes in the analysis.*
As the referee points out the few genes that are specifically expressed in the orange group are indeed typical microglial markers. However, the orange group also expresses other genes not specific for microglial cells and we are therefore uncertain if these cells are truly microglia. With this in mind we have chosen to exclude this cell group, as our main question was to identify specific genes characterizing mature astrocytes and oligodendrocytes.

*2) The referee points out that it is hard from Fig. 2A (now Fig. 2C), to tell whether there is Sox9 expression in the day 4 culture. The referee request us to show single channel images when presenting SOX3 and SOX9 expression.*
In the revised manuscript we are showing single channel images for SOX3 and SOX9 (Fig. 2C). Moreover, in the revised manuscript we have analyzed *Sox3* and *Sox9* expression levels in NPCs and GPCs using RNA-seq (Fig. 2B).

*3) The referee points out that we in our model Fig. 6E indicate that SOX3 binds both SOX9 and NFIA, but that we do not present any experimental evidence for this. The referee further points that we should indicate in the text that we do not know if SOX3 and SOX9 simultaneously co-bind the same DNA-regions.*
In the revised manuscript, we present a model figure in which we depict co-binding of SOX3, SOX9 and NFIA to enhancer regions around astrocyte genes in GPCs. This model is based on the finding that SOX3 and SOX9 bind to the same astrocytic enhancer regions in the same cell population (GPCs) and that these enhancer regions are enriched for *NFI* motifs. In the revised manuscript we also demonstrate that NFIA binding is enriched at SOX9 targeted regions.

We agree with the referee that we do not present any evidence for a direct interaction between these transcription factors or that they co-target DNA. Ultimately these experiments should have been performed at single cell resolution, a technique which not yet has been established for transcription factors. Nevertheless, these limitations have been clearly outlined in the revised legend for Fig. 6E *".....Neuronal and glial gene expression become activated when SOX3 is downregulated and a cellular context of activating transcription factors has developed, which according to the model can consist of SOX11 in neurons, SOX10 in oligodendrocytes and NFIA and SOX9 in astrocytes. It should be noted that we do not provide any evidence that SOX3/SOX9 or NFIA/SOX9 physically interact or co-bind DNA".*

*4) The referee points out that we should specify the sequence or specific motifs we used for the fgfbp3 and slc1a3 genes in the luciferase assay in Figure 5.*
The regions used for functional experiments in the luciferase assays in Fig. 5 are derived from specific ChIP-seq peak. We have specified the DNA-regions used for the functional luciferase assays in the Material and Methods section of the revised manuscript.

**Referee #3:**

*1) To ensure that the ESC derived NPCs and GPCs used in the ChIP-seq experiments are comparable to the spinal cord cells used in our scRNA-seq experiments, the referee requests that we examine our ESC derived cells using RNA-seq.*
In the revised manuscript RNA-seq analysis of the ESC derived NPCs and GPCs have been included and the comparison between RNA-seq and scRNA-seq data sets are presented in Supplementary Fig. 3A. Moreover, the revised manuscript now also incorporates *in vivo* ChIP-seq analysis of SOX2 (SOX3 homolog) in the mouse spinal cord. The SOX2 ChIP-seq analysis is presented in the revised Figure 2E and clearly shows that neuronal- and glial-specific genes are prebound also in the embryonic spinal cord.

*2) To confirm that NFIA targets SOX9 bound DNA-regions the referee requests we perform ChIP-seq or -qPCR analysis of NFIA.*
We have performed NFIA ChIP-seq experiments with GPCs and show that NFIA binding indeed is enriched at SOX9 targeted DNA-regions. These new data are presented in Supplementary Fig. 5. In

 

addition, the newly added RNA-seq analysis of NPC and GPC cultures confirms the expression of NFIA in GPCs.

*Minor comments*

*1) The referee wonders why we do not show tracks for Sox10 in oligodendrocytes in Fig. 4 and why the figure legend of Fig.4D only refers to the left panel?*

ChIP-experiment for SOX10 was performed on rat oligodendrocytes (Goméz-Lopez et al., 2011) and peaks refer to coordinates of rn5 genome assembly. For this study, peak coordinates from rn5 have been translated into the mm9 genome assembly by LiftOver (Hinrichs, A. NAR, 2006), which converts genome coordinates. However, mapped reads from SOX10 ChIP could not be translated into mm9 and tracks for SOX10 and are therefore not shown in Fig.4. In the revised manuscript the right panel of Fig. 4D has been removed.

*2. The referee points out that the data presented in Fig.3C-D would benefit from a more extensive description.*

In the revised manuscript data presented in Fig.3C-D has been more extensively described in the accompanying figure legend.

*3. The referee points out that we should briefly describe how the ES cells were differentiated.*

In the revised manuscript ES cell differentiation protocol has been briefly described in the Materials and Methods section.

The EMBO Journal - Editorial Decision 22 June 2018

Thank you for you for submitting a new version of your manuscript to The EMBO Journal and my apologies for the delay in communicating our decision to you. I have now had the chance to go through the new version, the original referee reports and your point-by-point response and to discuss this matter with my colleagues in the editorial team. I am sorry to say that we have to stand by our initial decision not to pursue publication of this manuscript in The EMBO Journal.

While the new data and the clarification you provide in the manuscript seem to address some of the criticisms the referees had of the original submission, the main concerns raised by the referees - the descriptive nature of the data and the lack of in vive loss-of-function data for Sox3 - remain unchanged from the previous round. Unfortunately, we will therefore be unable to proceed to another round of in-depth review of the manuscript.

However, given the extent and technical quality of the data provided, I have taken the liberty to discuss your study and the referee reports with my colleagues at our sister journal EMBO Reports. The responsible editor there, Esther Schnapp, would be happy to consider your revised manuscript for publication in EMBO reports and she has offered to go back to the same referees that saw the previous version for EMBO Journal. You can transfer the manuscript directly to EMBO Reports using the link provided below and there is no need to reformat anything. Please feel free to contact Esther directly for specific questions about the transfer process or the requirements for acceptance in EMBO Reports.

I am sorry that we cannot be more positive for The EMBO Journal in this case but I hope you will take the chance to transfer the revised version manuscript to EMBO reports instead.

1st Editorial Decision 11 July 2018

Thank you for the submission of your revised manuscript. We have now received the enclosed comments from the referees and I am happy to tell you that we can in principle accept your manuscript now.

Only a few minor changes are needed:

- Please add all author contributions, up to 5 keywords and a conflict of interest statement to the manuscript file.

- The ORCID IDs of both corresponding authors need to be added to their personal profile page in our online manuscript system. We can unfortunately not do this for you.

- The reference format needs to be changed to the numbered EMBO reports format (in EndNote) with up to ten authors listed.

- Please send us a completed authors checklist that can be found here:
http://embor.embopress.org/authorguide#revision
along with the final manuscript.

- All main and EV figures need to be uploaded as separate files.

- The 2 tables should be called Dataset EV1 and Dataset EV2. Please upload the tables as excel files and add legends to the first tab of the file.

- You can keep the Appendix with the 5 supplementary figures. In this case please add a table of content with page numbers to the Appendix file. Alternatively, you can upload the 5 Appendix figures as individual Expanded View (EV) figures that are embedded in the main manuscript text and expand when clicked online. In this case, please add the EV figure legends to the main manuscript file following the main figure legends. Please make sure to correct all callouts to the EV figures and tables in the manuscript text.

- Scale bars are missing for figures 2C, 4E and 6, please add.

- Fig 2B: please define the bars and error bars in the figure legend. Fig 3E, F: Please define "n" as the number of independently performed experiments the data are based on in the figure legend. Fig 4: Please define "n" and the error bars. SF 2D: Please define the box plots, error bars and "n". SF 3D,E: Please define "n".

EMBO press papers are accompanied online by A) a short (1-2 sentences) summary of the findings and their significance, B) 2-3 bullet points highlighting key results and C) a synopsis image that is 550x200-400 pixels large (the height is variable). You can either show a model or key data in the synopsis image. Please note that text needs to be readable at the final size. Please send us this information along with the revised manuscript.

I look forward to seeing a final version of your manuscript as soon as possible.

REFEREE REPORTS

Referee #1:

I already found this manuscript of high technical quality and interest when it was initially submitted to EMBO journal. Now the authors have addressed all my concerns and I think it makes it suitable for publication for EMBO reports.

Referee #2:

The authors have satisfactorily addressed all my concerns. I therefore support publication of this manuscript.

 

1st Revision - authors' response 24 July 2018

Thank you for your e-mail. The changes you suggested to the abstract are all fine and incorporated these into the revised manuscript (attached).

In Fig. 3E, F and EV3 we compare gene groups. The statistics (p-values) are based on the size of these groups and not the “n-value” as previously stated. The n-values indicate the number of ChIP-seq experiments, and we now describe this in the Materials and Methods section on p. 21 and in EV3.

In the revised manuscript we have in the legends (p. 32 and p. 37) clearly stated how the p-values in Fig. 3E, F and EV3 have been calculated. We have also stated how we calculated the p-value in Fig. 2B (p. 30).

 

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

| |
|---|
| Corresponding Author Name: Erik C. Böttger |
| Journal Submitted to: EMBO Reports |
| Manuscript Number: EMBOR-2018-46193 |

**Reporting Checklist For Life Sciences Articles (Rev. June 2017)**

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

## A- Figures

### 1. Data

**The data shown in figures should satisfy the following conditions:**

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

### 2. Captions

**Each figure caption should contain the following information, for each panel where they are relevant:**

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

**In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.**

## B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

| | |
|---|---|
| 1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size? | Sample sizes were chosen according to our previous experience and recommendations in Crawley JN. 2008. Behavioral phenotyping strategies for mutant mice. Neuron 57(6):809-818. |
| 1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used. | Sample size estimates were based on literature recommendations and Monte-Carlo simulations. |
| 2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established? | If necessary animals with incomplete or technically invalid data (e.g. failure of video-tracking) were excluded before disclosing genotypes and running the statistical analysis. |
| 3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe. | Genotypes cannot be randomized, but order of testing was predetermined based on a random sequence. |
| For animal studies, include a statement about randomization even if no randomization was used. | See above. |
| 4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe. | Scoring of behavioral data was either performed by persons unaware of genotype of animals or by fully automated software routines. |
| 4.b. For animal studies, include a statement about blinding even if no blinding was done | Persons performing experiments were not aware of the genotype of animals. |
| 5. For every figure, are statistical tests justified as appropriate? | Yes, detailed statistical results are given in figure legends, models and procedures are described in the Methods. |
| Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. | Using preset criteria based on data distributions known from all 2000-5000 mice tested in the respective paradigms to date, data in Fig. 5C,D,G,H and Fig. S5B,C,G,J,M,N were subjected to Box-Cox transformation before running ANOVA. For all other tests the data were assumed to meet assumptions for ANOVA. All graphs show untransformed data. |
| Is there an estimate of variation within each group of data? | SEM are shown as well as individual data points. |
| Is the variance similar between the groups that are being statistically compared? | Yes, where needed Box-Cox transforms were performed as described above. |

## C- Reagents

## USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

| | |
|---|---|
| 6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right). | Done. |
| 7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination. | Done. |

* for all hyperlinks, please see the table at the top right of the document

## D- Animal Models

| | |
|---|---|
| 8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals. | This is detailed in the methods section and in the respective figure legends. |
| 9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments. | Mouse experiments were approved by the Veterinary Office of the Canton of Zurich (licenses 29/2012 and 44/2015) and monitored by the Animal Welfare Officer of the University of Zurich. |
| 10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance. | Compliance is confirmed. |

## E- Human Subjects

| | |
|---|---|
| 11. Identify the committee(s) approving the study protocol. | NA |
| 12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report. | NA |
| 13. For publication of patient photos, include a statement confirming that consent to publish was obtained. | NA |
| 14. Report any restrictions on the availability (and/or on the use) of human data or samples. | NA |
| 15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable. | NA |
| 16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list. | NA |
| 17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines. | NA |

## F- Data Accessibility

| | |
|---|---|
| 18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'.<br><br>Data deposition in a public repository is mandatory for:<br>a. Protein, DNA and RNA sequences<br>b. Macromolecular structures<br>c. Crystallographic data for small molecules<br>d. Functional genomics data<br>e. Proteomics and molecular interactions | Done. |
| 19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right). | Done. |
| 20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right). | Done. |
| 21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information. | Done. |

## G- Dual use research of concern

| | |
|---|---|
| 22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could. | The study does not fall under dual use research restrictions. |