EMBO reports

TGF-β-activated LncRNA LINC00115 is a critical regulator of glioma stem-like cell tumorigenicity

Jianming Tang, Bo Yu, Yanxin Li, Weiwei Zhang, Angel A. Alvarez, Bo Hu, Shi-Yuan Cheng, and Haizhong Feng

Review timeline:

Submission date:	26 March 2019
Editorial Decision:	3 May 2019
Revision received:	4 August 2019
Editorial Decision:	3 September 2019
Revision received:	10 September 2019
Accepted:	16 September 2019

Editor: Achim Breiling/Esther Schnapp

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 3 May 2019

Thank you for the submission of your research manuscript to EMBO reports. We have now received reports from the three referees that were asked to evaluate your study, which can be found at the end of this email.

As you will see, all referees think the manuscript is of interest and has novelty, but requires a major revision to allow publication in EMBO reports. As the reports are below, and all of the points need to be addressed, I will not further detail them here. However, in particular the concerns and suggestions of referee #2 need attention. Moreover, please reorganize the manuscript as indicated by referees #2 and #3, and have the revised manuscript proofread by a native speaker.

Given the constructive referee comments, we would like to invite you to revise your manuscript with the understanding that all referee concerns must be addressed in the revised manuscript and in a detailed point-by-point response. Acceptance of your manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of revision only and acceptance or rejection of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision; they will otherwise be treated as new submissions. Please contact me if a 3-months time frame is not sufficient so that we can discuss the revisions further.

When submitting your revised manuscript, please carefully review the instructions that follow below. Failure to include requested items will delay the evaluation of your revision. When submitting your revised manuscript, we will require:

1) a .docx formatted version of the manuscript text (including legends for main figures, EV figures and tables). Please make sure that the changes are highlighted to be clearly visible.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure), of main figures and EV figures.

3) a .docx formatted letter INCLUDING the reviewers' reports and your detailed point-by-point responses to their comments. As part of the EMBO Press transparent editorial process, the point-by-point response is part of the Review Process File (RPF), which will be published alongside your paper.

4) a complete author checklist, which you can download from our author guidelines (http://embor.embopress.org/authorguide#revision). Please insert page numbers in the checklist to indicate where the requested information can be found in the manuscript. The completed author checklist will also be part of the RPF.

Please also follow our guidelines for the use of living organisms, and the respective reporting guidelines: http://embor.embopress.org/authorguide#livingorganisms

5) Please also note that we now mandate that all corresponding authors provide an ORCID digital identifier that is linked to his/her EMBO reports account.

6) Before submitting your revision, primary datasets produced in this study need to be deposited in an appropriate public database. See: http://embor.embopress.org/authorguide#datadeposition

Please remember to provide a reviewer password if the datasets are not yet public.

The accession numbers and database should be listed in a formal "Data Availability " section (placed after Materials & Methods) that follows the model below. Please note that the Data Availability Section is restricted to new primary data that are part of this study.

Data availability

The datasets produced in this study are available in the following databases:

- RNA-Seq data: Gene Expression Omnibus GSE46843 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46843)
- [data type]: [name of the resource] [accession number/identifier/doi] ([URL or identifiers.org/DATABASE:ACCESSION])

*** Note - All links should resolve to a page where the data can be accessed. ***

7) We now strongly encourage the publication of original source data with the aim of making primary data more accessible and transparent to the reader. The source data will be published in a separate source data file online along with the accepted manuscript and will be linked to the relevant figure. If you would like to use this opportunity, please submit the source data (for example scans of entire gels or blots, data points of graphs in an excel sheet, additional images, etc.) of your key experiments together with the revised manuscript. Please include size markers for scans of entire gels, label the scans with figure and panel number, and send one PDF file per figure.

8) Please format the references according to our journal style. See: http://embor.embopress.org/authorguide#referencesformat

9) Our journal encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at: http://embor.embopress.org/authorguide#referencesformat

10) The Expanded View format, which will be displayed in the main HTML of the paper in a collapsible format, has replaced the Supplementary information. You can submit up to 5 images as Expanded View. Please follow the nomenclature Figure EV1, Figure EV2 etc. The figure legend for these should be included in the main manuscript document file in a section called Expanded View Figure Legends after the main Figure Legends section. Additional Supplementary material should be supplied as a single pdf labeled Appendix. The Appendix includes a table of content on the first page, all figures and their legends. Please follow the nomenclature Appendix Figure Sx throughout the text and also label the figures according to this nomenclature.

For more details please refer to our guide to authors:
http://embor.embopress.org/authorguide#manuscriptpreparation

See also our guide for figure preparation:
http://www.embopress.org/sites/default/files/EMBOPress_Figure_Guidelines_061115.pdf

11) Regarding data quantification and statistics, can you please specify, where applicable, the number "n" for how many independent experiments (biological replicates) were performed, the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values in the respective figure legends. Please provide statistical testing where applicable. See:
http://embor.embopress.org/authorguide#statisticalanalysis

12) Please move the conflict of interest statement to the end of the manuscript text, above the acknowledgements. Please also add up to five key words to the title page, and remove 'Category: The molecular oncology of human tumors'.

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision.

REFEREE REPORTS

Referee #1:

The Authors identified the lncRNA LIN00115 as a novel potential target in anti-glioma therapy due to its ability to regulate glioma stem cells in vitro self-renewal and tumorigenicity. Moreover, they clarify LIN00115 mechanism of action. Indeed, they demonstrated that, sequestrating miRNA mir200-c, LIN00115 exerts a dual function on glioma stem cells.

As a first point, they found that LIN00115 positively regulates the expression of ZEB, a critical EMT regulator whose role in GSCs has been already clearly demonstrated in literature. Secondly, they showed that LIN00115 silencing decreases the expression of ZNF596, and, consequently, of one of its canonical targets: EZH2.

They demonstrated that this is sufficient to impair EZH2 physiological function in their models, since they documented a reduction is H3K27 tri-methylation levels and in the phosphorylation of STAT3.

Revision points:

1. In the introduction, authors should gently control this phrase, that seems incomplete. "In this study, we performed RNA-Seq analysis in GSCs and identified long intergenic non-protein coding RNA 115 (LINC00115) as a highly activated lncRNA by."

2. RT-qPCR used to validate RNA-seq results should be repeated on additional patient derived GSC samples, in order to take in account, the great inter-tumoral heterogenity of this tumors.

3. The in silico analysis done using data from TCGA RNAseq and Rembrandt database show actually a very weak correlation index (Fig. 4H-I, Fig. 5-J-K). Authors should repeat the analysis segregating patients on the base of glioma grading. This is also supported by the fact that Authors

show also a different expression levels of LIN00115 in low and high-grade gliomas.

4. Authors should better explain why they choose to analyze the expression changes of EXH2, among all the possible transcriptional targets of ZNF596.

5. In the Fig.7, Authors show the effect of GSK343 on tumorigenic properties of GSCs. They should also show a survival analysis of these mice, as done in the previous figures.

Referee #2:

Major findings: lncRNA LINC00115 is upregulated in response to TGF-beta in glioma sphere lines. Knockdown of LINC00115 impaired xenograft formation in mice. The authors provide mechanistic insight underlying the involvement of LINC00115 in tumourigenicity. Specifically, they show that LINC00115 regulates ZEB1 and ZNF596/EZH2/STAT3-signaling through competitively binding miR-200s.

Overall impression: The manuscript touches on interesting aspects but is difficult to follow. It seems to put different parts of projects together without a clear flow and in a superficial manner. TGFbeta, and modulation of EZH2 have many biological effects. Whereas the claims are novel, this reviewer is concerned about the scientific quality of the manuscript and overall clinical relevance of the findings. A reorganisation/integration of the project is necessary. The authors study two functional pathways (ZEB1 and ZNF596/EZH2/STAT3). Many parts are not relevant for the questions raised and should be removed.

Main questions:
1. It is not clear why 16h of TGFbeta treatment are chosen for the experiment, this will elicit also many secondary effects. The mechanism of induction is not clear. TGFbeta activates the SMAD signaling pathway etc.
2. It is not clear why the authors mention differential expression of LIN00115 in ADH high vs low and CD133 high vs low cells. Are they implying that different effects are induced by TGFbeta treatment in cells with different stemness features? This is not followed up on.
3. Differential expression of a gene in tumor cells vs "normal" brain does not make it a gene involved in progression. Furthermore, LGG II and III are different diseases from GBM, so most genes are differentially expressed. This is not an argument that a gene is involved in progression.
4. The in vivo experiments show reduced growth with knock-down of LINC00115. However, for the claim that LINC0015 is required for self-renewal, complete removal would be required (e.g. CRISPR Cas-9).
5. The outcome data presented in Fig 1 seems highly selected for showing the effect of interest. Using the TCGA GBM dataset (HG U133A) or the Murat GBM dataset there is no difference when comparing high versus low, while in the Frejje dataset longer survival is associated with high expression (explored with the GlioVis website http://gliovis.bioinfo.cnio.es/), thus the opposite of the datasets chosen for Figure 1. Furthermore, according to the GlioVis webpage the LINC00115 is not measured on the AFFY chip used by Rembrandt.
6. I don't understand the relevance of the comparison shown in Figure 1B, it does not contribute anything.
7. The experiment shown in Fig 4G, shows the effect of TGFbeta1 on EMT, and not as claimed the effect of LINC0015. Respective genetic studies need to be added, loss of the signature upon knock-down of LNC0015 in the TGFbeta treated cells, and gain of signature with ectopic expression. An inducible system for the sh against LINC0015 may be useful to address the question.
8. The tumourigenicity was impaired by constitutive sh against LINC00115. The manuscript would benefit from using inducible hairpins. It is known that the early phase of tumor take is very sensitive to any interference. The question is whether it contributes to tumor maintenance, only then it would be of interest to target
9. It is not clear why the authors look at EZH2 after CRISPR knock-out of ZNF596.
10. Negative controls for all immunohistochemistry (with just a secondary antibody) and RNAscope (sense) is needed. Figure 8 A can not be interpreted as shown.

The manuscript contains spelling mistakes and unfinished sentences. It should be proofread.

Referee #3:

The authors propose a mechanism of tumorigenesis in glioblastoma mediated by a lncRNA, LINC00115. They first analyzed alterations in the transcriptome induced by TGFbeta1 in GSCs, leading them to identify LINC00115 as a potential regulator of tumorigenesis. They validated the relevance of LINC00115 using patient survival data, patient tumor gene expression data, and in situ hybridization of tumor vs adjacent tissue. They evaluated the role of LINC00115 in GBM both in vitro and in vivo using knockdowns and found that LINC00115 had an effect on GSC expansion. Using a database search, they predicted that LINC00115 may bind to miR-200s, and they validated this using immunoprecipitation, affinity pull-down, and a luciferase reporter assay. Finally, they evaluated the relationship between LINC0015 expression and ZEB1 signaling using knockdowns and found that only re-expression of LINC00115 WT rescued the expression of ZEB1 and restored the animal survival time to baseline. Next, they found that ZNF596 may bind LINC00115, and they validated the importance of ZNF596 in the signaling pathway by demonstrating that it reversed the LINC00115-induced phenotype in vitro and that ZNF596 directly associates with miR-200s. Finally, using CRISPR knockouts, they found EZH2 and STAT3 signaling were transcriptionally regulated by LINC00115 and ZNF596 in GSCs, and that the impaired tumorigenicity phenotype in LINC00115 knockdown GSCs was reversed by ZNF596 overexpression. They hypothesized that ZNF596 acts as transcription factor downstream to LINC00115 to activate EZH2/STAT3 signaling, which promotes GSC tumorigenesis.

The authors propose a mechanism of GSC tumorigenesis, supported by an impressive set of experimental data. The introduction should be expanded to discuss the importance of miRNAs in cancer, and to highlight the importance of miRNA sponges. The results section is at times difficult to follow, particularly when ZEB1 was introduced as a signaling pathway. The authors may want to reorder the subsections to make the results easier to follow. Additionally, parts of the discussion focus too much on restating the results, rather than summarizing and discussing the importance. It may be beneficial to outline the results in terms of the tumorigenicity pathways outlined in Figure 8D.

Page 8: "These results suggest that LINC00115 is critical for glioma progression and GSC maintenance." This sentence is an overstatement. The results reported at this point simply demonstrate that this lncRNA is highly expressed in GBM and is correlated with patient survival.

Figure 1A: Why is MALAT1 highlighted?

Figure 3: Layout needs to be rearranged, specifically for 3C.

Figure 6: Layout needs to be rearranged, specifically for 6D-E.

1st Revision - authors' response 4 August 2019

Point-to-point response to the reviewers' comments

Reviewer #1:

1. *In the introduction, authors should gently control this phrase, that seems incomplete. "In this study, we performed RNA-Seq analysis in GSCs and identified long intergenic non-protein coding RNA 115 (LINC00115) as a highly activated lncRNA by."*

Response: We apologize for this mistake. We have revised it as "…activated lncRNA by TGF-b1".

2. *RT-qPCR used to validate RNA-seq results should be repeated on additional patient derived GSC samples, in order to take in account, the great inter-tumoral heterogenity of this tumors.*

Response: We appreciate this critical comment. As requested by the reviewer, we performed a new set of experiments in new 83 and 157 GSCs in the revised Fig. S2. LINC00115 also was upregulated by TGF-b1 stimulation in these GSCs.

3. *The in silico analysis done using data from TCGA RNAseq and Rembrandt database show actually a very weak correlation index (Fig. 4H-I, Fig. 5-J-K). Authors should repeat the analysis segregating patients on the base of glioma grading. This is also supported by the fact that Authors show also a different expression levels of LIN00115 in low and high-grade gliomas.*

Response: We appreciate this insightful comment by the reviewer. As suggested by the reviewer, we have segregated patients on GBM or Grade II-III from TCGA or REMBRANDT datasets and repeated the analysis in the revised Fig.4A, 4B, 5J, 5K, Fig. S3A, 3B, Fig. S6A, and S6B.

4. Authors should better explain why they choose to analyze the expression changes of EZH2, among all the possible transcriptional targets of ZNF596.

Response: We appreciate this critical comment. We generated CRISPR knockouts of *ZNF596* and performed RNA-Seq analysis in 1123 GSCs. GSEA showed that Enhancer of Zeste Homolog 2 (EZH2)-targeted gene signature were significantly altered in *ZNF596* knockout GSCs (the revised Figure 7A). Moreover, qRT-PCR and Western blotting assays validated that *ZNF596* knockout inhibited the expression of *EZH2* mRNA (the revised Figure 7B) and its protein (the revised Figure 7C) in 1123 and 528 GSCs.

5. In the Fig.7, Authors show the effect of GSK343 on tumorigenic properties of GSCs. They should also show a survival analysis of these mice, as done in the previous figures.

Response: We appreciate this important comment by the reviewer. We generated doxycycline-inducible LIN00115 shRNA 1123 and 528 GSCs, and then assessed the effects of EZH2 inhibitor GSK343 treatment on LINC00115-regulated GSC sphere formation, tumor growth, and animal survival in the revised Fig. 7G to 7I.

Reviewer #2:

1. *It is not clear why 16h of TGFbeta treatment are chosen for the experiment, this will elicit also many secondary effects. The mechanism of induction is not clear. TGFbeta activates the SMAD signaling pathway etc.*

Response: We apologize for this mistake. We have corrected and added data in the revised Supplementary Figure 1. 1123 and 528 GSCs were pre-cultured for 16 h in DMEM/F12 medium with EGF (2 ng/ml) and bFGF (2 ng/ml) and then followed by co-culturing with or without 20-μg/ml TGF-β1 for various times (Supplementary Figure 1). We found that TGF-β1-induced higher levels of SMAD3 phosphorylation (p-SMAD3) and a TGF-β1 signaling downstream effector, inhibitors of DNA-binding protein 1 (ID1) protein expression at 3 h compared with those at 0, 0.5, or 8-h time points. Thus, we performed RNA-Seq transcriptome analysis to compare lncRNA expression levels treated with or without TGF-b1 for 3 h in GSCs.

2. It is not clear why the authors mention differential expression of LIN00115 in ADH high vs low and CD133 high vs low cells. Are they implying that different effects are induced by TGFbeta treatment in cells with different stemness features? This is not followed up on.

Response: We appreciate this critical comment. We tried to show that LINC00115 is highly expressed in ADH^{high} or $CD133^{high}$ stemness cells. Since this is not followed up on the TGF-b treatment, we have removed this data in the revised manuscript.

3. Differential expression of a gene in tumor cells vs "normal" brain does not make it a gene involved in progression. Furthermore, LGG II and III are different diseases from GBM, so most genes are differentially expressed. This is not an argument that a gene is involved in progression.

Response: We appreciate and agree with this insightful comment by the reviewer. We have revised this as suggested by Reviewer #3: "These data suggested that LINC00115 is highly expressed in GBM and its expression is putatively correlated with patient survival."

4. The in vivo experiments show reduced growth with knock-down of LINC00115. However, for the claim that LINC0015 is required for self-renewal, complete removal would be required (e.g. CRISPR Cas-9).
Response: We appreciate this critical comment. We tried to generate GSCs with complete removal of LINC00115 using CRISPR, but we do not get one cell clone until now. However, we generated doxycycline-inducible LIN00115 shRNA 1123 and 528 GSCs, and then assessed LINC00115-regulated GSC sphere formation, tumor growth, and animal survival with or without EZH2 inhibitor GSK343 treatment in the revised Fig. 7G to 7I.

5. The outcome data presented in Fig 1 seems highly selected for showing the effect of interest. Using the TCGA GBM dataset (HG U133A) or the Murat GBM dataset there is no difference when comparing high versus low, while in the Frejje dataset longer survival is associated with high expression (explored with the GlioVis website http://gliovis.bioinfo.cnio.es/), thus the opposite of the datasets chosen for Figure 1. Furthermore, according to the GlioVis webpage the LINC00115 is not measured on the AFFY chip used by Rembrandt.
Response: We appreciate and agree with this insightful comment.
1) We performed survival analysis in a couple of glioma datasets and found that it is significant in REMBRANDT and GSE83300 datasets. We also found that it is significant in TCGA RNA-Seq dataset (the revised Figure 1G). It may be due to high tumoral heterogenity of gliomas or sample sizes. We revised our conclusion and showed that LINC00115 expression is putatively correlated with glioma patient survival.
2) We did not find LINC00115 or NCRNA00115 in GlioVis webpage. We downloaded the REMBRANDT dataset from http://www.betasta-sis.com/, and LINC00115 is named as NCRNA00115 in this dataset. We have linked this website in the revised text.

6. *I don't understand the relevance of the comparison shown in Figure 1B, it does not contribute anything.*
Response: We appreciate and agree with this insightful comment by the reviewer. Since this is not followed up on the TGF-b treatment, we have removed this data.

7. *The experiment shown in Fig 4G, shows the effect of TGFbeta1 on EMT, and not as claimed the effect of LINC0015. Respective genetic studies need to be added, loss of the signature upon knock-down of LNC0015 in the TGFbeta treated cells, and gain of signature with ectopic expression. An inducible system for the sh against LINC0015 may be useful to address the question.*
Response: We appreciate this critical comment. As suggested by the reviewer, we generated 1123 GSCs with doxycycline-inducible LIN00115 shRNA or ectopic expression LINC00115 wild type (WT). GSEA analysis showed EMT gene signature was markedly lost in TGF-b1-treated 1123/shLINC00115 GSCs compared with the control and gained with ectopic expression of LINC00115 WT (Fig. 4H and 4I).

8. The tumourigenicity was impaired by constitutive sh against LINC00115. The manuscript would benefit from using inducible hairpins. It is known that the early phase of tumor take is very sensitive to any interference. The question is whether it contributes to tumor maintenance, only then it would be of interest to target.
Response: We appreciate this insightful comment. We generated doxycycline-inducible LIN00115 shRNA 1123 and 528 GSCs, and performed a new set of experiments to assess the effects of EZH2 inhibitor GSK343 treatment on LINC00115-regulated GSC sphere formation, tumor growth, and animal survival in the revised Fig. 7G to 7I.

9. It is not clear why the authors look at EZH2 after CRISPR knock-out of ZNF596.
Response: We appreciate this critical comment by the reviewer. As response to Reviewer#1, we generated CRISPR knockouts of *ZNF596* and performed RNA-Seq analysis in 1123 GSCs. GSEA showed that Enhancer of Zeste Homolog 2 (EZH2)-targeted gene signature were significantly altered in *ZNF596* knockout GSCs (new Figure 7A). Moreover, qRT-PCR and Western blotting assays validated that *ZNF596* knockout inhibited the expression of EZH2.

10. *Negative controls for all immunohistochemistry (with just a secondary antibody) and RNAscope (sense) is needed. Figure 8 A can not be interpreted as shown.*

Response: As required by the reviewer, we performed new experiments and showed negative controls for RNAscope (sense) and IHC (IgG) in the revised Figure S8A and 8B, respectively.

Reviewer #3:

1. *The authors propose a mechanism of GSC tumorigenesis, supported by an impressive set of experimental data. The introduction should be expanded to discuss the importance of miRNAs in cancer, and to highlight the importance of miRNA sponges. The results section is at times difficult to follow, particularly when ZEB1 was introduced as a signaling pathway. The authors may want to reorder the subsections to make the results easier to follow. Additionally, parts of the discussion focus too much on restating the results, rather than summarizing and discussing the importance. It may be beneficial to outline the results in terms of the tumorigenicity pathways outlined in Figure 8D.*

Response: We appreciate this critical comment.
1) We have revised the introduction and discussed the importance of miRNAs in cancer and highlighted the importance of lncRNAs AS miRNA sponges in pages 3-4.
2) As suggested by the reviewer, we reordered Fig. 3, Fig. 4, and Fig. 6 to make the results easier to follow.
3) We also revised the discussion section as suggested by the reviewer. We discussed LINC00115 functions as a miRNA sponge to competitively bind miR-200s and activate ZEB1 signaling.

2. *Page 8: "These results suggest that LINC00115 is critical for glioma progression and GSC maintenance." This sentence is an overstatement. The results reported at this point simply demonstrate that this lncRNA is highly expressed in GBM and is correlated with patient survival.*

Response: We appreciate and agree with this insightful comment. We have revised this as: "These data suggested that LINC00115 is highly expressed in GBM and its expression is putatively correlated with patient survival."

Figure 1A: Why is MALAT1 highlighted?
Response: We want to show that the important LncRNA MALAT1 is upregulated by TGF-b1. Since this is not followed up and there are not more data to support this, we removed this in the revised manuscript.

Figure 3: Layout needs to be rearranged, specifically for 3C.
Response: We appreciate this critical comment. As required by the reviewer, we have rearranged Fig. 3.

Figure 6: Layout needs to be rearranged, specifically for 6D-E.
Response: We also have reordered Fig. 4 and Fig. 6 to make it clear.

2nd Editorial Decision 3 September 2019

Thank you for the submission of your revised manuscript. I have taken over its handling as Achim is currently not in the office. We have now received the enclosed reports from the referees that were asked to assess it, and I am happy to tell you that both support its publication now. Only a few more minor changes will be required before we can proceed with the official acceptance of your manuscript.

Nearly all of your main figures state that the statistical analyses are based on 2-3 independent experiments. If n=2 no statistics can be calculated. Please either repeat the experiments at least one more time, or remove the error bars for all cases where n=2. In these cases, you can show all data points in the graphs, but no error bars or p-values. It seems that figures 1-7 all need to be corrected.

Several supplementary figures do not state n, this information must be added to the figure legends (eg SF2, SF4).

I attach a manuscript word file to this email with comments by our data editors. Please address all comments in the final manuscript file.

You have uploaded 8 supplementary figures. 5 of these can be changed into EV figures that expand when clicked in the online version of the manuscript. All EV figures need to be uploaded as individual files and the EV figure legends need to be added after the main figure legends. All remaining figures need to be collected in an Appendix file with a table of content and page numbers. Please see our guide to authors for more information. The supplementary table can be an EV table, an Appendix table or a regular table in the methods section.

Please remember that all scale bars must be defined in the figure legends. Some are currently missing (eg in SF5).

Please make sure that all information on funding is added to the manuscript file.

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 new revised version of your manuscript as soon as possible. Please let me know if you have any questions.

REFEREE REPORTS

Referee #2:

The authors have answered all of my questions

Referee #3:

The authors have addressed my concerns. The manuscript is now suitable for publication.

2nd Revision - authors' response 10 September 2019

The authors performed all minor editorial changes.

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: Haizhong Feng
Journal Submitted to: EMBO Reports
Manuscript Number: EMBOR-2019-48170V1

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 size is described in each figure legends.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Sample size is described in each figure legends.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	We did not exclude any raw data.
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.	Mice were randomly allocated by one investigator and treated by another investigator.
For animal studies, include a statement about randomization even if no randomization was used.	Mice were randomly allocated.
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.	Mice were randomly allocated by one investigator and assessed by another investigator.
4.b. For animal studies, include a statement about blinding even if no blinding was done	Mice were randomly allocated by one investigator and assessed by another investigator.
5. For every figure, are statistical tests justified as appropriate?	Yes, the statistical tests are justified appropriatedly for every figure..
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes, the data meet the assumptions of the tests.

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/

Is there an estimate of variation within each group of data?	Yes, we have inserted the estimate of variation within each group of data.
Is the variance similar between the groups that are being statistically compared?	Yes, for statistical data in each data, no substantial difference has been observed among the variations.

C- Reagents

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).	Actin (I-19), ZEB1 (Cat#SC-10570), ZNF596 (Cat#SC-98284), STAT3 (H-190, Santa Cruz Biotechnology); Nestin (Cat#ABD69MI, Fisher Scientific); Flag (M2, Sigma-Aldrich); Vimentin (Cat#5741S), E-cadherin (Cat#3195S), EZH2 (Cat#5246S), Tri-Methyl-Histone H3 (Lys27)(Cat#9733S), Histone H3 (Cat#9715S), phospho-STAT3 (Y705) (D3A7, Cell Signaling Technology).
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	We reported the source of glioma cell lines, which were recently authenticated using STR DNA fingerprinting at Shanghai Biowing Applied Biotechnology Co., Ltd (Shanghai, China).As mycoplasma contamination was not tested recently, we did not report here.

* 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.	Athymic (NCr nu/nu) female mice at ages of 6-8 weeks (SLAC, Shanghai, China) were used for all animal experiments (page 24).
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	All experiments using animals were approved by Shanghai Jiao Tong University Institutional Animal Care and Use Committee (IACUC)(page 24).
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.	We confirmed compliance.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	In accordance to a protocol approved by Shanghai Jiao Tong University Institutional Clinical Care and Use Committee, according to the Declaration of Helsinki.
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.	In accordance to a protocol approved by Shanghai Jiao Tong University Institutional Clinical Care and Use Committee, according to the Declaration of Helsinki, clinical brain tissue specimens were collected at Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China. The investigators obtained informed written consent from all subjects (pages 23-24).
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	No patient photo.
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	The use of human data or samples needs the approvement from China government.
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	No clinicla trial.
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.	No clinicla trial.
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.	Yes, we have followed these guidelines.

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'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	Yes, the "Data availability" was provided.
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).	RNA-Seq data reported in this study have been deposited with the Gene Expression Omnibus under the accession GEO ID: GSE134595 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE134595).
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).	not related
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.	not related

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.	no