EMBO reports

DPPA2 and DPPA4 are necessary to establish a 2C-like state in mouse embryonic stem cells

Alberto De Iaco, Alexandre Coudray, Julien Duc and Didier Trono

Review timeline:

Submission date:	9 November 2018
Editorial Decision:	10 December 2018
Revision received:	24 January 2019
Editorial Decision:	7 February 2019
Revision received:	7 March 2019
Accepted:	11 March 2019

Editor: 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 10 December 2018

Thank you for the submission of your manuscript to EMBO reports. It was sent to three referees, and so far we have received reports from two of them, which I paste below. As both referees feel that the manuscript is interesting and recommend that you should be given a chance to revise it, I would like to ask you to begin revising your manuscript according to the referees' comments. Please note that this is a preliminary decision made in the interest of time, and that it is subject to change should the third referee offer very strong and convincing reasons for this.

I have gone through all referee comments and think that all make sense and should be addressed. If you would like to fast-track this manuscript, I suggest that you submit the revised version by the end of January, if you think this is possible. I will also forward the missing referee report as soon as we will have received it.

We would thus like to invite you to revise your manuscript with the understanding that the referee concerns must be fully addressed and their suggestions taken on board. Please address all referee concerns in a complete point-by-point response. Acceptance of the 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 us if a 3-months time frame is not sufficient for the revisions so that we can discuss this further. You can either publish the study as a short report or as a full article. For short reports, the revised manuscript should not exceed 27,000 characters (including spaces but excluding materials & methods and references) and 5 main plus 5 expanded view figures. The results and discussion sections must further be combined, which will help to shorten the manuscript text by eliminating some redundancy that is inevitable when discussing the same experiments twice. For a normal article there are no length limitations, but it should have more than 5 main figures and the results and discussion sections must be separate. In both cases, the entire materials and methods must be included in the main manuscript file.

Supplementary figures, tables and movies can be provided as Expanded View (EV) files, and we can offer a maximum of 5 EV figures per manuscript. EV figures are embedded in the main manuscript text and

expand when clicked in the html version. Additional supplementary figures will need to be included in an Appendix file. Tables can either be provided as regular tables, as EV tables or as Datasets. Please see our guide to authors for more information.

Regarding data quantification, please specify the number "n" for how many independent experiments were performed, the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values in the respective figure legends. This information must be provided in the figure legends. Please also include scale bars in all microscopy images.

Please note that the EMBO reports reference style is numbered. It can be found in EndNote.

We 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.

When submitting your revised manuscript, we will require:

- 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 in the manuscript the requested information can be found. The completed author checklist will also be part of the RPF (see below).
- a letter detailing your responses to the referee comments in Word format (.doc)
- a Microsoft Word file (.doc) of the revised manuscript text
- editable TIFF or EPS-formatted figure files in high resolution. In order to avoid delays later in the process, please read our figure guidelines before preparing your manuscript figures at: http://www.embopress.org/sites/default/files/EMBOPress_Figure_Guidelines_061115.pdf

We would also welcome the submission of cover suggestions, or motifs to be used by our Graphics Illustrator in designing a cover.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File (RPF) to accompany accepted manuscripts. This File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

You are able to opt out of this by letting the editorial office know (emboreports@embo.org). If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

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:

De Iaco et al. identify DPPA2 and DPPA4 as key activators of a 2-cell-like transcriptional programme in mESCs. The authors find that both DPPA2 and DPPA4 bind to and activate Dux and consequently its downstream targets. DPPA2/4 also have Dux-independent targets, including L1 elements.

This is a nicely put together study that sets the scene for testing the roles of maternally inherited DPPA2/4 in ZGA in vivo. The insights gained here will be extremely useful to ask targeted questions therein. Importantly, the authors make it clear (including on the title) that their findings only relate to ESCs and everything else is extrapolation that requires thorough testing. I have the following relatively minor comments/suggestions:

© European Molecular Biology Organization

1. Is Dppa2 expression Dux-dependent? The authors only show data for Dppa4 (Supp. Fig. 1D). A preprint from the Reik lab suggests a feedback loop by which Dux activates Dppa2 (but not Dppa4), although this is from an overexpression experiment.

2. Figure 2E shows no evidence of expression of exogenous Dux. Do the primers distinguish endogenous from exogenous? If so, could the authors also display levels of the exogenous Dux?

3. Is DPPA2/4 binding preference to high %GC in general, or do these proteins basically bind CpG islands (as in the examples on Supp. Fig. 3F)? The 5' UTR of L1s are also essentially CpG islands. This could suggest mechanisms regulating DPPA2/4 binding, such as DNA methylation or other CpG island-specific features.

4. Could the authors give a sense of what proportion of DPPA2/4-bound genes are affected at the transcriptional level in the KOs?

5. The following sentence could do with rephrasing, as it can easily be misread to give the opposite meaning to what is intended: "...the SAP domain is absolutely essential neither for the genomic recruitment ..." (line 151).

6. Supp. Fig. 3D would benefit from clarifying that "DPPA2 on Dppa2 KO" (and similar labels) refers to DPPA2-HA overexpression in Dppa2 KO cells. Readers may otherwise interpret it as a ChIP for a protein that isn't there.

7. I disagree that "Transcription of L1Md_T and L1Md_A, but not of MERVL, was partly rescued..." (line 167), given that the effect on L1A is very minor or absent (Supp. Fig. 4B).

8. It is perhaps also unwarranted to maintain that L1 expression is completely Dux-independent when there is a clear L1 downregulation in the Dux KO (Supp. Fig. 4C). Although relatively small, this effect is significant and larger than the effect of DPPA2/4 rescue on L1A (Supp. Fig. 4B).

Referee #2:

De Iaco et al. DPPA2 and DPPa4 are necessary to establish a 2C-like state in mouse embryonic stem cells.
The authors identify Dppa2 and Dppa4 as genes expressed around the 2-cell stage, based on public murine RNAseq datasets. The identify an isoform, which lacks the SAP domain of Dppa4, based on the same datasets. They then go on and show that 2CLC (Mervl+) expressing also Zscan4 have higher levels of Dppa2 and 4 transcripts, compared to MERVL-/Zscan4- cells. RNAi for either Dppa2 or Dppa 4 significantly reduces the number of MERVL+ cells. They generate Dppa2 and Dppa4 KO clones using CRISPR, and perform RNAseq, in which they observed downregulation of Dux-regulated genes and LINE-1 elements from the young families (Md_A and T), as well as MERVL. By complementation assays, they show that Dux target genes Zscan4 and Tdpoz4 are downregulated in either Dppa2 or 4 KO, but their expression is partially or fully restored, when Dux is expressed in these cells. They identify Dppa2 and 4 binding regions by ChIP, and show that they bind more GC regions, including the Dux gene regulatory region.
The paper presents novel findings on a very hot topic. Namely, potential regulators of a '2-cell-like-state'. I am therefore fully supportive for publication of this manuscript, but would need to see the following experiments and amendments added, as follows.
The manuscript title reflects very well the conclusions of the paper, but the text of the manuscript itself does not: namely, the authors show that Dppa2/Dppa4 are necessary for the 2C-state, but they do not show that they 'induce' or 'promote' the 2C-like state. My comments are therefore two-fold i) to add experiments to fully disentangle the relationship between Dppa2/ Dppa4 and Dux and ii) to rephrase their manuscript throughout accordingly.
Also, overall, I think the text writing could improve to make it easier for the reader to follow the story, since it reads a little broken throughout.
Additional experiments:
I. The authors are missing a key experiment, which is the overexpression of Dppa2 in combination and in absence of Dux. Can the authors:
a) Overexpress Dppa2 (or Dppa4) alone and in combination with Dux in wild type ES cells, and measure

© European Molecular Biology Organization

whether they i) impact % of MERVL positive cells (as in Fig. S2B) and ii) affect gene expression of the genes assessed in Fig. 3A.
b) Overexpress Dppa2 (or Dppa4) in the Dux KO clone and i) measure the number of MERVL positive cells (this can be done alternative by QPCR, but ideally with e.g. Immunostaining of Gag and counts, if the Dux KO cell line does not contain the MERVL reporter) ii) asses QPCR levels of the genes assessed in S3D (as in figure 3A, but the figure S3D is basically missing the Dux KO dataset)
In the discussion and throughout the text the authors' argument on DPPA2 and DPPA4 being inducers/upstream/activators of Dux, LINE-1 and 2CLCs is not fully demonstrated. I believe that without the above two experiments, the authors cannot really state whether Dppa2/ Dppa4 are upstream or downstream of Dux, especially since the 'rescue' experiments presented in Fig 3A and S3D are not really conclusive towards this regard (e.g. the rescues are not full, and I am unsure as to whether the levels of expression, as the authors explain, are sufficient to explain this phenotype).
II. The authors should include the WT control in S3D, otherwise the data is very difficult to interpret. The comparison with the WT situation could potentially explain why the rescue infigure 3A is not complete.
Additional analysis and/or text corrections:
I. Lines 74 to 76: "Here, we demonstrate that DPPA2 and DPPA4 act as inducers of Dux and LINE1 transcription in mESCs to promote the establishment of a 2C-like state". This is the main conclusion at the end of the introduction. Please correct, as per my point above, as the authors do not demonstrate DPPA2 and DPPA4 are inducers of either Dux or LINE1 to promote 2CLCs. From their data, they demonstrate DPPA2 and DPPA4 are necessary for either induction or maintenance of 2CLCs population (e.g. only for the spontaneously arising in ESC cultures), but no induction is formally demonstrated.
II. Lines 105 to 106: "DPPA2 and DPPA4 induce expression of DUX and other ZGA genes in 2C-like mESCs". As above, the statement is not founded, as the authors do not demonstrate that DPPA2 and DPPA4 induce Dux expression. Please rephrase.
III. Lines 121 to 125: "Expression of some of these genes (Zscan4 and Tdpoz4) was rescued when Dux was ectopically expressed in Dppa2 or Dppa4 KO mESCs (Figure 2E), suggesting that DPPA2 and DPPA4 act upstream of DUX in the establishment of a 2C-like state in mESCs". The expression was not completely rescued, this suggests that there might be some role for DPPA2 and DPPA4 downstream or in parallel with Dux. This is an interesting observation, the authors should discuss this.
IV. Lines 129 to 131: "Nevertheless, only 1/10 of these genes were transcribed specifically at ZGA (Expanded View Figure 2G), pointing to differences in the transcriptomes of 2C embryos and 2C-like ES cells (Madanet al. 2009; Nakamura et al. 2011)". Unless I am mistaken, the figure does not show an overlap with 2C-like cell genes, therefore this conclusion cannot be made. An overlap with the 2CLC genes described in either Macfarlan et al or Ishiuchi et al should be done.
V. Line 140: "DPPA2 and DPPA4 bind the promoters of Dux and other ZGA genes". As above, I was unable to find the information of which ZGA gene promoters are bound by DPPA2 and DPPA4. Is this analysis missing? (e.g. ZGA gene promoters extracted from their ChIPseq analysis?). Please correct, add the analysis and/or clarify.
VI. Can the authors ask if the 'Dppa-dependent'- genes have a different expression dynamics (temporal) in the embryos, compared to the 'Dux-dependent' genes? (using the same datasets as those in Figure 1A)?. This will add a little bit more embryo biology to their manuscript, without the need of having to perform embryo work, which I understand requires skills and is heavy.
VII. Data in Figure S2A and S2B are rather important, I suggest these 2 panels be moved to the main manuscript (Figure 2B in the main manuscript should be removed accordingly, it does not add anything to the text and takes unnecessary space)
VIII. Throughout the manuscript, I find the term 'Dux-dependent' genes confusing. Are Dux-dependent genes also directly bound by Dux (e.g. the comparison between Dux KO and Dux ChIP seq from Hendrikson et al may be informative). Also, what is the difference between 'Dux-dependent' genes and 'Dux-regulated' genes? Is there a difference between the Dux bound genes and the Dppa2 (or Dppa4) bound genes, in terms of distance to LTRs (mt2_mm or MERVL_int)?

1st Revision - authors' response 24 January 2019

Referee #1:

De Iaco et al. identify DPPA2 and DPPA4 as key activators of a 2-cell-like transcriptional programme in mESCs. The authors find that both DPPA2 and DPPA4 bind to and activate Dux and consequently its downstream targets. DPPA2/4 also have Dux-independent targets, including L1 elements.

© European Molecular Biology Organization

This is a nicely put together study that sets the scene for testing the roles of maternally inherited DPPA2/4 in ZGA in vivo. The insights gained here will be extremely useful to ask targeted questions therein. Importantly, the authors make it clear (including on the title) that their findings only relate to ESCs and everything else is extrapolation that requires thorough testing. I have the following relatively minor comments/suggestions:

1. Is Dppa2 expression Dux-dependent? The authors only show data for Dppa4 (Supp. Fig. 1D). A preprint from the Reik lab suggests a feedback loop by which Dux activates Dppa2 (but not Dppa4), although this is from an overexpression experiment.
We find that DUX is not necessary for Dppa2 transcription, but high levels of the transcription factor lead to an increased expression of Dppa2 *in vitro*. Notably, expression of Dppa2 is not significantly decreased in Dux KO mESCs compared to WT cells, but its transcription is increased by 2-fold when an HA tagged version of DUX is ectopically expressed in the Dux KO cell line. We now added Figure EV1D and EV1E showing the expression of Dppa2 in WT, Dux KO mESCs and Dux KO mESC everexpressing DUX-HA.

2. Figure 2E shows no evidence of expression of exogenous Dux. Do the primers distinguish endogenous from exogenous? If so, could the authors also display levels of the exogenous Dux?
The cDNA encoding for exogenous DUX protein is codon optimized whereas the primers used in our work target the endogenous Dux ORF. It is thus technically difficult to compare the two. However, we know from previous immunofluorescence experiments directed to the HA-tagged form of DUX (De Iaco 2017 Nature Genetics) that the level of DUX expression reached using this plasmid is sufficient to get around 50% of the cells to express readily detectable levels of DUX. Furthermore, WT mESCs transfected with this vector display increased levels of DUX targets (Zscan4 and Tdpoz4) confirming that the exogenous protein can add to the effect of its endogenous counterpart.

3. Is DPPA2/4 binding preference to high %GC in general, or do these proteins basically bind CpG islands (as in the examples on Supp. Fig. 3F)? The 5' UTR of L1s are also essentially CpG islands. This could suggest mechanisms regulating DPPA2/4 binding, such as DNA methylation or other CpG island-specific features.
As suggested by the reviewer, DPPA2/4 not only bind high GC-rich sequences, but they associate preferably with CpG island as shown by new Figure EV4C.

4. Could the authors give a sense of what proportion of DPPA2/4-bound genes are affected at the transcriptional level in the KOs?
In new Figure EV3F, we now illustrate the expression of genes bound by DPPA2 and DPPA4 at their promoter in mESCs depleted of the two proteins. We found that a large fraction of these genes is downregulated in the KO cells.

5. The following sentence could do with rephrasing, as it can easily be misread to give the opposite meaning to what is intended: "...the SAP domain is absolutely essential neither for the genomic recruitment ..." (line 151).
We modified the formulation.

6. Supp. Fig. 3D would benefit from clarifying that "DPPA2 on Dppa2 KO" (and similar labels) refers to DPPA2-HA overexpression in Dppa2 KO cells. Readers may otherwise interpret it as a ChIP for a protein that isn't there.
We modified the figure.

7. I disagree that "Transcription of L1Md_T and L1Md_A, but not of MERVL,

© European Molecular Biology Organization

was partly rescued..." (line 167), given that the effect on L1A is very minor or absent (Supp. Fig. 4B).
We modified our statement.

8. It is perhaps also unwarranted to maintain that L1 expression is completely Dux-independent when there is a clear L1 downregulation in the Dux KO (Supp. Fig. 4C). Although relatively small, this effect is significant and larger than the effect of DPPA2/4 rescue on L1A (Supp. Fig. 4B).
We modified our statement.

Referee #2:

De Iaco et al. DPPA2 and DPPa4 are necessary to establish a 2C-like state in mouse embryonic stem cells.
The authors identify Dppa2 and Dppa4 as genes expressed around the 2-cell stage, based on public murine RNAseq datasets. The identify an isoform, which lacks the SAP domain of Dppa4, based on the same datasets. They then go on and show that 2CLC (Mervl+) expressing also Zscan4 have higher levels of Dppa2 and 4 transcripts, compared to MERVL-/Zscan4- cells. RNAi for either Dppa2 or Dppa 4 significantly reduces the number of MERVL+ cells. They generate Dppa2 and Dppa4 KO clones using CRISPR, and perform RNAseq, in which they observed downregulation of Dux-regulated genes and LINE-1 elements from the young families (Md_A and T), as well as MERVL. By complementation assays, they show that Dux target genes Zscan4 and Tdpoz4 are downregulated in either Dppa2 or 4 KO, but their expression is partially or fully restored, when Dux is expressed in these cells. They identify Dppa2 and 4 binding regions by ChIP, and show that they bind more GC regions, including the Dux gene regulatory region.

The paper presents novel findings on a very hot topic. Namely, potential regulators of a '2-cell-like-state'. I am therefore fully supportive for publication of this manuscript, but would need to see the following experiments and amendments added, as follows.

The manuscript title reflects very well the conclusions of the paper, but the text of the manuscript itself does not: namely, the authors show that Dppa2/Dppa4 are necessary for the 2C-state, but they do not show that they 'induce' or 'promote' the 2C-like state. My comments are therefore two-fold i) to add experiments to fully disentangle the relationship between Dppa2/ Dppa4 and Dux and ii) to rephrase their manuscript throughout accordingly. Also, overall, I think the text writing could improve to make it easier for the reader to follow the story, since it reads a little broken throughout.
In the previous version of the manuscript, we transduced cells with lentiviral vectors encoding for the standard ORF of DPPA2 and DPPA4 to rescue their expression in the KO cells. Using this technique, we were able to express DPPA2 or DPPA4 in all cells of the mESC population, but to lower levels compared to WT cells. This was probably one of the reasons why we could not rescue the activity of DPPA2 or DPPA4 in the corresponding KO cells. We now transfected the KO cells with plasmids encoding for codon optimized cDNAs. By using this strategy, we obtained a higher expression of the transgenes in a lower number of cells (around 35% using a GFP reporter as a control), with a consequent mild rescue of DPPA2 and DPPA4 activity in the respective KO cell line.
Another possibility for the lack of rescue of DPPA2 or DPPA4 activity in the KO cells is that expression of the two proteins has to be equimolar, since they work as a heterodimer. To address this hypothesis, we transfected DPPA2/DPPA4 double KO mESCs with the plasmids encoding codon optimized DPPA2 and DPPA4 and found full rescue of transactivation of DUX expression.

© European Molecular Biology Organization

In new Figure 3A, we now show rescue of DPPA2 and DPPA4 activity using transfection instead of transduction. Therefore we are now confident that DPPA2 and DPPA4 are not only “necessary” for the 2C-like cells, but they “promote” this state.

Additional experiments:

I. The authors are missing a key experiment, which is the overexpression of Dppa2 in combination and in absence of Dux. Can the authors:

a) Overexpress Dppa2 (or Dppa4) alone and in combination with Dux in wild type ES cells, and measure whether they i) impact % of MERVL positive cells (as in Fig. S2B) and ii) affect gene expression of the genes assessed in Fig. 3A.
We tested whether DPPA2 and DPPA4 in combination with DUX had an effect on the % of MERVL positive cells and expression of DUX-regulated genes. In new Figure EV 3AB we show that transfection of DPPA2 and DPPA4 simultaneously with DUX does not alter the expression of downstream targets of DUX.

b) Overexpress Dppa2 (or Dppa4) in the Dux KO clone and i) measure the number of MERVL positive cells (this can be done alternative by QPCR, but ideally with e.g. Immunostaining of Gag and counts, if the Dux KO cell line does not contain the MERVL reporter) ii) asses QPCR levels of the genes assessed in S3D (as in figure 3A, but the figure S3D is basically missing the Dux KO dataset)
In new Figure EV 3CD we now show no effect on the expression of DUX downstream targets when DPPA2 and DPPA4 are overexpressed in *Dux* KO mESCs.

In the discussion and throughout the text the authors' argument on DPPA2 and DPPA4 being inducers/upstream/activators of Dux, LINE-1 and 2CLCs is not fully demonstrated. I believe that without the above two experiments, the authors cannot really state whether Dppa2/ Dppa4 are upstream or downstream of Dux, especially since the 'rescue' experiments presented in Fig 3A and S3D are not really conclusive towards this regard (e.g. the rescues are not full, and I am unsure as to whether the levels of expression, as the authors explain, are sufficient to explain this phenotype).
We hope that our responses / new data now convince the reviewer of the soundness of our general statements.

II. The authors should include the WT control in S3D, otherwise the data is very difficult to interpret. The comparison with the WT situation could potentially explain why the rescue infigure 3A is not complete.
We now added the WT control.

Additional analysis and/or text corrections:

I. Lines 74 to 76: "Here, we demonstrate that DPPA2 and DPPA4 act as inducers of Dux and LINE1 transcription in mESCs to promote the establishment of a 2C-like state". This is the main conclusion at the end of the introduction. Please correct, as per my point above, as the authors do not demonstrate DPPA2 and DPPA4 are inducers of either Dux or LINE1 to promote 2CLCs. From their data, they demonstrate DPPA2 and DPPA4 are necessary for either induction or maintenance of 2CLCs population (e.g. only for the spontaneously arising in ESC cultures), but no induction is formally demonstrated.

II. Lines 105 to 106: "DPPA2 and DPPA4 induce expression of DUX and other ZGA genes in 2C-like mESCs". As above, the statement is not founded, as

© European Molecular Biology Organization

the authors do not demonstrate that DPPA2 and DPPA4 induce Dux expression. Please rephrase.
We believe that our new data allow us to demonstrate that DPPA2 and DPPA4 induce DUX expression and as a consequence 2C-state.

III. Lines 121 to 125: "Expression of some of these genes (Zscan4 and Tdpoz4) was rescued when Dux was ectopically expressed in Dppa2 or Dppa4 KO mESCs (Figure 2E), suggesting that DPPA2 and DPPA4 act upstream of DUX in the establishment of a 2C-like state in mESCs". The expression was not completely rescued, this suggests that there might be some role for DPPA2 and DPPA4 downstream or in parallel with Dux. This is an interesting observation, the authors should discuss this.
New EV Figure 3AB shows that DPPA2 and DPPA4 do not favor establishment of a 2C-state when DUX is ectopically expressed. This evidence suggests that DPPA2 and DPPA4 do not act downstream or in parallel with DUX.

IV. Lines 129 to 131: "Nevertheless, only 1/10 of these genes were transcribed specifically at ZGA (Expanded View Figure 2G), pointing to differences in the transcriptomes of 2C embryos and 2C-like ES cells (Madanet al. 2009; Nakamura et al. 2011)". Unless I am mistaken, the figure does not show an overlap with 2C-like cell genes, therefore this conclusion cannot be made. An overlap with the 2CLC genes described in either Macfarlan et al or Ishiuchi et al should be done.
New EV Figure 2E depicts the overlap of ZGA, DPPA2/4 regulated, DUX regulated and 2C-like genes. It reveals that a large fraction of 2C-like genes are not ZGA-specific, including many (612) genes regulated by DPPA2/4 in mESCs. Thus, the transcriptomes 2C-like cells and ZGA have common features but are distinct.

V. Line 140: "DPPA2 and DPPA4 bind the promoters of Dux and other ZGA genes". As above, I was unable to find the information of which ZGA gene promoters are bound by DPPA2 and DPPA4. Is this analysis missing? (e.g. ZGA gene promoters extracted from their ChIPseq analysis?). Please correct, add the analysis and/or clarify.
Our mistake. Indeed we did not find promoters of other ZGA genes to be bound by DPPA2 and DPPA4. We modified this statement.

VI. Can the authors ask if the 'Dppa-dependent'- genes have a different expression dynamics (temporal) in the embryos, compared to the 'Duxdependent' genes? (using the same datasets as those in Figure 1A)?. This will add a little bit more embryo biology to their manuscript, without the need of having to perform embryo work, which I understand requires skills and is heavy.
New EV Figure 2F shows the dynamics of expression of DPPA2/4-dependent and DUX-dependent genes during pre-implantation embryo development. While the majority of DUX-dependent genes are clearly expressed at ZGA, their DPPA2/4-dpendent counterparts have different patterns of expression, with the most enriched fraction displaying the properties of “maternal genes”, that is, expressed in the zygote and degraded at ZGA.
VII. Data in Figure S2A and S2B are rather important, I suggest these 2 panels be moved to the main manuscript (Figure 2B in the main manuscript should be removed accordingly, it does not add anything to the text and takes unnecessary space)
Old EV Figure 2AB are now Figure 2BC.

VIII. Throughout the manuscript, I find the term 'Dux-dependent' genes confusing. Are Dux-dependent genes also directly bound by Dux (e.g. the comparison between Dux KO and Dux ChIP seq from Hendrikson et al may be informative). Also, what is the difference between 'Dux-dependent' genes

© European Molecular Biology Organization

and 'Dux-regulated' genes? Is there a difference between the Dux bound genes and the Dppa2 (or Dppa4) bound genes, in terms of distance to LTRs (mt2_mm or MERVL_int)?
It is difficult to determine precisely which genes are bound by DUX at their promoter because this activator often triggers expression of its targets from alternative unannotated promoters, some of which are localized far from the known TSS. We therefore call DUX-dependent all the genes that are downregulated in DUX KO cells, irrespectively of the ChIP-seq data. To address the reviewer's final point, we asked whether the annotated TSSs of genes downregulated in DPPA2, DPPA4 or DUX KO cells were close to MT2_mm, and found that . In new EV Figure 2G, we now demonstrate that DUX- but not DPPA2- or DPPA4-regulated genes were enriched in the proximity of MT2_mm sequences.

2nd Editorial Decision 7 February 2019

Thank you for the submission of your revised manuscript. I am glad to say that both referees support its publication now. Only a few more minor changes are needed before we can proceed with its official acceptance.

- as I mentioned in my last letter, the manuscript needs to be layed out as a scientific report (given its 4 main figures) with combined Results and Discussion sections. Please combine both sections, which will also help to reduce some redundancy that is inevitable when discussing the same experiments twice.

- please upload all figure files, including the EV figures, as individual files

- Fig 3 E,F,G,H are called out but these panels do not exist

- the EV figures need to be called out. Please correct the name of the EV figures (currently called supplementary figures). EV figure 5 is called EV figure 4, please correct.

- "Table S1" and "Table S2" are called out, but there is no Table S1. Table S2 uploaded as "Table 1". Please rename to Table 1 or Table EV1.

- please add the GEO accession numbers for the RNA-seq and ChIP-seq data to the materials and methods section.

- please correct the reference style to the numbered EMBO reports style that can be found in EndNote

- please add the ORCID IDs of the corresponding authors to your personal profile page of our online manuscript tracking system. We can unfortunately not do this for you.

- please provide up to 5 keywords

- please provide a running title

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 would like to suggest a few minor changes to the abstract. Please let me know whether you agree with the following:

After fertilization of the transcriptionally silent oocyte, expression from both parental chromosomes is

© European Molecular Biology Organization

launched through zygotic genome activation (ZGA), occurring in the mouse at the 2-cell (2C) stage. Amongst the first elements to be transcribed is the Dux gene, the product of which induces a wide array of ZGA genes and a subset of evolutionary recent LINE-1 retrotransposons that regulate chromatin accessibility in the early embryo. The maternally-inherited factors that activate Dux and LINE-1 transcription have so far remained unknown. Mouse embryonic stem cells (mESCs) recapitulate some aspects of ZGA in culture, owing to their ability to cycle through a 2C-like stage when Dux, its target genes and LINE-1 integrants are expressed. Here we identify the paralog proteins DPPA2 and DPPA4 as necessary for the activation of Dux and LINE-1 expression in mESCs. Since their encoding RNAs are maternally transmitted to the zygote, it is likely that these factors are important upstream mediators of murine ZGA.

I look forward to seeing a final version of your manuscript as soon as possible. Please let me know if you have any questions.

REFEREE REPORTS

Referee #1:

The authors have adequately addressed my comments.

Referee #2:

The authors did not address all the points I raised in my original review. The paper is important and should nevertheless be published.

2nd Revision - authors' response 7 March 2019

The authors performed all minor editorial changes.

© European Molecular Biology Organization

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: Didier Trono
Journal Submitted to: EMBO reports
Manuscript Number: EMBOR-2018-47382

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 χ2 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?	n = 3 is the minimum nuber of independent replicates that can be used to caluculate statistics in the experiments that we performed in this manuscript and was sufficient to demonstrate our hypotheses
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
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.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
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.	No
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes. t-test
Is there an estimate of variation within each group of data?	Yes
Is the variance similar between the groups that are being statistically compared?	Yes

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).	NA
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.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
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.	NA

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'. 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	Depositing data on GEO
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).	Datasets are shared on GEO
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).	NA
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.	We described the computational methods in Material and Methods

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.	NA