THE EMBO JOURNAL

Manuscript EMBO-2013-86537

Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal

Brad Lackford, Chengguo Yao, Georgette M. Charles, Lingjie Weng, Xiaofeng Zheng, Eun-A Choi, Xiaohui Xie, Ji Wan, Yi Xing, Johannes M. Freudenberg, Pengyi Yang, Raja Jothi, Guang Hu & Yongsheng Shi

Corresponding author: Yongsheng Shi, University of California, Irvine

Review timeline:

Submission date:	14 August 2013
Editorial Decision:	26 September 2013
Revision received:	14 December 2013
Editorial Decision:	14 January 2014
Revision received:	30 January 2014
Accepted:	03 February 2014

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.)

Editor: Anne Nielsen

1st Editorial Decision 26 September 2013

Thank you for submitting your manuscript for consideration by The EMBO Journal and my apologies for the unusually long duration of the review period. Your study has now been seen by three referees whose comments are shown below.

As you will see from the reports, all referees express high interest in the findings reported in your manuscript; however they also raise a number of specific concerns that you will have to address in full before the referees can support publication of a revised manuscript. While we in principle agree with all points raised by the referees, I would suggest you to particularly focus your efforts on clarifying the following points:

-> Could the outcome of Fip1 depletion impair general polyadenylation rather than cause a specific effect on APA? (ref#1)

-> Is altered APA usage the cause or consequence of differentiation? (ref#2).

-> Conduct additional experiments to validate the quantification of APA usage given the discrepancy between direct RNA seq and Q-PCR data (refs #1 and #2)

-> Adjust for 3'UTR length difference in expression assays (ref #1 and #3)

-> Assay for expression of pluripotency markers to demonstrate differentiation stage (Refs #2 and #3) and include colony-forming assays for self-renewal.

-> Expand the discussion to place the effects for Fip1 depletion in the contexts of effects reported for other components of the CPSF machinery (refs #1 and #3)

Given the referees' overall positive recommendations, we offer you the opportunity to submit a revised version of the manuscript, addressing the comments of all three reviewers. I should add that it is EMBO Journal policy to allow only a single round of revision, and acceptance or rejection of your manuscript will therefore depend on the completeness of your responses to the full satisfaction of the referees in this revised version. Please do not hesitate to contact me if you have questions related to the review process and the requests made by the referees.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Peer-Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website: http://www.nature.com/emboj/about/process.html

We generally allow three months as standard revision time. As a matter of policy, competing manuscripts published during this period will not negatively impact on our assessment of the conceptual advance presented by your study. However, we request that you contact the editor as soon as possible upon publication of any related work, to discuss how to proceed.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

REFEREE REPORTS:

Referee #1:

Charles et al., knock-down the essential polyadenylation factor Fip1 in mouse ES cells and study the phenotypic changes with respect to self-renewal. They focus on correlating the observed phenotypic changes to alterations in the relative expression of alternative mRNA isoforms due to alternative polyadenylation (APA).

Overall, the study does not convincingly show that KD of Fip1 specifically targets genes involved in APA. Furthermore, the cause of the phenotype could as well be due to a general defect in polyadenylation.

Fip1 is a subunit of the essential polyadenylation factor Cpsf. Knock-down of Fip1 has consequences on many polyadenylation sites (PAS). In addition to affecting genes with several PAS, it also affects genes with one PAS (as shown in Fig. 5). Less recognition of a PAS because of a weaker polyadenylation machinery leads to read-through transcription. If these read-through transcripts are not being cleaved and polyadenylated at a downstream PAS, the result is a reduction in mRNA levels and thus in protein levels. If there is another functional PAS downstream, the

transcripts are being cleaved and polyadenylated at the downstream PAS (as it is the case for genes with two PAS). The mRNAs with extended 3'UTRs often generate less protein. Therefore, in both cases the functional consequence of a weaker PAS recognition is a reduction in protein levels no matter if genes with one or two PAS are investigated. Thus, the observed phenotype cannot be attributed to changes in APA. Many transcription factors that are essential for self-renewal have only one PAS (e.g. Oct-4). An alternative explanation of the phenotype could be a reduction in mRNA levels as a result of a weaker polyadenylation machinery. The authors have not ruled out that the observed phenotype is due to a general polyadenylation defect that affects genes with one PAS that are important for pluripotency or self-renewal.

Unless, the authors show that expression of the extended 3'UTR isoforms are necessary for impaired self-renewal and have a specific effect (and not just loss of protein expression), this paper cannot claim that APA plays a specific role during ES cell self-renewal. For example, it could be shown that there are genes whose extended 3'UTR isoforms produce more protein. If these genes are involved in differentiation and are shown to be necessary for the observed phenotype, it would demonstrate a more specific role of APA in this context.

Specific points:

1. Control experiments for genes with one PAS are missing. Genes with one PAS are as likely as genes with several PAS to be targets of Fip1. Therefore, four groups of genes should be examined: A: genes with one PAS whose mRNA levels decrease upon Fip1 KD (potential Fip1 targets), B: genes with one PAS whose mRNA levels do not decrease upon Fip1 KD, C: genes with several PAS that are Fip1 targets, D: genes with several PAS, that are not responsive to Fip1 KD. For the genes in group A reads downstream of the PAS (within 5 or 10 kb) should be found.

How many genes are in each group? Are genes involved in self-renewal overrepresented in group C or are they similarly present in other groups? What is the distribution of GO terms in the four groups? How does the CLIP signal look in the four groups? Can the four groups be distinguished with respect to PAS signals, motifs or expression level?

2. There are discrepancies between the data obtained by sequencing and by qRT-PCR. For example, when comparing Fig. 1B and Fig. S4, the % distal PAS usage for Ahctf1 is 60 and 90 in the sequencing data, but 20 and 60 in the qRT-PCR; for Ncaph2 the % distal PAS usage according to sequencing is 45 and 80, but measured by qRT-PCR it is 10 and 40. In Fig. S3, the % distal usage of Ncl is 10% after Fip1 KD, but in Fig. S5, it is 70-90%.
This demonstrates that one of their methods (qRT-PCR or sequencing) is not quantitative. The validation of the sequencing method is at best qualitative. The sentence (page 8): "These and other APA changes identified by DRS analysis were validated by RT-qPCR assays (Supplementary Figure 4), confirming the accuracy of our sequencing analysis" is not justified.

3. I am not convinced that the expression changes seen in genes with several PAS are due to the changes in mRNA ratios and not due to changes in mRNA abundance levels. Especially KD with siRNA-2 leads to reduction in mRNA levels between 0.6 and 0.8 (Fig. S6). This can explain downregulation at the protein level. Furthermore, downregulation of protein is shown for WWP2 (Fig. S7) but neither the qRT-PCR nor the luciferase assay (Fig. 3) that tests the contribution of the individual mRNA isoforms is shown for WWP2. Therefore, it is unclear if the protein changes are due to changes in mRNA ratios.

4. For the luciferase assay in Fig. 3A, the length of the 3'UTR is different. The difference in length needs to be adjusted for by transfecting higher amounts of the constructs (normalized to the same molar amounts of the plasmids used). The amount of luciferase construct transfected influences the amount of luciferase activity measured and could explain the lower amount of luciferase activity obtained with the extended UTR.

5. It is unclear how much sequence surrounding the PAS was included in the luciferase assay in Fig. 5C. The regulatory elements surrounding the proximal PAS might be missing and therefore the proximal PAS seems to be weaker.

6. Fig. 5D. It is unclear what the figure shows. Is this the CLIP profile for all genes or for Fip1 target genes or for Fip1 target genes with a canonical PAS? The CLIP profiles should be shown for all four groups (A-D).

7. The CLIP signal in Fig. 5E should be shown normalized to the number of genes in each group. Furthermore, the CLIP signal is detected across the whole 3'UTR (Fig. S12); therefore, the representation in Fig. 5E is misleading. The CLIP signal across the whole length of the UTR should be shown for all 4 groups.

8. -The figure legends should explain all the abbreviations used in the figures (Fig. S10, NE, Fig. 5C, SVL, wm).

-The axes should be labeled better.
-The legend for Fig. 1C is wrong, shown is not the fold change but the absolute expression, yellow corresponds to high expression.
-A scale should be included in the sequencing data to judge sequencing depth.
-Some of the supplementary figures are out of order in the text (Fig. S9, S10, S11).
-Statistics should be included on the sequencing data. How many reads were obtained per sample? How many are mapping to the genome, to the 3'UTR? How many genes are expressed? What are the cut-offs used for including a gene as expressed? What are the cut-offs used for calling a peak as an alternative isoform? How many genes have more than one PAS in the 3'UTR?
-The time points used in the study need to be spelled out in the figure legends and in the main text.
-Some of the supplementary figures (Fig. S4, S6, S7) are more important than some of the figures shown in the main text (Fig. 1C-E).

Referee #2:

Charles et al., examine the role of Fip1 in the control of alternative polyadenylation (APA) in mouse embryonic stem cells (ESCs). It has previously been reported that the length of many 3'UTRs is developmentally regulated with shorter 3'UTRs being predominant in ESCs compared with longer 3'UTRs observed for the corresponding mRNAs in differentiated cells. Although APA has been implicated in this dynamic 3'UTR regulation the mechanistic details have remained unknown. This manuscript very nicely addresses the role of Fip1 in the control of APA in ESCs. The main findings of this manuscript are that 1) Fip1 is highly expressed in undifferentiated ESCs relative to differentiated cells and Fip1 depletion leads to compromised ESC self-renewal 2) Direct RNA sequencing reveals that Fip1 knockdown in ESCs leads to changes in APA with a general shift in towards more distal PAS usage. 3) Reporter assays demonstrate that Fip1 can directly impact APA for selected 3'UTRs. 4) Fip1 is required for efficient reprogramming of MEFs to iPSCs 5) Fip1 depletion leads to compromised 3' processing activity in cells and in biochemical assays 6) Finally, a model is proposed whereby higher levels of Fip1/ promotes ESC self-renewal by stimulating the production of mRNAs with shorter 3' UTRs for regulated expression of important self-renewal genes. This is an interesting story with overall very good quality data. The manuscript is well written. If the specific points below are addressed then this manuscript is certainly suitable for publication in EMBO Journal.

Specific points:

1) The authors show that Fip1 knockdown leads to decreased expression of pluripotency genes and a corresponding increase in markers of differentiation (Figure 1D), this is consistent with a

requirement of Fip1 in ESC self-renewal. However the data presented in supplementary figure 2 indicate that some of the same markers of differentiation are expressed at a lower level in the Fip1 knockdown cells compared to control siRNA during ESC differentiation into embryoid bodies (EBs). These data seem contradictory to the model that loss of Fip1 leads to loss of self-renewal and precocious cell differentiation. Also, the most dramatic effects of Fip1-depletion on marker gene expression are observed at days 4, and 6 of EB formation (Figure S2) - a time-point where Fip1 expression is already downregulated (Figure 1E). It is not clear from the presented data how the authors can conclude that Fip1 is required for ESC pluripotency.

2) Figure 2. Since Fip1 knockdown leads to ESC differentiation (Figure 1) it is unclear whether the observed changes in alternative polyadenylation (APA) are directly related to the role of Fip1 in the PAS selection or may instead be indirect effects due to ESC differentiation. The authors attempt to address this by also measuring APA at an earlier time-point (48hrs) after Fip1 knockdown (Figure S5). However the authors should measure Fip1 knockdown and Oct4 expression (as well as a selection of differentiation genes) by western blot to confirm that Fip1 protein is effectively depleted and to help support the claim that the changes in APA are directly related to Fip1 function.

3) How well established is this direct RNA sequencing technology for monitoring alternative polyadenylation events? Perhaps the authors should use this method to monitor changes in PAS during ESC differentiation to a) confirm the changes that have been previously reported using alternative approaches and b) to examine the relative contribution of Fip1 in this reported phenomenon? Are the same distal sites utilized upon Fip1 knockdown that would normally be selected in differentiated cells?

4) What is the consequence for the PtoD switch upon Fip1 knockdown? The authors analyze expression of a small subset of proteins (Ncaph2, WWP2, Etf1) by Western blot (Figure S7). This is an important result that should be included in the main figures (maybe Figure 3).

5) The authors could further test their model that the abundance of Fip1 is an important determinant of APA by performing Fip1 overexpression assays in MEFs and examining the DtoP expected for the more sensitive 3'UTRs. An extension of this would be to test whether Fip1 expression can increase the efficiency of iPSC generation (in combination with other reprogramming factors).

Referee #3:

In the manuscript "Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal," Charles et al expand on how Fip1 and APA regulate ESC differentiation. Fip1 had been previously identified in screens as important in maintaining pluripotency. Fip1 has also been shown to be a component of the CPSF complex that direct cleavage/polyadenylation of transcripts. Furthermore, levels of the CPSF complex have been shown to regulate APA by initially binding strong canonical sites, but expanding to weaker non-canonical sites when present at higher levels leading to overall shorter 3'UTRs. This previous work has been appropriately referenced in the manuscript. They advance this knowledge by further characterizing the effect of Fip1 knockdown on ESC self-renewal and the induction of pluripotency. They detail the effects on targets by combination of direct RNA sequencing using Helicos and iCLIP sequencing. Bioinformatically, they make the interesting observation that distance between potential cleavage sites significantly impacts the consequence of differential Fip1 levels. Overall, this is a well-written paper that presents a number of interesting concepts. Given additional data supporting several of their conclusions and other small revisions, I would highly recommend publication in EMBO.

General comments.

1) The paper would be strengthened by a more detailed comparison of their data to that of the Jenal, Martin, and Yao papers which similarly evaluate how components of the CPSF complex influence APA.

2) Westerns throughout should be quantified, including a median, SD, number of repeats, and significance of difference.

3) Note there are some mis-referencing of supplemental figures in text.

4) Add significance values to all figures where difference is claimed.

Specific comments.

1) Figure 1: 1b) should show FACs plots including definitions of GFP- versus GFP+. 1c) While Fip1 has an obvious morphological effects in these images, statement of decreased AP is not clear and may reflect density of colonies in control rather than AP/cell. 1d: based on this data which is done in self-renewal conditions along with EB data in sup fig. 2, it does not appear that Fip1 knockdown is inducing differentiation in any sort of normal sense. Instead it appears to be leading to abnormal cellular phenotypes. Therefore, authors should use concept of induction of differentiation more sparingly. Instead, they can say FIP1 is required for normal ESC morphology and gene expression. To show diminished self-renewal, they need to perform colony forming assays. Also, they need to add additional pluripotency markers such as Nanog, Sox2, Klf4, Rex1, Sall4, etc to support disruption of the pluripotency network.

2) Figure 2: 2c) analysis is poorly described. How are genes normalized? Argument that FIP1 targets are going down is not convincing. Need some sort of analysis that supports significance. 2d) not sure how this figure supports argument that FIP1 regulates self-renewal/pluripotency. Its most dramatic effect seems to be proliferation. 2e) like 2c, poor description making figure impossible to interpret. Are these a mix of genes up and down in comparisons of Neuron to ESC and Flp1KD to control? Separate out into genes up and down.

3) Figure 3: 3a) in what cellular context were these experiments performed. Authors should control for length effects (i.e. independent of particular 3'UTRs chosen). That is would random sequences of equal length similarly reduce relative expression of luciferase. 3b) Evaluation of effect of kd of Fip1 targets is much too limited. All that is shown is morphology. At minimum should do PCR for representative pluripotency markers.

4) Fig. 4: This data does not support that FIP1 is required for reprogramming directly. The findings could be and likely are secondary to the differential effect of FIP1 on ESC and MEF proliferation (fig. 4c). The reduced AP intensity and definitely the number of flow positive Oct4-GFP could be ascribed to decreased proliferation of late stage reprogramming cells. Therefore, should be cautious of over-interpretation of data. Why did the authors not count Oct4-GFP colonies rather than flow? A&B) show levels in iPSCs as comparison. Would also be important to show that Flp1 kd influences shift toward proimal usage in a subset of targets during reprogramming (flip side of 4b).

5) Authors should be more consistent about choice of targets that they follow in 3a, 4b, , sup 5c, and sup 7. Otherwise, it appears that authors are only choosing those targets that support their argument for each independent experiment.

6) Fig. 5: Interesting findings in this figure. How do the proximal PASs differ in targets versus non-targets. Also, what is the authors' source of 3'UTR annotation including the defining of proximal and distal PASs.

We have carried out new experiments and analyses and revised the text to address all the questions raised, and we believe the manuscript has been substantially improved. Attached is our point-to-point response to the reviewers' comments. We look forward to hearing from you soon on your decision on our manuscript. Thank you.

-> Could the outcome of Fip1 depletion impair general polyadenylation rather than cause a specific effect on APA? (ref#1)
-> Is altered APA usage the cause or consequence of differentiation? (ref#2).
-> Conduct additional experiments to validate the quantification of APA usage given the discrepancy between direct RNA seq and Q-PCR data (refs #1 and #2)
-> Adjust for 3'UTR length difference in expression assays (ref #1 and #3)
-> Assay for expression of pluripotency markers to demonstrate differentiation stage (Refs #2 and #3) and include colony-forming assays for self-renewal.
-> Expand the discussion to place the effects for Fip1 depletion in the contexts of effects reported for other components of the CPSF machinery (refs #1 and #3)

<u>Referee #1 (Remarks to the Author):</u>
Charles et al., knock-down the essential polyadenylation factor Fip1 in mouse ES cells and study the phenotypic changes with respect to self-renewal. They focus on correlating the observed phenotypic changes to alterations in the relative expression of alternative mRNA isoforms due to alternative polyadenylation (APA).
Overall, the study does not convincingly show that KD of Fip1 specifically targets genes involved in APA. Furthermore, the cause of the phenotype could as well be due to a general defect in polyadenylation.
Fip1 is a subunit of the essential polyadenylation factor Cpsf. Knock-down of Fip1 has consequences on many polyadenylation sites (PAS). In addition to affecting genes with several PAS, it also affects genes with one PAS (as shown in Fig. 5). Less recognition of a PAS because of a weaker polyadenylation machinery leads to read-through transcription. If these read-through transcripts are not being

cleaved and polyadenylated at a downstream PAS, the result is a reduction in mRNA levels and thus in protein levels. If there is another functional PAS downstream, the transcripts are being cleaved and polyadenylated at the downstream PAS (as it is the case for genes with two PAS). The mRNAs with extended 3'UTRs often generate less protein. Therefore, in both cases the functional consequence of a weaker PAS recognition is a reduction in protein levels no matter if genes with one or two PAS are investigated. Thus, the observed phenotype cannot be attributed to changes in APA. Many transcription factors that are essential for self-renewal have only one PAS (e.g. Oct-4). An alternative explanation of the phenotype could be a reduction in mRNA levels as a result of a weaker polyadenylation machinery. The authors have not ruled out that the observed phenotype is due to a general polyadenylation defect that affects genes with one PAS that are important for pluripotency or self-renewal.

Unless, the authors show that expression of the extended 3'UTR isoforms are necessary for impaired self-renewal and have a specific effect (and not just loss of protein expression), this paper cannot claim that APA plays a specific role during ES cell self-renewal. For example, it could be shown that there are genes whose extended 3'UTR isoforms produce more protein. If these genes are involved in differentiation and are shown to be necessary for the observed phenotype, it would demonstrate a more specific role of APA in this context.

Response:

We thank reviewer-1 for the constructive comments and suggestions.

We have carried out additional experiments and analyses to address whether the effect of Fip1 knockdown (KD) on ESC self-renewal was due to a general polyadenylation defect. **First**, we have compared the gene expression profiles in control and Fip1 KD ESCs based on microarray analyses (with three biological replicates). We found that less than 1.5% of the single-poly(A) site (PAS) genes showed a significant decrease in mRNA levels (by 50% or more, *FDR* < 0.05) upon Fip1 KD (please see response to specific point 1 for more details). Thus the expression of the vast majority of single-PAS genes is unaltered in Fip1-depleted cells, which strongly argues against a general polyadenylation defect. **Second**, we have carried out gene ontology analyses of the single-PAS genes that showed significant decrease in their mRNA levels and did not find any enriched functional categories. **Third**, if the loss of self-renewal in Fip1-depleted

ESCs was due to a general polyadenylation defect, one would predict that knockdown of other essential mRNA 3' processing factors should have similar effects. To test this, we have depleted two other essential mRNA 3' processing factors, CPSF30 (another subunit of the CPSF complex) and CF Im25 (a subunit of the CF Im complex). Interestingly, depletion of these factors did not result in similar APA changes in Fip1 target mRNAs tested (Supplementary Figure S28). More importantly, no apparent loss of self-renewal was observed in these cells based on cell morphology (Supplementary Figure S27) and RT-qPCR analyses of ESC marker genes (Supplementary Figure S28). Together these new results provided compelling evidence that loss of ESC self-renewal in ESCs upon Fip1 depletion is unlikely to be caused by a general polyadenylation defect.

Our main conclusion is that Fip1-mediated APA regulation promotes stem cell self-renewal. This conclusion is well supported by our observations that Fip1 regulates the APA of many critical self-renewal factors and helps to maintain their optimal expression in ESCs and iPSCs. We did not, nor did we intend to, rule out the possibility that Fip1 may also contribute to self-renewal by regulating the mRNA levels of certain genes. In fact, these two mechanisms are not mutually exclusive. This is analogous to the role of APA in cancer development: the observation that APA changes contribute to proto-oncogene activation does not rule out the possibility that mRNA levels also play an important role in this process (Mayr & Bartel, 2009). Recognizing other possible mechanisms, we concluded that "Fip1 promotes ESC self-renewal and reprogramming, **in part**, by maintaining the ESC-specific APA profiles…" (Abstract).

Specific points:

1. Control experiments for genes with one PAS are missing. Genes with one PAS are as likely as genes with several PAS to be targets of Fip1. Therefore, four groups of genes should be examined: A: genes with one PAS whose mRNA levels decrease upon Fip1 KD (potential Fip1 targets), B: genes with one PAS whose mRNA levels do not decrease upon Fip1 KD, C: genes with several PAS that are Fip1 targets, D: genes with several PAS, that are not responsive to Fip1 KD. For the genes in group A reads downstream of the PAS (within 5 or 10 kb) should be found.

How many genes are in each group? Are genes involved in self-renewal overrepresented in group C or are they similarly present in other groups? What is

the distribution of GO terms in the four groups? How does the CLIP signal look in the four groups? Can the four groups be distinguished with respect to PAS signals, motifs or expression level?

Response:

For classifying single- and multiple-PAS genes, we first used all direct RNA sequencing (DRS) peaks in the ESC sample without any read count filter. We have divided all genes into four groups according to the classification system suggested by the reviewer. Group A: single-PAS genes whose expression decreased significantly in Fip1 KD cells (by at least 50%, FDR < 0.05); Group B: single-PAS genes without significant decrease; Group C: multi-PAS genes with significant APA changes; Group D: multi-PAS genes without significant APA changes. Based on microarray data and our APA analyses, the numbers of genes in each group are: (A) 9 genes; (B) 1,559 genes; (C) 374 genes; (D) 9577 genes. Therefore, ~0.6% of genes with a single PAS showed significant decrease in their mRNA levels while ~3.8% of multi-PAS genes displayed significant APA changes upon Fip1 depletion. For the 9 genes in group A, we have searched in Fip1 KD DRS data for new peaks in the downstream 10kb region or until the next annotated gene. Within this region, we detected 13 new DRS peaks in total for 5 genes, and 12 peaks have only one read and one peak has 2 reads.

If we only use high-confidence DRS peaks (read count >= 10) for this analysis, the numbers of genes in each group are: (A) 77 genes; (B) 5,870 genes; (C) 374 genes; (D) 2,529 genes. Under this condition, ~1.3% of single-PAS genes showed significant decrease in mRNA levels while ~13% of multi-PAS genes displayed significant APA changes. When we looked for novel DRS peaks downstream of the 77 group A genes in Fip1 KD data, we found 130 peaks, 90% of which have only one read and the highest peak has 3 reads. Thus, using two different read count cut-offs, we observed that polyadenylation still occurs at the same sites for group A genes and we did not detect significant usage of novel downstream PASs. This is consistent with a recent study showing that mammalian genes use the same PASs across different tissues and under different conditions (Lianoglou et al, 2013). Together, our analyses strongly suggest that Fip1 KD in ESCs does not lead to a general polyadenylation defect.

As suggested by the reviewer, we have carried out gene ontology (GO) analyses of the group A-D genes, but we failed to detect any enriched functional categories in group A genes either with or without the read count cut-off (group A-B in Supplementary Figure

S25, group C in Fig. 2D).

We have also compared the iCLIP map of group A-D. Following the reviewer's suggestions, we have modified our method for normalizing iCLIP results: iCLIP signals are first normalized by transcript number (based on DRS read counts) and then summation of the iCLIP signals within each group are normalized by the number of genes in each group (details described in Methods). Using this method, we found that the mRNAs of group A genes had significantly lower Fip1 iCLIP signals near their PASs or in their 3' UTRs ($p=2.4 \times 10^{-62}$). In addition, the PASs of group A genes contain less canonical PAS features, such as the AAUAAA hexamer and the downstream UG-rich sequence (Supplementary Figure S25B), but there is no significant difference in the expression levels between group A and B genes. For group C and D genes, there are significantly more Fip1 iCLIP signals in the 3' UTRs of group C mRNAs than the group D mRNAs ($p = 1.1 \times 10^{-18}$, Figure 5E). Motif analyses detected enrichments of canonical PAS features such as AAUAAA and UG-rich sequences in the proximal PASs of group C genes compared to those of group D genes (Supplementary Figure S22). Our analyses showed that group C genes are expressed at higher levels in ESCs (Figure 2C), consistent with the notion that Fip1 KD preferentially regulates APA on ESC genes. Together, these data suggest that Fip1-RNA interactions and PAS sequences play a role in differentiating Fip1 targets from non-targets. We thank the reviewer for these suggestions.

2. There are discrepancies between the data obtained by sequencing and by qRT-PCR. For example, when comparing Fig. 1B and Fig. S4, the % distal PAS usage for Ahctf1 is 60 and 90 in the sequencing data, but 20 and 60 in the qRT-PCR; for Ncaph2 the % distal PAS usage according to sequencing is 45 and 80, but measured by qRT-PCR it is 10 and 40. In Fig. S3, the % distal usage of Ncl is 10% after Fip1 KD, but in Fig. S5, it is 70-90%.

This demonstrates that one of their methods (qRT-PCR or sequencing) is not quantitative. The validation of the sequencing method is at best qualitative. The sentence (page 8): "These and other APA changes identified by DRS analysis were validated by RT-qPCR assays (Supplementary Figure 4), confirming the accuracy of our sequencing analysis" is not justified.

Response:

We have carried out additional experiments/analyses to address this issue. **First**, we

have compared the gene expression profiling results of control ESCs and Fip1 KD cells based on microarray analyses and those based on DRS analyses. The results from these two independent methods showed excellent agreement (r^2=0.94) (Supplementary Figure S6), suggesting that our DRS analyses were highly quantitative. **Second**, we calculated the aUTR usage fold changes between Fip1 KD and control ESC samples based on RT-qPCR data and compared them with the results based on our DRS analyses for all 22 Fip1 targets tested, and we observed a high correlation (r^2=0.79, Supplementary Figure S8B-C). Therefore, the RT-qPCR and DRS analyses are consistent in relative quantification between different samples. However, although we always carefully test all the primers for our RT-qPCR analyses to ensure near perfect amplification efficiency (>0.95), we agree with the reviewer that this method may have technical limitations in the absolute quantification of the percentage of APA isoforms. Together, we conclude that our DRS analyses are highly quantitative and that RT-qPCR is a reliable method for relative quantification of APA changes between different samples. In a separate note, we would like to respectfully point out that the reviewer's estimation of aUTR usage based on the DRS tracks may not be accurate, as some of the genes contain lower DRS peaks that are not clearly visible in those figures.

3. I am not convinced that the expression changes seen in genes with several PAS are due to the changes in mRNA ratios and not due to changes in mRNA abundance levels. Especially KD with siRNA-2 leads to reduction in mRNA levels between 0.6 and 0.8 (Fig. S6). This can explain downregulation at the protein level. Furthermore, downregulation of protein is shown for WWP2 (Fig. S7) but neither the qRT-PCR nor the luciferase assay (Fig. 3) that tests the contribution of the individual mRNA isoforms is shown for WWP2. Therefore, it is unclear if the protein changes are due to changes in mRNA ratios.

Response:

We have now included the RT-qPCR (Supplementary Figure S8B) and luciferase assay results (Fig. 3A) for Wwp2. Consistent with other Fip1 targets, the extended 3' UTR of Wwp2 also suppresses reporter gene expression. We have also quantified the Western blotting analyses results for all three targets. Our data showed that Fip1 KD only led to subtle changes in the mRNA levels of Fip1 targets (20% decrease on average and statistically insignificant, Supplementary Figure 12A), but caused much greater and statistically significant changes in their protein levels (60% decrease on average, Fig. 3B

and Supplementary Figure 12B). Therefore, the reduction in Fip1 target protein levels cannot be solely attributed to changes in their mRNA levels. Together with our 3' UTR reporter assay results showing that aUTRs suppress gene expression, we conclude that Fip1-mediated APA regulation plays an important role in the regulation of the expression of its target genes.

4. For the luciferase assay in Fig. 3A, the length of the 3'UTR is different. The difference in length needs to be adjusted for by transfecting higher amounts of the constructs (normalized to the same molar amounts of the plasmids used). The amount of luciferase construct transfected influences the amount of luciferase activity measured and could explain the lower amount of luciferase activity obtained with the extended UTR.

Response:

Following the reviewer's suggestion, we have repeated the luciferase assays in Figure 3A by carefully transfecting the same molar amount of DNA for the cUTR and c+aUTR reporter constructs. The new results are similar to our previous data and consistent with our conclusion that "Fip1 depletion-induced 3' UTR extension silence gene expression".

5. It is unclear how much sequence surrounding the PAS was included in the luciferase assay in Fig. 5C. The regulatory elements surrounding the proximal PAS might be missing and therefore the proximal PAS seems to be weaker.

Response:

For all analyses of PAS strength by in vitro assays or in vivo dual luciferase assays, the region from -100 nt to +100 nt relative to the cleavage sites were used. This region was chosen for the following reasons. **First**, all the known key cis-elements for PAS are found within this region, including the AAUAAA hexamer (-35 to -10nt), the U/GU-rich downstream element (+10 to +40nt), and several other auxiliary sequence elements (Colgan & Manley, 1997; Zhao et al, 1999). **Second**, as shown in Supplementary Figure 20A, the nucleotide composition around the cleavage sites at both proximal and distal PASs shows that the core sequence elements, including the U-rich and A-rich upstream elements and the U/UG-rich downstream elements are found with this the -100nt to +100nt region. Outside this region, the nucleotide distribution appears random (~25% for each nucleotide). **Third**, we have analyzed the sequence conservation near PASs. As

shown in Supplementary Figure 20B, for both proximal and distal PAS, this 200 nt region is the most conserved region and thus likely contain all the core sequence elements. **Fourth**, our Fip1 iCLIP analyses and previous CLIP-seq analyses of CstF64 and CF Im all showed that these core 3' processing factors bind to PASs within this region (Martin et al, 2012; Yao et al, 2012). **Finally**, since the same 200nt region was used for both proximal and distal PAS, our results suggest that the core sequences of the proximal and distal PASs have significant differences in cleavage/polyadenylation efficiency. We agree with the reviewer that it is possible that both proximal and distal PAS can be influenced by sequences located outside this region. We have clearly noted the region analyzed in the manuscript and concluded that "…they contain more canonical sequence features in their **core** region and thus have higher affinity for the 3' processing machinery …" (page 15).

6. Fig. 5D. It is unclear what the figure shows. Is this the CLIP profile for all genes or for Fip1 target genes or for Fip1 target genes with a canonical PAS? The CLIP profiles should be shown for all four groups (A-D).

Response:

Fig. 5D shows the iCLIP signal distribution for all genes with A(A/U)UAAA signal. The iCLIP profiles for all groups are included (groups A and B: in Supplementary Figure 26A; groups C and D: Fig. 5E)

7. The CLIP signal in Fig. 5E should be shown normalized to the number of genes in each group. Furthermore, the CLIP signal is detected across the whole 3'UTR (Fig. S12); therefore, the representation in Fig. 5E is misleading. The CLIP signal across the whole length of the UTR should be shown for all 4 groups.

Response:

Following the reviewer's suggestion, we have re-generated iCLIP maps by normalizing the iCLIP signals by transcript number for each isoform and the gene numbers in each group. Also the new iCLIP maps show the entire 3' UTRs as suggested by the reviewer.

8. -The figure legends should explain all the abbreviations used in the figures (Fig. S10, NE, Fig. 5C, SVL, wm).

Response:

The explanations for all the abbreviations have been added to the figure legends.

9. -The axes should be labeled better.

Response:

More detailed description of the axes has been added.

10. -The legend for Fig. 2C is wrong, shown is not the fold change but the absolute expression, yellow corresponds to high expression.

Response:

The label has been corrected.

-A scale should be included in the sequencing data to judge sequencing depth.

Response:

Scales have been added.

-Some of the supplementary figures are out of order in the text (Fig. S9, S10, S11).

Response:

These have been corrected.

-Statistics should be included on the sequencing data. How many reads were obtained per sample? How many are mapping to the genome, to the 3'UTR? How many genes are expressed? What are the cut-offs used for including a gene as expressed? What are the cut-offs used for calling a peak as an alternative isoform? How many genes have more than one PAS in the 3'UTR?

Response:

The statistics have been added (Supplementary Material, Direct RNA Sequencing section).

-The time points used in the study need to be spelled out in the figure legends and in the main text.

Response:

All time points have been specified.

-Some of the supplementary figures (Fig. S4, S6, S7) are more important than some of the figures shown in the main text (Fig. 1C-E).

Response:

We thank the reviewer for the suggestion. The original Supplementary Fig. S7 have been moved to the main figures (Figure 3B) while Fig. S4 and S6 are kept in Supplementary data due to space constraint.

Referee #2 (Remarks to the Author):

Charles et al., examine the role of Fip1 in the control of alternative polyadenylation (APA) in mouse embryonic stem cells (ESCs). It has previously been reported that the length of many 3'UTRs is developmentally regulated with shorter 3'UTRs being predominant in ESCs compared with longer 3'UTRs observed for the corresponding mRNAs in differentiated cells. Although APA has been implicated in this dynamic 3'UTR regulation the mechanistic details have remained unknown. This manuscript very nicely addresses the role of Fip1 in the control of APA in ESCs. The main findings of this manuscript are that 1) Fip1 is highly expressed in undifferentiated ESCs relative to differentiated cells and Fip1 depletion leads to compromised ESC self-renewal 2) Direct RNA sequencing reveals that Fip1 knockdown in ESCs leads to changes in APA with a general shift in towards more distal PAS usage. 3) Reporter assays demonstrate that Fip1 can directly impact APA for selected 3'UTRs. 4) Fip1 is required for efficient reprogramming of MEFs to iPSCs 5) Fip1 depletion leads to compromised 3' processing activity in cells and in biochemical assays 6) Finally, a model is proposed whereby higher levels of Fip1/ promotes ESC self-renewal by stimulating the production of mRNAs with shorter 3' UTRs for regulated expression of important self-renewal genes. This is an interesting story with overall very good quality data. The manuscript is well written. If the specific points below are addressed then this manuscript is certainly suitable for publication in EMBO Journal.

We thank reviewer-2 for the enthusiasm and positive comments.

Specific points:

1) The authors show that Fip1 knockdown leads to decreased expression of pluripotency genes and a corresponding increase in markers of differentiation (Figure 1D), this is consistent with a requirement of Fip1 in ESC self-renewal.

However the data presented in supplementary figure 2 indicate that some of the same markers of differentiation are expressed at a lower level in the Fip1 knockdown cells compared to control siRNA during ESC differentiation into embryoid bodies (EBs). These data seem contradictory to the model that loss of Fip1 leads to loss of self-renewal and precocious cell differentiation. Also, the most dramatic effects of Fip1-depletion on marker gene expression are observed at days 4, and 6 of EB formation (Figure S2) - a time-point where Fip1 expression is already downregulated (Figure 1E). It is not clear from the presented data how the authors can conclude that Fip1 is required for ESC pluripotency.

Response:

We thank reviewer-2 for raising this point. We have carried out RT-qPCRs on additional markers. Our data showed that Fip1 KD initially led to abnormal differentiation kinetics and propensity, as indicated by the enhanced up-regulation of Sox1, Sox17, Gata6, Kdr and reduced up-regulation of Fgf5 on day-2 (Supplementary Figure S4). At later time points, Fip1 KD led to reduced differentiation capacity, as indicated by the reduced up-regulation of all the markers tested on day-4 and day-6 (Supplementary Figure S4). These results suggest that Fip1 KD impaired the normal course of differentiation of ESCs. Thus, we concluded that Fip1 plays a role in pluripotency. We would also like to point out that Figure 1E shows changes in Fip1 levels after LIF withdrawal, and its kinetics may not be directly comparable to that during EB formation.

2) Figure 2. Since Fip1 knockdown leads to ESC differentiation (Figure 1) it is unclear whether the observed changes in alternative polyadenylation (APA) are directly related to the role of Fip1 in the PAS selection or may instead be indirect effects due to ESC differentiation. The authors attempt to address this by also measuring APA at an earlier time-point (48hrs) after Fip1 knockdown (Figure S5). However the authors should measure Fip1 knockdown and Oct4 expression (as well as a selection of differentiation genes) by western blot to confirm that Fip1 protein is effectively depleted and to help support the claim that the changes in APA are directly related to Fip1 function.

Response:

We have addressed this point in two ways. **First**, as suggested by the reviewer, we have included the Fip1 and Oct4 western blots at 24 and 48 hrs after Fip1 siRNA transfection (Supplementary Figure S9). We have also added detailed analysis of cell morphology,

lineage marker expression, and APA changes at these time points (Supplementary Figure S9 and S10). Our data showed that the decrease in Fip1 protein levels and the associated APA changes were detected as early as 24 hrs after siRNA transfection (Supplementary Figure S9 and S10A). On the contrary, no obvious signs of differentiation were detected at this time point based on cell morphology and the overall lineage marker expression (Supplementary Figure S9 and S10B). Although Fip1 KD caused a subtle reduction in Oct4 mRNA levels at these time points (Supplementary Figure S10B), the Oct4 protein levels were less affected (Supplementary Figure S9). Additionally it has been shown that small decrease in Oct4 level by itself is not sufficient to cause ESC differentiation and may even be beneficial for self-renewal and pluripotency (Karwacki-Neisius et al, 2013). These results suggest that APA changes take place before ESC maintenance defects can be detected following Fip1 KD. **Second**, we have carried out DRS analysis of HeLa cells in which Fip1 have been depleted by RNAi and compared the APA profile to that of the control HeLa cells. As shown in Supplementary Figure S29C, Fip1 KD in HeLa cells led to significant APA changes in the mRNAs of 241 genes, and in the majority (71%) of these genes the APA changes were PtoD. Thus Fip1 KD caused predominantly PtoD APA shifts in HeLa cells, highly similar to what was observed in ESCs (Fig. 2A). However, Fip1 KD had no effect on the proliferation of HeLa cells (Supplementary Figure S29B). Together these data provided strong evidence that APA changes in Fip1 KD cells are directly related to Fip1 function in PAS selection and are not indirect effect of differentiation.

3) How well established is this direct RNA sequencing technology for monitoring alternative polyadenylation events? Perhaps the authors should use this method to monitor changes in PAS during ESC differentiation to a) confirm the changes that have been previously reported using alternative approaches and b) to examine the relative contribution of Fip1 in this reported phenomenon? Are the same distal sites utilized upon Fip1 knockdown that would normally be selected in differentiated cells?

Response:

Direct RNA sequencing using the Helicos platform is a well-established and highly quantitative method for both digital gene expression profiling and alternative polyadenylation analyses. For examples, DRS has been used in the following publications for APA analyses:

1. Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, John B, Milos PM. (2010) Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. ***Cell*** 143(6):1018-29.

2. Lin Y, Li Z, Ozsolak F, Kim SW, Arango-Argoty G, Liu TT, Tenenbaum SA, Bailey T, Monaghan AP, Milos PM, John B. (2012) An in-depth map of polyadenylation sites in cancer. ***Nucleic Acids Res.*** 40(17):8460-71.

3. Sherstnev A, Duc C, Cole C, Zacharaki V, Hornyik C, Ozsolak F, Milos PM, Barton GJ, Simpson GG. (2012) Direct sequencing of Arabidopsis thaliana RNA reveals patterns of cleavage and polyadenylation. ***Nat Struct Mol Biol.*** 19(8):845-52.

4. Dittmar KA, Jiang P, Park JW, Amirikian K, Wan J, Shen S, Xing Y, Carstens RP. (2012) Genome-wide determination of a broad ESRP-regulated posttranscriptional network by high-throughput sequencing. ***Mol Cell Biol.*** 32(8):1468-82.

5. Ji X, Wan J, Vishnu M, Xing Y, Liebhaber SA. αCP Poly(C) binding proteins act as global regulators of alternative polyadenylation. ***Mol Cell Biol.*** 2013 Jul;33(13):2560-73.

6. Yao C, Biesinger J, Wan J, Weng L, Xing Y, Xie X, Shi Y. Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. ***Proc Natl Acad Sci U S A.*** 2012 Nov 13;109(46):18773-8.

7. Yao C, Choi EA, Weng L, Xie X, Wan J, Xing Y, Moresco JJ, Tu PG, Yates JR 3rd, Shi Y. (2013) Overlapping and distinct functions of CstF64 and CstF64τ in mammalian mRNA 3' processing. ***RNA*** 19: 1781-90.

8. Moqtaderi Z, Geisberg JV, Jin Y, Fan X, Struhl K. (2013) Species-specific factors mediate extensive heterogeneity of mRNA 3' ends in yeasts. ***Proc Natl Acad Sci U S A.*** 110(27): 11073-8.

To test the quantitative accuracy of DRS analyses on our samples, we have compared gene expression profiling results of control ESCs and Fip1 KD cells using DRS analyses and another well-established method, microarray. When the results from these two methods were compared, the data showed excellent consistency (r^2=0.94) (Supplementary Figure S6), strongly suggesting that our DRS analyses were highly quantitative.

Previously we have characterized the APA changes during ESC differentiation into neurons by high throughput-sequencing (Shepard et al, 2011). To address the possible role of Fip1-mediated APA regulation in differentiation, we have compared the Fip1 KD-induced APA changes detected in current study with the APA changes that occur during

neuronal differentiation identified in our previous study (data reanalyzed using the same statistical criteria), and made several interesting observations. First, PtoD shifts are the predominant type of APA change in Fip1-depleted ESCs and during neuronal differentiation: 88% and 83% of APA changes are PtoD for neuronal differentiation and for Fip1 KD respectively. Second, ~25% of the PtoD APA changes induced by Fip1 depletion in ESCs were observed during ESC differentiation into neurons ($p<10^{-40}$) (Figure 2E) (Shepard et al, 2011). In 86% of these genes, polyadenylation shifted to the same distal PASs during neuronal differentiation and in Fip1-depleted ESCs. Finally we have performed RT-qPCR analyses of the APA profiles of eight Fip1 targets during retinoic acid-induced neural differentiation. Our results showed that the mRNAs of four of these genes (50%) showed PtoD APA changes, similar to those induced by Fip1 KD in ESCs (Supplementary Figure S12). Together, the large overlap between Fip1-regulated and differentiation-regulated APA events and the differentiation-associated Fip1 protein level change strongly suggest that Fip1 plays a role in regulating APA during ESC differentiation. These points have been added to the manuscript.

4) What is the consequence for the PtoD switch upon Fip1 knockdown? The authors analyze expression of a small subset of proteins (Ncaph2, WWP2, Etf1) by Western blot (Figure S7). This is an important result that should be included in the main figures (maybe Figure 3).

Response: We thank the reviewer for the suggestion. We have moved the data to Fig. 3B in our revised manuscript and added quantification of the Western analyses in Supplementary Figure S14.

5) The authors could further test their model that the abundance of Fip1 is an important determinant of APA by performing Fip1 overexpression assays in MEFs and examining the DtoP expected for the more sensitive 3'UTRs. An extension of this would be to test whether Fip1 expression can increase the efficiency of iPSC generation (in combination with other reprogramming factors).

Response:

We thank reviewer-2 for this insightful comment and suggestion. We have tried many times to overexpress Fip1 in MEFs with different approaches, including plasmid transfections and lentiviral transductions. Unfortunately, we have not been able to overexpress Fip1 to a significant level, and the exogenous Fip1 we were able to

introduce only represented a very small fraction of the endogenous Fip1. Thus, we are not able to carry out the experiment as suggested. There are at least two possible explanations for this. First, negative auto-regulation has been reported for a number of mRNA processing factors. Excess mRNA processing factors, such as U1A and SR proteins, modulate the processing and/or stability of their own mRNAs, thereby repressing their own expression and maintaining homeostasis (Boelens et al, 1993; Lareau et al, 2007). Similar negative feedback mechanisms may exist for Fip1 and prevent Fip1 overexpression. Alternatively, as Fip1 is a known intrinsically disordered/unstructured protein (Meinke et al, 2008), Fip1 may require other CPSF subunits for correct folding and/or stability. Like other disordered proteins (Fink, 2005), Fip1, when present at excess levels without sufficient amounts of other CPSF subunits to form complexes, may be unstable and thus quickly degraded. We are currently investigating the exact mechanisms.

Referee #3 (Remarks to the Author):

In the manuscript "Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal," Charles et al expand on how Fip1 and APA regulate ESC differentiation. Fip1 had been previously identified in screens as important in maintaining pluripotency. Fip1 has also been shown to be a component of the CPSF complex that direct cleavage/polyadenylation of transcripts. Furthermore, levels of the CPSF complex have been shown to regulate APA by initially binding strong canonical sites, but expanding to weaker non-canonical sites when present at higher levels leading to overall shorter 3'UTRs. This previous work has been appropriately referenced in the manuscript. They advance this knowledge by further characterizing the effect of Fip1 knockdown on ESC self-renewal and the induction of pluripotency. They detail the effects on targets by combination of direct RNA sequencing using Helicos and iCIIP sequencing. Bioinformatically, they make the interesting observation that distance between potential cleavage sites significantly impacts the consequence of differential Fip1 levels. Overall, this is a well-written paper that presents a number of interesting concepts. Given additional data supporting several of their conclusions and other small revisions, I would highly recommend publication in EMBO.

We thank reviewer-3 for the enthusiasm and positive comments.

General comments.

1) The paper would be strengthened by a more detailed comparison of their data to that of the Jenal, Martin, and Yao papers which similarly evaluate how components of the CPSF complex influence APA.

Response:

We thank the reviewer for this suggestion. We have carried out additional analyses to compare the target specificity and directionality of APA regulation by these factors and added the results to the Supplementary Figure S30 and the Discussion.

2) Westerns throughout should be quantified, including a median, SD, number of repeats, and significance of difference.

Response:

Quantification for all western analyses has been added. Median, SD, and number of repeats, and significance of difference are included in the figures and/or figure legends.

3) Note there are some mis-referencing of Supplementary figures in text.

Response: These errors have been corrected.

4) Add significance values to all figures where difference is claimed.

Response: Significance values are included in figures and/or figure legends.

Specific comments.

1) Figure 1: 1b) should show FACs plots including definitions of GFP- versus GFP+. 1c) While Fip1 has an obvious morphological effects in these images, statement of decreased AP is not clear and may reflect density of colonies in control rather than AP/cell. 1d: based on this data which is done in self-renewal conditions along with EB data in sup fig. 2, it does not appear that Fip1 knockdown is inducing differentiation in any sort of normal sense. Instead it appears to be leading to abnormal cellular phenotypes. Therefore, authors should use concept of induction of differentiation more sparingly. Instead, they can say

FIP1 is required for normal ESC morphology and gene expression. To show diminished self-renewal, they need to perform colony forming assays. Also, they need to add additional pluripotency markers such as Nanog, Sox2, Klf4, Rex1, Sall4, etc to support disruption of the pluripotency network.

Response:

Figure 1B) As suggested by the reviewer, we have added FACS plots (Supplementary Figure S2) and added definition of GFP- in the figure legend of Supplementary Figure S2.

Figure 1C) The reduction in AP staining is consistent with the results from the Oct4GiP reporter assay (Figure 1B), the morphology changes (Figure 1C and Supplementary Figure S3A), and the expression analysis of lineage markers (Figure 1D). There were comparable numbers of colonies in the control and Fip1 KD images in Fig. 1C. But as Fip1 KD led to loss of ESC identity, Fip1 KD cells were more flat and appeared to be more confluent. The actual number of cells in Fip1 KD was usually lower compared to control (Fig. 1C).

Figure 1D) Following the reviewer's suggestion, we have performed colony formation assays in control or Fip1 KD ESCs (Supplementary Figure S3B). Our results showed that Fip1 KD led to significant increase in the percentage of partially differentiated and differentiated cells ($p = 0$, Pearson's chi-square test). We have also included RT-qPCR data on additional pluripotency markers as suggested by the reviewer (Fig. 1D). Our data showed that Fip1 is important for ESC maintenance, and we have changed the wording in our description of Fip1 KD phenotype to "Fip1 KD led to impaired ESC maintenance and identity".

2) Figure 2: 2c) analysis is poorly described. How are genes normalized? Argument that FIP1 targets are going down is not convincing. Need some sort of analysis that supports significance. 2d) not sure how this figure supports argument that FIP1 regulates self-renewal/pluripotency. Its most dramatic effect seems to be proliferation. 2e) like 2c, poor description making figure impossible to interpret. Are these a mix of genes up and down in comparisons of Neuron to ESC and Flp1KD to control? Separate out into genes up and down.

Response:

Figure 2C) Raw mRNA expression data of 11-Point Time Course Study of Differentiating J1 Embryoid Bodies was downloaded from GEO (accession GSE3749) and re-

processed using the RMA methodology and Entrez gene based re-annotated CDFs (version 13) to summarize probes. Expression data from both the A and B chip was merged averaging duplicate probe sets that were represented on both chips. Next, using the limma package, a liner model fit with the 11 time points as factors was computed and variance estimates were adjusted using an empirical Bayes method. To calculate the statistical significance, we performed a pairwise comparison of Fip1-APA target genes and the set of all other genes through each time point using Wilcox sum rank test, and added the result to Supplementary Figure S11.

Figure 2D) Our gene ontology analyses showed that genes that function in regulating cellular, embryonic, and tissue development are enriched in Fip1 targets (Fig. 2D). Although clearly not a definitive proof by itself, this observation provides supportive evidence that Fip1 regulates self-renewal through these targets. This conclusion is further supported by our experimental evidence (Fig. 3C). As cell cycle and proliferation genes are also enriched, we mentioned their enrichment in the text to faithfully describe the result.

Figure 2E) This figure compares the genes that showed significant APA changes in Fip1-depleted cells and those that showed significant APA changes during ESC differentiation into neurons as reported by our previous study (Shepard et al, 2011). As the vast majority of the APA changes in both cases were proximal to distal shifts (PtoD), we only showed the overlap for PtoD genes in the previous submission. Following the reviewer's suggestion, we have now included both PtoD and DtoP genes in the comparison. By comparing APA changes in these two studies, we found the following. First, PtoD shifts are the predominant type of APA change in Fip1-depleted ESCs and during neuronal differentiation: 88% and 83% of APA changes are PtoD shifts for neuronal differentiation and for Fip1 KD respectively. Second, ~25% of the PtoD APA changes induced by Fip1 depletion in ESCs were observed during ESC differentiation into neurons ($p<10^{-40}$) (Figure 2E) (Shepard et al, 2011). In 86% of these genes, polyadenylation shifted to the same distal PASs during neuronal differentiation and in Fip1-depleted ESCs. Together, the large overlap between Fip1-regulated and differentiation-regulated APA events and the differentiation-associated Fip1 protein level change strongly suggest that Fip1 plays a role in regulating APA during ESC differentiation. These points have been added to the manuscript.

3) Figure 3: 3a) in what cellular context were these experiments performed.

Authors should control for length effects (i.e. independent of particular 3'UTRs chosen). That is would random sequences of equal length similarly reduce relative expression of luciferase. 3b) Evaluation of effect of kd of Fip1 targets is much too limited. All that is shown is morphology. At minimum should do PCR for representative pluripotency markers.

Response:

Figure 3A): The reporter assays in Fig. 3A were carried out in ESCs. Equal molar amounts of the constitutive (cUTR) or cUTR plus alternative 3' UTRs (c+aUTR)-containing reporter DNAs were transfected. Following the reviewer's suggestion, we have tested whether the aUTR sequences are important for the observed gene silencing effect. For "random" sequence controls, we tested the anti-sense sequences of the c+aUTRs (as-(c+aUTRs)) of *Wdr18*, *Etf1*, *Wwp2*, and *Ncaph2* in the same assay. In all cases, the as-(c+aUTR)s had significantly different effect on reporter expression than the corresponding c+aUTRs. These results suggest that the effect of aUTRs on gene expression is not simply due to their lengths, but is highly dependent on the specific sequences.

3B): We have performed RT-qPCR analyses of both ESC and differentiation marker genes in control and Fip1 target-depleted ESCs (Fig. 3D). Consistent with the morphological changes in these cells, our RT-qPCR results showed that the expression of ESC marker genes (*Oct4*, *Nanog*, *Sox2*, *Klf4*, or *Esrrb*) decreased while differentiation markers (*Cdx2*, *Gata3, Nestin, Sox17, Fgf5,* or *FoxA15*) increased (Supplementary Figure S16). Together we provided evidence at both molecular and morphological levels showing that depletion of Fip1 APA target genes led to the loss of self-renewal and ESC identity.

4) Fig. 4: This data does not support that FIP1 is required for reprogramming directly. The findings could be and likely are secondary to the differential effect of FIP1 on ESC and MEF proliferation (fig. 4c). The reduced AP intensity and definitely the number of flow positive Oct4-GFP could be ascribed to decreased proliferation of late stage reprogramming cells. Therefore, should be cautious of over-interpretation of data. Why did the authors not count Oct4-GFP colonies rather than flow? A&B) show levels in iPSCs as comparison. Would also be important to show that Flp1 kd influences shift toward primal usage in a subset of targets during reprogramming (flip side of 4b).

Response:

We thank reviewer-3 for the insightful comments. We have now added RT-qPCR data on marker genes at earlier time points (day 3 and day 6) during reprogramming. Our results showed that, as early as day 3, Fip1 KD led to reduced expression of early iPSC markers (*SSEA1*, *Cdh1, Alp1,* and *Fbxo15*) without affecting cell cycle/proliferation marker (*Cdkn2b*). These data provided additional evidence that Fip1 KD had a negative impact on reprogramming early on. As pointed out by the reviewer, it is possible that Fip1 also impairs iPSC maintenance. This point has been added to the manuscript.

To monitor the reprogramming efficiency, we used two assays. First, we performed AP staining 12 days after induction and counted AP positive colonies (Fig. 4C). Second, we counted the number of GFP positive cells by FACS in Oct4GFP reporter MEFs. We did not count GFP+ colonies because it is equivalent to counting AP positive colonies. Results from both assays showed a significant decrease in reprogramming efficiency in Fip1-depleted cells. We have also included RT-qPCR analysis of aUTR usage in Fip1 APA targets during reprogramming of control or Fip1 KD MEFs as suggested. Our results showed that Fip1 KD led to increased aUTR usage (Supplementary Figure S17B), which is consistent with our conclusion that Fip1 promotes an ESC-like APA pattern during reprogramming. We used ESCs instead of iPSCs for comparison in Fig. 4A-B to avoid variations in individual iPSC lines and to demonstrate that Fip1 levels and Fip1 APA profiles are restored to an ESC-like state.

5) Authors should be more consistent about choice of targets that they follow in 3a, 4b, sup 5c, and sup 7. Otherwise, it appears that authors are only choosing those targets that support their argument for each independent experiment.

Response:

In the revised manuscript, we have consistently characterized a set of three Fip1 targets, including *Ncaph2, Etf1*, and *Wwp2* in all assays. These include ***1)*** validation of APA change in Fip1 KD samples by RT-qPCR (Supplementary Figure S8); ***2)*** luciferase reporter assays to determine the impact of cUTR and c+aUTR on gene expression (Figure 3A); ***3)*** Western analyses of endogenous proteins in Fip1 KD samples (Figure 3B); ***4)*** knockdown in ESCs with two distinct siRNAs and determine the effect on ESC self-renewal by cell morphology and RT-qPCR of marker gene expression (Figure 3C and Supplementary Figure S15-16). The results in all assays are consistent with our model. As it is impractical to carry out all these experiments for all targets, we included

additional targets in each assay to further strengthen our model.

6) Fig. 5: Interesting findings in this figure. How do the proximal PASs differ in targets versus non-targets. Also, what is the authors' source of 3'UTR annotation including the defining of proximal and distal PASs.

Response: We have compared the proximal PASs (-100nt to +100nt) in Fip1 APA target vs. non-target mRNAs to identify significantly enriched sequence motifs in Fip1 target mRNAs. Within 100nt upstream of the cleavage sites (-100nt to 0 relative to the cleavage site), the top two most enriched sequence motifs in Fip1 target mRNAs are AAUAAA and UUUUGU with z-scores of 5.2 and 4.6 respectively. Within 100nt downstream of the cleavage sites (0 to +100nt), the top two most enriched motifs are UCUGUG and CUGUGG (their respective z-scores are 5.2 and 4.4). The full results are included in Supplementary Figure S22.

We used the Ensembl annotation of 3' UTRs in our analyses. In our APA analyses, we first identify two PASs with the most significant changes in usage for each gene (with the lowest p values based on Fisher's exact test) upon Fip1 KD. According to the positions of these two PASs relative to the transcript start site, the upstream PAS is designated as proximal and the downstream site as distal. This information is included in the Supplementary Information.

Reference:

Boelens WC, Jansen EJ, van Venrooij WJ, Stripecke R, Mattaj IW, Gunderson SI (1993) The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell* **72**(6)**:** 881-892

Colgan DF, Manley JL (1997) Mechanism and regulation of mRNA polyadenylation. *Genes Dev* **11**(21)**:** 2755-2766

Fink AL (2005) Natively unfolded proteins. *Curr Opin Struct Biol* **15**(1)**:** 35-41

Jenal M, Elkon R, Loayza-Puch F, van Haaften G, Kuhn U, Menzies FM, Oude Vrielink JA, Bos AJ, Drost J, Rooijers K, Rubinsztein DC, Agami R (2012) The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **149**(3)**:** 538-553

Karwacki-Neisius V, Goke J, Osorno R, Halbritter F, Ng JH, Weisse AY, Wong FC, Gagliardi A, Mullin NP, Festuccia N, Colby D, Tomlinson SR, Ng HH, Chambers I (2013)

Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. *Cell Stem Cell* **12**(5)**:** 531-545

Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE (2007) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* **446**(7138)**:** 926-929

Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C (2013) Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev* **27**(21)**:** 2380-2396

Martin G, Gruber AR, Keller W, Zavolan M (2012) Genome-wide Analysis of Pre-mRNA 3' End Processing Reveals a Decisive Role of Human Cleavage Factor I in the Regulation of 3' UTR Length. *Cell Rep* **1**(6)**:** 753-763

Mayr C, Bartel DP (2009) Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**(4)**:** 673-684

Meinke G, Ezeokonkwo C, Balbo P, Stafford W, Moore C, Bohm A (2008) Structure of yeast poly(A) polymerase in complex with a peptide from Fip1, an intrinsically disordered protein. *Biochemistry* **47**(26)**:** 6859-6869

Shepard PJ, Choi E, Lu J, Flanagan LA, Hertel KJ, Shi Y (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA* **17**(4)

Yao C, Biesinger J, Wan J, Weng L, Xing Y, Xie X, Shi Y (2012) Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. *Proc Natl Acad Sci U S A* **109**(46)**:** 18773-18778

Zhao J, Hyman L, Moore C (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* **63**(2)**:** 405-445

2nd Editorial Decision 14 January 2014

Thank you for submitting your revised manuscript for consideration by the EMBO Journal. It has now been seen by two of the original referees whose comments are enclosed below.

As you will see, both referees find that the revised manuscript has been significantly improved, highlight the strong impact of your findings and support publication in The EMBO Journal in principle; however, they do still raise a number of minor points that will have to be addressed before the manuscript can be officially accepted for publication. While most of these concerns can be addressed by textual changes to specify experimental procedures and data analysis, we do have to ask that you expand the computational analysis to further address the two major points raised by ref #1. In addition, we would ask you to include a second control siRNA in assaying the morphological changes in ESCs following Fip1 depletion and to better explain the use of different pluripotency markers in these assays (ref #2).

Given the referees' positive recommendations, I would like to invite you to submit a final revised version of the manuscript, addressing the comments of both reviewers. When preparing this revised version I would also ask you to consider the following editorial points:

-> make sure that the revised manuscript contains statements for author contribution and conflict of interest.
-> indicate basis for statistics and scale bar size for all relevant figures in figure legend (including supplementary material).
-> include a more extensive description of materials and methods in the main manuscript file.

As of Jan 1st 2014 every paper published in The EMBO Journal includes a 'Synopsis' to further enhance its discoverability. Synopses are displayed on the html version of the paper and are freely accessible to all readers. The synopsis will include a short standfirst - written by the handling editor - as well as 2-5 one sentence bullet points that summarise the paper and are provided by the authors. These bullet points should be complementary to the abstract - i.e. not repeat the same text. We encourage inclusion of key acronyms and quantitative information. I would therefore ask you to include your suggestions for bullet points when you submit the final manuscript.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

REFEREE REPORTS:

Referee #1:

Although, the revised manuscript is substantially improved there are still a number of issues with the manuscript.

1. I am still not convinced that Fip1 KD does not lead to decreased recognition of PAS of genes with single 3'UTRs.

A. The authors show strong evidence (Fig. 5A, Suppl. Fig. S19) that genes with single 3'UTRs show decreased recognition of their PAS. They tested the strength of PAS of three genes with single 3'UTRs and in all of them they detect decreased recognition of PAS. These findings are not discussed. Instead, they claim that genes with single 3'UTRs do not change mRNA levels. Their findings are based on the analysis of microarray data. Microarrays are not suitable to measure absolute mRNA levels; they only measure relative levels. Therefore, the analysis of microarray data is not useful in this context.

B. In Fig. 5E, they show that Fip1 targets have significantly higher Fip1 iCLIP signals than non-targets. This is convincing, because the difference is about 10-fold. However, in Fig. S26, they show Fip1 iCLIP signals of genes with single 3'UTRs and find even higher levels than in Fip1 APA target genes (about 20-fold higher than in Fip1 APA targets). This is a very significant finding which is not discussed. This is another hint that Fip1 seems to target genes with single 3'UTRs.

2. They authors claim several times throughout the manuscript that they identified the mechanism of regulation of APA by Fip1. The biggest question is how Fip1 only has an effect on about 300 genes out of 3000 genes with several PAS. This question was not addressed.

A. If the level of Fip1 binding to the 3'UTR is the most important determinant, then also genes with single 3'UTRs are targeted.
B. If the distance between the proximal and distal PAS is an absolute requirement, the distance between the proximal and distal PAS of APA target and non-target genes needs to be compared.
C. If the number of Fip1 binding sites (6-mer of Us) is important, it needs to be shown that non-targets do not have these binding sites in their 3'UTRs.
With the presented data, it is not clear how Fip1 identifies its targets.

Minor:

1. The y-axis in Fig. 1D is show to be log2 expression. In the initial submission of the manuscript it was not log2. If it is log2, then the values need to be changed because the bars on the negative side of the y-axis indicate values below 1, but for example log2 of 0.4 is 1.32.

2. Fig. S6. R2=0.94, but in the title it says r2=-0.94. I'm not sure if I understand what was done to the values on the y-axis, because in the figure legend it was stated that these are "predicted" values. By reading the text it is suggested that the microarray analysis serves as a validation of the DRS. But if the values on the y-axis are "predicted", then this is not a validation. This should be clarified.

3. In the 'Introduction' it reads as if ESC have overall shorter 3'UTRs, whereas differentiated cells have longer 3'UTRs. However, there have been recent publications that question this statement. The 'Introduction' should give a more balanced view about the current knowledge about APA.

4. In the microarray result from Fig. 2C, they show that Fip1 targets decrease in their mRNA abundance during ESC differentiation. This finding is used as confirmation that Fip1 targets play an important role in ESC. However, throughout the manuscript, the authors claim that Fip1 KD does not change mRNA levels of single UTR genes or of APA targets. This seems like a contradiction to me.

5. In the 'Discussion' it is stated that in the extended UTR of Fip1 targets there are 2.5x more miRNA target sites than in the common UTR. I was not able to find how this analysis was done, but the analysis needs to be adjusted for the different 3'UTR lengths. The more sequence is analyzed the higher is the probability of finding a miRNA binding site. Extended UTRs are much longer than common UTRs and therefore have a much higher probability of having miRNA binding sites.

Referee #2:

The revised manuscript is much improved. The findings are interesting and certainly worthy of publication. However, there remain some important concerns that should be addressed, largely through changes to the text.

Figures 1c and S3c: The alkaline phosphatase staining in these figures does not convincingly show a cell by cell decrease following knockdown. The strong staining of colonies in mock and control is likely largely due to the compactness of colonies, i.e. increased number of cells per unit surface area rather than increased AP activity in individual cells.

S3b: Would be useful to include colony number. Is there a decrease in number of colonies that form per number of cells plated following knockdown (kd). Such a result would provide the strongest

evidence for decreased self-renewal.

Fig. 1d: Identify statistically significant changes.

Fig. S4: Could use more explanation. It seems that the conclusion would be that kd promotes differentiation under self-renewal conditions, but inhibits differentiation under differentiation conditions. Important and somewhat confusing finding.

Fig S10B: Also could use more explanation. Oct4 is down, while Nanog, Sox2, Esrbb are all up by day 2. How does this fit with model of kd promoting differentiation? Is Oct4 directly targeted by Fip1?

Bottom p. 9 to top of p. 10: Authors write "In 86% of these genes, polyadenylation shifted to the same distal PASs during neuronal differentiation and in Fip1-depleted ESCs." However, it is unclear how often a single 3'UTR within this group has more than two PAS and, therefore, can use a distinct PAS following Fip1 KD versus neuronal differentiation.

Fig. S13: This result does not support the author's point and is superficially addressed in the text. In contrast to what the authors expected, two of the control (antisense) 3'UTRs did destabilize the mRNA levels, a third even enhanced destabilization, and the fourth somehow promoted stabilization. These results more strongly supports a role for 3'UTR length itself, rather than specific sequences in the long 3'UTRs underlying destabilization. It certainly does not support the opposite as the authors suggest.

Fig. 3c, S15B, and S16: It is hard to interpret when every target they knockdown leads to morphological changes and decreased levels of pluripotency genes. Would like to see more than than single siControl. For example, would be nice to see siRNAs to a couple mRNAs that are expressed in ESCs, but are not targets of Flp1 as controls. Otherwise, it is worrisome that any gene knockdown leads to this phenotype and hence is non-specific. Along similar lines, it is very strange that in S16, they use different markers for pluripotency for each individual knockdown. It makes it look like the authors selected the marker that supported their conclusion for each target and ignored the ones that did not support their conclusion. I am sure that was not their intent, but it needs explanation.

Fig. S17B: Changes do not look significant. Again, the results appear to go against authors' interpretation.

Fig. 5E: Confusing results. What underlies the striking differences in patterns between PtoD and DtoP tracks? The difference is ignored in the text.

Fig. S22: Provide more explanation in legend.

Fig. 6: Says very little. Fig S24, although highly speculative, is a more interesting and telling model.

Page 19: Sentence: "At lower Fip1 levels, the binding of Fip1/CPSF to the distal PASs decreases, which leads to de-repression of the proximal PASs" does not make sense. Maybe just needs to be re-written.

Referee #1:

Although, the revised manuscript is substantially improved there are still a number of issues with the manuscript.

1. I am still not convinced that Fip1 KD does not lead to decreased recognition of PAS of genes with single 3'UTRs.

A. The authors show strong evidence (Fig. 5A, Suppl. Fig. S19) that genes with single 3'UTRs show decreased recognition of their PAS. They tested the strength of PAS of three genes with single 3'UTRs and in all of them they detect decreased recognition of PAS. These findings are not discussed. Instead, they claim that genes with single 3'UTRs do not change mRNA levels. Their findings are based on the analysis of microarray data. Microarrays are not suitable to measure absolute mRNA levels; they only measure relative levels. Therefore, the analysis of microarray data is not useful in this context.

-Response:

Our data suggest that Fip1 KD leads to decreased recognition of PAS of both single-PAS genes and APA genes (Fig. 5A, Suppl. Fig. S19). For most Fip1 APA target genes, lower cleavage/polyadenylation efficiency at proximal PASs results in more transcription read-through and a shift of polyadenylation to the stronger distal PASs (Fig. 5A and C). For single-PAS genes, however, polyadenylation still occurs at the same PAS (see our previous response to Reviewer 1's comments). To understand why polyadenylation pattern did not change for single-PAS genes, we searched for potential PASs in the 5 kb

region downstream of the annotated PAS using a previously published support vector machine (SVM)-based method (Cheng et al. 2006). We found that the potential PASs downstream of single-PAS genes are significantly weaker than the mapped PAS of both single-PAS and APA genes based on their SVM scores ($p=1.8x10^{-55}$ and $p<1x10^{-100}$ respectively) (Supplementary Figure S26). These results indicate that, although the cleavage/polyadenylation efficiency decreases at the PAS of single-PAS genes, polyadenylation still occurs at the same sites as there are no other stronger PAS further downstream. This provides an explanation for our microarray analysis results that the mRNA levels of single PAS genes did not change following Fip1 depletion. It should be pointed out that microarray analyses are suitable for determining whether the mRNA levels of single-PAS genes change or not upon Fip1 KD.

B. In Fig. 5E, they show that Fip1 targets have significantly higher Fip1 iCLIP signals than non-targets. This is convincing, because the difference is about 10-fold. However, in Fig. S26, they show Fip1 iCLIP signals of genes with single 3'UTRs and find even higher levels than in Fip1 APA target genes (about 20-fold higher than in Fip1 APA targets). This is a very significant finding which is not discussed. This is another hint that Fip1 seems to target genes with single 3'UTRs.

-Response:

The strong Fip1 iCLIP signals at the PAS of single 3' UTR genes suggest that these PASs are targeted by Fip1. This is consistent with our results that Fip1 KD leads to decreased recognition of PAS of both single-PAS genes and APA genes (Fig. 5A, Suppl. Fig. S19). Discussion has been added to clarify this point.

2. They authors claim several times throughout the manuscript that they identified the mechanism of regulation of APA by Fip1. The biggest question is how Fip1 only has an effect on about 300 genes out of 3000 genes with several PAS. This question was not addressed.

A. If the level of Fip1 binding to the 3'UTR is the most important determinant, then also genes with single 3'UTRs are targeted.

B. If the distance between the proximal and distal PAS is an absolute requirement, the distance between the proximal and distal PAS of APA target and non-target genes needs to be compared.

C. If the number of Fip1 binding sites (6-mer of Us) is important, it needs to be shown that non-targets do not have these binding sites in their 3'UTRs. With the presented data, it is not clear how Fip1 identifies its targets.

-Response:

As summarized in the Abstract, "the specificity and the mode of Fip1-mediated APA regulation depend on multiple factors, including Fip1-RNA interactions and the distance between alternative polyadenylation sites" (i.e. a combination of mechanisms A and B as suggested by the reviewer). Our results suggest that Fip1-RNA interactions play an important role in determining Fip1 target specificity (Fig. 5E). Indeed our results showed that single-PAS genes are also targeted by Fip1 (Supplementary Figure S26), but their polyadenylation pattern does not change upon Fip1 depletion due to reasons described above (see Response to the first point). For Fip1 APA target genes, the direction of the APA changes is influenced by the distance between the proximal and distal PASs (Fig. 5F). Following the reviewer's suggestion, we have added the distance between proximal and distal PASs for the non-target genes (Fig. 5F).

Minor:

1. The y-axis in Fig. 1D is show to be log2 expression. In the initial submission of the manuscript it was not log2. If it is log2, then the values need to be changed because the bars on the negative side of the y-axis indicate values below 1, but for example log2 of 0.4 is 1.32.

Response: We thank the reviewer for spotting this error and "log2" has been removed from the y-axis label.

2. Fig. S6. R2=0.94, but in the title it says r2=-0.94. I'm not sure if I understand what was done to the values on the y-axis, because in the figure legend it was stated that these are "predicted" values. By reading the text it is suggested that the microarray analysis serves as a validation of the DRS. But if the values on the y-axis are "predicted", then this is not a validation. This should be clarified.

Response:

The R^2 value is 0.94 and this typo in the chart title has been corrected. We agree with the reviewer that it is confusing to call the values on the y-axis "predicted" as they were calculated based on experimental data (DRS and microarray). We have removed the

word “predicted” and added more detailed description in the figure legend on how the values on the y-axis are calculated.

3. In the 'Introduction' it reads as if ESC have overall shorter 3'UTRs, whereas differentiated cells have longer 3'UTRs. However, there have been recent publications that question this statement. The 'Introduction' should give a more balanced view about the current knowledge about APA.

Response:

We thank the reviewer for this point. We have added sentences to emphasize that the relationship between the global APA profile and cellular proliferation may be more complicated and subject to tissue-specific regulation. A recent publication is cited to support this point (Lianoglou et al. 2013).

4. In the microarray result from Fig. 2C, they show that Fip1 targets decrease in their mRNA abundance during ESC differentiation. This finding is used as confirmation that Fip1 targets play an important role in ESC. However, throughout the manuscript, the authors claim that Fip1 KD does not change mRNA levels of single UTR genes or of APA targets. This seems like a contradiction to me.

Response:

In Fig. 2C, the down-regulation of Fip1 target genes is caused by ESC differentiation, but not necessarily by Fip1. Our DRS and microarray data demonstrated that Fip1 KD in ESC led to changes in APA patterns but not in the mRNA levels of most Fip1 target genes (Fig. 2A). We have specifically pointed out that “In addition, these results also suggest that Fip1 APA genes are regulated at both transcriptional and post-transcriptional levels (such as APA) during differentiation.” (Results) and that APA is a fine-tuning mechanism for gene expression that works with other mechanisms (Discussion).

5. In the 'Discussion' it is stated that in the extended UTR of Fip1 targets there are 2.5x more miRNA target sites than in the common UTR. I was not able to find how this analysis was done, but the analysis needs to be adjusted for the different 3'UTR lengths. The more sequence is analyzed the higher is the probability of finding a miRNA binding site. Extended UTRs are much longer than common

UTRs and therefore have a much higher probability of having miRNA binding sites.

Response:

In our manuscript, we only described the total number of predicted miRNA target sites, and we did not discuss the density of these sites. Following the reviewer's suggestion, we have added Supplementary Figure S31 to show both the number and density of miRNA target sites in the cUTRs and aUTRs of Fip1 APA target genes. We have also added detailed description of the miRNA analyses in the Supplementary Information.

Referee #2:

The revised manuscript is much improved. The findings are interesting and certainly worthy of publication. However, there remain some important concerns that should be addressed, largely through changes to the text.

Figures 1c and S3c: The alkaline phosphatase staining in these figures does not convincingly show a cell by cell decrease following knockdown. The strong staining of colonies in mock and control is likely largely due to the compactness of colonies, i.e. increased number of cells per unit surface area rather than increased AP activity in individual cells.

Response:

We thank reviewer-2 for raising this point. We have now repeated AP staining six days after the transfection of Fip1 siRNAs, and the new result more clearly demonstrated that Fip1 KD cells have reduced AP staining (Supplementary Figure S3A).

S3b: Would be useful to include colony number. Is there a decrease in number of colonies that form per number of cells plated following knockdown (kd). Such a result would provide the strongest evidence for decreased self-renewal.

Response:

We thank reviewer-2 for this suggestion. We have counted the colony numbers per 10,000 plated cells seven days after transfection of control or Fip1 siRNAs and found that Fip1 KD led to a significant decrease in colony numbers ($p<0.01$), providing further support for a critical role for Fip1 in ESC self-renewal. These results have been included in the Supplementary Figure S3D in the revised manuscript.

Fig. 1d: Identify statistically significant changes.

Response:

Statistically significant changes are marked in the revised Fig. 1D.

Fig. S4: Could use more explanation. It seems that the conclusion would be that kd promotes differentiation under self-renewal conditions, but inhibits differentiation under differentiation conditions. Important and somewhat confusing finding.

Response:

Our data suggest that Fip1 is required for ESC self-renewal, and thus its KD leads to differentiation under normal ESC culture conditions (Fig. 1). In addition, Fip1 may also play important roles in pluripotency, and its KD results in abnormal differentiation kinetics, as suggested by the marker gene expression (Supplementary Fig. S4), and thus impaired developmental potential of ESCs. The main text has been modified to better explain this point.

Fig S10B: Also could use more explanation. Oct4 is down, while Nanog, Sox2, Esrbb are all up by day 2. How does this fit with model of kd promoting differentiation? Is Oct4 directly targeted by Fip1?

Response:

The RT-qPCR analyses shown in Fig S10B were carried out at early time points after Fip1 KD. All ESC markers, including *Oct4*, *Nanog*, *Sox2*, and *Esrbb*, showed relatively small changes (~1.5 fold or less), indicating that gross differentiation has not occurred at these time points. However, significant APA changes in many Fip1 APA target genes were observed (Supplementary Figure S10A). More significant changes in ESC and differentiation marker gene expression were observed later after Fip1 KD (Fig. 1D). Together, these results suggest that APA changes precede gross ESC differentiation. Our Fip1 iCLIP-seq analyses detected little, if any, Fip1 binding at *Oct4* PAS (from -100nt to +100nt), indicating that Oct4 PAS is indirectly or weakly targeted by Fip1. Our results support a model in which Fip1 KD led to significant changes in the APA profiles (Supplementary Figure S8B) and decreased expression of many critical self-renewal factors (Fig. 3A and B), which in turn leads to Oct4 downregulation.

Bottom p. 9 to top of p. 10: Authors write "In 86% of these genes, polyadenylation shifted to the same distal PASs during neuronal differentiation and in Fip1-depleted ESCs." However, it is unclear how often a single 3'UTR within this group has more than two PAS and, therefore, can use a distinct PAS following Fip1 KD versus neuronal differentiation.

Response:

It should be clarified that all genes within this group have multiple PASs and their APA showed PtoD changes, and there is no single 3' UTR genes (Fig. 2E). In 86% of these genes, polyadenylation shifted to the same distal PASs during neuronal differentiation and in Fip1-depleted ESCs. For the rest of the genes, most of them (64%) have three or more PASs and polyadenylation shifted from the proximal to both distal PASs. This point has been added to the text.

Fig. S13: This result does not support the author's point and is superficially addressed in the text. In contrast to what the authors expected, two of the control (antisense) 3'UTRs did destabilize the mRNA levels, a third even enhanced destabilization, and the fourth somehow promoted stabilization. These results more strongly supports a role for 3'UTR length itself, rather than specific sequences in the long 3'UTRs underlying destabilization. It certainly does not support the opposite as the authors suggest.

Response:

Experiments shown in Fig. S13 were designed to test whether the inhibitory effect of aUTRs was due to their length. Our results showed that the antisense sequences of selected aUTRs could have similar (*Wdr18*, *Etf1*, and *Wwp2*) or opposite (*Ncaph2*) effect on reporter gene expression compared to the original aUTRs. As the aUTRs and antisense aUTRs have the same length but different effects on reporter gene expression, we concluded that "…the inhibitory effect of aUTRs was not simply due to their lengths". This conclusion was fully supported by our data and did not rule out 3' UTR lengths as a contributing factor of the observed effect. To clarify this point, we have rephrased our conclusion to: "…the inhibitory effect of aUTRs may be due to their lengths and/or sequences".

Fig. 3c, S15B, and S16: It is hard to interpret when every target they knockdown leads to morphological changes and decreased levels of pluripotency genes.

Would like to see more than single siControl. For example, would be nice to see siRNAs to a couple mRNAs that are expressed in ESCs, but are not targets of Flp1 as controls. Otherwise, it is worrisome that any gene knockdown leads to this phenotype and hence is non-specific. Along similar lines, it is very strange that in S16, they use different markers for pluripotency for each individual knockdown. It makes it look like the authors selected the marker that supported their conclusion for each target and ignored the ones that did not support their conclusion. I am sure that was not their intent, but it needs explanation.

Response:

In the previous version of our manuscript, we showed that the depletion of CPSF30 and CFIm 25 did not lead to differentiation and loss of self-renewal as observed when Fip1 or Fip1 target genes were knocked down (Supplementary Figure S27). These results suggest that not all gene knockdowns lead to differentiation and that the phenotypes we observed were gene-specific. To further address the reviewer's concern, we have knocked down by RNAi *Apex1* and *Nasp,* two highly expressed genes in ESCs. As shown in Supplementary Fig. S15, depletion of these genes did not cause ESC differentiation based on cell morphology, furthering supporting the specificity of our assay.

Although Fip1 target genes are important for ESC self-renewal, they may regulate self-renewal in different ways and their depletion may lead to differentiation into different lineages. Therefore, silencing Fip1 target genes can result in different differentiation phenotype and marker expression profiles. We included specific markers that showed significant changes upon KD of each Fip1 target genes to show that these genes are required for ESC maintenance (Supplementary Figure S16).

Fig. S17B: Changes do not look significant. Again, the results appear to go against authors' interpretation.

Response:

These changes are indeed moderate, especially compared to the effect of Fip1 KD on APA in mESCs (Supplementary Figure S8B). This is consistent with our observation that the influence of Fip1 on APA and cellular proliferation may be cell type-specific and is more pronounced in stem cells. For example, we have shown that Fip1 KD in HeLa cells also led to predominantly PtoD APA changes (Supplementary Figure S29C), similar to its effect in mESCs. But the number of affected genes is significantly smaller

(Supplementary Figure S29C). During somatic reprogramming, only a small percentage (<1%) of cells are reprogrammed into iPSCs, thus the effect of Fip1 KD on the global APA profiles in the bulk population of MEFs, including cells that are undergoing reprogramming and those that are not, is expected to be less pronounced than that in mESCs. These points have been added to the text.

Fig. 5E: Confusing results. What underlies the striking differences in patterns between PtoD and DtoP tracks? The difference is ignored in the text.

Response:

The major difference between the PtoD and DtoP tracks is the density of the iCLIP tags. This is likely due to the difference in the numbers of PtoD (311 genes) and DtoP (63 genes) genes. Given the lower number of DtoP genes, iCLIP tags are detected at less positions and thus appear more sparse in Fig. 5E. This explanation has been added to the figure legend for Fig. 5E.

Fig. S22: Provide more explanation in legend.

Response:

More detailed explanation has been added to the figure legend.

Fig. 6: Says very little. Fig S24, although highly speculative, is a more interesting and telling model.

Response:

Following the reviewer's suggestion, we have moved the original Fig. S24 to the main figure as the new Fig. 6. We thank the reviewer for this suggestion.

Page 19: Sentence: "At lower Fip1 levels, the binding of Fip1/CPSF to the distal PASs decreases, which leads to de-repression of the proximal PASs" does not make sense. Maybe just needs to be re-written.

Response:

More description has been added to more clearly spell out the direct competition between Fip1/CPSF and CstF binding and the recognition of the distal and proximal PAS.

Reference:

Cheng Y, Miura RM, Tian B (2006) Prediction of mRNA polyadenylation sites by support vector machine. Bioinformatics 22 (19):2320-2325. doi:10.1093/bioinformatics/btl394

Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C (2013) Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. Genes Dev 27 (21):2380-2396.