THE EMBO JOURNAL

Regulation of RNA polymerase II processivity by Spt5 is restricted to a narrow window during elongation

Johanna Fitz, Tobias Neumann and Rushad Pavri.

Review timeline:

Submission date:	9th August 2017
Editorial Decision:	6th September 2017
Revision received:	10th December 2017
Editorial Decision:	4th January 2018
Revision received:	10th January 2018
Accepted:	15th January 2018

Editor: Anne Nielsen

Transaction Report:

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

1st Editorial Decision — 6th September 2017

Thank you for submitting your manuscript for consideration by The EMBO Journal. It has now been seen by three referees whose comments are shown below.

As you will see from the reports, our referees all express interest in the findings reported in your manuscript although they also raise a number of points that you will have to address before they can support publication here. I would particularly ask you to focus your efforts on the following points:

-> Refs #1 and #3 are both concerned about the long duration of Spt5 depletion and the potential side-effects it could have on other transcription regulators. It will therefore be important that you address the expression levels of relevant factors and comment on how the long depletion of Spt5 may impact the results.

-> Refs #1 and #3 are also both curious about the nature/features associated with the downstream regions that require Spt5 for effective processitivity. The referees make constructive suggestions on how to address this and I'll encourage you to follow them.

-> Finally, the referees point out that the manuscript is difficult to read in places and that a more extensive discussion/comparison to the existing literature is needed. Please incorporate this in the revised study.

Given the referees' overall positive recommendations, I would like to invite you 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 of your manuscript will therefore depend on the completeness of your responses in this revised version.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the 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://emboj.embopress.org/about#Transparent_Process

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. Should you foresee a problem in meeting this three-month deadline, please let us know in advance and we may be able to grant an extension.

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

REFEREE REPORTS

Referee #1:

This manuscript describes a lot of data that focuses on effects of mammalian Spt5 in pol II transcription. Spt5 is inducibly depleted by 4-HT treatment, and the bulk of the data are from MEFs treated with 4-HT for 72h. The results are interesting and thorough, although of course there are always more experiments to be completed. Although the data generally coincide with findings from other labs, the results here expand existing understanding of this important pol II regulatory factor. Below I point out some areas and experiments that are important to address in a revision.

1. It is important to further verify the depletion of Spt5. At the least, this will require additional information from the authors. Westerns were completed that showed reduced full-length Spt5 protein, and that's good. However, the authors also have RNA-Seq data from the Spt5-depleted cells. Can they confirm that Spt5 mRNA reads were absent/depleted in 4-HT treated cells? Could they map reads to exon 14? I did not see any mention of this. Furthermore, what antibodies were used to detect Spt5 by western? Does the epitope target the deleted region? It is mentioned that antibody information is provided in Table S2 but I could not find Table S2 anywhere among the mansucript materials. Ideally, the authors could probe Spt5 with multiple epitopes to confirm that no truncated products are generated.

2. The recent report about Spt5 depletion in yeast from the Winston lab (Mol Cell 2017 77) is mentioned but I did not see a specific mention of the issue with downstream antisense transcription. The authors should check their GRO-Seq data to see if unexpected antisense transcription is similarly observed upon Spt5 depletion in MEFs. I did not see any mention of this in the current version of the manuscript. As I recall, the Winston group observed widespread antisense transcription upon degron depletion of yeast Spt5 that initiated about 500 bp downstream of the TSS. Of course, the potential location of antisense transcripts may be different in mammalian cells, but is there any evidence at all that within-gene antisense transcription is similarly induced upon Spt5 depletion in MEFs?

3. Related to point 2, the Lis lab (Booth et al. Genome Res. 2016 799) observed global decreases in antisense transcription upon depletion of Spt4 in S. pombe. Similar results are not seen in S. cerevisiae. Once the authors assess antisense transcription from their GRO-Seq data, the authors should comment on this work as well.

4. The Spt5 depletion occurs over a long time frame, 72h. It is important that the authors confirm how/whether global levels of pol II and Spt4 undergo changes during this time frame, using quantitative westerns. Of course, dozens of factors could be tested but these two would be a good start. It is highly plausible that similar reduction in Spt4 will be noted upon Spt5 depletion. In that case, the authors can simply change their description to be reporting on the effects of depletion of the DSIF complex.

5. The Ansari lab reported loss of pol II processivity/premature termination in yeast, upon selective

inhibition of Kin28 (Rodriguez-Molina et al. Mol Cell 2016 433). These results have some similarities with what is shown here in Spt5 depleted cells. Because Spt5 is a putative substrate of Kin28 (Rob Fisher's lab), these results should be mentioned.

Other points
6. Pol II pausing is believed to serve as a checkpoint to enable 5'-capping of the transcript. Given that pausing is reduced in Spt5-depleted cells, do the authors observe changes in mRNA stability? Capping defects? I understand that there are always more experiments to propose and always more work to do, so perhaps all of these questions can remain for future work, but at minimum the authors should acknowledge this and discuss it as one potential manifestation of Spt5 depletion.

7. Can the authors draw any further conclusion from the 15-20kb window downstream of the promoter? Have they attempted to see whether these may correlate with other features, such as long 5'-UTR introns?

8. In describing the flavopiridol (FP) experiments, it is stated that FP "inhibits Ser2 phosphorylation of Pol II by P-TEFb." This is true, but misleading. FP is a pan-kinase inhibitor and treatment of cells with FP will effectively inhibit dozens of kinases (e.g. see Asghar et al. Nat Rev Drug Disc. 2015 130). This is especially true with 4h treatment at a concentration of 2 μM. This statement should be re-phrased.

Referee #2:

Fitz et al have used a 2-step Flox system to make MEFs that have only very low levels of Spt5 and importantly, Spt5 is depleted from genes. GRO-seq analysis indicates that Spt5 depletion causes a drop in pol II at the TSS and an increase after that at the 5' end of genes, as would be expected. Spt5 depletion causes a promoter-proximal increase in pol II CTD Ser2 phosphorylation. Interestingly, Spt5 depletion causes a transcription elongation decrease only after about 15-20kb and a termination defect, as measured by GRO-seq. This is validated with single gene ChIP and RT-qPCR of pre-mRNA. The authors conclude that Spt5 plays a role at this transition point in continued productive elongation.

Flavopiridol experiments show convincingly that elongation rates are not changed by Spt5 depletion. The authors then conclude that increased intitiation contributes to increases in transcription up to 15kb from the TSS.

I think this is a well thought-out and well executed study and the results are interesting and important

Minor point.
I'm not sure I really understood the section on elongation rates calculated from nascent and mRNA levels-last paragraph of page 12 to third paragraph of page 14. Perhaps the authors could make this a bit clearer?

Non-essential
It would be nice to know what is happening to RNA processing when Spt5 is depleted.

Referee #3:

In the present manuscript, Fitz at al. use genome-scale approaches along with single-gene assays to investigate the role of the highly conserved elongation factor Spt5 during RNA polymerase II (Pol II) transcription in mammalian cells. The authors deplete Spt5 from mouse embryonic fibroblasts using the Cre-loxP system and study the genome-wide consequences on Pol II occupancy, the Pol II CTD phosphorylation state, nascent transcription and total RNA levels. ChIP-seq and GRO-seq analyses reveal a strong decrease of Pol II occupancy at the promoter-proximal regions and an

increase over the proximal gene body regions genome-wide. The authors further show by ChIP-seq that the CTD-Ser5P are decreased and CTD-Ser2P levels are increased at the 5'-end of genes upon Spt5 depletion. Furthermore, the authors provide evidence for Pol II occupancy to be strongly decreased at the distal gene body region (>20 kb) of long genes. Since the authors found no evidence for reduced Pol II transcription elongation rates upon Spt5 depletion, the decrease in Pol II occupancy rather reflects a reduction of transcribing Pol II complexes. These observations led to the conclusion that Spt5 is important for the processivity of RNA polymerase transcription >20 kb downstream of the TSS at long mammalian genes.

This study uncovers new interesting aspects of the role of Spt5 during Pol II transcription elongation. The finding that Spt5 is required for processive transcription 20 kb downstream of the TSS at long mammalian genes is a significant result and will be of broad interest for the transcription community. This work will inspire future studies to further characterize the Spt5-dependent proposed processivity checkpoint. However, in its present form the manuscript is in parts hard to read and needs to be improved.

Major comments:
(1) Spt5 is a key elongation factor that closely interacts with many other elongation and termination factors during Pol II transcription, including Spt4 and Paf1. Thus depletion of Spt5 per se will trigger pleiotropic effects. A main weakness of the present study is the long Spt5 depletion time (72 hours). This long treatment time further increases the risk of indirect effects and cellular adjustments, and complicates the assignment of Spt5-specific functions. The authors provide no information on whether other components of the Pol II elongation complex, such as Spt4 and Paf1, are affected upon Spt5 depletion. The author should address these limitations in the main text of this study. Conclusions, such as on page 20 „defined the precise role of Spt5", should be softened accordingly.

(2) The authors conclude on Page 8 that upon Spt5 depletion "Pol II undergoes normal initiation". This is not in line with several observations presented in the manuscript. ChIP-seq analyses clearly show a strong decrease of Pol II and Pol II Ser5P occupancy upon Spt5 depletion around the TSS. This suggests that transcription initiation seems also to be affected when Spt5 is depleted. This should be addressed in the main text and the conclusion needs to be softened.

(3) A major finding by Fitz et al. is that Spt5 is important for the processivity of Pol II transcription past approximately 20 kb. This transition in Pol II occupancy (Fig. 3A) should be characterized in more depth. How many active genes show this transition in Pol II occupancy? What is the exact location (distance to TSS on average) of this transition? How does the location vary between different genes?

(4) As mentioned earlier, the manuscript is at parts hard to read and lengthy. For instance, the authors conclude from pages 12 to 14 that Spt5 depletion does not affect Pol II transcription elongation rates. This is then again presented and now in a more convincing way on pages 14 to 16. Both sections (pages 12 to 16) should be combined and at the same time shortened.

(5) Figure 1 A, B: the decrease in Pol II levels is much stronger in case of ChIP-seq as compared to GRO-seq measurements. What is the reason for this discrepancy? This needs to be explained in the main text.

Minor comments:
(1) Formatting error on Page 2, first paragraph: 5.6-dichloro-1-__-D-ribofuranosylbenzimidazole; similar formatting errors also on page 22

(2) Page 2, second parapgraph: „Promoters initially recruit Pol II to assamble...". This statement is confusing since general transcription factors including TBP recruit Pol II to the promoter region. This statement needs to be changed.

(3) Typo on Page 2, second paragraph: „P-TEFb inhibition leads a genome-wide accumulation..." needs to be replaced by „P-TEFb inhibition leads to a genome-wide accumulation..."

(4) Page 2/3 first paragraph: „Importantly, Spt5 remains associated with Pol II until termination...". The original papers need to be cited that support this statement.

(5) Page 5, second paragraph: the correlation coefficients between the replicates need to be added.

(6) Page 8, second paragraph: CTD Ser5P is not a mark of „early initiation", as stated by the authors, but rather of early elongation. This needs to be changed.

(7) Page 19, first paragraph: „no further loss of Pol II ... was discernable until the TTS, even at very long genes." This statement is too vague. Please add gene lenghtes in kb.

(8) Page 20, third paragraph: „Since yeast genes are much shorter than mammalian genes...". This is again too vague. Please add average gene lengthes in kb.

(9) Figure 3A: an arrow at the 20 kb region right at the transition of the Pol II occupancy would be helpful.

(10) Figure 3B (upper panel) and 6B: Differences between WT and SPT5dep can hardly be seen. Zoom-in views would be helpful.

(11) Table S1 and S2 as stated in the main text are missing.

(12) No ArrayExpress or GEO identifier for the genome-wide data is provided.

1st Revision - authors' response 10th December 2017

We thank the reviewers for their time, and for their insightful and constructive criticisms and suggestions. We have performed additional experiments and analyses to address the concerns. We hope that the revised manuscript will meet their approval for publication in The EMBO Journal. Below, we provide a point-by-point response to the reviewers' comments.

Referee #1:

This manuscript describes a lot of data that focuses on effects of mammalian Spt5 in pol II transcription. Spt5 is inducibly depleted by 4-HT treatment, and the bulk of the data are from MEFs treated with 4-HT for 72h. The results are interesting and thorough, although of course there are always more experiments to be completed. Although the data generally coincide with findings from other labs, the results here expand existing understanding of this important pol II regulatory factor. Below I point out some areas and experiments that are important to address in a revision.

1. It is important to further verify the depletion of Spt5. At the least, this will require additional information from the authors. Westerns were completed that showed reduced full-length Spt5 protein, and that's good. However, the authors also have RNA-Seq data from the Spt5-depleted cells. Can they confirm that Spt5 mRNA reads were absent/depleted in 4-HT treated cells? Could they map reads to exon 14? I did not see any mention of this. Furthermore, what antibodies were used to detect Spt5 by western? Does the epitope target the deleted region? It is mentioned that antibody information is provided in Table S2 but I could not find Table S2 anywhere among the mansucript materials. Ideally, the authors could probe Spt5 with multiple epitopes to confirm that no truncated products are generated.

We now show that read mapping to exons 14-16 from mRNAseq, GROseq and ChIPseq data (Figure S1E-F). Although the Spt5 mRNA is down-regulated ~4-fold in the Spt5-depleted cells, exons 14-16 are significantly more depleted than the remaining exons. In GROseq and ChIPseq, we find a very dramatic loss of reads in exons 14-16. The discrepancy between these techniques is likely due to the fact that GROseq and ChIPseq measure the signals at the gene at the final time-point of the experiment when the floxed exons are deleted (we confirm this also by genotyping in Figure S1D), whereas mRNAseq reports on the steady-state pool of mRNA that would include full-length mRNA synthesized prior to complete deletion of the floxed segment in all cells. This also explains, in part, the fact that ~10% of Spt5 protein is still detected in nuclear extracts at the end-point of culturing.

Regarding Spt5 western blots, we specifically used an antibody recognizing an N-terminal portion of Spt5 that is retained in the truncated protein and thus should be able to detect the truncated form. We now show the whole blot probed with this antibody in Figure S1G to clarify that the truncated form in not detected. We have also added Tables EV2 and EV3 listing the primers and antibodies. We apologize for not providing this before.

2. The recent report about Spt5 depletion in yeast from the Winston lab (Mol Cell 2017 77) is mentioned but I did not see a specific mention of the issue with downstream antisense transcription. The authors should check their GRO-Seq data to see if unexpected antisense transcription is similarly observed upon Spt5 depletion in MEFs. I did not see any mention of this in the current version of the manuscript. As I recall, the Winston group observed widespread antisense transcription upon degron depletion of yeast Spt5 that initiated about 500 bp downstream of the TSS. Of course, the potential location of antisense transcripts may be different in mammalian cells, but is there any evidence at all that within-gene antisense transcription is similarly induced upon Spt5 depletion in MEFs?

We now provide a new figure dedicated to antisense transcript analysis (Figure S5). Within gene bodies, there appears to be no clear trend towards up-or down regulation and the location of such changes is also highly varied within the gene (Figure S5A). In the promoter-proximal region (TSS -/+500 bp), we detect global decrease in antisense transcription in the TSS-upstream region (TSS - 500 bp; also seen in the metagene plots in Figures 1B and 2A) but no bias towards up- or down-regulation in the TSS-downstream region (TSS + 500 bp; Figure S5C-D). This suggests that the role of Spt5 in regulating antisense transcription may differ between mammalian cells and yeast.

3. Related to point 2, the Lis lab (Booth et al. Genome Res. 2016 799) observed global decreases in antisense transcription upon depletion of Spt4 in S. pombe. Similar results are not seen in S. cerevisiae. Once the authors assess antisense transcription from their GRO-Seq data, the authors should comment on this work as well.

As requested, we have commented on this in the Discussion.

4. The Spt5 depletion occurs over a long time frame, 72h. It is important that the authors confirm how/whether global levels of pol II and Spt4 undergo changes during this time frame, using quantitative westerns. Of course, dozens of factors could be tested but these two would be a good start. It is highly plausible that similar reduction in Spt4 will be noted upon Spt5 depletion. In that case, the authors can simply change their description to be reporting on the effects of depletion of the DSIF complex.

This is an important point and we now show new western blot analyses of several elongation factors including the Pol II phosphorylated forms, Spt4, Paf1, Spt16 (FACT component) and H3K36me3, a mark associated with elongation. None of these factors undergoes significant changes in protein levels upon loss of Spt5 (Figure S3A-B).

5. The Ansari lab reported loss of pol II processivity/premature termination in yeast, upon selective inhibition of Kin28 (Rodriguez-Molina et al. Mol Cell 2016 433). These results have some similarities with what is shown here in Spt5 depleted cells. Because Spt5 is a putative substrate of Kin28 (Rob Fisher's lab), these results should be mentioned.

Thank you for raising this point. We have discussed the Rodriguez-Molina et al. paper and the potential contribution of Spt5 phosphorylation by Cdk7/Kin28 to the phenotype of Spt5-depleted cells.

Other points
6. Pol II pausing is believed to serve as a checkpoint to enable 5'-capping of the transcript. Given that pausing is reduced in Spt5-depleted cells, do the authors observe changes in mRNA stability? Capping defects? I understand that there are always more experiments to propose and always more work to do, so perhaps all of these questions can remain for future work, but at minimum the authors should acknowledge this and discuss it as one potential manifestation of Spt5 depletion.

Although we have not directly analyzed mRNA capping or stability in this study, we note that there is good agreement between the changes in nascent transcription and mRNA levels upon Spt5 depletion (Figure 5D), and further, that global splicing efficiency in Spt5-depleted cells is comparable to WT cells (new Figure S5B). This suggests that mRNA processing and turnover is not significantly affected by the depletion of Spt5 on a global scale, although we cannot rule out that a subset of down-regulated mRNAs may be regulated via Spt5-dependent capping or processing (a project currently ongoing in our lab as a follow-up study). Finally, mRNAseq analysis showed only a relatively small number of differentially expressed genes (591 genes >1.5-fold changed; Figure 5A) upon Spt5 depletion with no bias for up- or down-regulated mRNAs. In sum, our data suggest that on a global scale, Spt5 depletion under our conditions does not have major effects on mRNA processing and turnover. We have addressed these points in the revised manuscript.

7. Can the authors draw any further conclusion from the 15-20kb window downstream of the promoter? Have they attempted to see whether these may correlate with other features, such as long 5'-UTR introns?

We have performed a new analysis looking into the possible correlation of our elongation phenotype with the presence of 5'UTR introns. As seen in Figure 3E, the 15-20 kb transition phenotype is observed to a similar extent between genes that have 5'UTR introns and those that do not. As mentioned above, there is also no evidence of any changes in antisense transcription in the TSS +15-20 kb region.

8. In describing the flavopiridol (FP) experiments, it is stated that FP "inhibits Ser2 phosphorylation of Pol II by P-TEFb." This is true, but misleading. FP is a pan-kinase inhibitor and treatment of cells with FP will effectively inhibit dozens of kinases (e.g. see Asghar et al. Nat Rev Drug Disc. 2015 130). This is especially true with 4h treatment at a concentration of 2 μM. This statement should be re-phrased.

Thank you for pointing this out. We have re-phrased this sentence.

Referee #2:

Fitz et al have used a 2-step Flox system to make MEFs that have only very low levels of Spt5 and importantly, Spt5 is depleted from genes. GRO-seq analysis indicates that Spt5 depletion causes a drop in pol II at the TSS and an increase after that at the 5' end of genes, as would be expected. Spt5 depletion causes a promoter-proximal increase in pol II CTD Ser2 phosphorylation. Interestingly, Spt5 depletion causes a transcription elongation decrease only after about 15-20kb and a termination defect, as measured by GRO-seq. This is validated with single gene ChIP and RT-qPCR of pre-mRNA. The authors conclude that Spt5 plays a role at this transition point in continued productive elongation.

Flavopiridol experiments show convincingly that elongation rates are not changed by Spt5 depletion. The authors then conclude that increased intitiation contributes to increases in transcription up to 15kb from the TSS.

I think this is a well thought-out and well executed study and the results are interesting and important

Minor point.
I'm not sure I really understood the section on elongation rates calculated from nascent and mRNA levels-last paragraph of page 12 to third paragraph of page 14. Perhaps the authors could make this a bit clearer?

We have modified this section to make this clearer, a point also raised by reviewer 3. We hope the new version rectifies this problem. In brief, we have used the rationale (used by the Lis group in Booth et al, Genome Research 2016) that mRNA levels should generally correlate with nascent transcript levels if rates of elongation are unchanged. Major changes in elongation rate can affect the correlation between nascent transcription and mRNA expression, as suggested in Booth et al in

the context of Spt4 knockout S. pombe. Our point here is that in Spt5-depleted MEFs, changes in nascent transcription and mRNA expression were still well-correlated suggesting that a major global change in elongation was not occurring in these cells. Since this was only a suggestive analysis, we performed the rate experiments to directly assess the elongation rate at several genes.

Non-essential
It would be nice to know what is happening to RNA processing when Spt5 is depleted.

Although we have not undertaken an analysis of mRNA capping or mRNA stability, we discuss in the revised manuscript that a global alteration in mRNA processing and turnover is unlikely, at least under our experimental set-up. Several observations suggest this: (1) GROseq (nascent transcription) and mRNAseq changes in Spt5-depleted cells are generally well correlated (Figure S5D), (2) our new analysis of splice junction usage shows no significant change upon Spt5 depletion, and (3) the relatively small number of differentially expressed genes (591 genes, Figure 5A). However, we cannot rule out that some mRNAs will be subjected to Spt5-mediated processing and we are currently investigating this as a follow-up project.

Referee #3:

In the present manuscript, Fitz at al. use genome-scale approaches along with single-gene assays to investigate the role of the highly conserved elongation factor Spt5 during RNA polymerase II (Pol II) transcription in mammalian cells. The authors deplete Spt5 from mouse embryonic fibroblasts using the Cre-loxP system and study the genome-wide consequences on Pol II occupancy, the Pol II CTD phosphorylation state, nascent transcription and total RNA levels. ChIP-seq and GRO-seq analyses reveal a strong decrease of Pol II occupancy at the promoter-proximal regions and an increase over the proximal gene body regions genome-wide. The authors further show by ChIP-seq that the CTD-Ser5P are decreased and CTD-Ser2P levels are increased at the 5'-end of genes upon Spt5 depletion. Furthermore, the authors provide evidence for Pol II occupancy to be strongly decreased at the distal gene body region (>20 kb) of long genes. Since the authors found no evidence for reduced Pol II transcription elongation rates upon Spt5 depletion, the decrease in Pol II occupancy rather reflects a reduction of transcribing Pol II complexes. These observations led to the conclusion that Spt5 is important for the processivity of RNA polymerase transcription >20 kb downstream of the TSS at long mammalian genes.

This study uncovers new interesting aspects of the role of Spt5 during Pol II transcription elongation. The finding that Spt5 is required for processive transcription 20 kb downstream of the TSS at long mammalian genes is a significant result and will be of broad interest for the transcription community. This work will inspire future studies to further characterize the Spt5-dependent proposed processivity checkpoint. However, in its present form the manuscript is in parts hard to read and needs to be improved.

Major comments:
(1) Spt5 is a key elongation factor that closely interacts with many other elongation and termination factors during Pol II transcription, including Spt4 and Paf1. Thus depletion of Spt5 per se will trigger pleiotropic effects. A main weakness of the present study is the long Spt5 depletion time (72 hours). This long treatment time further increases the risk of indirect effects and cellular adjustments, and complicates the assignment of Spt5-specific functions. The authors provide no information on whether other components of the Pol II elongation complex, such as Spt4 and Paf1, are affected upon Spt5 depletion. The author should address these limitations in the main text of this study. Conclusions, such as on page 20 „defined the precise role of Spt5", should be softened accordingly.

We have performed protein analysis of Spt4, Paf1, Spt16 (FACT), H3K36me3 (a modification associated with elongation) as well as the major initiation factors, TBP and TFIIB. We find no significant change in their levels upon Spt5 depletion (Figure S3). We also note that our mRNAseq analysis did not reveal any known elongation factors, chromatin remodelers or histone modifying enzymes to be differentially expressed upon Spt5 depletion. We also note that the phenotype seen in Spt5-depleted cells is different from that reported upon depletion of the PAF complex (Chen et al.,

Cell 2016) wherein Pol II occupancy was not altered in the promoter-proximal region but was increased throughout the gene body in a subset of genes. Nevertheless, we agree with the reviewer that we cannot rule out that some other factor(s) not probed in this study at the protein level could contribute to the elongation phenotype. Accordingly, we have softened the conclusion, as requested.

(2) The authors conclude on Page 8 that upon Spt5 depletion "Pol II undergoes normal initiation". This is not in line with several observations presented in the manuscript. ChIP-seq analyses clearly show a strong decrease of Pol II and Pol II Ser5P occupancy upon Spt5 depletion around the TSS. This suggests that transcription initiation seems also to be affected when Spt5 is depleted. This should be addressed in the main text and the conclusion needs to be softened.

We apologize for the lack of clarity here and we have elaborated on this in the text. The decrease in Pol II in the TSS region is best explained by the reduced pause duration caused by the loss of Spt5 and NELF. This is in line with similar observations and conclusions from multiple studies where Spt5 or NELF were depleted. The loss of Ser5P-Pol II signal in the TSS region correlates with the loss of Pol II (Figure 2) which, along with the fact that S5P-Pol II protein levels were not altered (Figure S3), led us to conclude that the reduced Ser5P-Pol II ChIP signal around the TSS reflects the reduction in steady-state paused Pol II rather than a defect in phosphorylation of Ser5P-Pol II. Regarding initiation, we note that if initiation was decreased, that is, if less Pol II were recruited to the PIC, then we would have also observed decreased Pol II levels in the TSS-proximal gene body and, consequently, decreased transcription throughout the gene. Instead, we see increased Pol II and nascent transcription in the TSS-proximal gene bodies genome-wide, which is expected if pause residence time is reduced leading to more rounds of initiation. This inverse relationship between pausing and initiation has recently been elucidated in detail by two elegant studies (Shao & Zeitlinger 2017 and Gressel et al 2017), which we discuss in the manuscript. The reviewer is right, however, that "normal initiation" is an inaccurate conclusion to draw and we have changed this. Thank you for pointing it out.

(3) A major finding by Fitz et al. is that Spt5 is important for the processivity of Pol II transcription past approximately 20 kb. This transition in Pol II occupancy (Fig. 3A) should be characterized in more depth. How many active genes show this transition in Pol II occupancy? What is the exact location (distance to TSS on average) of this transition? How does the location vary between different genes?

Thank you for raising this point. We now show more details of this transition zone in new Figure 3C-D. As seen in Figure 3C, the average $Spt5^{dep}$/WT ratio clearly decreases between ~13 kb and 23 kb from the TSS (the region within the gene used for this analysis), and this is observed at 70% of genes >50 kb in length. However, there is considerable variation (spikiness) in the signal in this region making it difficult to assign an exact location for the transition, although we assign an average value of ~17 kb (Figure 3C). The variation most likely arises from the fact that the signal decreases gradually rather than at a single point, and that the change is not uniform at all points between different genes. For this reason, we called it a transition <u>zone</u> occurring between 15-20 kb from the TSS.

(4) As mentioned earlier, the manuscript is at parts hard to read and lengthy. For instance, the authors conclude from pages 12 to 14 that Spt5 depletion does not affect Pol II transcription elongation rates. This is then again presented and now in a more convincing way on pages 14 to 16. Both sections (pages 12 to 16) should be combined and at the same time shortened.

We agree with reviewer and we have both shortened and combined this section of the results. We hope the revised version is clearer.

(5) Figure 1 A, B: the decrease in Pol II levels is much stronger in case of ChIP-seq as compared to GRO-seq measurements. What is the reason for this discrepancy? This needs to be explained in the main text.

It is true that there is not an absolute correlation in the metagene analysis between the fold-change in GROseq and Pol II ChIPseq promoter-proximal pause peak upon Spt5 depletion. In contrast, the fold-changes in gene bodies are more comparable between these assays (as seen in Figure 2A). This is likely due to the technical differences between GROseq and ChIPseq. Whereas GROseq measures

nascent transcription, Pol II ChIPseq measures Pol II occupancy. In other words, GROseq signal derives from all forms of transcribing Pol II whereas ChIP is biased towards paused or very slow-moving Pol II (as in the TSS region) since ChIP is highly dependent on the cross-linking efficiency of protein to DNA.
A high correlation between the two techniques is expected if reduced Pol II occupancy in a region results in a corresponding reduction in nascent transcription. This is not the case in Spt5-depleted cells where reduced pause duration is associated with increased initiation and increased nascent transcription. Here, the Pol II complexes entering the gene spend less time at the pause site. This could cause the lack of correlation between the decrease in ChIPseq versus GROseq signal in the paused region. In contrast, in gene bodies, where a high occurrence of pausing is rare and Pol II moves rapidly, the changes in Pol II occupancy and nascent transcription are comparable between WT and Spt5-depleted cells. We have explained this in the main text.

Minor comments:
(1) Formatting error on Page 2, first paragraph: 5.6-dichloro-1-__-D-ribofuranosylbenzimidazole; similar formatting errors also on page 22

This has been rectified, thank you.

(2) Page 2, second parapgraph: „Promoters initially recruit Pol II to assamble...". This statement is confusing since general transcription factors including TBP recruit Pol II to the promoter region. This statement needs to be changed.
This has been rectified, thank you.

(3) Typo on Page 2, second paragraph: „P-TEFb inhibition leads a genome-wide accumulation..." needs to be replaced by „P-TEFb inhibition leads to a genome-wide accumulation..."
This has been rectified, thank you.

(4) Page 2/3 first paragraph: „Importantly, Spt5 remains associated with Pol II until termination...". The original papers need to be cited that support this statement.

The citations have now been included.

(5) Page 5, second paragraph: the correlation coefficients between the replicates need to be added.

This has now been included.

(6) Page 8, second paragraph: CTD Ser5P is not a mark of „early initiation", as stated by the authors, but rather of early elongation. This needs to be changed.

Thank you for this point – we have made the change.

(7) Page 19, first paragraph: „no further loss of Pol II ... was discernable until the TTS, even at very long genes." This statement is too vague. Please add gene lenghtes in kb.

We agree with the reviewer and we have performed a new quantitative analysis looking at fold-changes ($Spt5^{dep}$/WT) of Pol II between different regions within genes (new Figure 3D). This analysis shows that beyond the transition point, the fold-changes in Pol II remain relatively unchanged up to the TTS.

(8) Page 20, third paragraph: „Since yeast genes are much shorter than mammalian genes...". This is again too vague. Please add average gene lengthes in kb.

This sentence has been modified to include average gene lengths.

(9) Figure 3A: an arrow at the 20 kb region right at the transition of the Pol II occupancy would be helpful.

We have done as requested.

(10) Figure 3B (upper panel) and 6B: Differences between WT and SPT5dep can hardly be seen. Zoom-in views would be helpful.

We have done as requested. The zoom-in views are presented as insets in Figure 3B and 6B. We note that the Chd2 gene in Figure 6B overlaps another gene and hence two peaks are seen. Upon further investigation, we realized that as per the latest annotation of the mouse genome (mm10), Chd2 has a longer isoform that begins precisely where the first Pol II peak is seen in our data. Since we have used the previous mm9 mouse genome annotation for all our analyses in this study, we have not shown the isoform seen in mm10 version.

(11) Table S1 and S2 as stated in the main text are missing.

We apologize for this omission – the tables are now included (Tables EV2 and EV3).

(12) No ArrayExpress or GEO identifier for the genome-wide data is provided.

We have uploaded all raw sequencing data to GEO (accession number GSE106313). We apologize for not including it before.

2nd Editorial Decision 4th January 2018

Thank you for submitting a revised version of your manuscript to The EMBO Journal. It has now been seen by two of the original referees and their comments are shown below.

As you will see, the referees both find that all major criticisms have been sufficiently addressed and they recommend the manuscript for publication, pending a few minor clarifications. I would therefore like to invite you to submit a final revision of the study in which you clarify/elaborate on these few remaining points.

Thank you again for giving us the chance to consider your manuscript for The EMBO Journal, I look forward to receiving your final revision.

REFEREE REPORTS

Referee #1:

I am happy with the revisions and comments and clarifications made by the authors. I just have a few additional suggestions that I think the authors could use to improve the final version.

1. In EV Figure 1G, it appears that the order of the lanes are switched for the Spt5 blots on the left side of the image. Strange, but the trend for WT Spt5 is the opposite of that seen for the control, histone H3.

2. In EV Figure 4 and elsewhere, it is still very hard to visualize any changes in GRO-Seq reads or RNAPII occupancy in the gene bodies, because the signal is so small due to the scale being defined by the TSS region. In a few places a magnified image is shown and perhaps this can be done throughout.

Referee #3:

The authors have addressed all of my concerns. Most notably, Fitz et al. performed additional experiments to show that protein levels of other integral subunits of the Pol II transcription elongation complex, including the levels of the SPT5 interactor SPT4 and PAF1, are not significantly changed upon SPT5 depletion (Fig. EV3). This suggests that indirect effects that may arise from the SPT5 depletion are minimal and that the observed transcriptional changes are most

likely SPT5-specific. Furthermore, the authors performed addtional bioinformatics analyses to reveal more interesting details regarding the transition of Pol II occupancy upon SPT5 depletion (Fig. 3 C,D). For instance, the authors now show that SPT5 is required for processive transcription elongation at the majority (70%) of long genes.

The authors now also added the supplementary tables (EV1: deep sequencing statistics; EV2: qPCR primer sequences; EV3: information on antibodies) and provide a GEO identifier for the genome-wide data.

Overall, I feel that the additional analyses and clarifications by Fitz et al. have strongly improved the manuscript.

2nd Revision - authors' response 10th January 2018

We thank the reviewers for their positive comments on the revision of our manuscript. Below, we have addressed the additional comments and suggestions.

Referee #1:

I am happy with the revisions and comments and clarifications made by the authors. I just have a few additional suggestions that I think the authors could use to improve the final version.

1. In EV Figure 1G, it appears that the order of the lanes are switched for the Spt5 blots on the left side of the image. Strange, but the trend for WT Spt5 is the opposite of that seen for the control, histone H3.

It is true that the titration looks odd but we have not flipped it or made any error in loading. The issue is a recurrent one only when we load higher amounts of extract to probe for Spt5, and that too, only in WT cells (note that we don't see this in the Spt5-dep lanes). You can notice the streaking in the higher titration lane in Figure EV1 which we think is indicative of some loss of protein within the gel. We also occasionally observe this issue with Pol II, as seen in Figure EV3A. We can (and have) loaded less extract which resolves this problem but then we don't see any Spt5 in the Spt5-dep sample giving the false impression that Spt5 has been completely depleted. We know from the ChIP-qPCR data in Figure EV1H that very low amounts of Spt5 are still chromatin-bound. Thus, it was essential to do the Western blot analysis with higher amounts of extract to detect the residual Spt5 and present the data more accurately. Importantly, what this effectively means is that the magnitude of Spt5 depletion is actually greater than what we measure from this blot.

2. In EV Figure 4 and elsewhere, it is still very hard to visualize any changes in GRO-Seq reads or RNAPII occupancy in the gene bodies, because the signal is so small due to the scale being defined by the TSS region. In a few places a magnified image is shown and perhaps this can be done throughout.

We have now shown magnified images throughout.

Referee #3:

The authors have addressed all of my concerns. Most notably, Fitz et al. performed additional experiments to show that protein levels of other integral subunits of the Pol II transcription elongation complex, including the levels of the SPT5 interactor SPT4 and PAF1, are not significantly changed upon SPT5 depletion (Fig. EV3). This suggests that indirect effects that may arise from the SPT5 depletion are minimal and that the observed transcriptional changes are most likely SPT5-specific. Furthermore, the authors performed addtional bioinformatics analyses to reveal more interesting details regarding the transition of Pol II occupancy upon SPT5 depletion (Fig. 3 C,D). For instance, the authors now show that SPT5 is required for processive transcription elongation at the majority (70%) of long genes.

The authors now also added the supplementary tables (EV1: deep sequencing statistics; EV2: qPCR primer sequences; EV3: information on antibodies) and provide a GEO identifier for the genome-wide data.

Overall, I feel that the additional analyses and clarifications by Fitz et al. have strongly improved the manuscript.

Accepted 15th January 2018

Thank you for submitting the final revision of your manuscript to The EMBO Journal, I am pleased to inform you that the study has now been officially accepted for publication here.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Rushad Pavri
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2017-97965

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

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

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

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

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

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

B- Statistics and general methods

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

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	Sample size was not pre-chosen
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA. Animals were only used to generate mouse primary embryonic fibroblasts from d13.5 embryos
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	No animals were excluded for mouse primary embryonic fibroblasts preparation
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA. Animals were not subjected to any treatment.
For animal studies, include a statement about randomization even if no randomization was used.	No randomization used
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	No blinding was used
4.b. For animal studies, include a statement about blinding even if no blinding was done	No blinding was used
5. For every figure, are statistical tests justified as appropriate?	Yes. We use non-parametric two-tailed t tests which are appropriate for analyses of differences in gene expression and transripton factor occupancy between different genotypes of mice. This is stated in the manuscript and figure legends
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes.
Is there an estimate of variation within each group of data?	Yes
Is the variance similar between the groups that are being statistically compared?	Yes

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	This information is in Table EV2 of the manuscript
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

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

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

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	This information is detailed in the Materials and Methods section.
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	All animal experiments were conducted as per the ethical guidelines laid down by the animal facility in compliance with all Austrian and EU law.
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

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

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	All raw next-generation sequencing data (GROseq, ChIPseq and mRNAseq) have been deposited in GEO under accession number GSE106313.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	Data is deposited in GEO (see above)
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA