THE EMBO JOURNAL

Manuscript EMBO-2011-77186

Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription

M. Behfar Ardehali, Amanda Mei, Katie L. Zobeck, Matthieu Caron, John T. Lis and Thomas Kusch

Corresponding author: Thomas Kusch, Rutgers University

Review timeline:

Submission date:	02 February 2011
Editorial Decision:	09 March 2011
Additional correspondence:	14 March 2011
Additional correspondence:	17 March 2011
Revision received:	11 April 2011
Editorial Decision:	10 May 2011
Revision received:	15 May 2011
Editorial Decision:	20 May 2011
Accepted:	20 May 2011

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 09 March 2011

Thank you for submitting your manuscript for consideration by The EMBO Journal. It has been now been evaluated by two referees and I enclose their reports below. As you will see the referees find the identification of dSet1 as the major functional H3K3 methyltransferase in Drosophila to be potentially interesting, however, they require further experimental analysis to make it suitable for The EMBO Journal. Both referees raise a number of important issues regarding controls that are required to strengthen the current data in the manuscript. Referee #2 also suggests that global expression profiling be performed to reveal the extent of dSet1 function in effecting gene expression, I believe that this would also significantly strengthen the main conclusion of the study. Given the interest in the study, should you be able to address the concerns we would be happy to consider a revised version of the manuscript for publication in The EMBO Journal.

I should remind you that it is EMBO Journal policy to allow a single round of revision only and that, therefore, acceptance or rejection of the manuscript will depend on the completeness of your responses included in the next, final version of the manuscript. When you submit a revised version to the EMBO Journal, please make sure you upload a letter of response to the referees' comments. Please note that when preparing your letter of response to the referees' comments 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 initiative, please visit our website: http://www.nature.com/emboj/about/process.html

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

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1 (Remarks to the Author):

The work of Ardehali et al. reports the identification and characterization of dSet1, a new Drosophila Set1-related H3K4 methyltransferase. Starting from database searches, the authors identified one ORF, CG40351, as a putative Set1 homolog in Drosophila. The identification of the associated factors after purification from nuclear extracts revealed that CG40351 encode a component of a complex similar to hCOMPASS, the human Set1-containing complex. Using a combination of RNAi-mediated knockdown studies and in vitro enzymatic activity assays, the authors conclude that the dSet1 plays a major role in H3K4 trimethylation in Drosophila. Colabeling experiments on polytene chromosomes with specific antibodies shows an extensive overlap of dSet1 and H3K4me3 and that many of the sites positive for transcribing Pol II get dSet1 signals. Moreover, the association of dSet1 with chromatin appears to require dCfp1, one component of the dCOMPASS. The functional impact of the RNAi-mediated knockdown of dSet1 was appreciated through the significant decrease of the RNA levels of several genes which correlates with the diminution of the amount of H3K4me3 at the transcription start site. The link between dSet1 and transcriptional activation is further documented by showing the recruitment of an EGF-dSet1 fusion protein to the hsp70 loci after induction through heat shock. FRAP studies suggest that the association of dSet1 with different active transcription loci is dynamic. Last, the effect of dSet1 depletion on the distribution of Pol II along activated hsp70 loci, with a relative increase and reduction at the hsp70 5'-end region and downstream regions respectively, let the authors to propose that dSet1 positively controls the rate of Pol II release into productive elongation.

Outside the identification and characterization of a new Set1 H3K4 methyltransferase, this paper addresses important questions about the specific roles of the various type-2 histone lysine methyltransferases present in Drosophila and the functional role of the H3K4 trimethylation mediated by dSet1 during transcription. This all the more important that the functional significance of H3K4 methylation is hampered by the existence of several genes encoding H3K4 trimethyltransferases.

However, in its state, many conclusions are not supported by the data, mainly because the experiments are poorly controlled. I have listed a list of major and minor points that should be addressed to strengthen the conclusions drawn by the authors.

Major points

1) The fact that the protein corresponding to CG40351 is the Drosophila homolog of Set1 is based only on the fact that a tagged version copurifies with proteins that are themselves homologs of the components of the human COMPASS. Some additional informations are required about dSet1 itself to be more consistent about this homology. Thus, the phylogenetic comparison (Fig. 1A) is clearly not sufficient and a sequence alignment of dSet1 with other Set1 proteins is missing, with a special emphasis on regions, such as the RRM domain, which are characteristics of the Set1 family of histone H3K4 methyltransferases (see Lee et al, JBC 2007, and Tresaugues et al, J Mol Biol, 2006).

2) In Figure 1, the gel of the final purification should be shown.

3) Many experiments rely on the use of the dSet1 antibody. This antibody should be better characterized in Western blots and in IF. For instance, Figure S1 only suggests that the antibody is

specific. To be sure that the antiserum is specific, they should test the anti-dSet1 in cells transfected by FH-Set1 and by FH-Set1 carrying a deletion.

The specificity of the antibody in IF is characterized in Figure 5A. However, a control showing that the signal disappears when dSet1 is KD is missing.
The tested antibody has been raised against which antigen ?

4) It is very difficult to obtain even semi-quantitative data from the Figure 2 (panels A and B). For Figure 2A, would it be possible to have an estimation of the RNAi-mediated depletion ?

For Figure 2B, I don't understand why the signals obtained with the anti-H3K4me1, me2 and me3 are so weak in the lacZ control. It is well known that these antibodies are very sensitive to the level of histone loaded on the gel. Even if the effect of the various KD are visible, it is difficult to obtain any quantitive data from the Western blot.

Does the « coomassie » panel in figure 2B represent the histone extract ? In such a case, the signal with the anti-methyl H3K4 should be really stronger (compare for instance the native fly histones (nCH) of figure 2C).

5) Page 7 bottom and top of page 8, why do the authors restrict the role of dSet1 in H3K4me3 while KD of dSet1 also strongly affects H3K4me2 ? See also figure 3.

6) For Figure 3 see my previous comment on the characterization of the dSet1 anti-serum. Also, one can wonder whether the resolution level of polytene chromosome analysis (Figure 3) is sufficient to definitively conclude as the authors do. It appears that the colocation of dSet1 and active Pol II is not so evident as dSet1 signals exceed massively those of Pol II. The description of the results is too much qualitative : are quantitative data conceivable?

When the authors compare the distributions of dSet1 and transcriptionally active Pol II on polytene chromosomes by indirect immunofluorescence microscopy (Figure 3D), how do they know that PolII is active or poised. Do they use H5 or H14 antibodies ?

The fact that dSet1 is "found in a large number of DAPI-negative interband regions that do not contain detectable amounts of Pol IIo" is not discussed. Does this means that dSet1 could be recruited to chromatin independently of Pol II and/or that a consequent fraction of H3K4 is trimethylated independently of transcription ?

7) All the conclusions made from Figure 3E (eg dCfp1 is required for dSet1 association with chromatin and H3K4 trimethylation at transcription puffs) cannot be done from the data shown in Figure 3E. This panel should be removed or properly controlled.

Figures 4 and 5 are nice figures.

8) It would be interesting and informative to perform the same experiment shown in Figure 6C with PolII phospho-specific antibodies.

Figure 6C suggests that dSet1 might have a positive role in the release of Pol II into productive elongation. In yeast, it has been reported that inactivating Set1 catalytic activity results in changes in the distribution of RNAPII on the MET16 coding région (Santos-Rosa et al., Mol Cell, 2003) but in the opposite way than those shown in Figure 6C (which is rather convincing). Could the authors comment.
Figure 6 is difficult to follow. A color figure would help.

Minor points :

1) page 3, line 4 : the reference for the distribution of H3K4me3 in yeast (Pokholok et al., 2005) is missing. In general, reference to yeast Set1 is made many times, however the MS does not contain any reference about yeast Set1.

2) page 3, line 8-9 : it is said that in S. cerevisiae the mutation of SET1 does not cause obvious transcription defects. This is true in mitotic cells but not in meiotic cells where the expression of a whole set of genes, i.e. the middle meiotic genes, is affected by the loss of Set1 (Sollier et al., 2004).

3) The generation of the cell lines expressing expressing an N- or C-terminally FLAG/HA-tagged CG40351 should be described somewhere in the MS.

4) It is stated that the combined theoretical mass of all dSet1 complex subunits is 632.8 kDa. Do the authors assume that each protein of the complex is present as a monomer.

5) Considering the data presented on Figure 2B, the consequences of Trr and Trx depletion look relatively similar with a defect in the three forms of H3K4 methylation, so the differences underlined by the authors on page 6 (lines 18-19) appear exaggerated.

6) In vitro the FHdSet1 is able to monomethylate H3K4 (Figure 2C) while its depletion in vivo has no effect. Please comment.

7) When the authors observe that « the catalytic SET domain of dSet1, unlike the complete complex, is ineffective as KMT », they may cite the corresponding references that show this in other species.

8) In figure 4, could the authors describe the expression status of the chosen expressed genes (strong expression?).

9) Figure 3D : the signification of the two arrowheads on the left part is not specified.

10) Figure 6B, bottom : the squares must be ordered as on Fig. 6C. Figure 6C : why the results of the control NHS not presented as on Figure 6B.

11) One main conclusion is that, among all the described histone H3K4 methyltransferases, dSet1 is responsible for the bulk of H3K4 trimethylation in a way that is independent from the activity of Trr and Trx but we think that alternative view is conceivable. In fact the depletion of dSet1 leads to a stringent decrease of both H3K4me2 and -me3, with H3K4me1 staying unchanged, whereas all the methylated forms are affected through Trr or Trx depletion (Fig. 2B). This could be interpreted as if the various KMT2 act successively with the product of Trr and/or Trx (H3K4me1) being the substrate for dSet1. This fits with the fact that the recombinant Trr GST-Set domain display only a monomethylase activity in vitro (Fig. S2). The in vitro monomethylation of H3K4 in the presence of purified dSet1 complex (Fig. 2C) could result from the presence of copurified monomethylase (Trr or Trx) activity. It would be interesting to mix the recombinant dSet1 and Trr GST-Set domain to see if some di-/tri-methylated H3K4 could be detected in this case. Also the analysis of polytene chromosome, similar to that presented on Fig. 3, in conditions where either Trr or Trr expression is affected, if the corresponding flies exist in the collection, would be informative. Similarly, the effect of Trr or trx depletion on the transcription levels of the genes tested on Figure 4 is required.

Referee #2 (Remarks to the Author):

Review of "Drosophila dSET1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription" submitted by Ardehali et al.

In their study the authors show for the first time that dSET1 and not Trx or Trr is the predominant histone H3K4 trimethyltransferase in Drosophila. Association of dSET1 with actively transcribed genes, binding to promoter regions, involvement in heat shock response and possibly also dependence of Pol II elongation on dSet1 are the main findings supporting a central role of dSet1 in H3K4 trimethylation in Drosophila.

Homology between the Drosophila CG40351 gene and the yeast SET1 or human SetB genes suggested that CG40351 encodes the Drosophila homolog of SET1. Purification of the dSet1 complex revealed the presence of a complex in Drosophila, which is comparable in composition with the human COMPASS complex.

Comparative analysis of RNAi-mediated knock down of Trx, Trr and dSet1 suggests preferential control of H3K4me2 and H3K4me3 by dSET1 in S2 cells. Polytene chromosome analysis further indicates predominant overlapping of dSet1 and H3K4me3. It is important to show by DNA staining that the overlap appears to occur preferentially in interbands. Such data are presented with Fig. S3 but this Figure is not cited in the text. This staining pattern also raises the question whether dSET1 actively associates or even binds with H3K4me2 and me3, which would be expected if it has a maintenance function. Consequently presence of H3K4me3 and binding of dSet1 would not indicate genetic activity. Preassembly of Pol II with interband regions was already shown long ago and this might be correlated with discrete H3K4me3 and dSet1 at interbands although no transcriptional activity is found. However colocalization of dSet1 with Pol II sites at polytene chromosomes might also simply be due to its abundance and chromosome wide binding (low resolution of immunostaining). Binding of dSet1 to puffs is demonstrated in Fig. S4 (narrower signal than Pol II band) but again the Figure is not cited in the text. Consequently the authors should recheck whether they indicated correctly all the very valuable supplementary Figures in the text of the manuscript.

The authors do not provide sufficient data demonstrating the mutant nature of the P element insertion for dCfp1 or any convincing data proving that this mutant abolishes chromosomal association of dSet1.

Effects on optimal expression in dSet1 KD cells are demonstrated for 7 randomly chosen genes. Although the data suggest a general role of dSet1 on efficient transcription of a wide set of genes finally genome wide expression data would have provided a much deeper insight in the suggested central role of dSet1 in control of gene activities.

Significant reduction of promoter associated H3K4me3 is clearly found in the four genes selected for the studies. It is not explained why the genes have been selected for expression and for ChIP analysis of promoter specific effects of dSet1 KD on H3K4me3.

Heterochromatin localization of EGFP-dSet1 is not seen in Fig. S8. Do the authors indeed see any association with chromocenter heterochromatin? Fig. S8 suggests just the opposite. The quality of Fig. 5C to show definitely recruitment of EGFP-dSet1 to hsp70 after heat shock is not sufficient. This would need at least good chromosome squashes. However the data presented by Fig. 6 are clear-cut and strongly support recruitment of dSet1 to the hsp70 and hsp26 promoters.

In summary the manuscript provides interesting new data about epigenetic control of H3K4 methylation by dSet1 in Drosophila and shows for first time a predominant role of the dSet1 methyltransferase in control of H3K4me3 dependent gene activity. The epigenetic mechanisms uncovered are of general interest and the data presented are novel and have a general scientific interest.

Additional correspondence (author) 14 March 2011

Thank you for your email inviting us to review a manuscript (EMBOJ-2011-77376) that is highly complementary to our submitted manuscript characterizing the main H3K4 methyltransferase, dSet1 (EMBOJ-2011-77186). Thank you also for overseeing the review of our paper and providing us the reviews that were generally positive and contained useful criticisms and suggestions. Upon reading this new manuscript, we realized how important it has become that the scientific community be made aware of the major role of dSet1 in H3K4me3 in flies (see also point 2 below, which provides another, most recently published study that would have greatly benefited from our findings). The manuscript as such contains major misinterpretations regarding the main H3K4 methyltransferases in flies, which would be eliminated if our data was available to the authors. We hope this underscores the impact of our work for many other labs and the scientific community.

The reason why we are writing to you is that we realize that we will be able to address all criticisms of the reviewers within less than one month, but not one suggestion by Reviewer 2. However, because you emphasized the reference of reviewer 2, to his suggestion of a genome-wide transcription study in cells depleted for dSet1, and indicated that you believe this would significantly

strengthen the study, we would like to outline why we feel that this would require a set of complementary experiments, which would require considerable delay and would go beyond the scope, length, and figure capacity of this paper.

We are of the opinion that the genome-wide transcription data would contribute to the manuscript's main 'take-home' message only if performed with additional genome-wide factor distribution experiments. The latter would require at least several months to obtain and analyze and would generate massive amounts of important additional data. We also would like to point out that a genome-wide study was published during the review process on Feb 9 (see point 2 below for details), which deals with the Ash2 component of the dSet1 complex subunit that we identified by proteomics. The data fully supports (but does not scoop) our findings and makes our detailed mechanistic studies even more interesting.
Since it is obvious that a swift publication of our work is of highest priority, would you consider our manuscript with all requested improvements and the discussion of the recently published transcriptome studies?

Here a more detailed list of arguments that we hope are helpful in your decision:

1) RNAi/transcription studies require dsRNA treatments for several days to effectively knock down dSet1. This will cause indirect transcription deregulation of a substantial number of genes. Only in combination with Chromatin Immunoprecipitation experiments with anti-dSet1 and -H3K4me3 antibodies, the data could be interpreted accurately. Furthermore, without ChIPs for Pol II, Pol II Ser5P, and Pol II Ser2P, a transcription study also would not really add to our main "take-home message" (H3K4me3 by dSet1 is required for efficient release of Pol IIo from promoters). We hope that you agree with us that a transcriptome analysis together with about 12-14 confirming ChIP experiments (without controls and duplicates) goes far beyond the requested additions and would also generate data sufficient for a 2nd manuscript.

2) A most recent genome-wide publication in Nucleic Acid Research (March 2011 issue; "Genome-wide chromatin occupancy analysis reveals a role of ASH2 in transcriptional pausing" Perez-Lluch et al., NAR, 39(4), online Feb 9; in press Mar 11, 2011) describes the genome-wide distribution of Ash2. We confirmed that Ash2 is a subunit of the fly dSet1 complex. The authors determined on genome-wide levels transcription defects and H3K4me3 changes in Ash2 mutants. When comparing the data to already existing transcriptome studies, they observed extremely high correlation between transcription defects of high level expressed genes and Ash2 dependency for H3K4me3. This is fully consistent with our findings on genes that we selected based on their different expression levels (we btw did not only study 7 genes as suggested by reviewer 2, but analyzed 9 more in the supplement. In addition, we performed detailed activation time course studies on 3 heat shock loci, bringing the total number to 19 genes). Most importantly, the authors misinterpreted these defects assuming that Ash2 (a confirmed Set1-binding partner from yeast and humans), might associate with Trx or Trr, although this interaction was never shown. As already mentioned, we actually show for the first time that the dSet1 complex contains Ash2 (Fig. 1C), and that dSet1 is the main H3K4 trimethyltransferase. Therefore, the observed global defects in Ash2 mutants are most likely due to defects in dSet1 complexes. This makes our mechanistic studies even more interesting as this work is completely independent, yet fully consistent, with our findings.

We hope that you agree with us that a speedy publication of the work would be of great benefit for the scientific community and we are looking forward to hearing back from you soon.

Additional correspondence (editor) 17 March 2011

Thank you for your letting regarding your revised manuscript at The EMBO Journal, I am sorry it has taken longer than I would have liked to respond to it.

I have read through the letter in detail and in short I am happy for you to discuss the recently published data on Ash2 instead of performing the transcriptome analysis and required genome-wide binding analysis. The main reason why I stressed the importance of this was that it would have been good to show that there

are global effects on gene expression consistent with dSet1 being the main H3K4 methyltransferase, but I agree that additional experiments that would have to be done to demonstrate that these are direct effects which would take a long time to acquire.

So I agree with your proposal for the revised manuscript.

Yours sincerely,

Editor
The EMBO Journal

1st Revision - authors' response 11 April 2011

Referee #1:

Major points

1) The fact that the protein corresponding to CG40351 is the Drosophila homolog of Set1 is based only on the fact that a tagged version copurifies with proteins that are themselves homologs of the components of the human COMPASS. Some additional informations are required about dSet1 itself to be more consistent about this homology. Thus, the phylogenetic comparison (Fig. 1A) is clearly not sufficient and a sequence alignment of dSet1 with other Set1 proteins is missing, with a special emphasis on regions, such as the RRM domain, which are characteristics of the Set1 family of histone H3K4 methyltransferases (see Lee et al, JBC 2007, and Tresaugues et al, J Mol Biol, 2006).

We added a comparison of the domain structure of all factors shown in the phylogenetic tree in supplemental figure S1. This figure shows the conserved RRM and postSet domains, which also was added to the main text. The suggested literature was cited.

2) In Figure 1, the gel of the final purification should be shown.

Due to spatial constraints, we decided to include this figure in the supplement (Fig. S2). It also must be emphasized that the mass-spectrometric was performed as mudPit analysis, and not by excising and analyzing single bands. Therefore, the Silverstain is of limited information value, and it would be inaccurate to assign identified proteins to visualized bands.

3) Many experiments rely on the use of the dSet1 antibody. This antibody should be better characterized in Western blots and in IF. For instance, Figure S1 only suggests that the antibody is specific. To be sure that the antiserum is specific, they should test the anti-dSet1 in cells transfected by FH-Set1 and by FH-Set1 carrying a deletion. The specificity of the antibody in IF is characterized in Figure 5A. However, a control showing that the signal disappears when dSet1 is KD is missing. The tested antibody has been raised against which antigen ?

Figures further confirming the specificity of the antisera in both IF and immunoblotting have been added (Fig. S3, S6). Apart from the other RNAi data (Fig. 2A, 5, 7) we added a new figure (Figure 4) showing that dSet1 signals are eliminated on polytene chromosomes from dCfp1 mutants. Finally, we rephrased the materials and methods section describing the generation of recombinant proteins for better understanding (p. 22).

4) It is very difficult to obtain even semi-quantitative data from the Figure 2 (panels A and B). For Figure 2A, would it be possible to have an estimation of the RNAi-mediated depletion? For Figure 2B, I don't understand why the signals obtained with the anti-H3K4me1, me2 and me3 are so weak in the lacZ control. It is well known that these antibodies are very sensitive to the level of histone loaded on the gel. Even if the effect of the various KD are visible, it is difficult to obtain any quantitive data from the Western blot.

Does the « coomassie » panel in figure 2B represent the histone extract ? In such a case, the signal with the anti-methyl H3K4 should be really stronger (compare for instance the native fly histones (nCH) of figure 2C).

The band intensities were quantified by intensity scanning and the numbers were added. Also, the differences in the exposures between figures 2B and 2C were intentional. Our goal in the RNAi experiments was to avoid saturation effects by using strong exposures. This allowed for the detection of differences in H3K methylation changes in KDs for Trx or Trr, which were in some cases rather mild. By contrast, long exposures were chosen in Fig. 2C to ensure that none of the signals were due to contaminating KMTs. To further address this important point, we now added data in the supplement, which was generated using other methyl-H3K4 specific antibodies in dilutions (Fig. S4).

5) Page 7 bottom and top of page 8, why do the authors restrict the role of dSet1 in H3K4me3 while KD of dSet1 also strongly affects H3K4me2 ? See also figure 3.

We discussed this issue before in the text. Due to the presence of histone demethylases in these cells, any changes in mono- or dimethyl-H3K4 cannot be directly attributed to the loss of dSet1. In fact, our unpublished data indicates that these KDMs not only target a large number of dSet1-regulated genes, but also are responsible for the generation of the bulk of H3K4me2 from H3K4me3 generated by dSet1. Therefore, the loss of H3K4me2 in dSet1-depleted cells is not indicative for a direct role of dSet1 in the generation of H3K4me2. We also would like to emphasize that our intentions were to study transcription-linked H3K4me3 at TSS.

6) For Figure 3 see my previous comment on the characterization of the dSet1 anti-serum. Also, one can wonder whether the resolution level of polytene chromosome analysis (Figure 3) is sufficient to definitively conclude as the authors do. It appears that the colocation of dSet1 and active Pol II is not so evident as dSet1 signals exceed massively those of Pol II. The description of the results is too much qualitative : are quantitative data conceivable?

When the authors compare the distributions of dSet1 and transcriptionally active Pol II on polytene chromosomes by indirect immunofluorescence microscopy (Figure 3D), how do they know that PolII is active or poised. Do they use H5 or H14 antibodies ? The fact that dSet1 is "found in a large number of DAPI-negative interband regions that do not contain detectable amounts of Pol IIo" is not discussed. Does this means that dSet1 could be recruited to chromatin independently of Pol II and/or that a consequent fraction of H3K4 is trimethylated independently of transcription ?

A quantitative assessment is unfortunately not possible due to the limitations in the dynamic range of digital cameras, exposure times etc. However, we now added separate panels for Pol IIo(Ser5P) and Pol IIo (Ser2P), which show differences in colocalization (Fig. 3D, 3E). We also would like to point out that S5P is not indicative for paused polymerase alone, but also is found in elongating Pol IIo in higher eukaryotes (also see major point 8). The data also revealed that Pol IIo intensity levels are extremely dynamic, while dSet1 levels are more uniform. Therefore, the visualization of a complete overlap would require the overexposure of many transcription puffs. To avoid this, we magnified regions of moderate and low intensities, and enhanced the Pol II channels (see legend Fig. 3). Our new data now indicates that the vast majority of dSet1-positive bands indeed have weak, but detectable Pol IIo signals, which only could be visualized by adjustment of the Pol IIo channels.

7) All the conclusions made from Figure 3E (eg dCfp1 is required for dSet1 association with chromatin and H3K4 trimethylation at transcription puffs) cannot be done from the data shown in Figure 3E. This panel should be removed or properly controlled. Figures 4 and 5 are nice figures.

The previously not shown data has been added and completed and is shown in Figure 4 now.

8) It would be interesting and informative to perform the same experiment shown in Figure 6C with PolII phospho-specific antibodies. Figure 6C suggests that dSet1 might have a positive role in the release of Pol II into productive elongation. In yeast, it has been reported that inactivating Set1 catalytic activity results in changes in the distribution of RNAPII on the MET16 coding région (Santos-Rosa et al., Mol Cell, 2003) but in the opposite way than those shown in Figure 6C (which is rather convincing). Could the authors comment.

Figure 6 is difficult to follow. A color figure would help.

We attempted these experiments, but retrieved uninformative results. We would like to emphasize that in higher eukaryotes, the Pol II phosphorylation state of elongating and promoter-occupied Pol IIo is S5P. S2P is predominantly enriched in productively elongating Pol IIo with higher signals near the 3' end (e.g., see Boehm et al, MCB 23, pp. 7628-37 or Espinosa Genes Dev. 2006 Mar 1;20(5):601-12). We have now added color figures and discussed the work of Santos-Rosa et al.

Minor points :

1) page 3, line 4 : the reference for the distribution of H3K4me3 in yeast (Pokholok et al., 2005) is missing. In general, reference to yeast Set1 is made many times, however the MS does not contain any reference about yeast Set1.

These publications have now been cited.

2) page 3, line 8-9 : it is said that in S. cerevisiae the mutation of SET1 does not cause obvious transcription defects. This is true in mitotic cells but not in meiotic cells where the expression of a whole set of genes, i.e. the middle meiotic genes, is affected by the loss of Set1 (Sollier et al., 2004).

We clarified this.

3) The generation of the cell lines expressing expressing an N- or C-terminally FLAG/HA-tagged CG40351 should be described somewhere in the MS.

The description has been added to the material and methods section.

4) It is stated that the combined theoretical mass of all dSet1 complex subunits is 632.8 kDa. Do the authors assume that each protein of the complex is present as a monomer.

This has been clarified.

5) Considering the data presented on Figure 2B, the consequences of Trr and Trx depletion look relatively similar with a defect in the three forms of H3K4 methylation, so the differences underlined by the authors on page 6 (lines 18-19) appear exaggerated.

To address this important issue, we added quantitative measurements (Fig. 2, text) and new data with various methyl-H3K4-specific antibodies (figure S4).

6) In vitro the FHdSet1 is able to monomethylate H3K4 (Figure 2C) while its depletion in vivo has no effect. Please comment.

We discussed in greater detail that the purified KMT2 complexes are usually not efficient in fully trimethylating recombinant nucleosomal arrays and tend to generate substantial amounts of intermediates like H3K4me1 or –me2.

7) When the authors observe « the catalytic SET domain of dSet1, unlike the complete complex, is ineffective as KMT », they may cite the corresponding references that show this in other species.

We added these references.

8) In figure 4, could the authors describe the expression status of the chosen expressed genes (strong expression?).

We now mention this in the text and figure legend.

9) Figure 3D : the signification of the two arrowheads on the left part is not specified.

The arrows were duplicated during the generation of the pdf due to a glitch in the software. They have been removed in the updated figure.

10) Figure 6B, bottom : the squares must be ordered as on Fig. 6C. Figure 6C : why the results of the control NHS not presented as on Figure 6B.

The squares have been rearranged. We also added NHS data to this figure (now Fig. 7).

11) One main conclusion is that, among all the described histone H3K4 methyltransferases, dSet1 is responsible for the bulk of H3K4 trimethylation in a way that is independent from the activity of Trr and Trx but we think that alternative view is conceivable. In fact the depletion of dSet1 leads to a stringent decrease of both H3K4me2 and -me3, with H3K4me1 staying unchanged, whereas all the methylated forms are affected through Trr or Trx depletion (Fig. 2B). This could be interpreted as if the various KMT2 act successively with the product of Trr and/or Trx (H3K4me1) being the substrate for dSet1. This fits with the fact that the recombinant Trr GST-Set domain display only a monomethylase activity in vitro (Fig. S2). The in vitro monomethylation of H3K4 in the presence of purified dSet1 complex (Fig. 2C) could result from the presence of copurified monomethylase (Trr or Trx) activity. It would be interesting to mix the recombinant dSet1 and Trr GST-Set domain to see if so me di-/tri-methylated H3K4 could be detected in this case. Also the analysis of polytene chromosome, similar to that presented on Fig. 3, in conditions where either Trr or Trr expression is affected, if the corresponding flies exist in the collection, would be informative. Similarly, the effect of Trr or trx depletion on the transcription levels of the genes tested on Figure 4 is required.

We thank the reviewer for this interesting interpretation and experimental suggestions. Unfortunately, these experiments were conducted without any success. We also used purified Trr-monomethylated arrays in subsequent assays with GST-dSet1SET and did not observe any H3K4me2 or –me3. This is consistent with structure-function assays from others supporting that these domains at best can fit dimethylated H3K4 in their catalytic domains (e.g., Cosgrove and Patel, FEBS J 277, pp. 1832-42). In fact, loss-of-function and in vitro assays support that these enzymes depend on their full H3K4 trimethyltransferase activity on other complex subunits.

Referee #2 (Remarks to the Author):

Review of "Drosophila dSET1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription" submitted by Ardehali et al.

In their study the authors show for the first time that dSET1 and not Trx or Trr is the predominant histone H3K4 trimethyltransferase in Drosophila. Association of dSET1 with actively transcribed genes, binding to promoter regions, involvement in heat shock response and possibly also dependence of Pol II elongation on dSet1 are the main findings supporting a central role of dSet1 in H3K4 trimethylation in Drosophila.

Homology between the Drosophila CG40351 gene and the yeast SET1 or human SetB genes suggested that CG40351 encodes the Drosophila homolog of SET1. Purification of the dSet1 complex revealed the presence of a complex in Drosophila, which is comparable in composition with the human COMPASS complex.

Comparative analysis of RNAi-mediated knock down of Trx, Trr and dSet1 suggests preferential control of H3K4me2 and H3K4me3 by dSET1 in S2 cells. Polytene chromosome analysis further indicates predominant overlapping of dSet1 and H3K4me3. It is important to show by DNA staining that the overlap appears to occur preferentially in interbands. Such data are presented with Fig. S3 but this Figure is not cited in the text. This staining pattern also raises the question whether dSET1 actively associates or even binds with H3K4me2 and me3, which would be expected if it has a maintenance function. Consequently presence of H3K4me3 and binding of dSet1 would not indicate genetic activity. Preassembly of Pol II with interband regions was already shown long ago and this might be correlated with discrete H3K4me3 and dSet1 at interbands although no transcriptional activity is found. However colocalization of dSet1 with Pol II sites at polytene chromosomes might also simply be due to its
abundance and chromosome wide binding (low resolution of immunostaining). Binding of dSet1 to puffs is demonstrated in Fig. S4 (narrower signal than Pol II band) but again the Figure is not cited

in the text. Consequently the authors should recheck whether they indicated correctly all the very valuable supplementary Figures in the text of the manuscript.

We revised the manuscript and now mention all supplemental data including several newly added figures.

The authors do not provide sufficient data demonstrating the mutant nature of the P element insertion for dCfp1 or any convincing data proving that this mutant abolishes chromosomal association of dSet1.

This data was added as a new figure 4.

Effects on optimal expression in dSet1 KD cells are demonstrated for 7 randomly chosen genes. Although the data suggest a general role of dSet1 on efficient transcription of a wide set of genes finally genome wide expression data would have provided a much deeper insight in the suggested central role of dSet1 in control of gene activities.
We thank the reviewer for making this valuable experimental suggestion. We; however, are concerned that the addition of genome-wide transcription data would be inconclusive without additional dSet1-, H3K4me3, and Pol II-ChIP experiments, which would exceed the scope and volume of this publication. We also would like to point out that a very recent publication reported the genome-wide distribution of Ash2 and changes of H3K4me3 in an ash2 mutant (Perez-Lluch et al, Nucleic Acids Res. Feb 9, 2011, epub ahead of print). We showed that Ash2 is a subunit of the dSet1 complex (Fig. 1C). In full support of our findings, the authors observed a high positive correlation between expression levels of affected genes and Ash2-dependent H3K4me3, which we discussed in our revised manuscript. We hope that this independent study published during the review of our manuscript as well as data on 9 more genes (figure S9) further supports our conclusions that highly expressed genes depend more heavily on H3K4me3 by dSet1.

Significant reduction of promoter associated H3K4me3 is clearly found in the four genes selected for the studies. It is not explained why the genes have been selected for expression and for ChIP analysis of promoter specific effects of dSet1 KD on H3K4me3.

This has been added to the figure legend and text.

Heterochromatin localization of EGFP-dSet1 is not seen in Fig. S8. Do the authors indeed see any association with chromocenter heterochromatin? Fig. S8 suggests just the opposite.

We clarified this by rephrasing these paragraphs.

The quality of Fig. 5C to show definitely recruitment of EGFP-dSet1 to hsp70 after heat shock is not sufficient. This would need at least good chromosome squashes. However the data presented by Fig. 6 are clear-cut and strongly support recruitment of dSet1 to the hsp70 and hsp26 promoters.

We added a figure showing that eGFP-dSet1 accumulates at the activated hsp70 clusters on polytene chromosomes (Fig. S14).

In summary the manuscript provides interesting new data about epigenetic control of H3K4 methylation by dSet1 in Drosophila and shows for first time a predominant role of the dSet1 methyltransferase in control of H3K4me3 dependent gene activity. The epigenetic mechanisms uncovered are of general interest and the data presented are novel and have a general scientific interest.

2nd Editorial Decision 10 May 2011

Sorry for not responding to your previous email but I have just returned to the office today. I have received the final report form the two referees who have evaluated the study, both find the revised version to be very much improved and recommend publication after the minor issues below have

been accommodated.

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://www.nature.com/emboj/about/process.html

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

Yours sincerely,

Editor
The EMBO Journal

REFEREE REPORTS

Referee #1 (Remarks to the Author):

The MS has been significantly improved by the revisions. I am rather satisfied by the point by point response made by the authors. The retention of PolII at promoters in d-Set1 RNAi-treated cells is clearly an important point.

I still have a few minor concerns.

The comment to point 2 is weird. Even if the mass-spec is performed as mudPit analysis, we expect from the purification to see dSet1 and its associated proteins.

I still don't understand how Figure 2B was performed. If the figure represents immunoblots of histone extracts (eg histones extracted from cell extracts), what is the relevance of the Coomassie control?

In the legend of the new Figure 4 B, Pol II 0 is (green).

point 5. The author's explanation should be mentioned in the text.

I still have some problems with the way they refer to previous work. For instance why the PNAS paper of Miller is cited for the purification of COMPASS in yeast and not those of F. Stewart and P Nagy in EMBO J and PNAS, respectively.

Referee #2 (Remarks to the Author):

Comments on the revised manuscript of "Drosophila dSET1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription" submitted by Ardehali et al.

The author's respond carefully on all critical points raised in sufficient detail and also changed the manuscript in the most sensitive parts as suggested.

The authors provide a plausible answer on the concern about the selected genes to visualize the role of dSET1 on efficient transcription of a large set of genes. The data of a recent publication on genome wide distribution of Ash2 and about changes in H3K4me3 in an ash2 mutant support strongly the data presented especially because the authors could be shown that Ash2 is associated with the dSet1 complex. The publication by Perez-Lluch et al. is cited and discussed. This is finally a fine complementary result and strengthens substantially the data presented by the authors in the their manuscript. In addition now data for 9 more genes are presented in Supplementary Figure S9.

The effect of the P element insertion on dCfp1 is now presented in Figure 4A and clearly proves that the mutant is a null or at least a very strong amorphic allele.

The additional Figure S14 now clarifies the concerns about recruitment of EGFP-dSET1 to hsp70 loci after heat shock. It might have been in addition possible to compare the normal situation without heat shock with the situation after heat shock. However this is only a very minor point also because such analysis might not allow differentiating between the possible resting state and the active state of dSET1 at these sites.

Because all the data of the studies in total plausibly support the main conclusions drawn I finally would like to recommend the manuscript for publication.

2nd Revision - authors' response 15 May 2011

Point-by-point response to reviewers' comments:

Reviewer 1:

"The MS has been significantly improved by the revisions. I am rather satisfied by the point by point response made by the authors. The retention of PolII at promoters in d-Set1 RNAi-treated cells is clearly an important point.

I still have a few minor concerns.

The comment to point 2 is weird. Even if the mass-spec is performed as mudPit analysis, we expect from the purification to see dSet1 and its associated proteins.

As we stated in our previous response, no band was excised from the gel for our mass-spec analysis. To address the remaining concern of the reviewer and thereby facilitate the assessment of the purity and quality of the two isolated dSet1 complexes, we now list in the figure legend of figure S2 the predicted molecular masses of all potential isoforms of the identified subunits. We also cited two publications reporting the unusual migratory properties for one subunit (Ash2) as well as the proteolytic processing of another (dHcf). This added information should emphasize that the assignment of bands to identified subunits based on their migration in the SDS PAGE gel would be rather speculative.

I still don't understand how Figure 2B was performed. If the figure represents immunoblots of histone extracts (eg histones extracted from cell extracts), what is the relevance of the Coomassie control?

The Coomassie-stained control panel was added to Figure 2B to demonstrate that the histone extractions from the different samples yielded equal amounts. We feel it is helpful to the reader to present this. We would also like to point out that the antibody stainings against H3K9me3 serve as an additional control in this figure.

In the legend of the new Figure 4 B, Pol II 0 is (green).

We have now changed the label in the legend of Figure 4.

The author's explanation should be mentioned in the text.

In our revised manuscript, we now cite a very recent publication addressing the functional interaction between the identified H3K4-KDMs from flies in global H3K4me3 demethylation (di

Stefano et al, 2011). We feel that the citation of this publication is more informative for the readers than citing unpublished work from our laboratory.

I still have some problems with the way they refer to previous work. For instance why the PNAS paper of Miller is cited for the purification of COMPASS in yeast and not those of F. Stewart and P Nagy in EMBO J and PNAS, respectively.

We now cite the suggested publications.

Referee #2 (Remarks to the Author):

Comments on the revised manuscript of "Drosophila dSET1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription" submitted by Ardehali et al.

The author's respond carefully on all critical points raised in sufficient detail and also changed the manuscript in the most sensitive parts as suggested.

The authors provide a plausible answer on the concern about the selected genes to visualize the role of dSET1 on efficient transcription of a large set of genes. The data of a recent publication on genome wide distribution of Ash2 and about changes in H3K4me3 in an ash2 mutant support strongly the data presented especially because the authors could be shown that Ash2 is associated with the dSet1 complex. The publication by Perez-Lluch et al. is cited and discussed. This is finally a fine complementary result and strengthens substantially the data presented by the authors in the their manuscript. In addition now data for 9 more genes are presented in Supplementary Figure S9.

The effect of the P element insertion on dCfp1 is now presented in Figure 4A and clearly proves that the mutant is a null or at least a very strong amorphic allele.

The additional Figure S14 now clarifies the concerns about recruitment of EGFP-dSET1 to hsp70 loci after heat shock. It might have been in addition possible to compare the normal situation without heat shock with the situation after heat shock. However this is only a very minor point also because such analysis might not allow differentiating between the possible resting state and the active state of dSET1 at these sites.

We have now added to supplemental figure S14 a panel showing the dSet1/eGFP-dSet1 staining at the major hsp70 shock loci (87A & 87C) before heat shock.

Because all the data of the studies in total plausibly support the main conclusions drawn I finally would like to recommend the manuscript for publication.

3rd Editorial Decision 20 May 2011

Sorry for the delay but I have just returned from a small RNA meeting in VIenna. I have looked through the revised version of your manuscript and find that you have addressed all the remaining concerns of the referees and I am happy to accept it for publication in The EMBO Journal. You will receive the official acceptance letter in the next day or so.

