THE EMBO JOURNAL

Manuscript EMBO-2015-93534

Parental Epigenetic Asymmetry of PRC2-Mediated Histone Modifications in the Arabidopsis Endosperm

Jordi Moreno-Romero, Hua Jiang, Juan Santos-González and Claudia Köhler

Corresponding author: Claudia Köhler, Swedish University of Agricultural Sciences

Review timeline:

Submission date:	20 November 2015
Editorial Decision:	22 December 2015
Revision received:	26 January 2016
Editorial Decision:	18 February 2016
Revision received:	14 March 2016
Accepted:	17 March 2016

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 22 December 2015

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, all threes referees express interest in the findings reported in your manuscript, although they do also raise a number of concerns regarding both the conclusiveness and the presentation of the data that will have to be addressed before they can support publication of your study in The EMBO Journal.

For such a revised manuscript I would particularly ask you to focus your efforts on the following points:

-> Please elaborate on the TE expression analysis as well as the presence of H3K27me3 on DME target genes in the paternal genome as suggested by ref#1

-> In addition, both refs #1 and #2 ask you to provide more insight on the interplay between H3K9me2 and H3K27me2 in controlling gene expression at the pericentromeric regions (with ref #2 pointing out the need to better support the causality of the presented model)

-> Refs #2 and #3 also both comment on a number of sections in the manuscript that would benefit from rephrasing and clarification. Referee #3 furthermore emphasizes the need for you to keep the more generalist reader in mind when describing the findings in this study. I would strongly encourage you to follow these suggestions.

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 or rejection of your manuscript will therefore depend on the completeness of your responses in this revised version. I realise that each of the reports raise a large number of individual points to address in this case, but at the same time they do reflect overlapping concerns and we find them to be overall reasonable and constructive.

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

REFEREE COMMENTS

Referee #1:

PRC2-Mediated Histone Modifications Maintain Parental Epigenetic Asymmetry in the Endosperm

Summary
Imprinting in Arabidopsis relies on epigenetic differences between the parental genomes. DNA methylation, histone methylation and more recently, small RNAs have been implicated in this process. The primary histone methylation mark responsible for imprinting is H3K27me3 which is established through activity of the PRC. This manuscript is the first to address parental specific differences in H3K27me3 marks in the Arabidopsis endosperm in a genome-wide manner. Previous studies had investigated individual loci, or genome-wide in a non-parental-specific manner. This study reveals an inverse correlation between DNA methylation/H3K9me2 marks and H3K27me3 marks which has been established in other tissues, illustrating the interplay between the DNA demethylase DEMETER and PRC activity as a means for imprinting. In addition, H3K27me3 is reported to act in lieu of H3K9me2 at the paternal pericentromeric heterochromatin. Transcriptional effects of H3K27me3 are investigated using data from fie mutant endosperm to determine the functional effects of the mark, observing that paternally expressed imprinted genes are heavily reliant on maternal H3K27me3 marks for repression. This study advances our knowledge of imprinting and would be of interest to a wide audience.

The manuscript is well written, and the figures are presented in a manner that is easy to interpret. Supplemental datasets are given where necessary and provide additional context or confirmatory data. The material and methods section is also comprehensive and would allow for other laboratories to perform similar investigations.

Major concerns:

The results from figure 2C and D suggest that many TEs are protected from repression by H3K27me3, given the anti-correlation of TE expression in fie and H3K27me3 marks. This might suggest that DNA methylation or other repressive marks are applied to these regions in the absence of fie activity/H3K27me3. While a genome-wide analysis of methylation in this class of TE in fie endosperm is not necessary, a subset of TEs could be analyzed using McrBC or alternative low-throughput techniques. A comparison of fie expression levels for the classes of genes and TEs identified in Figure 1 would be very interesting - ie. paternally marked, maternally marked, shared, and unmarked from the Ler x Col cross.

One surprising result is the presence of H3K27me3 on the paternal genome at DME target regions. This is noted by the authors but not expanded upon, however it is quite surprising. Typically, the paternal allele of DME targets remains methylated as DME is not expressed in sperm cells, which should result in the absence of H3K27me3. This effect appears to be slightly more prominent in the Col accession than the Ler accession. Data presented later in the manuscript shows that differential H3K27me3 correlates with imprinted gene expression, although presumably there are a number of

differentially H3K27me3 marked genes that are not imprinted - it is possible that these genes are simply not expressed at high levels in the endosperm, and are therefore not readily detected in imprinting screens. The identity and expression levels of these genes would be of interest to readers and could be included as a supplementary dataset. If possible, the "imprinting scores" for these genes should also be provided as it is possible that they do show some parental bias, despite not being called imprinted.

The author's assertion that the absence of H3K9me2 in the pericentromere results in the presence of H3K27me3 at gypsy elements could be strengthened by observing parental differences in H3K9me2 ChIP data at gypsy and non-gypsy TEs (as shown for H3K27me3 in Figure 4E). The authors could also look at H3K27me1 and CG/CHG/CHH methylation in the same regions as a supplemental dataset.

Minor:
In the supplementary tables S2, S3 and S4 the "worksheets" in the Excel documents are named incorrectly.

Line 82: "be" should be "being"
Line 166: A new paragraph should start here
Line 185: "loose" should be "lose"

Referee #2:

The manuscript "PRC2-mediated Histone Modifications Maintain Parental Epigenetic Asymmetry in the Endosperm" by Moreno-Romero et al. describes the profiles of chromatin modifications of the maternally and paternally derived genomes in the endosperm. It is in this tissue, which surrounds and nourishes the embryo, that most if not all genomic imprinting takes place in plants. The authors made use of the INTACT system to isolate endosperm nuclei 4 days after fertilization in reciprocal crosses between the Col and Ler accessions in order to discriminate the maternal and paternal genomes. The distribution of DNA methylation and H3K27me3, H3K27me1 and H3K9me2 is analyzed and compared with published transcriptome and methylome data obtained for the same tissue later during development or from seedlings. The main results are that: (1) CHH methylation is almost absent in the early endosperm before cellularization, suggesting a lack of de novo DNA methylation activity via RdDM; (2) H3K27me3 is redistributed towards pericentromeric regions compared to its typical euchromatic localization in seedlings; (3) this relocalization is most apparent on the paternally-derived genome and over regions that show DEMETER-dependent reduction in DNA methylation; (4) H3K27me1 and H3K9me2 show parent-of-origin specific distribution patterns in the endosperm; and (5) genes showing preferential paternal expression in the endosperm are globally marked by H3K27me3 over the maternal allele, yet the reverse is not true for preferentially maternal expressed genes (this has already been shown for maize endosperm: Zhang et al, Genome Research 2014).
This is an important study, which provides several novel mechanistic insights into the differential marking of the maternal and paternal genomes in the endosperm and their relationship to genomic imprinting. The data and experimental approaches appear sound and the conclusions justified. However, there are several points that need to be addressed before this manuscript can be accepted for publication in the EMBO Journal.

1) The tile is misleading, as the results do not provide any evidence that PRC2-mediated histone modifications are responsible for maintaining the DNA methylation asymmetry between the two parental genomes in the endosperm. As a matter of fact, this asymmetry is preserved in fie mutant endosperm (Ibarra et al, Science 2012).
2) Both in the abstract and the text, the authors state that "de novo DNA methylation is impaired in the early endosperm". The phrasing is incorrect. At best, the data suggests "an absence of de novo DNA methylation in the early endosperm".
3) Line 46: the phrase is awkward as by definition, epigenetic differences involve no DNA sequence changes.
4) For the ChIP-seq experiments further controls are required to insure specificity. For instance, the loss of PRC2 was shown to result in unspecific binding of the anti-H3K27me3 antibody to

H3K27me2 and H3K27me1 (Bouyer et al, PLoS Genetics 2011). To make sure that the apparent relocalization of H3K27me3 to pericentromeric regions in early endosperm is real and does not result from cross-reaction of the antibody with H3K27me1/2, the authors should perform a western blot analysis.
5) The authors have previously published data indicating H3K27me3 enrichment over TEs in the endosperm (Weinhofer et al, PLoS Genetics 2010). The two data sets should therefore be compared to demonstrate consistency.
6) The authors claim that "H3K27me3 at pericentromeric regions may compensate for the reduced H3K9me2..." (line 180). This should be tested directly by showing the overlap between TEs that lose H3K9me2 and those that gain H3K27me3 over the paternal genome in the endosperm. The authors should also test if the regions that gain H3K27me3 in the endosperm overlap with those that show ectopic H3K27me3 and /or loss of H3K9me2 in met1 seedlings (Deleris et al, PLoS Genetics 2012). These comparisons are necessary to determine if the occupancy of H3K27me3 over regions that lose DNA methylation is really a consequence of this loss or results instead from a specific targeting of PRC2 during this developmental stage.
7) Gypsy LTRs show strong H3K27me3 enrichment over the paternal allele. As these TEs are preferentially targeted by CMT2 rather than RdDM, it is important to know if these sequences actually lose DNA methylation over the paternal genome in the endosperm. And what about H3K9me2? Moreover, H3K27me3 enrichment differs depending on the direction of the cross (compare Figure 4E and 4F with appendix Figure S5C and 5D). This difference needs to be analyzed in more detail or at least discussed.
8) The H3K27me3 enrichment over pericentromeric sequences seen in the endosperm does not appear to correlate well with derepression of these sequences in fie mutant endosperm. Indeed, the highest abundance of TE transcripts in fie endosperm seems to be unrelated to H3K27me3 enrichment over these sequences in wild type endosperm (Figure 2C). Therefore, this questions the extent to which TE activation in fie endosperm is a consequence of H3K27me3 loss. Could this activation be instead an indirect consequence of deregulation of other chromatin regulators or developmental defects in fie? As maternal and paternal alleles show different H3K27me3 enrichment, the transcriptome data should be analyzed separately for the two parental alleles.
9) The authors refer to Figure 2D for more quantitative analysis of the deregulation of TEs in fie, yet in the figure legend the analysis refers to genes, not TEs. Please clarify. In the figure itself, the color code does not match between panels 2C and 2D. Quartiles appear to be organized in the reverse order (1st quartile of upregulated transcripts is given as "low fie - wt", although this seems to correspond to "high fie - wt").
10) The global ChIP-profiles (chromosomal distribution of z-score normalized H3K27me3; Figure 1D) do not appear to match the heat map views (Figure 2B). For example, the global H3K27me3 profiles look almost identical for the paternal and maternal genomes (Figure 1D), yet the heat map views look dramatically different, indicating strong abundance of paternal H3K27me3 over pericentromeric regions and a broader distribution for the maternal genome (Figure 2B). Similarly, profiles and heat map views don't seem to match for H3K9me2 (Figure 4A). Can this be explained? A heat map view for leaf tissue should also be included.
11) Line 200: The conclusion that "H3K27me3 does not necessarily cause complete silencing of the targeted allele" is misleading as the levels of H3K27me3 here is very low and hence only few nuclei are marked.

Referee #3:

The paper concerns parent of origin specific differences in DNA methylation and histone methylation patterns in Arabidopsis endosperm and is relevant for mechanisms of genomic imprinting. Technically the paper is extremely impressive, with the use of the INTACT method to purify endosperm nuclei from reciprocal crosses and to map histone methylation and DNA methylation to male and female parental alleles. There is a considerable amount of novel genomic data. The main points I took from the paper are that imprinted genes with paternal expression are predominantly marked in maternal specific fashion with H3K27me3, and that regions that specifically lack DNA methylation in maternal alleles tend to have H3K27me3 on maternal or paternal alleles. In general I think the paper is suitable for EMBO J and will be of interest for the genomic imprintin and epigenetics communities. Although the paper is well written, I found it extremely hard to digest and feel that in its present form it will only be accessible to quite

specialised audiences, i.e. it could perhaps be improved if the authors made more effort to explain the various figures more fully or included explanatory figures showing the different aspects they are distinguishing. I have a few minor comments as follows:

1. DME targetted regions seem to be defined as ones with less CG DNA methylation on the maternal allele than on the paternal allele. This is a very indirect marker for DME targetting and it is not clear how well it correlates with e.g. DME localisation to genes. For example, it is known that MET1 is down regulated in the central cell, so some differences between male and female alleles may be passive rather than a result of active demethylation by DME. Comparisons of how well the inferred DME targetted regions correspond with data from other studies involving methylome analysis in dme mutant endosperm would be helpful here. The model seems to be that DME demethylates maternal alleles in the central cell and this directs H3K27me3 methylation to demethylated regions; are there good reasons to exclude the alternative i.e. that H3K27me3 methylation changes in the central cell and this directs DME targetting?

2. Fig 2C. I did not find this expression analysis easy to follow. Figure 2C plots H3K27me3 levels but without making clear whether which particular type, ie elsewhere they distinguish maternal and paternal allele specific H3K27me3 as well as accession specific (Col vs Ler), it is not clear which type is being shown here or if some kind of average is being plotted. Expression levels in the fie mutant seem anti-correlated with H3K27me3 levels throughout the genome rather than only in pericentromeric regions as claimed. I don't understand the conclusion that "H3K27me3 is not sufficient to maintain transcriptional silencing of a subset of gene preferentially located at pericentromeric regions". Looking at the data in 2C one could almost argue that H3K27me3 is required for transcriptional activity of TE.

3. Given the very poor correlation between H3K27me3 methylation levels at genes or TE and their expression levels in fie mutants, how is this reconciled with the loss of imprinting of PEGs in fie mutants where presumably maternal specific H3K27me3 is repressing expression of the maternal allele?

4. It wasn't clear from the methods whether the tissue used in the RNA seq expression analysis of fie mutants was from a mixture of fie mutant and FIE+ endosperm, as would typically be obtained from pollinating a fie/FIE female, or whether rare fie homozygous females were obtained.

5. Figure 3 A and B are complicated and not very well explained in the legend, so it is difficult to work out exactly what is being plotted on the y axis and for which subset of genes.

6. If DME targetted regions are those with more CG methylation on the paternal allele, and some of these are also paternally specifically H3K27me3 methylated alleles, does this mean that the H3K27me3 and mCG methylation co-occur on these paternal alleles, ie the normal anti-correlation breaks down?

7. Fig 4A - it wasn't wholly clear to me why the top and bottom panels of each figure didn't seem to agree, i.e. in the plots in the upper part male and female H3K9me2 marked alleles seem to have similar profiles, in the lower part the paternally specifically marked alleles generally have less H3K9me2.

8. Figure 5 - what are the p values actually testing? It seems from the p values in part B that these are also very low and hence significant, i.e. does this imply that a greater proportion of MEGs are enriched for H3K27me3 marked genes than would be expected by chance, contrary to the text? What are the colours in the Venn diagrams depicting, this is not shown in legend or figure?

9. Line 103 states about 70% if TE and genes marked with H3K27me3 overlapped, it looks to me that figure is more like 55% overall (62% for genes and 52% for TE based on Venn diagram in Figure 1A).

10. Line 113 - I calculate 40% for maternal and 25% for paternal based on Fig S2B.

11. A few minor typos as follows - Line 74 mentions PEG for first time but the abbreviation is not defined until later in text. Line 82 should read "to being located". Line 124 delete of to read "three

previously published..." Line 174 insert space in "OfH3". Line 185 and elsewhere "lose" not "loose". Line 269 should read "explanation of how"

1st Revision - authors' response 26 January 2016

Response to reviewer's comments
Referee #1:

Summary
Imprinting in Arabidopsis relies on epigenetic differences between the parental genomes. DNA methylation, histone methylation and more recently, small RNAs have been implicated in this process. The primary histone methylation mark responsible for imprinting is H3K27me3 which is established through activity of the PRC. This manuscript is the first to address parental specific differences in H3K27me3 marks in the Arabidopsis endosperm in a genome-wide manner. Previous studies had investigated individual loci, or genome-wide in a non-parental-specific manner. This study reveals an inverse correlation between DNA methylation/H3K9me2 marks and H3K27me3 marks which has been established in other tissues, illustrating the interplay between the DNA demethylase DEMETER and PRC activity as a means for imprinting. In addition, H3K27me3 is reported to act in lieu of H3K9me2 at the paternal pericentromeric heterochromatin.
Transcriptional effects of H3K27me3 are investigated using data from fie
mutant endosperm to determine the functional effects of the mark, observing that paternally expressed imprinted genes are heavily reliant on maternal H3K27me3 marks for repression. This study advances our knowledge of imprinting and would be of interest to a wide audience.

The manuscript is well written, and the figures are presented in a manner that is easy to interpret. Supplemental datasets are given where necessary and provide additional context or confirmatory data. The material and methods section is also comprehensive and would allow for other laboratories to perform similar investigations.

We thank the reviewer for the careful evaluation of our manuscript and the helpful suggestions for improvement.

Major concerns:
1. The results from figure 2C and D suggest that many TEs are protected from repression by H3K27me3, given the anti-correlation of TE expression in fie and H3K27me3 marks. This might suggest that DNA methylation or other repressive marks are applied to these regions in the absence of fie activity/H3K27me3. While a genome-wide analysis of methylation in this class of TE in fie endosperm is not necessary, a subset of TEs could be analyzed using McrBC or alternative low-throughput techniques.
As suggested by the reviewer, we analyzed the DNA methylation level of TEs affected by the loss of FIE function using previously published data (Ibarra et al., 2012, Science, 337, 1360-1364). This analysis revealed that indeed TEs that remained silenced in *fie* were marked by DNA methylation and even gained CG DNA methylation, explaining that they remained silenced upon loss of FIS-PRC2 function. We included these data as Appendix Fig 6 and added the following text to the manuscript:
"We addressed the question whether TEs that remained silent in *fie* mutants were additionally marked by DNA methylation. Indeed, TEs with low expression in *fie* had higher levels of DNA methylation in all sequence contexts compared to TEs that became upregulated upon lack of FIS-PRC2 function (Appendix Fig S6). Consistent with the reported upregulation of genes mediating CG methylation in *fie* endosperm (Hsieh et al, 2011), the CG methylation level at TEs was increased in *fie*, while the CHG and CHH methylation level decreased, in agreement with previously published data (Ibarra et al, 2012). We thus conclude that H3K27me3 is may be functionally relevant for TE silencing in the endosperm, but that the additional regulation by DNA methylation potentially masks the effect of H3K27me3 depletion."
A comparison of fie expression levels for the classes of genes and TEs identified in Figure 1 would be very interesting - ie. paternally marked, maternally marked, shared, and unmarked from the Ler x Col cross.

In agreement with our previously published work (Weinhofer et al., 2010, PLoS Genet, 6, e1001152), we found only few genes marked by H3K27me3 to be deregulated in *fie*. We therefore decided not to include this analysis into the revised manuscript.

2. One surprising result is the presence of H3K27me3 on the paternal genome at DME target regions. This is noted by the authors but not expanded upon, however it is quite surprising. Typically, the paternal allele of DME targets remains methylated as DME is not expressed in sperm cells, which should result in the absence of H3K27me3. This effect appears to be slightly more prominent in the Col accession than the Ler accession.
Regions enriched for H3K27me3 at paternal DME target regions differ from those enriched for H3K27me3 at maternal DME target regions. While the first are predominantly located at pericentromeric regions, the latter are distributed throughout the genome. Pericentromeric regions are usually densely methylated in CHH context (Lister et al., 2008, Cell, 133, 523-536). The near absence of CHH methylation in the 4 DAP endosperm (Figure 3C) could promote targeting of H3K27me3 to those regions on the paternal genome. Furthermore, at paternal pericentromeric regions we find a depletion of H3K9me2 (see Figure 4A and 5C), which could explain the specific presence of H3K27me3 at paternal DME target regions.
3. Data presented later in the manuscript shows that differential H3K27me3 correlates with imprinted gene expression, although presumably there are a number of differentially H3K27me3 marked genes that are not imprinted - it is possible that these genes are simply not expressed at high levels in the endosperm, and are therefore not readily detected in imprinting screens. The identity and expression levels of these genes would be of interest to readers and could be included as a supplementary dataset. If possible, the "imprinting scores" for these genes should also be provided as it is possible that they do show some parental bias, despite not being called imprinted.
As suggested by the reviewer, we included data showing the imprinting scores of all genes together with the information whether those genes have parental-specific H3K27me3 (new Appendix Table S5 and Supplementary Figure S11). We also included the following text to the manuscript: "We furthermore tested whether genes with parentally-biased expression that were however not passing all criteria used to define imprinted genes (Pignatta et al, 2014), were enriched for H3K27me3 on the silent allele. We found that about half of all genes with paternally-biased expression (186 genes) were marked by maternal H3K27me3 in either Col or Ler accessions, while only 8-9% of all genes with maternally-biased expression (1233 genes) had paternal H3K27me3 (Appendix Figure S11 and Appendix Table S5). Thus, paternally-biased expression is strongly associated with H3K27me3 on maternal alleles, while maternally-biased expression and paternal H3K27me3 rarely coincide."

4. The author's assertion that the absence of H3K9me2 in the pericentromere results in the presence of H3K27me3 at gypsy elements could be strengthened by observing parental differences in H3K9me2 ChIP data at gypsy and non-gypsy TEs (as shown for H3K27me3 in Figure 4E). The authors could also look at H3K27me1 and CG/CHG/CHH methylation in the same regions as a supplemental dataset.
As suggested by the reviewer, we analyzed the parental-specific distribution of H3K9me2 and H3K27me1 on gypsy and non-gypsy TEs and included these data as new Figure 5 into the manuscript and added the following text: "Consistent with the idea that paternal-specific H3K27me3 may compensate for reduced H3K9me2, the parental distribution of H3K9me2 on gypsy elements was distinct to that of H3K27me3 with increased levels of H3K9me2 on maternal compared to paternal alleles (Fig 5C and D and Appendix Fig S8C and D). Also the parental distribution of H3K27me1 on gypsy elements was similar to that of H3K9me2, contrasting the paternal bias of this modification on non-gypsy TEs (Fig 5E and F and Appendix Fig S8E and F)."
The parental-specific distribution of DNA methylation on gypsy and non-gypsy TEs was similar to the general pattern shown in Figure 3C and we therefore did not include a separate analysis of DNA methylation on those regions into the manuscript.
Minor:
In the supplementary tables S2, S3 and S4 the "worksheets" in the Excel documents are named incorrectly.
We have corrected this mistake.

Line 82: "be" should be "being"
Line 166: A new paragraph should start here
Line 185: "loose" should be "lose"
Corrected.

Referee #2:

The manuscript "PRC2-mediated Histone Modifications Maintain Parental Epigenetic Asymmetry in the Endosperm" by Moreno-Romero et al. describes the profiles of chromatin modifications of the maternally and paternally derived genomes in the endosperm. It is in this tissue, which surrounds and nourishes the embryo, that most if not all genomic imprinting takes place in plants. The authors made use of the INTACT system to isolate endosperm nuclei 4 days after fertilization in reciprocal crosses between the Col and Ler accessions in order to discriminate the maternal and paternal genomes. The distribution of DNA methylation and H3K27me3, H3K27me1 and H3K9me2 is analyzed and compared with published transcriptome and methylome data obtained for the same tissue later during development or from seedlings. The main results are that: (1) CHH methylation is almost absent in the early endosperm before cellularization, suggesting a lack of de novo DNA methylation activity via RdDM; (2) H3K27me3 is redistributed towards pericentromeric regions compared to its typical euchromatic localization in seedlings; (3) this relocalization is most apparent on the paternally-derived genome and over regions that show DEMETER-dependent reduction in DNA methylation; (4) H3K27me1 and H3K9me2 show parent-of-origin specific distribution patterns in the endosperm; and (5) genes showing preferential *paternal expression in the endosperm are globally marked by H3K27me3 over the maternal allele, yet the reverse is not true for preferentially maternal expressed genes (this has already been shown for maize endosperm: Zhang et al, Genome Research 2014).*
This is an important study, which provides several novel mechanistic insights into the differential marking of the maternal and paternal genomes in the endosperm and their relationship to genomic imprinting. The data and experimental approaches appear sound and the conclusions justified. However, there are several points that need to be addressed before this manuscript can be accepted for publication in the EMBO Journal.
We thank the reviewer for the careful evaluation of our manuscript and the helpful suggestions for improvement.

1) The tile is misleading, as the results do not provide any evidence that PRC2-mediated histone modifications are responsible for maintaining the DNA methylation asymmetry between the two parental genomes in the endosperm. As a matter of fact, this asymmetry is preserved in fie mutant endosperm (Ibarra et al, Science 2012).
We have changed the title as follows: "Parental Epigenetic Asymmetry of PRC2-Mediated Histone Modifications in the Arabidopsis Endosperm"

2) Both in the abstract and the text, the authors state that "de novo DNA methylation is impaired in the early endosperm". The phrasing is incorrect. At best, the data suggests "an absence of de novo DNA methylation in the early endosperm".
We have rephrased as suggested by the reviewer.

3) Line 46: the phrase is awkward as by definition, epigenetic differences involve no DNA sequence changes.
We agree and have removed the phrase "despite their genetic resemblance".

4) For the ChIP-seq experiments further controls are required to insure specificity. For instance, the loss of PRC2 was shown to result in unspecific binding of the anti-H3K27me3 antibody to H3K27me2 and H3K27me1 (Bouyer et al, PLoS Genetics 2011). To make sure that the apparent relocalization of H3K27me3 to pericentromeric regions in early endosperm is real and does not result from cross-reaction of the antibody with H3K27me1/2, the authors should perform a western blot analysis.
We used the anti-H3K27me3 antibody from Millipore (Cat. # 07-449), which is a widely used antibody that based on dot-blot analyses does not cross-react with H3K27me2 and H3K27me1 peptides (see customer information: http://www.merckmillipore.com/DE/en/product/Anti-trimethyl-Histone-H3-%28Lys27%29-Antibody,MM_NF-07-449?bd=1#documentation). Furthermore, parental-specific differences of H3K27me3 (Figure 2B) and H3K27me1 (Figure 4B) are clearly distinct, making it unlikely that the observed localization of H3K27me3 to paternal pericentromeric regions is a consequence of unspecific antibody binding. Lastly, we performed ChIP experiments using leaf tissues with the same antibody under the same conditions as we used for the endosperm ChIP experiments and our data are in close agreement with previously published data (Appendix

Figure S3), adding additional support for our conclusion that our observations are not a consequence of unspecific binding of the H3K27me3 antibody.

5) The authors have previously published data indicating H3K27me3 enrichment over TEs in the endosperm (Weinhofer et al, PLoS Genetics 2010). The two data sets should therefore be compared to demonstrate consistency.
In the previous study we identified only 385 transposable elements (TEs) as targets for H3K27me3 in the endosperm compared to more than 10.000 TEs identified in the new study. In the previous study we used Affymetrix Tiling Arrays, where low-complexity sequences and simple repeats are excluded from probe design and thus the detection of TEs is very limited, explaining the low number of identified TEs. Nevertheless, we identified 64% of those 385 previously detected TEs in our new dataset ($p<1.5E-60$). This is a significant overlap in particular considering that we used different stages of endosperm development in both studies (1-4 days after pollination in Weinhofer et al. compared to 4 DAP in this study), different accession combinations (Ler in Weinhofer et al. compared to Ler x Col and Col x Ler hybrids in this study), different methods to isolate the endosperm (FACS versus INTACT), different amplification kits (WGA-4 single cell amplification kit in Weinhofer et al. versus Microplex Library Preparation kit (Diagenode; this study) and different detection methods (Tiling arrays versus NGS). Despite those differences between our previous and our current study, both studies revealed that TEs are targeted by H3K27me3 in the endosperm, demonstrating consistency between the different studies.

6) The authors claim that "H3K27me3 at pericentromeric regions may compensate for the reduced H3K9me2..." (line 180). This should be tested directly by showing the overlap between TEs that lose H3K9me2 and those that gain H3K27me3 over the paternal genome in the endosperm.
We have analyzed the distribution of H3K9me2 over gypsy elements and the corresponding distribution of H3K27me3 on the same elements and find that both marks have opposite trends, strongly supporting our conclusion. We observed the same trend as for H3K9me2 also for H3K27me1 and also included this analysis into the manuscript. We included this new data as Figure 5 in the manuscript and added the following text: "Consistent with the idea that paternal-specific H3K27me3 may compensate for reduced H3K9me2, the parental distribution of H3K9me2 on gypsy elements was distinct to that of H3K27me3 with increased levels of H3K9me2 on maternal compared to paternal alleles (Fig 5C and D and Appendix Fig S8C and D). Also the parental distribution of H3K27me1 on gypsy elements was similar to that of H3K9me2, contrasting the paternal bias of this modification on non-gypsy TEs (Fig 5E and F and Appendix Fig S8E and F)."
Testing for overlap of gypsy TEs gaining or losing parental specific H3K27me3 and H3K9me2 would require thresholding the differences on maternal and paternal alleles, which is inherently difficult and prone to bias; we therefore decided not to include this type of analysis into the manuscript.

The authors should also test if the regions that gain H3K27me3 in the endosperm overlap with those that show ectopic H3K27me3 and /or loss of H3K9me2 in met1 seedlings (Deleris et al, PLoS Genetics 2012). These comparisons are necessary to determine if the occupancy of H3K27me3 over regions that lose DNA methylation is really a consequence of this loss or results instead from a specific targeting of PRC2 during this developmental stage.
We found a significant overlap of those TEs that have been reported to gain H3K27me3 in *met1* and those being marked by H3K27me3 in the endosperm. We included this analysis as Appendix Fig S9B and added the following text: "Consistently, 52% of those TEs that gained H3K27me3 in the *met1* mutant overlapped with those marked by H3K27me3 in the endosperm (Hypergeometric test $p<1E-20$; Appendix Fig S9B)."

7) Gypsy LTRs show strong H3K27me3 enrichment over the paternal allele. As these TEs are preferentially targeted by CMT2 rather than RdDM, it is important to know if these sequences actually lose DNA methylation over the paternal genome in the endosperm. And what about H3K9me2? Moreover, H3K27me3 enrichment differs depending on the direction of the cross (compare Figure 4E and 4F with appendix Figure S5C and 5D). This difference needs to be analyzed in more detail or at least discussed.
As shown in Figure 3C, maternal and paternal genomes are almost completely devoid of CHH methylation in the 4 DAP endosperm. This is contrasted by the presence of H3K9me2 on maternal and paternal genomes and parental-specific differences in the H3K9me2 distribution, with reduced H3K9me2 on the paternal genome (Figure 4A).

As suggested by the reviewer, we discussed the accession-specific differences in H3K27me3 as follows in the manuscript: "The level of H3K27me3 on gypsy elements was substantially higher on the paternal Col compared to the paternal L*er* genome (Fig 5A and B and Appendix Fig S8A and B). Nevertheless, parental differences in H3K27me3 on gypsy elements compared to non-gypsy TEs were significant in both accessions (Fig 5B and Appendix Fig S8B)."

8) The H3K27me3 enrichment over pericentromeric sequences seen in the endosperm does not appear to correlate well with derepression of these sequences in fie mutant endosperm. Indeed, the highest abundance of TE transcripts in fie endosperm seems to be unrelated to H3K27me3 enrichment over these sequences in wild type endosperm (Figure 2C). Therefore, this questions the extent to which TE activation in fie endosperm is a consequence of H3K27me3 loss. Could this activation be instead an indirect consequence of deregulation of other chromatin regulators or developmental defects in fie? As maternal and paternal alleles show different H3K27me3 enrichment, the transcriptome data should be analyzed separately for the two parental alleles.
As suggested by the reviewer, we included a parental-specific analysis of TE expression in *fie* as Appendix Fig S5. This analysis did not reveal parental-specific effects on TE expression caused by the loss of FIE function. We agree with the reviewer that we cannot exclude that TE activation in *fie* could be a consequence of additional defects, as e.g. changes in DNA methylation, as previously reported by Ibarra et al., Science 2012. To account for this possibility, we tested DNA methylation changes at regions marked by H3K27me3 and included this analysis as Appendix Fig S6 in the manuscript and added the following text: "We addressed the question whether TEs that remained silent in the *fie* mutant were additionally marked by DNA methylation. Indeed, TEs with low expression in *fie* had higher levels of DNA methylation in all sequence contexts compared to TEs that became upregulated upon lack of FIS-PRC2 function (Appendix Fig S6). Consistent with the reported upregulation of genes mediating CG DNA methylation in fie endosperm (Hsieh et al, 2011), the CG DNA methylation level at TEs was increased in fie, while the CHG and CHH methylation level decreased, in agreement with previously published data (Ibarra et al, 2012). We thus conclude that H3K27me3 may be functionally relevant for TE silencing in the endosperm, but that the additional regulation by DNA methylation potentially masks the effect of H3K27me3 depletion."

9) The authors refer to Figure 2D for more quantitative analysis of the deregulation of TEs in fie, yet in the figure legend the analysis refers to genes, not TEs. Please clarify. In the figure itself, the color code does not match between panels 2C and 2D. Quartiles appear to be organized in the reverse order (1st quartile of upregulated transcripts is given as "low fie - wt", although this seems to correspond to "high fie - wt").
We corrected the mistake in the figure legend; the figure does indeed show TEs. The order of quantiles in Fig. 2D is correct and does thus correspond to the color code used in Fig. 2C; low changes in expression are in green and high expression changes in red.

10) The global ChIP-profiles (chromosomal distribution of z-score normalized H3K27me3; Figure 1D) do not appear to match the heat map views (Figure 2B). For example, the global H3K27me3 profiles look almost identical for the paternal and maternal genomes (Figure 1D), yet the heat map views look dramatically different, indicating strong abundance of paternal H3K27me3 over pericentromeric regions and a broader distribution for the maternal genome (Figure 2B). Similarly, profiles and heat map views don't seem to match for H3K9me2 (Figure 4A). Can this be explained? A heat map view for leaf tissue should also be included.
The global ChIP profiles show z-score normalized histone modifications on maternal and paternal genomes; while the heat maps show parental-specific (exclusively only on the maternal or paternal genome) histone marks, explaining the difference between the global ChIP profiles and the heat maps. We used heatmaps to visualize positions of parental-specific differences of histone modifications in the endosperm. We did not generate parental-specific profiles of histone modifications in leaves and therefore prefer not to include heat maps of leaves.

11) Line 200: The conclusion that "H3K27me3 does not necessarily cause complete silencing of the targeted allele" is misleading as the levels of H3K27me3 here is very low and hence only few nuclei are marked.
As suggested by the reviewer we removed this conclusion.

Referee #3:

The paper concerns parent of origin specific differences in DNA methylation and histone methylation patterns in Arabidopsis endosperm and is relevant for mechanisms of genomic imprinting. Technically the paper is extremely impressive, with the use of the INTACT method to purify endosperm nuclei from reciprocal crosses and to map histone methylation and DNA methylation to male and female parental alleles. There is a considerable amount of novel genomic data. The main points I took from the paper are that imprinted genes with paternal expression are predominantly marked in maternal specific fashion with H3K27me3, and that regions that specifically lack DNA methylation in maternal alleles tend to have H3K27me3 on maternal or paternal alleles. In general I think the paper is suitable for EMBO J and will be of interest for the genomic imprintin and epigenetics communities. Although the paper is well written, I found it extremely hard to digest and feel that in its present form it will
only be accessible to quite specialised audiences, i.e. it could perhaps be improved if the authors made more effort to explain the various figures more fully or included explanatory figures showing the different aspects they are distinguishing. I have a few minor comments as follows:

We thank the reviewer for the careful evaluation of our manuscript and the helpful suggestions for improvement. As suggested by the reviewer we added additional explanations at several positions in the manuscript to improve readability and understanding.

1. DME targetted regions seem to be defined as ones with less CG DNA methylation on the maternal allele than on the paternal allele. This is a very indirect marker for DME targetting and it is not clear how well it correlates with e.g. DME localisation to genes. For example, it is known that MET1 is down regulated in the central cell, so some differences between male and female alleles may be passive rather than a result of active demethylation by DME. Comparisons of how well the inferred DME targetted regions correspond with data from other studies involving methylome analysis in dme mutant endosperm would be helpful here.
To infer DME target regions we reanalyzed previously published data of the *dme* mutant (Ibarra et al., Science 337, 1360-1364 (2012)) and defined DME target regions as described in the methods: "DEMETER targeted regions were defined as those 50bp windows were the log 2 fold change in CG methylation between the *dme* mutant (Ibarra et al, 2012) and the wild type is > 1.5."

The model seems to be that DME demethylates maternal alleles in the central cell and this directs H3K27me3 methylation to demethylated regions; are there good reasons to exclude the alternative i.e. that H3K27me3 methylation changes in the central cell and this directs DME targetting?
We accounted for this possibility as follows in the discussion: "It is formally possible that differences in H3K27me3 direct differences in DNA methylation. However, endosperm lacking the main FIS-PRC2 component FIE maintain parental-specific differences in CG methylation (Ibarra *et al,* 2012), arguing in favor of the idea that differences in DNA methylation direct differences in H3K27me3."

2. Fig 2C. I did not find this expression analysis easy to follow. Figure 2C plots H3K27me3 levels but without making clear whether which particular type, ie elsewhere they distinguish maternal and paternal allele specific H3K27me3 as well as accession specific (Col vs Ler), it is not clear which type is being shown here or if some kind of average is being plotted. Expression levels in the fie mutant seem anti-correlated with H3K27me3 levels throughout the genome rather than only in pericentromeric regions as claimed. I don't understand the conclusion that "H3K27me3 is not sufficient to maintain transcriptional silencing of a subset of gene preferentially located at pericentromeric regions". Looking at the data in 2C one could almost argue that H3K27me3 is required for transcriptional activity of TE.
Expression data of *fie* endosperm have been used from previously published data (Hsieh et al., Proc Natl Acad Sci U S A 108, 1755-1762 (2011)) that was generated from Ler x Col crosses. We therefore only analyzed the cross Ler x Col. We included this information in the figure legend. To account for the reviewer's concern, we included parental-specific information of this cross as Appendix Fig S5. We also agree that our conclusion was too strong and rephrased as follows: "Indeed, TEs in *fie* endosperm were upregulated (Fig 2C), but the level of deregulation anticorrelated with the level of H3K27me3, which was particularly pronounced at pericentromeric regions (Fig 2C and D)."

We furthermore included an analysis of DNA methylation changes in *fie* at TEs marked by H3K27me3 and included this analysis as follows into the manuscript: "We addressed the question whether TEs that remained silent in the *fie* mutant were additionally marked by DNA methylation. Indeed, TEs with low expression in *fie* had higher levels of DNA methylation in all sequence contexts compared to TEs that became upregulated upon lack of FIS-PRC2 function (Appendix Fig S6). Consistent with the reported upregulation of genes mediating CG DNA methylation in *fie* endosperm (Hsieh et al, 2011), the CG DNA methylation level at TEs was increased in *fie*, while the CHG and CHH methylation level decreased, in agreement with previously published data (Ibarra et al, 2012). We thus conclude that H3K27me3 may be functionally relevant for TE silencing in the endosperm, but that the additional regulation by DNA methylation potentially masks the effect of H3K27me3 depletion."

3. Given the very poor correlation between H3K27me3 methylation levels at genes or TE and their expression levels in fie mutants, how is this reconciled with the loss of imprinting of PEGs in fie mutants where presumably maternal specific H3K27me3 is repressing expression of the maternal allele?
The expression status of TEs close to imprinted genes is probably not decisive for their effect on the expression of the neighboring gene. While TEs are often silenced by redundantly acting epigenetic mechanisms (Baubec et al., Plant Cell 22, 34–47 (2010)) and also shown in this manuscript (Appendix Figure S6), this does not apply for the neighboring genes. Therefore, loss of H3K27me3 may not effect expression of many TEs, but will cause activation of the genes flanked by TEs.

4. It wasn't clear from the methods whether the tissue used in the RNA seq expression analysis of fie mutants was from a mixture of fie mutant and FIE+ endosperm, as would typically be obtained from pollinating a fie/FIE female, or whether rare fie homozygous females were obtained.
The RNAseq data from *fie* were used from a previously published study ((Hsieh et al., Proc Natl Acad Sci U S A 108, 1755-1762 (2011)) as referenced in the results and in the methods ("Data from Ler x Col and Ler fie x Col crosses (Hsieh et al, 2011) were downloaded from the GEO omnibus. Reads were mapped and expression normalized as described in (Pignatta et al, 2014)."). Based on the information provided by the authors, they dissected endosperm of *fie* mutant seeds, which were identified by the abnormal endosperm.

5. Figure 3 A and B are complicated and not very well explained in the legend, so it is difficult to work out exactly what is being plotted on the y axis and for which subset of genes.
We included the information that kernel density plots visualize the frequency distribution of DNA methylation (3A) and H3K27me3 (3B) and also included the full wording for MSRs and PSRs to avoid misunderstanding.

6. If DME targetted regions are those with more CG methylation on the paternal allele, and some of these are also paternally specifically H3K27me3 methylated alleles, does this mean that the H3K27me3 and mCG methylation co-occur on these paternal alleles, ie the normal anti-correlation breaks down?
Regions enriched for H3K27me3 at paternal DME target regions differ from those enriched for H3K27me3 at maternal DME target regions. While the first are predominantly located at pericentromeric regions, the latter are distributed throughout the genome. Pericentromeric regions are usually densely methylated in CHH context (Lister et al., 2008, Cell, 133, 523-536). The near absence of CHH methylation in the 4 DAP endosperm (Figure 3C) could promote targeting of H3K27me3 to those regions on the paternal genome. Furthermore, at paternal pericentromeric regions we find a depletion of H3K9me2 (see Figure 4A and 5C), which could explain the specific presence of H3K27me3 at paternal DME target regions.

7. Fig 4A - it wasn't wholly clear to me why the top and bottom panels of each figure didn't seem to agree, i.e. in the plots in the upper part male and female H3K9me2 marked alleles seem to have similar profiles, in the lower part the paternally specifically marked alleles generally have less H3K9me2.
The global ChIP profiles (top panels) show z-score normalized histone modifications on maternal and paternal genomes; while the heat maps (bottom panels) show parental-specific (exclusively only on the maternal or paternal genome) histone marks, explaining the difference between the global ChIP profiles and the heat maps.

8. Figure 5 - what are the p values actually testing? It seems from the p values in part B that these are also very low and hence significant, i.e. does this imply that a greater proportion of MEGs are enriched for H3K27me3 marked genes than would be expected by chance, contrary to the text? What are the colours in the Venn diagrams depicting, this is not shown in legend or figure?
The p values shown in Figure 5 (now Fig 6) test the significance of the overlap of parental-specific marks in Col and L*er* accessions; e.g. Fig 5B (now Fig 6B) upper left Venn diagram shows that out of 14 maternal-specific H3K27me3 marks in Col eight overlap with maternal-specific H3K27me3 marks in Ler. While this is significant, the overlap of MEGs with parental-specific marks is not, as stated in the text: "In contrast to the strong overlap of PEGs with maternal H3K27me3 regions, out of 131 MEGs only nine were commonly marked by paternal H3K27me3 in Col and L*er* accessions (Hypergeometric test $p>0.1$; Fig 5B and Appendix Table S4), suggesting that H3K27me3 on paternal alleles is not a major determinant for genomic imprinting in Arabidopsis."
The color code used in Fig 6 corresponds to the same used in Fig 1B and Fig 4C. We have added this information to the legend.

9. Line 103 states about 70% if TE and genes marked with H3K27me3 overlapped, it looks to me that figure is more like 55% overall (62% for genes and 52% for TE based on Venn diagram in Figure 1A).
The overlap of genes marked by H3K27me3 in Col x Ler crosses with those marked in Ler x Col crosses is 77% (3177 out of 4116). The overlap of genes marked by H3K27me3 in Ler x Col crosses with those marked in Col x Ler crosses is 75% (3177 out of 4232). The overlap of TEs marked by H3K27me3 in Col x Ler crosses with those marked in Ler x Col crosses is 67% (5239 out of 7829). The overlap of TEs marked by H3K27me3 in Ler x Col crosses with those marked in Col x Ler crosses is 69% (5239 out of 7588). To ease understanding of the figure and the calculation, we have now included the percent of overlap into the figure.

10. Line 113 - I calculate 40% for maternal and 25% for paternal based on Fig S2B.
The overlap of genes marked by maternal-specific H3K27me3 in Col with those marked in Ler is 65% (1861 out of 2850). The overlap of genes marked by maternal-specific H3K27me3 in Ler with those marked in Col is 63% (1861 out of 2942). The overlap of TEs marked by maternal-specific H3K27me3 in Col with those marked in Ler is 53% (2683 out of 5054). The overlap of TEs marked by maternal-specific H3K27me3 in Ler with those marked in Col is 58% (2683 out of 5054).
The overlap of genes marked by paternal-specific H3K27me3 in Col with those marked in Ler is 32% (769 out of 2390). The overlap of genes marked by paternal-specific H3K27me3 in Ler with those marked in Col is 30% (769 out of 2585). The overlap of TEs marked by paternal-specific H3K27me3 in Col with those marked in Ler is 42% (2030 out of 4838). The overlap of TEs marked by paternal-specific H3K27me3 in Ler with those marked in Col is 46% (2030 out of 4430).
To ease understanding of the figure and the calculation, we have now included the percent of overlap into the figure. We also amended the text as follows: "About 60% of maternal-specific genes and about 50% of maternal-specific TEs overlapped between the accessions (Appendix Fig S2B), as well as about 30% of paternal-specific genes and about 40% of paternal-specific TEs (Appendix Fig S2B)."

11. A few minor typos as follows - Line 74 mentions PEG for first time but the abbreviation is not defined until later in text. Line 82 should read "to being located". Line 124 delete of to read "three previously published..." Line 174 insert space in "OfH3". Line 185 and elsewhere "lose" not "loose". Line 269 should read "explanation of how"

We have corrected all typos.

2nd Editorial Decision 18 February 2016

Thank you for submitting a revised version of your manuscript. It has now been seen by two of the original referees whose comments are shown below. As you will see the referees both acknowledge that the manuscript has improved relative to the original version; however, they do still raise a

number of remaining points that would have to be clarified/discussed further before they can recommend publication here.

I realize that the list of points is rather long but in my view this reflects genuine interest and appreciation from the referees of the value and quality of the presented dataset, although they would like to see a deeper and more conclusive analysis of this data. While involving further efforts on your side this additional rewriting/discussion would in our opinion strengthen the manuscript and help bring your message across also to the non-specialist reader.

I would therefore invite you to submit a final revision in which you address the remaining points raised by the referees. In addition, I would ask you to address the following editorial issues:

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 COMMENTS

Referee #1:

The revised manuscript and rebuttal addresses many points that were raised in the initial review, some figures have been added and others have been slightly improved. The writing remains strong and the figures are visually pleasing. My main concern is that data pointing to CG and H3K27me3 methylation co-occurring on the paternal pericentromere, in contrast to what is seen in the euchromatic regions is somewhat overlooked in the conclusions.

The data surrounding TE activation in fie endosperms is still confusing. While the authors now include methylation data on the TEs of interest in fie mutant endosperm, it is still not clear why certain TEs are upregulated in fie mutant endosperm. In all 4 groups of TEs (high fie-wt...low fie-wt) an increase in CG methylation is observed and a decrease in CHG/CHH is seen in fie mutants. This doesn't fully explain why expression of "group one" is reduced and "group four" is increased. For example, group four has low H3K27me3, and gains CG methylation in fie mutants, yet still has increased expression in fie. It appears that H3K27me3 allows some transcription and excludes other silencing marks when paired with high CG methylation in the pericentromere (group 1), whereas TEs not targeted by H3K27me3 nor CG (group 4) are upregulated in fie endosperm through an unknown mechanism, perhaps as a result of aberrant endosperm growth.

The authors add the line "We addressed the question whether TEs that remained silent in the fie mutant were additionally marked by DNA methylation." However, the implication from the heatmap is that several TEs are in fact down regulated in the fie mutant, suggesting that they were expressed at some level in the wild-type endosperm.

The methylation data shown for the four groups of TEs also suggests that H3K27me3 may be localized to CG containing regions in the pericentromere - the 1st group is highest for CG and H3K27me3 (Fig 2D and Fig S6). Several pieces of data in the paper suggest that pericentromeric TEs are targeted by H3K27me3 and CG methylation on the paternal genome, which should be mentioned in the conclusions.

The targeting of paternal H3K27me3 to "DME target regions" is not explained sufficiently, given that DME target regions are defined as regions containing CG in the paternal genome, and that this is the opposite of the authors' conclusions regarding eucrhomatic regions. While examining potential causes for this difference, the authors investigate H3K9me2 levels, and discover increased levels at the pericentromere on the maternal genome compared to the paternal genome. This suggests that pericentromeric "DME targets" are marked by H3K27me3 on the paternal side, and H3K9me2 on the maternal side, however this is not directly shown. Gypsy elements are shown instead which also fit the criteria of "DME targets" as they have increased paternal CG methylation, according to Figure 3C and the author's rebuttal. Showing "pericentromeric DME targets" would give better continuity to this portion of the manuscript.

The authors mention in their rebuttal that Gypsy elements have similar methylation status as the global TE analysis shown in Figure 3C. This is worth mentioning in the paper at least, as it fits with the idea that CG and H3K27me3 co-occur on the paternal genome in the pericentromere.

Minor
Line 118 - needs a space at "S2B).Accession-specific..."
Line 228 - "this regions" should be "these regions"

Referee #2:

The revised version of the manuscript has been modified only slightly. I feel that many of the points raised in my initial review have not been addressed in sufficient depth. First of all, it is still not clear if the levels of H3K27me3 methylation are comparable between endosperm and sporophytic tissues. The issue of cross-reactivity still needs to be properly investigated since it has been established by others that the antibody used to detect H3K27me3 does cross-react with H3K27me2 and H3K27me1 in the absence of H3K27me3 (Bouyer et al, PLoS Genetics 2011). This point is particularly important given that the authors claim that in the endosperm both H3K27me1 and H3K27me3 are present over heterochromatic regions, which typically are devoid of H3K27me3 in other tissues. I therefore still believe that a western blot analysis or ChIP-qPCR on individual loci, preferentially transposable elements (TEs) versus genes is required, to compare vegetative and endosperm tissues. By the same token, I still believe that the comparison between the previously (Weinhofer et al, PLoS Genetic 2010) and presently identified H3K27me3 marked TEs in the endosperm is important to reveal the consistency of this H3K27me3 marking, irrespective of the methods used. Therefore this information should be part of the current manuscript and not only provided to the reviewer. Although this manuscript uncovers very interesting chromatin dynamics in the endosperm, the functional relevance of the observations is less clear. Moreover, several conclusions are not actually supported by the data and should therefore be toned down or removed.

1) It is stated several times that H3K27me3 might replace H3K27me1/H3K9me2 over paternal TEs. Admittedly, although H3K27me3 is globally enriched and H3K27me1/H3K9me2 reduced over paternal vs. maternal alleles of GYPSY elements, there is no data indicating that H3K27me3 actually replaces the other two marks over these elements. The observed effects could result from different GYPSY copies being differently marked or from different marking over distinct parts of the elements.

2) The notion that H3K27me3 over paternal pericentromeric regions might contribute to the silencing of TEs located there is very unlikely as these elements tend to be not affected in fie endosperm. Moreover the authors show in the revised version of the manuscript that TEs that are reactivated most strongly in fie endosperm are only marginally marked by H3K27me3 and DNA methylation and that all TEs gain CG methylation in fie endosperm, irrespective of their level of deregulation. Also, DNA methylation and H3K27me3 are highly abundant over TEs that show no activation upon loss of H3K27me3 in the endosperm. In summary, these results do not allow the conclusion that H3K27me3 contributes to TE silencing in this tissue.

3) The authors suggest that DNA demethylation by DME allows the deposition of H3K27me3 over hypomethylated regions of paternally expressed imprinted genes. Yet H3K27me3 redistribution over DME-dependent loci is also observed to the same extent over paternal alleles that are generally not affected by DME. This complicates the general conclusion that it is the simple DNA demethylation that allows PRC2 to target these regions. The Discussion should therefore be modified accordingly.

4) An important finding of this work is that H3K27me3 and DNA methylation co-localize over TEs in the endosperm, in a parent-specific manner. These two chromatin modifications are largely mutually exclusive in vegetative tissues in plants and mutual exclusion also appears to be the norm in mammals (but with some notable exceptions). The authors should make better use of their epigenomic data to determine if co-localization is real or reflect deposition of these two marks over distinct parts of the same TEs or differential marking of the paternal and maternal alleles of the same TEs. Indeed, the coincidence of H3K27me3 and 5mC is mainly found over Col alleles and is almost absent in Ler alleles. Thus, the conclusion "that accession-specific differences in DNA methylation

can give rise to accession-specific differences in H3K27me3, providing an explanation for the large number of accession specific H3K27me3 marked regions" (Line 275) is erroneous. Reduced DNA methylation itself is unlikely the primary determinant of H3K27me3 deposition as this should be reflected in reduced DNA methylation over H3K27me3 marked regions, which is not the case (Col showing elevated and Ler low DNA methylation levels, respectively). Therefore DNA methylation differences between Ler and Col cannot explain the differential targeting by PRC2 between accessions.

Minor comments:
Line 68: "...PRC2 () modifies histones by applying histone trimethylation marks ()." The term 'applying' in the context of depositing covalent histone modifications seems awkward.
Line 96: "To understand the () biological significance of parental-specific () epigenetic modifications in the () endosperm, ..." What is meant by the term 'biological significance'? This needs to be indicated more precisely or changed.

Line 190: "... we found a marked reduction of paternal-specific H3K9me2 marked regions..." It is not indicated in which respect the reduction takes place. What is used as the 'reference' distribution? The large-scale comparison (genome-wide view) looks almost identical for the maternal and paternal genome in the endosperm as well as for vegetative tissue (Figure 4A).

Line 232: "() imprinting expression of all PEGs broke down in fie ()". What about non-H3K27me3 marked PEGs? These genes should be less or not affected in fie in case there is a direct relationship between H3K27me3 and the imprinting status of the underlying genes.

Line 273: "() H3K27me3 () is localized to DNA methylated regions targeted by DME." Regions targeted by DME are supposedly hypomethylated, therefore the sentence is misleading.

Line 306: "() at later stages of endosperm development the hypomethylated status of the maternal genome is maintained ()." In fact the authors clearly demonstrate that there is an increase in DNA methylation between 4DAP and 6DAP endosperm, although the maternal genome persistently shows less CG methylation compared to the paternal genome. Nonetheless, there is substantial remethylation over both, maternal and paternal alleles, within few days in all sequence contexts. This should be taken into account.

Line 327: "() our analysis reveals that FIS-PRC2 maintains epigenetic asymmetry between parental genomes during endosperm development as a response to differential epigenetic states established during female and male gametogenesis." This statement is far-fetched, given the lack of information/analysis of gametophytic chromatin modification profiles.

2nd Revision - authors' response 14 March 2016

Response to reviewers comments

Referee #1:

The revised manuscript and rebuttal addresses many points that were raised in the initial review, some figures have been added and others have been slightly improved. The writing remains strong and the figures are visually pleasing. My main concern is that data pointing to CG and H3K27me3 methylation co-occurring on the paternal pericentromere, in contrast to what is seen in the euchromatic regions is somewhat overlooked in the conclusions.
We have included a discussion about the co-occurrence of DNA methylation and H3K27me3 at pericentromeric regions in the manuscript, including in the abstract: "…H3K27me3 and DNA methylation are not necessarily exclusive marks at pericentromeric regions in the endosperm."
In the results: "Furthermore, the concomitant presence of H3K27me3 and CG DNA methylation at pericentromeric regions suggests that both marks are not necessarily exclusive."
In the discussion: "As shown in this study, H3K27me3 in the endosperm is localized to DNA demethylated regions targeted by DME but also co-localizes with DNA methylated pericentromeric regions, revealing that both modifications are not necessarily exclusive. Co-localization of both

modifications also occurs in animal cells but both modifications are mutually exclusive in CpG-dense regions (Statham *et al*, 2012; Brinkman *et al*, 2012)."

The data surrounding TE activation in fie endosperms is still confusing. While the authors now include methylation data on the TEs of interest in fie mutant endosperm, it is still not clear why certain TEs are upregulated in fie mutant endosperm. In all 4 groups of TEs (high fie-wt...low fie-wt) an increase in CG methylation is observed and a decrease in CHG/CHH is seen in fie mutants. This doesn't fully explain why expression of "group one" is reduced and "group four" is increased. For example, group four has low H3K27me3, and gains CG methylation in fie mutants, yet still has increased expression in fie. It appears that H3K27me3 allows some transcription and excludes other silencing marks when paired with high CG methylation in the pericentromere (group 1), whereas TEs not targeted by H3K27me3 nor CG (group 4) are upregulated in fie endosperm through an unknown mechanism, perhaps as a result of aberrant endosperm growth.
We accounted for this concern by including the following explanatory paragraph: "Those TEs with low expression in *fie* were preferentially localized in pericentromeric regions (Fig 2C) and densely marked by CG DNA methylation that furthermore increased upon FIE depletion (Appendix Fig S6), correlating with reduced TE expression in *fie*. The most highly upregulated TEs in *fie* were neither substantially marked by H3K27me3 nor by DNA methylation (Fig 2D and Appendix Fig S6), suggesting an unknown mechanism responsible for their upregulation, possibly connected to increased endosperm growth. "
We furthermore have toned down our conclusion about the effect of H3K27me3 in silencing of TEs: "In summary, our data suggest that H3K27me3 is not a major silencing mark for TEs in the endosperm; however, the additional regulation by DNA methylation and unknown factors may potentially mask the effect of H3K27me3 depletion."

The authors add the line "We addressed the question whether TEs that remained silent in the fie mutant were additionally marked by DNA methylation." However, the implication from the heatmap is that several TEs are in fact down regulated in the fie mutant, suggesting that they were expressed at some level in the wild-type endosperm.
We agree with the reviewer and made the following addition: "We addressed the question whether TEs that remained silent **or were downregulated** in the *fie* mutant were additionally marked by DNA methylation."

The methylation data shown for the four groups of TEs also suggests that H3K27me3 may be localized to CG containing regions in the pericentromere - the 1st group is highest for CG and H3K27me3 (Fig 2D and Fig S6). Several pieces of data in the paper suggest that pericentromeric TEs are targeted by H3K27me3 and CG methylation on the paternal genome, which should be mentioned in the conclusions.
See response to comment 1.

The targeting of paternal H3K27me3 to "DME target regions" is not explained sufficiently, given that DME target regions are defined as regions containing CG in the paternal genome, and that this is the opposite of the authors' conclusions regarding eucrhomatic regions. While examining potential causes for this difference, the authors investigate H3K9me2 levels, and discover increased levels at the pericentromere on the maternal genome compared to the paternal genome. This suggests that pericentromeric "DME targets" are marked by H3K27me3 on the paternal side, and H3K9me2 on the maternal side, however this is not directly shown. Gypsy elements are shown instead which also fit the criteria of "DME targets" as they have increased paternal CG methylation, according to Figure 3C and the author's rebuttal. Showing "pericentromeric DME targets" would give better continuity to this portion of the manuscript.
As suggested by the reviewer we included the analysis of chromatin modifications on DME target regions as Fig 5G-I into the manuscript and included the following text. "While pericentromeric DME regions had similar levels of H3K27me3 as gypsy elements (Figure 5A and G), pericentromeric DME target regions were not substantially enriched for H3K9me2 and H3K27me1 on both parental alleles (Fig 5H and I), which may promote their targeting by H3K27me3. Depletion of heterochromatic modifications on DME target sites is in agreement with previously published data (Ibarra et al, 2012)."
The authors mention in their rebuttal that Gypsy elements have similar methylation status as the global TE analysis shown in Figure 3C. This is worth mentioning in the paper at least, as it fits with the idea that CG and H3K27me3 co-occur on the paternal genome in the pericentromere.

We agree with the reviewer and have emphasized this as follows: "Together, we conclude that specific localization of H3K27me3 on maternal alleles occurs on sites that undergo DME-mediated hypomethylation in the central cell, while H3K27me3 at paternal pericentromeric regions co-occurs with DNA methylation."

Minor
Line 118 - needs a space at "S2B).Accession-specific..."
Line 228 - "this regions" should be "these regions"
We corrected those mistakes.

Referee #2:

The revised version of the manuscript has been modified only slightly. I feel that many of the points raised in my initial review have not been addressed in sufficient depth. First of all, it is still not clear if the levels of H3K27me3 methylation are comparable between endosperm and sporophytic tissues. The issue of cross-reactivity still needs to be properly investigated since it has been established by others that the antibody used to detect H3K27me3 does cross-react with H3K27me2 and H3K27me1 in the absence of H3K27me3 (Bouyer et al, PLoS Genetics 2011). This point is particularly important given that the authors claim that in the endosperm both H3K27me1 and H3K27me3 are present over heterochromatic regions, which typically are devoid of H3K27me3 in other tissues. I therefore still believe that a western blot analysis or ChIP-qPCR on individual loci, preferentially transposable elements (TEs) versus genes is required, to compare vegetative and endosperm tissues.
To account for this concern, we included an Appendix Figure 12 showing qPCR data of ChIP experiments using antibodies against H3K27me3, H3k9me2 and H3K27me1. We tested enrichment of those marks on genes and TEs in endosperm and leaf tissues. The data confirm that TEs which are specifically marked by H3K27me3 in the endosperm are not marked in leaves, while genes that have been published to be marked in the endosperm and in leaves have comparable levels for both modifications. The data furthermore show that TEs marked by H3K27me3 are also marked by heterochromatic H3K9me2 and H3K27me1, consistent with our genome-wide ChIP data.
By the same token, I still believe that the comparison between the previously (Weinhofer et al, PLoS Genetic 2010) and presently identified H3K27me3 marked TEs in the endosperm is important to reveal the consistency of this H3K27me3 marking, irrespective of the methods used. Therefore this information should be part of the current manuscript and not only provided to the reviewer.
We have included this comparison as follows in the text: Targeting of TEs by H3K27me3 has been previously reported (Weinhofer et al, 2010) and 64% of previously identified TEs were also identified in our new dataset ($p<1.5E-60$; Table EV2).

Although this manuscript uncovers very interesting chromatin dynamics in the endosperm, the functional relevance of the observations is less clear. Moreover, several conclusions are not actually supported by the data and should therefore be toned down or removed.

1) It is stated several times that H3K27me3 might replace H3K27me1/H3K9me2 over paternal TEs. Admittedly, although H3K27me3 is globally enriched and H3K27me1/H3K9me2 reduced over paternal vs. maternal alleles of GYPSY elements, there is no data indicating that H3K27me3 actually replaces the other two marks over these elements. The observed effects could result from different GYPSY copies being differently marked or from different marking over distinct parts of the elements.
As requested by the reviewer, we removed the statement that H3K27me3 compensates for reduced H3K9me2 and H3K27me1. We rephrased that parental-specific differences in the distribution of H3K27me3 at pericentromeric regions correlate with parental-specific differences in H3K9me2 and H3K27me1 distribution.

2) The notion that H3K27me3 over paternal pericentromeric regions might contribute to the silencing of TEs located there is very unlikely as these elements tend to be not affected in fie endosperm. Moreover the authors show in the revised version of the manuscript that TEs that are reactivated most strongly in fie endosperm are only marginally marked by H3K27me3 and DNA methylation and that all TEs gain CG methylation in fie endosperm, irrespective of their level of deregulation. Also, DNA methylation and H3K27me3 are highly abundant over TEs that show no activation upon loss of H3K27me3 in the endosperm. In summary, these results do not allow the conclusion that H3K27me3 contributes to TE silencing in this tissue.

As suggested by the reviewer we rephrased our conclusion as follows: "In summary, our data suggest that H3K27me3 is not a major silencing mark for TEs in the endosperm; however, the additional regulation by DNA methylation and unknown factors may potentially mask the effect of H3K27me3 depletion."

3) The authors suggest that DNA demethylation by DME allows the deposition of H3K27me3 over hypomethylated regions of paternally expressed imprinted genes. Yet H3K27me3 redistribution over DME-dependent loci is also observed to the same extent over paternal alleles that are generally not affected by DME. This complicates the general conclusion that it is the simple DNA demethylation that allows PRC2 to target these regions. The Discussion should therefore be modified accordingly.
We have accounted for this concern and have discussed the difference of DME-target loci enriched for maternal H3K27me3 compared to those enriched for paternal H3K27me3. The latter are located in pericentromeric regions and are depleted for H3K9me2 and H3K27me1, which may allow them to be targeted by H3K27me3. We also discussed that "the concomitant presence of H3K27me3 and CG DNA methylation at pericentromeric regions suggests that both marks are not necessarily exclusive."

4) An important finding of this work is that H3K27me3 and DNA methylation co-localize over TEs in the endosperm, in a parent-specific manner. These two chromatin modifications are largely mutually exclusive in vegetative tissues in plants and mutual exclusion also appears to be the norm in mammals (but with some notable exceptions). The authors should make better use of their epigenomic data to determine if co-localization is real or reflect deposition of these two marks over distinct parts of the same TEs or differential marking of the paternal and maternal alleles of the same TEs. Indeed, the coincidence of H3K27me3 and 5mC is mainly found over Col alleles and is almost absent in Ler alleles. Thus, the conclusion "that accession-specific differences in DNA methylation can give rise to accession-specific differences in H3K27me3, providing an explanation for the large number of accession specific H3K27me3 marked regions" (Line 275) is erroneous. Reduced DNA methylation itself is
unlikely the primary determinant of H3K27me3 deposition as this should be reflected in reduced DNA methylation over H3K27me3 marked regions, which is not the case (Col showing elevated and Ler low DNA methylation levels, respectively). Therefore DNA methylation differences between Ler and Col cannot explain the differential targeting by PRC2 between accessions.
We have rephrased accordingly: "Accession-specific H3K27me3 marked regions differed substantially in their DNA methylation levels between the accessions (Appendix Fig S2C) and revealed high CG methylation levels at paternal H3K27me3 marked regions in the Col accession." And in the discussion we rephrased as follows: "As shown in this study, H3K27me3 in the endosperm is localized to DNA demethylated regions targeted by DME but also co-localizes with DNA methylated pericentromeric regions. We observed substantial differences in DNA methylation at accession-specific H3K27me3 marked regions with increased DNA methylation in the Col compared to the Ler accession. Whether the accession-specific differences in DNA methylation can give rise to accession-specific differences in H3K27me3 remains to be tested." We like to point out that the co-localization of DNA methylation and H3K27me3 also occurs in animals and both marks are only exclusive at CpG islands. We have included this information in the discussion: "Co-localization of both modifications also occurs in animal cells but both modifications are mutually exclusive in CpG-dense regions (Statham et al, 2012; Brinkman et al, 2012)."

Minor comments:
Line 68: "...PRC2 () modifies histones by applying histone trimethylation marks ()." The term 'applying' in the context of depositing covalent histone modifications seems awkward.
We have changed to "depositing histone trimethylation marks"

Line 96: "To understand the () biological significance of parental-specific () epigenetic modifications in the () endosperm, ..." What is meant by the term 'biological significance'? This needs to be indicated more precisely or changed.
We changed to "biological importance".

Line 190: "... we found a marked reduction of paternal-specific H3K9me2 marked regions..." It is not indicated in which respect the reduction takes place. What is used as the 'reference' distribution? The large-scale comparison (genome-wide view) looks almost identical for the

maternal and paternal genome in the endosperm as well as for vegetative tissue (Figure 4A).

The comparison is to maternal-specific H3K9me2 regions (lower panel in Figure 4A), which has now been indicated.

Line 232: "() imprinting expression of all PEGs broke down in fie ()". What about non-H3K27me3 marked PEGs? These genes should be less or not affected in fie in case there is a direct relationship between H3K27me3 and the imprinting status of the underlying genes.

There are only six PEGs without maternal H3K27me3 and of those three are affected in fie. We included this information as follows into the manuscript: "The maternal alleles of three out of six PEGs that were not marked by maternal H3K27me3 became upregulated in *fie* (Table EV3). Whether distantly located H3K27me3-marked regions establish imprinted expression of those genes remains to be investigated."

Line 273: "() H3K27me3 () is localized to DNA methylated regions targeted by DME." Regions targeted by DME are supposedly hypomethylated, therefore the sentence is misleading.
We corrected to "is localized to DNA demethylated regions targeted by DME"

Line 306: "() at later stages of endosperm development the hypomethylated status of the maternal genome is maintained ()." In fact the authors clearly demonstrate that there is an increase in DNA methylation between 4DAP and 6DAP endosperm, although the maternal genome persistently shows less CG methylation compared to the paternal genome. Nonetheless, there is substantial remethylation over both, maternal and paternal alleles, within few days in all sequence contexts. This should be taken into account.
To address this concern we have rephrased as follows: "At later stages of endosperm development reduced DNA methylation of the maternal compared to the paternal genome is maintained, even though *de novo* DNA methylation has resumed, as judged by increased DNA methylation levels in all sequence contexts and expression of *DRM1* and *DRM2* (Belmonte *et al*, 2013)."

Line 327: "() our analysis reveals that FIS-PRC2 maintains epigenetic asymmetry between parental genomes during endosperm development as a response to differential epigenetic states established during female and male gametogenesis." This statement is far-fetched, given the lack of information/analysis of gametophytic chromatin modification profiles.
We toned down the conclusion as follows: "In conclusion, our analyses reveal that FIS-PRC2 maintains epigenetic asymmetry between parental genomes during endosperm development leading to paternal-specific expression of imprinted genes."

Acceptance 17 March 2016

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

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: Claudia Köhler
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2015-93534R

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

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.

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

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?	We have generated ChIP-sequencing and bisulfite sequencing data in biological duplicates. For ChIP experiments Arabidopsis siliques were pooled from a minumum of 300 crosses of 5-10 individual plants, over different days till obtaining the neessary quantity to perform the experiment. The experiment was repeated with differetn plants grown at a differetn timepoint. For bisulfite sequencing experiments half the number of crosses was used.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Only endosperm purifications with low seed coat/embryo contamination were analysed, estimated based on the expected endosperm maternal:paternal ratio 2:1. Contamination levels did not exceed 12% (see Material and Methods, INTACT Nuclei Purification section)
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	The statistical tests used in this manuscript do not assume normally distributed data.
Is there an estimate of variation within each group of data?	A replicability assessment of our sequencing data is shown in Table EV6.
Is the variance similar between the groups that are being statistically compared?	Yes.

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/

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).	All antibodies used are listed in the Material and Methods indicating catalog number.
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

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

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

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

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under '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	Sequencing data generated in that work has been deposited in Gene Expression Omnibus (GSE66585) as referenced in Materials and Methods. Previosly published data used has been also referenced in the paper.
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 have been deposited in public repositories and accession numbers indicated in the manuscript.
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. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: **Primary Data** Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in Shewanella oneidensis MR-1. Gene Expression Omnibus GSE39462 **Referenced Data** Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	Data generated and data already published used in the paper has been referenced in Materials and Methods.
22. 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

23. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	No