THE EMBO JOURNAL

Manuscript EMBO-2012-84234

# *Arabidopsis* MSI1 connects LHP1 to PRC2 complexes

Maria Derkacheva, Yvonne Steinbach, Thomas Wildhaber, Iva Mozgova, Walid Mahrez, Paolo Nanni, Sylvain Bischof, Wilhelm Gruissem and Lars Hennig

*Corresponding author: Lars Hennig, Swedish University of Agricultural Sciences*

**Review timeline:**

| | |
|---|---|
| Submission date: | 19 December 2012 |
| Editorial Decision: | 29 January 2013 |
| Revision received: | 23 April 2013 |
| Editorial Decision: | 13 May 2013 |
| Revision received: | 15 May 2013 |
| Accepted: | 17 May 2013 |

## Transaction Report:

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

Editor: Karin Dumstrei

1st Editorial Decision 29 January 2013

Thank you for submitting your manuscript to the EMBO Journal. Your study has now been three referees and their comments are provided below.

As you can see below the referees appreciate the analysis. However revisions are also needed in order to consider publication here. In particular the link between MSI1-LHP1 needs to be significantly strengthen. Should you be able to address the concerns raised in full then we would consider a revised version. I should add that it is EMBO Journal policy to allow only a single round of revision, and acceptance of your manuscript will therefore depend on the completeness of your responses in this revised version.

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

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

REFEREE REPORTS

Referee #1

Polycomb-group proteins in plants likely exist in different complexes which control different developmental transitions. While proteins of the so called Polycomb repressive complex 2 (PRC2) are conserved in plants, several members of PRC1 are lacking in plants. The study by Derkacheva et

 

al. focuses on the PRC2 component MSI1 which is likely present in all different PRC2 complex. So far, a role for MSI1 in embryo/endosperm development and repression of flowering is well established while it is unclear whether MSI1 also acts in the vernalization pathway (induction of flowering after prolonged cold treatment). MSI1 has four homologues in the Arabidopsis genome, however, only a link to Polycomb proteins was previously revealed for MSI4 (and not for MSI2, 3 and5).
To reveal the complexes MSI1 is part of the authors now performed IP with FLAG-EMF2 (a PRC2 component) and MSI1-GFP experiments and consecutive mass spectrometry. While some expected candidates are found in both IPs several are missing, this is CLF, the histone methyltransferase of the PRC2 (homologous to SWN), and proteins of the CAF complexes and histone deacetylase complex members (for MSI1). Surprisingly, MSI1 co-purifies with LHP1, a protein considered to be part of PRC1.
Subsequently, the authors reveal a role for MSI1 in repression of several novel Pc-G targets in leaves and in the vernalization response. The major novelty of this manuscript is the (direct) interaction of MSI1/PRC2 and LHP1 on which the authors base an interesting model. This suggests that LHP1 may be required for reestablishing H3K27me3 after replication as LHP1 binds to H3K27me3 and may recruit PRC2 to establish H3K27me3 on the newly synthesized DNA/Histones. However, this remains rather speculative and additional experiments are required to find further evidence for this model. As a whole, the manuscript provides interesting and novel data; however, most of the data confirm and extend previous knowledge on MSI1, while the novel data (MSI1-LHP1) appears premature.
Major points:
- if LHP1 is required for re-establishing H3K27me3, then a reduction of H3K27me3 would be expected which is apparently not the case (see Turck et al. 2007). The authors should also provide immunoblot analyses of H3K27me3 in lhp1 to check this possibility as the previous data were based on ChIP-chip results (which will not detect a small decrease in H3K27me3).
- The authors discuss that the low number of LHP1 peptides in the IP experiment suggests that LHP1 is only transiently associated with PRC2/MSI1, however, CLF and the CAF complex are not found at all and FIE only in two of the replicates, suggesting that the IP experiment has even not recovered always the expected candidates. It cannot be ruled out that LHP1 may be a PRC2 member and has no function in a putative plant PRC1, therefore the reciprocal IP-massspectrometry with LHP1-GFP or a similar line is essential to see whether LHP1 co-purifies with PRC2 or a putative PRC1. In addition, a Co-IP MSI1 and CLF should be performed as the IP experiments suggest that CLF and MSI1 may not be in the same complex.

- In the methods section, the authors write: "Proteins identified with at least two unique peptides in at least two replicates but never in control samples were taken into account." Therefore, LHP1 should not have been included in the list or all proteins following the LHP1 criterion should be provided. I therefore suggest that the full dataset is provided in a supplementary table/file.
- The analysis of expression changes in lhp1 and clf (fig. 4) and msi1-cs and emf2 (Fig. 1) appears strange. It is well established that MSI, CLF, EMF2 and LHP1 repress common targets (e.g.Hennig et al. 2003, Turck et al. 2007, Farrona et al. 2011)? Do we need that information? I am not convinced that comparing a lhp1 ft double with clf in SD will diminish variability. On the other hand, for figure 1 they use datasets from very differently grown plants. Therefore, the four transcriptomic datasets should be compared with each other and a venn diagram provided for the overlap of genes.

Minor points:
- Fig. 1 and 4: The blots have been cropped a lot, so please provide a less cropped version (at least in the sup data)
- Fig. 1: Why was a flag-emf2 line used for the IP experiment and an yfp-emf2 line for the co-IP? Why was not the same line for everything?
- It is hard to follow from which promoters the tagged cDNAs where expressed. Please specify this in the legends and the methods.
- From which region is the intergenic region analysed in figure 2D?

- The immunoblot should be provided for figure 2B and not only the graph.

 

Referee #2

Polycomb mediated repression plays an essential role in the development of many higher eukaryotes, but despite the conservation of the polycomb system between organisms there is still some debate about the composition and function of the PRC2 complex in plants. This manuscript reports the composition of the PRC 2 complexes in somatic tissues in Arabidopsis. The authors show that plant PRC2 complexes contain MSI1, a homolog of the Drosophila protein p55. MSI1 was shown to be essential for the function of the EMF2 and VRN2 complexes, playing a role in the trimethylation of H3K27 and in gene repression. MSI1 interacts with both EMF2 and LHP1, and the authors propose that the LHP1-MSI1 interaction forms a positive feedback loop that recruits PRC2 complexes to target chromatin modified by H3K27me3 facilitating the inheritance of this modification.
This manuscript provides strong evidence for the role of MSI1 in PRC2 function and shows that none of the other p55 homologs in Arabidopsis can complement for loss of MSI1 function. This is important because the role of MSI2-5 in PRC2 repression has not been clear. However, there are some inconsistencies in the data presented (see below) that I found puzzling and that I think that the authors need to discuss. I have outlined my criticisms below for the authors' consideration:
1. Tables 1 & 2 present the data identifying the proteins immunoprecipitated with EMF2 and MSI1 respectively. LHP1 does not feature in Table 1, even though EMF2 is apparently strongly co-precipitated with LHP1-GFP (Figure 4C). In contrast, LHP1 is represented in the complex immunoprecipitated with MSI1-GFP, albeit in only three of the four experiments where it is represented by only 1 peptide. When looking at Figure 4 A & B it is clear that there is an interaction between LHP1 and MSI1, but the apparent enrichment of the co-immunoprecipitated protein (either MSI1 in panel A or LHP1 in panel B) is much lower than for EMF2 in panel C, suggesting that the interaction between MSI1 and LHP1 is weaker than that between EMF2 and LHP1. Why then is LHP1 not represented in the complex found in pulldown with FLAG-EMF2 (Table 1)?
2. Figure 2; the estimated reduction of H3K27me3 globally is very different from that at the individual gene level. Is this just an artefact of the different techniques used to estimate the level of change or do the authors think that there may be regions which do not rely on MSI1 for trimethylation of H3K27?
3. Figure 3; it is a little puzzling to see that FLC expression is partially reduced by vernalization of msiI-as plants, but that this is not reflected in the flowering time which is unchanged. Any comment?
4. Figure 4E; I do not understand this figure - for example, the Y axis is labelled "Maximal gene expression in leaves" , but it is not clear to me what the box-whiskers plot actually shows. Is it the average level of maximal expression across all expressed genes? How meaningful is this? I also note that the plots for PcG and "not up" look identical - is this a coincidence or a mistake?
5. Discussion p12 line 13; "but the presence of a PRC1 function in plants has not yet been firmly established". I don't agree with this comment as I believe that there is now sufficient data to support the existence and function of PRC1 in plants; for example, I refer the authors to Chen et al (2010 ) Cell Research 20: 1332 and Bratzel et al (2010) Curr Biol. 20: 1853 amongst others.
6. The other piece of data that is apparently inconsistent with published literature is the failure to detect CLF in the EMF2 complex - the major sporophytic PRC2 in plants. I find this puzzling because clf mutants have quite a strong phenotype, but swn mutants show no obvious developmental defects (Chanvivattana et al [2004] Devel. 131:5263), so this doesn't quite fit. The authors have discussed this to some extent on p10 -11 but not in the context of the phenotypes of the clf and swn mutants.
7. If the model presented in Figure 5 is correct, then one would predict that there would be a reduction of H3K27me3 in an lhp1 mutant at least at genes whose repression depends on LHP1. Is this the case? It has previously been shown that the distribution of H3K27me3 is unchanged in an lhp1 mutant (Turck et al., 2007), but I don't believe the level of the modification has been examined in the mutant background so this would be a useful contribution to our understanding.
Minor criticisms:
8. Legend to Figure 1 needs some reorganization. (A) includes the sentence "Lack of MSI1 and lack of EMF2 cause similar changes in the transcriptome" but this information relates to panel (B) and should appear after not before the letter (B).
9. The reference Schatlowski et al is missing from the bibliography

 

Referee #3

The m/s describes mass spectrometry analysis of immunoprecipitated (MS IP) plant Polycomb group (PcG) protein complexes. They show that MSI1 is associated in Arabidopsis with EMF2 and with LHP1 and provide some co-IP, ChIP and gene expression analysis to support that the interactions are functionally relevant. Although there has been much analysis of animal PcG complexes, there has been relatively little in vivo analysis of the plant complexes. As such, I think the paper is of general interest for the plant epigenetics community and is suitable for publication in EMBO J. The study I think is well done and I have few comments, which are minor ones.

- It is not always made clear in methods what tissues are being used for the various IP's and mass spec analysis. In some cases inflorescence tissue is stated, but in many of the experiments this is not clear.
- The MS and microarray experiments are presented in highly summarised forms in results. It would be useful for the community to have the data presented more fully in supplementary information rather than just a graph (4D) or the top hits (MS data). Although the raw data has been submitted to databases, it would be useful to have fuller analyses presented in the SI.
- The interaction between MSI1 and EMF2 found by MS-IP is validated by co-IP from transgenic tobacco plants whereas that between LHP1 and MSI1 by co-IP from transgenic Arabidopsis. Although this is slightly inconsistent and it would be preferable to have done both in Arabidopsis, I think it is acceptable given that they have done reciprocal MS-IP with both tagged EMF2 and with tagged MSI1 and have four replicates for each experiment.
- I found figure 5E very hard to follow and feel that it would better benefit from better explanation in the main text and legend.
- In Figure 2A I think they should show the western the data is based as well as (or instead of) the graph summary, which gives little idea of what the blot it is based on looks like.
- The authors do not find a role for MSI4 in that their MS IP of EMF2 and MSI1 do not reveal MSI4 but rather MSI1. As cited by the authors, the Genshik group have published IP's showing that MSI4 and CLF co-IP from Arabidopsis. The authors here conclude MSI4 indirectly affects PRC2 activity via its role in other complexes. I think this is premature and they should leave the possibility that MSI4 is present in PRC2 complexes open. What they see may well depend on the tissue used, the extraction buffers, the tags and the component used for IP. After all, CLF which almost certainly is a PRC2 member, is also not seen in their IP's with EMF2 or MSI1, so the negative results do not necessarily mean that MSI4 is not a member of some PRC2 complexes.

1st Revision - authors' response 23 April 2013

Referee #1

- if LHP1 is required for re-establishing H3K27me3, then a reduction of H3K27me3 would be expected which is apparently not the case (see Turck et al. 2007). The authors should also provide immunoblot analyses of H3K27me3 in lhp1 to check this possibility as the previous data were based on ChIP-chip results (which will not detect a small decrease in H3K27me3).
**We agree with the reviewer that this is an important test of our hypothesis. We performed the proposed ChIP experiment and the results fully support our hypothesis.**

- The authors discuss that the low number of LHP1 peptides in the IP experiment suggests that LHP1 is only transiently associated with PRC2/MSI1, however, CLF and the CAF complex are not found at all and FIE only in two of the replicates, suggesting that the IP experiment has even not recovered always the expected candidates. It cannot be ruled out that LHP1 may be a PRC2 member and has no function in a putative plant PRC1, therefore the reciprocal IP-mass spectrometry with LHP1-GFP or a similar line is essential to see whether LHP1 co-purifies with PRC2 or a putative PRC1.
**In order to substantiate the interaction of MSI1 and LHP1 we performed additional immunoaffinity purifications with MS1-GFP using a modified protocol that involved cross-linking of proteins prior to immunoprecipitation. This experiment, which is summarized in the new Table 3, resulted in a robust identification of LHP1. We agree with the reviewer that our data do not allow us to discuss whether LHP1 interacts transiently or not with PRC2, and therefore removed this text from the discussion.**

 

**The additional data strongly support our conclusion that LHP1 interacts with PRC2 in vivo. The focus of our study is the interaction of LHP1 and PRC2, which we have now firmly established. Although a model that LHP1 functions only with PRC2 but not with PRC1 is formally possible, published evidence supports a role of LHP1 in a plant PRC1-like complex. Therefore, we prefer to avoid speculation whether LHP1 is only a PRC2 or also a PRC1 subunit. The test of this hypothesis is out of the focus of our current study.**

In addition, a Co-IP MSI1 and CLF should be performed as the IP experiments suggest that CLF and MSI1 may not be in the same complex.
**The proposed Co-IP experiment was done to test MSI1-CLF interactions in vivo. The new panel B in Fig. 1 shows that MSI1 and CLF do interact in this *in vivo* assay.**

- In the methods section, the authors write: "Proteins identified with at least two unique peptides in at least two replicates but never in control samples were taken into account." Therefore, LHP1 should not have been included in the list or all proteins following the LHP1 criterion should be provided. I therefore suggest that the full dataset is provided in a supplementary table/file.
**We agree with the reviewer that LHP1 did not pass the criteria for reliable protein identification in the original MSI1 pulldown and removed LHP1 from the original Table 1. The MS/MS identification of LHP1 using the modified Co-IP method is now shown in the new Table 3. In addition, we include the full list of all identified interacting proteins in Supplementary Tables S1 and S2.**

- The analysis of expression changes in lhp1 and clf (fig. 4) and msi1-cs and emf2 (Fig. 1) appears strange. It is well established that MSI, CLF, EMF2 and LHP1 repress common targets (e.g. Hennig et al. 2003, Turck et al. 2007, Farrona et al. 2011)? Do we need that information? I am not convinced that comparing a lhp1 ft double with clf in SD will diminish variability. On the other hand, for figure 1 they use datasets from very differently grown plants. Therefore, the four transcriptomic datasets should be compared with each other and a venn diagram provided for the overlap of genes.
**Our transcriptome data analysis provides novel information and is not redundant with published reports. Concerning MSI1, RT-PCR data for only two PcG target genes were previously reported (AG and AP2). The strong genome-wide correlation shown in Fig. 1 provides very different evidence. Similarly, a genome-wide comparison of transcriptional changes in *lhp1* vs. other Arabidopsis PcG mutants has not been reported. Therefore, we are convinced that the presented genome-wide transcriptomes add substantial evidence to link LHP1 and MSI1 to PRC2 function. Our own work and that of other groups (e.g., from the Goto and Goodrich laboratories) has shown that *clf* and *lhp1* developmental phenotypes strongly depend on ectopic *FT* expression. Conditions that prevent *FT* upregulation make the morphology of the mutants much more similar to WT and help to reduce the number of genes with altered expression due to secondary effects. Therefore, we are convinced that our experimental design helps to focus on PcG function. For an alternative representation of the transcriptome data we added the proposed Venn diagrams (Supplementary Fig. S2).**

Minor points:

- Fig. 1: Why was a flag-emf2 line used for the IP experiment and an yfp-emf2 line for the co-IP? Why was not the same line for everything?
**The *FLAG-EMF2* construct was used to generate an Arabidopsis line for the IP. The Co-IP was performed in tobacco in which the *FLAG-EMF2* construct was not expressed and we therefore changed to an alternative tag.**

- It is hard to follow from which promoters the tagged cDNAs where expressed. Please specify this in the legends and the methods.
**The manuscript has been revised and now includes all promoter information.**

- From which region is the intergenic region analysed in figure 2D?
**The reviewer probably refers to Figure 1D. The information is now provided in the manuscript. The intergenic region is on chromosome 1 between 8383019- 8383083 (between *At1G23700* and *At1G23710*).**

 

- The immunoblot should be provided for figure 2B and not only the graph.
**The immunoblot is now included as Supplemental Figure 1.**

Referee #2

1. Tables 1 & 2 present the data identifying the proteins immunoprecipitated with EMF2 and MSI1 respectively. LHP1 does not feature in Table 1, even though EMF2 is apparently strongly co-precipitated with LHP1-GFP (Figure 4C). In contrast, LHP1 is represented in the complex immunoprecipitated with MSI1-GFP, albeit in only three of the four experiments where it is represented by only 1 peptide. When looking at Figure 4 A & B it is clear that there is an interaction between LHP1 and MSI1, but the apparent enrichment of the co-immunoprecipitated protein (either MSI1 in panel A or LHP1 in panel B) is much lower than for EMF2 in panel C, suggesting that the interaction between MSI1 and LHP1 is weaker than that between EMF2 and LHP1. Why then is LHP1 not represented in the complex found in pulldown with FLAG-EMF2 (Table 1)?
**The original experimental set-up does not allow quantitative statements about subunit abundance. In order to collect additional evidence for the LHP1-PRC2 interaction, new co-immunoprecipitation experiments were performed with MSI1-GFP using a modified protocol involving cross-linking of proteins. The new data add very strong additional support to our original conclusions.**
2. Figure 2; the estimated reduction of H3K27me3 globally is very different from that at the individual gene level. Is this just an artefact of the different techniques used to estimate the level of change or do the authors think that there may be regions which do not rely on MSI1 for trimethylation of H3K27?
**The msi1 knockdown line used in our study maintains about 5-10% of wild-type MSI1 levels (Hennig et al., 2003). Thus, some complete PRC2 complex can likely be formed even in this line. Efficacy of PcG protein targeting likely differs between target genes and therefore H3K27me3 reduction will likely differ as well between target genes that compete for the residual PRC2 complex available in this line. Future experiments will need to test this hypothesis.**

3. Figure 3; it is a little puzzling to see that FLC expression is partially reduced by vernalization of msiI-as plants, but that this is not reflected in the flowering time which is unchanged. Any comment?
**We think that the partial reduction of *FLC* expression is not sufficient to cause a substantial acceleration of flowering.**

4. Figure 4E; I do not understand this figure - for example, the Y axis is labelled "Maximal gene expression in leaves" , but it is not clear to me what the box-whiskers plot actually shows. Is it the average level of maximal expression across all expressed genes? How meaningful is this? I also note that the plots for PcG and "not up" look identical - is this a coincidence or a mistake?
**We apologize for the original description of the figure and added additional explanation to the legend and main text. Hopefully this analysis is now clear. The comparison of expression potential (i.e. the maximal expression that a gene can obtain in an organ) across groups of genes is meaningful because it is a generally accepted method of identifying enrichment for genes with organ-specific expression.**
**The similar plots are indeed a coincidence because the majority of PcG target genes are not upregulated.**

5. Discussion p12 line 13; "but the presence of a PRC1 function in plants has not yet been firmly established". I don't agree with this comment as I believe that there is now sufficient data to support the existence and function of PRC1 in plants; for example, I refer the authors to Chen et al (2010 ) Cell Research 20: 1332 and Bratzel et al (2010) Curr Biol. 20: 1853 amongst others.
**We agree with the reviewer and revised the text accordingly. The sentence now reads "… but the identity of PRC1 complexes in plants has not yet been fully established".**

6. The other piece of data that is apparently inconsistent with published literature is the failure to detect CLF in the EMF2 complex - the major sporophytic PRC2 in plants. I find this puzzling because clf mutants have quite a strong phenotype, but swn mutants show no obvious developmental

 

defects (Chanvivattana et al [2004] Devel. 131:5263), so this doesn't quite fit. The authors have discussed this to some extent on p10 -11 but not in the context of the phenotypes of the clf and swn mutants.
**We agree with the reviewer that this is an unexpected finding. In order to test MSI1-CLF interactions *in vivo*, the proposed Co-IP experiment was done and the discussion was extended as suggested. The new panel B in Fig. 1 shows that MSI1 and CLF interact in this *in vivo* assay. Therefore, the absence of CLF from the original immunoprecipitations is not good evidence against a function of CLF in Arabidopsis PRC2 complexes. Note that also de Lucia et al 2008 failed to detect CLF in the VRN complex despite convincing other evidence for CLF function in this complex.**

7. If the model presented in Figure 5 is correct, then one would predict that there would be a reduction of H3K27me3 in an lhp1 mutant at least at genes whose repression depends on LHP1. Is this the case? It has previously been shown that the distribution of H3K27me3 is unchanged in an lhp1 mutant (Turck et al., 2007), but I don't believe the level of the modification has been examined in the mutant background so this would be a useful contribution to our understanding.
**We agree with the reviewer that this is an important test of our hypothesis and performed the proposed ChIP experiment. The results fully support our hypothesis.**

Minor criticisms:

8. Legend to Figure 1 needs some reorganization. (A) includes the sentence "Lack of MSI1 and lack of EMF2 cause similar changes in the transcriptome" but this information relates to panel (B) and should appear after not before the letter (B).
**This has been corrected.**

9. The reference Schatlowski et al is missing from the bibliography
**This has been corrected.**

Referee #3

- It is not always made clear in methods what tissues are being used for the various IP's and mass spec analysis. In some cases inflorescence tissue is stated, but in many of the experiments this is not clear.
**This information is now provided in all relevant figure legends and table headings.**

- The MS and microarray experiments are presented in highly summarised forms in results. It would be useful for the community to have the data presented more fully in supplementary information rather than just a graph (4D) or the top hits (MS data). Although the raw data has been submitted to databases, it would be useful to have fuller analyses presented in the SI.
**In addition to the submission to public databases, processed microarray data and MS data are now included in the Supplemental information (Tab. S1 and S4).**

- The interaction between MSI1 and EMF2 found by MS-IP is validated by co-IP from transgenic tobacco plants whereas that between LHP1 and MSI1 by co-IP from transgenic Arabidopsis. Although this is slightly inconsistent and it would be preferable to have done both in Arabidopsis, I think it is acceptable given that they have done reciprocal MS-IP with both tagged EMF2 and with tagged MSI1 and have four replicates for each experiment.
**We appreciate that the reviewer finds our data convincing. Our rational was that consistent results from independent experimental systems give additional weight to the conclusions.**

- I found figure 5E very hard to follow and feel that it would better benefit from better explanation in the main text and legend.
**The reviewer probably referred to Figure 4E. We apologize for original description of the figure and added additional explanation to the legend and main text. Hopefully this analysis is now clear.**

- In Figure 2A I think they should show the western the data is based as well as (or instead of) the graph summary, which gives little idea of what the blot it is based on looks like.

 

**The immunoblot is now included as Supplemental Figure 1.**

- The authors do not find a role for MSI4 in that their MS IP of EMF2 and MSI1 do not reveal MSI4 but rather MSI1. As cited by the authors, the Genshik group have published IP's showing that MSI4 and CLF co-IP from Arabidopsis. The authors here conclude MSI4 indirectly affects PRC2 activity via its role in other complexes. I think this is premature and they should leave the possibility that MSI4 is present in PRC2 complexes open. What they see may well depend on the tissue used, the extraction buffers, the tags and the component used for IP. After all, CLF which almost certainly is a PRC2 member, is also not seen in their IP's with EMF2 or MSI1, so the negative results do not necessarily mean that MSI4 is not a member of some PRC2 complexes.

**We agree with the reviewer that more research is needed to establish the precise nature of the MSI4 – PRC2 interaction and toned down this aspect in the discussion.**

2nd Editorial Decision 13 May 2013

Thank you for submitting your revised manuscript to The EMBO Journal. Your study has now been re-reviewed by referees #1 and 2. As you can see below, both referees appreciate the introduced changes and support publication here. They have just a few minor concerns that need to be resolved before acceptance.

REFEREE REPORTS

Referee #1

The authors have now addressed all of my concern and most of them satisfactory.
I have only one last concern before I can recommend this manuscript for publication.
Supplementary Figure S2 appears a bit unclear. The authors state that the "most strongly up-regulated genes" were analyzed. How were they chosen, what was the threshold? For lhp1 - clf the data looks very concincing, however, for msi - lhp1 and msi - emf2 the overlap is relatively little, possibly not significant. Here I would like to see statistical analyses whether the overlap between the different backgrounds is significantly enriched and therefore supports the authors´ hypotheses.

Referee #2

The authors have satisfactorily addressed all my original criticisms and I commend them for the substantial work in revising the manuscript. The data showing the composition of the main somatic PRC2 is now more consistent with the known phenotypes of the clf and swn mutants. This, together with the new data supporting a role for LHP1 in maintaining H3K27me3 marking of at least some Arabidopsis loci, makes a substantial contribution to our understanding of the workings of polycomb complexes in plants.
I have just a couple of minor comments that I think would improve the manuscript:
1. p6 line3 27-28; the authors state that "the PHD-domain proteins VRN5 and VEL1 appear to function mainly together with the VRN complex. " This is not entirely consistent with Table 3 which shows that VEL1 is clearly immunopreciptated along with EMF2 rather than VRN2.
2. p7 line 22; "At3g28007 remains transcriptionally silent in msi1-cs" - it might be more accurate to say that there in no increase in expression in msi1-cs.

2nd Revision - authors' response 15 May 2013

Referee #1

The authors have now addressed all of my concern and most of them satisfactory.
I have only one last concern before I can recommend this manuscript for publication.
Supplementary Figure S2 appears a bit unclear. The authors state that the "most strongly up-regulated genes" were analyzed. How were they chosen, what was the threshold? For lhp1 - clf the

 

data looks very concincing, however, for msi - lhp1 and msi - emf2 the overlap is relatively little, possibly not significant. Here I would like to see statistical analyses whether the overlap between the different backgrounds is significantly enriched and therefore supports the authors´ hypotheses.

**Reply:**
**We have indicated selection thresholds and p-values for statistical significance of overlaps in the legend of Supplemental Figure S2.**

Referee #2

The authors have satisfactorily addressed all my original criticisms and I commend them for the substantial work in revising the manuscript. The data showing the composition of the main somatic PRC2 is now more consistent with the known phenotypes of the clf and swn mutants. This, together with the new data supporting a role for LHP1 in maintaining H3K27me3 marking of at least some Arabidopsis loci, makes a substantial contribution to our understanding of the workings of polycomb complexes in plants.
I have just a couple of minor comments that I think would improve the manuscript:
1. p6 line3 27-28; the authors state that "the PHD-domain proteins VRN5 and VEL1 appear to function mainly together with the VRN complex. " This is not entirely consistent with Table 3 which shows that VEL1 is clearly immunopreciptated along with EMF2 rather than VRN2.

**Reply:**
**We have deleted this sentence .**
**2. p7 line 22; "At3g28007 remains transcriptionally silent in msi1-cs" - it might be more accurate to say that there in no increase in expression in msi1-cs.**

**Reply:**
**We have changed the text as suggested.**

 