nature portfolio

Peer Review File

Upstream regulator of genomic imprinting in rice endosperm is a small RNA-associated chromatin remodeler

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this manuscript, the authors identified OsCLSY3 as a functional ortholog of the Arabidopsis CLASSY3 gene. They found that in contrast to Arabidopsis CLSY3, OsCLSY3 is a maternally expressed imprinted gene and the imprinting status depends on the presence of MITE elements in the promoter region. To dissect the function of OsCLSY3 during endosperm development, the authors generated CRISPR-Cas9 based knock-out lines (KO), RNAi-based knockdown lines (kd) and overexpression lines (OE). They found that pollen development was drastically affected in KO lines, resulting in severe sterility; seed size was reduced in kd and seed gemination speed was delayed in OE, together suggesting that OsCLSY3 regulates reproductive development. To understand OsCLSY3 function in detail, the authors made transcriptome analysis on kd endosperm. They found a clear reduction of sRNAs, many of which belong to endosperm-specific loci. Reduction of 23-24 nt sRNAs was mapped to gene regions as well as different TE families, correlating with altered gene expression. The authors further showed that reduced 23-24 nt sRNAs were also mapped to 9 of 25 maternally imprinted sRNA loci and all 16 paternally imprinted sRNA loci. Finally, the authors found that many imprinted genes flanked by imprinted sRNAs loci had increased expression and reduced gene methylation in kd. Taken together, the authors propose that the maternally imprinted OsCLSY3 control sRNA generation in rice endosperm, which mediates methylation of genes and TEs, in turn regulating gene expression and grain size.

This manuscript provides interesting and relevant data advancing our understanding of siRNA formation and function in plants and the role of CLASY proteins in this process. In particular the finding that OsCLSY3 is imprinted and specifically expressed in the endosperm is of high relevance, given that sirenRNA production in Brassica was proposed to be determined by the maternal seed coat. Furthermore, the functional relevance of sirenRNA production for pollen and seed development is relevant and interesting. Nevertheless, to fully support the claims of the authors, some additional experiments and analyses are required. In particular the connection between sirenRNAs and DNA methylation is insufficient and should be backed up with whole genome DNA methylation analyses.

Main comments:

1. Line 374: Based on the PCA analysis one cannot conclude that a similar set of sRNAs is affected in nrpd1 and clsy3 mutants. This should be directly tested by generating Venn diagrams testing the overlap of affected sRNAs.

2. L392: The selection of loci tested for changes in DNA methylation is unclear. The authors should generate genome-wide bisulfite data and test DNA methylation at all loci targeted by CLSY3-dependent sRNAs.

3. L411: To test the ability of sirenRNAs to guide trans methylation, the authors should look at DNA methylation in CLSY3kd endosperm and specifically look at loci that can be targeted by sirenRNAs in trans.

4. L428: Based on this data no conclusion can be drawn regarding antagonistic regulation of sRNAs and gene expression. This would require a correlation plot showing reduction of sRNAs in kd lines and corresponding expression differences.

5. Fig 7D: It would be interesting to test whether reduced expression of paternally expressed sRNAs in kd lines would also be observed in the cross ko or kd lines x wt.

6. The authors show in Fig 1F,1G and S2B that OsCLSY3 was expressed in reproductive organs but not in vegetative tissues. However, in Fig 2B, OsCLSY3 expression was detected in young seedlings without AZA treatment; in Fig 2J, MITE siRNAs were weakly present in vegetative leaves, flag leaves and endosperm, suggesting that OsCLSY3 was expressed in vegetative leaves and flag leaves in addition to the endosperm. Similarly, Fig 2G shows that the OsCLSY3 promoter was active in young leaves even without AZA treatment. If OsCLSY3 is not expressed in vegetative tissues, the authors need to explain potential contradictions of data shown in Fig 1 and Fig 2.

7. clsy3-KO showed severe pollen defects, indicating its relevance for pollen development. Since OsCLSY3 is a MEG in the endosperm, the function of OsCLSY3 in the endosperm can could be investigated by crossing of clsy3-KO (mother) with wild type (father). The authors should do this cross and check seed phenotypes at different developmental stages.

8. In clsy3-kd mutants, seed length, width and weight were reduced. The author should cross clsy3-kd with wild type and check parental effects of OsCLSY3 on seed phenotypes.

9. The authors performed transcriptome analysis on endosperm of 20-25 DAP. However, only mature grain length, width, weight and chalkiness were measured in clsy3-kd and OsCLSY3-OE, which cannot clearly manifest OsCLSY3 function. The authors should check whether early endosperm development is affected in these lines with histological methods.

10. Data shown in Figures 1 and 2 are not really adding much to the main story of this manuscript. That OsCLSY3 is imprinted has been previously shown and that MITEs are associated with imprinting is also not new. I suggest to shorten this part of the manuscript and move data to the supplement.

Minor Comments:

11. Since OsCLSY3 regulates pollen development, does OsCLSY3-OE plants have normal fertility?

12. In Fig S2F, “PB-1” should be “PB1”.

13. In fig S3G, the figures are incorrectly labelled.

14. Line 49-52: In many flowering plants, the central cell is a diploid cell (the nuclei fuse before fertilization). However, in some other species, like waterlily, the central cell has only one haploid

nucleus, which results in a diploid endosperm after double fertilization. Please rephrase this part correctly.

15. Line 87: CHH methylation can also be established by CMT2.

16. poliv mutant should be replaced by nrpd1 mutant.

17. Line 395: should be hypermethylated.

18. Line77: AtNRPD1 is not an imprinted gene. It is a paternal biased gene (PMID: 35389984; PMID: 24994762).

19. Line 103-107: Although it has been shown that 24 nt sRNAs are maternally biased in young seeds of Arabidopsis, it has also been shown that the paternal genome contributes substantially to Arabidopsis endosperm sRNAs (DOI: 10.1016/j.celrep.2017.11.078=. . Please phrase more carefully.

20. Line 139: "Mutation in OsCLSY3 negatively affected endosperm formation". The endosperm can form normally, but the development of the endosperm seems affected. So "Mutation in OsCLSY3 negatively affected endosperm development" is more precise.

21. Line 144: "paternally imprinted OsCLSY3" should be "maternally imprinted OsCLSY3".

22. Line 233: "OsCLSY3 is a maternally expressed imprinted (MEG) in rice" should be "OsCLSY3 is a maternally expressed imprinted gene (MEG) in rice".

23. The authors showed that 21-22 nt sRNAs mapped to miRNA encoding loci were not significantly reduced in Fig 5G. Since previous work revealed that RdDM pathway components affect generation of different sized sRNAs form the same loci ("RNA Pol IV induces antagonistic parent-of-origin effects on Arabidopsis endosperm, PMID: 35389984" and "Polymerase IV plays a crucial role in pollen development in Capsella, PMID:31988265"), it would be relevant to check whether reduced 21-22 nt sRNAs also occur at loci losing 23-24 nt sRNAs.

24. The authors should refer to sirenRNAs and siren loci when referring to endosperm-specific sRNAs/loci.

Reviewer #2 (Remarks to the Author):

In this manuscript, one chromatin remodeling protein, OsCLSY3 was investigated in epigenetic views. They conclude that 1) OsCLSY3 is maternally expressed imprinted gene in endosperm; 2) RdDM controls the DNA methylation on the OsCLSY3 promoter and thus its transcription; 3) OsCLSY3 regulates seed development. The results might be real if the concerns below be fixed.

Major concerns:

One main result is that OsCLSY3 is one MEGs. This important conclusion is only derived from one cross of Whiteponni (WP) x PB1. What is the result of reciprocal cross?

For imprinted genes in rice, a lot of crosses have been published, for example, Nipponbare with 9311 (PLoS Genetics, 2011, 7:e1002125); Longtefu with 02428 (New Phytologist, 2017, 216:373-387), Liuqianxin, Rongfeng and Wufeng (Plant Physiology, 2018, 177:1754-1771); Nipponbare, Kitaake, 93-11 and IR64 (PNAS, 2021, 118:e2104445118). The expression pattern of OsCLSY3 in the above crosses will provide more solid conclusion.

Utilization of transgene with promoter driven GUS to examine its expression pattern is often wrong-directed. The promoter of endogenous OsCLSY3 contained specific DNA methylation and various histone modifications, how to make sure these epigenetic marks are precisely established de novo in transgenes plants? The transcriptional levels of OsCLSY3 is very low in leaves (Fig 1G), for OsCLSY3::GUS line, however, the GUS signal is quite high before AZA treatment (Fig 2G). The inconsistency of OsCLSY3 expression for endogenous gene and transgene undermine the conclusion.

To examine the effect of TE on the transcription of OsCLSY3, the authors also use transgene of OsCLSY3::GUS. The drawback also exists. Why not delete TE in endogenous OsCLSY3?

The DNA methylation on the promoter of OsCLSY3 was reduced in rdr2, pol iv, pol v nrpd/e2, and clsy4 (Figure 3). Beside in pol iv panicle, the transcriptional levels of OsCLSY3 in various tissues of nrpe1, nrpd/e2, rdr2, clsy4 could be easily examined in published data and in this study. The data are important to confirm that the DNA methylation via RdDM in its promoter control the transcripts in various tissues.

The off-target events of CRISPR/Cas9 often occur especially for the homologous genes. It is necessary to examine the sequence of OsCLSY4 and OsCLSY1 in clsy3-ko mutants. The transcriptional levels of OsCLSY4 and OsCLSY1 in clsy3-kd should be examined to make sure the effect of artificial microRNA only decrease the transcriptional levels of OsCLSY3 but not OsCLSY1 and OsCLSY4.

The phenotypic analysis was conducted only in one line of transgenic plants for OsCLSY3::GUS (Fig. 2F-G), clsy3-kd (Fig. 3E-F, H-M). The results from one line were very vulnerable. At least three independent lines are needed.

Minor suggestions:

In line 76-77, “AtFIS2, AtMEA, AtVIM5, AtNRPD” should be “FIS2, MEA, VIM5 and NRPD1”.

In line 19, “RNA directed DNA methylation” should be “RNA-directed DNA methylation”. The accumulation of 24-nt siRNAs, the methylome in other tissue of clsy3 mutant should be examined.

Reviewer #3 (Remarks to the Author):

In the current work, entitled "Upstream regulator of genomic imprinting in rice endosperm is a small RNA-associated chromatin remodeler", Pal and colleagues identify CLSY3 as an imprinted gene in rice and proceed to characterize how it is imprinted and what affects it has on siRNAs, DNA methylation, gene expression, and rice development. Their work represents several key advances as the data presented clearly demonstrate that CLSY3 is a maternally expressed gene that is regulated by RdDM-mediated methylation at two TEs located in its promoter. Furthermore, *clsy3* mutants were found to have reduced siRNA levels in endosperm and they displayed severe fertility defects demonstrating a critical role for CLSY3 in rice reproductive development. However, based on the data presented, it is difficult to assess (**1**) the role of CLSY4 in regulating CLSY3 expression (**2**) the strength of the *clsy3* mutant relative to *poliv*, (**3**) the affects of *clsy3* on DNA methylation, and (**4**) the associations between losses of siRNAs and the expression of genes related to imprinting and/or the observed developmental defects. As detailed below, clarifying these connections will provide additional support to the authors claims and raise the impact of their findings.

Major comments:

1. **The role of CLSY4 in regulating CLSY3 expression and in controlling siren siRNAs remains unclear.**
    a. The text mentions the generation of *clsy4*-kd lines with an amiR strategy, but there is no data presented to show these lines have lower CLSY4 levels. They show the amiR is expressed in Fig. 3F but not that CLSY4 transcript levels or protein are affected. Thus, it is not clear if the reduced siRNAs at mites in these lines is due to reductions in CLSY4.
    b. The text states "Using bisulfite sequencing (BS-PCR), we found a reduction of DNA methylation at the OsCLSY3 promoter in leaves of *clsy4*-kd, which indicated that OsCLSY4 controls expression of OsCLSY3 via RdDM (Fig. 3F)." However, the expression of CLSY3 in the *clsy4*-kd line was not assessed. To demonstrated a role for CLSY4 in regulating CLSY3 expression qPCR experiments should be conducted in the *clsy4*-kd lines (once they are vetted as mentioned in "a").
    c. In Fig. 6F the *clsy4*-kd lines are used again and based on siRNA blots it was suggested that CLSY4 doesn't affect siren loci. However, the quality of the blot is a bit low and it is difficult to assess from this single experiment the role of CLSY4 at siren loci. At a minimum this blot should be repeated. However, including blots at additional siren loci or conducting smRNA-seq experiments in *clsy4*-kd lines would be advised if the authors want to claim CLSY4 is not involved in siren siRNA production.

2. **Some clarifications on the role of CLSY3 in siRNA production are required assess its contributions to the endosperm small RNA landscape**.
    a. Its not clear what parameters were used to call *clsy3*-dependent siRNAs. The methods mention using DESseq2, but the fold change and FDR cutoffs were not specified. Please add this information to the methods.
    b. In Fig. 5D is hard to tell the global decrease in siRNAs from these tracks. With a 70% decrease in siRNA producing loci there should be a way to make this more clear.
    c. Fig. 5F and this text are confusing. "We observed around 70% of sRNA loci lost sRNAs in *clsy3*-kd compared to WT which were further called CLSY3-dependent sRNA loci (Fig. 5F and Supplemental dataset S6)." Fig. 5F shows the overlap of shortstack clusters showing

8,903 of the clusters called in the *clsy3*-kd were also called in the WT. However, the 70% comes from dataset S6 showing 21,653 of the 29,850 WT cluster are downregulated in *clsy3*. Please clarify the text/figure to make this point more clear.

d. Regarding the siRNA levels at siren loci, the data in Fig. 6A-F and Fig. S6 clearly show effects in both *clsy3*-kd and *poliv*-kd lines. However, several questions remain that can be addressed with additional data analysis.
    i. While in the majority of siren loci are reduced in the *clsy3*-kd lines based on the z-score heatmap (Fig. 6c) it is hard to know how robust the changes are. Please indicate what fraction of these loci were identified as/overlap with *clsy3*-depenend clusters from the earlier analysis. Depending on the numbers, it might also be useful to visualize the siren loci in subgroups rather than all together.
    ii. For the siren loci data are also shown for the *poliv*-kd, but always as separated z-score plots or IGV track with different scales than used for the *clsy3*-kd. Thus, its hard to assess how strong the *clsy3*-kd line is compared to the *poliv*-kd line. Rather than separating these data, please plot them together to allow for such comparisons.

3. The connections between reduced siRNA levels and altered gene expression are difficult to follow, making it hard to assess the impact of CLSY3 on gene regulation. This is true for siren adjacent genes, imprinted genes and other genes associated with seed/reproductive traits.

a. The main issue is that it is not clear if the comparisons are restricted to statistically significantly reduced siRNA clusters and DE genes in the *clsy3*-kd lines or rather selected genes of interest showing trends in siRNA and gene expression levels. Z-scores are a nice way to represent data with a wide spread in expression levels, as is often the case for siRNA expression levels, but with just two genotypes included this can also make subtle differences look significant. Understanding the connections between siRNAs and gene expression is further convoluted by the mixed usage of gene names and gene IDs between heatmaps and screenshots making it hard to know what to compare. Please clarify these issues throughout and include gene lists for the siren adjacent genes, seed development and yield related genes, maternally and paternally expressed siRNA loci etc. Some specific examples are highlighted below.

b. Line 427 mentions about 1000 loci with reduced siRNAs and increased expression in *clsy3*-kd mutants. What are these loci and are they *clsy3*-dependent siRNA clusters from the DESeq2 analysis? The example, in Fig 6J it doesn't look like siRNAs at the LTR Gypsy are reduced very much. If that is one of the better examples then the changes in gene expression don't appear to be well correlated with changes in siRNA levels.

c. Its not clear which if any of the loci shown in Fig. 6L are statistically significantly downregulated in the *clsy3*-kd lines. Please indicate which of these genes are part of the DEGs from Fig. 6H.

d. In addition to the heatmap and boxplots in Fig 7A and B, it would be good to note that 74 upregulated and 39 downregulated imprinted genes are changed enough in expression to be captured in the endosperm *clsy3*-kd DEGs in Fig 6H while the other are not. For B and C given there are up and downregulated genes and some are significant and others not perhaps a breakdown of these classes for the boxpots would be more informative.

e. For Fig 7D it is not clear how the 15 MEG siRNA loci and 16 PEG siRNA loci were defined or how they related to the siRNA clusters identified from the shortstack analysis. Are these MEG and PEG loci included in the 21,653 *clsy3*-dep clusters? Like for the effects at siren loci, having the *poliv*-kd data visualized together, on the same scale, with the *clsy3*-kd data instead of in the supplement would be helpful to determine the relative strength of *clsy3* versus *poliv*.

f. In lines 463-465 the text states "However, most of the upregulated maternally expressed sRNA loci in *clsy3*-kd were downregulated in *poliv*-kd, probably due to a compensatory effect likely involving other CLSYs (Supplemental Fig. S8A and B)." However, as Fig. S8A only shows data for the *poliv*-kd and in Fig. S8B siRNA levels are not clearly higher in the *clsy3*-kd at either of the maternally expressed siRNA loci shown, it is not clear why these figures are referenced. Please clarify. Are the heatmaps for Fig. S8A and Fig. 7D in the same order? If so, please mention this and also indicate which genes are shown in the screen shots.

g. Are the imprinted genes in Fig 7E part of the 74 imprinted DEG from the mRNAseq? And are some of the genes from 7E shown in 7F? If so, please use a common naming system to make this more clear, if not, why were the genes in Fig. 7F chosen?

4. **Throughout the paper the connections between losses of siRNAs in the *clsy3*-kd line and changes in DNA methylation are weak.**

a. As for the links to expression, more transparency in why specific loci were selected to show changes in methylation between the endosperm and embryo and/or for targeted for BS sequencing would be helpful. Furthermore, transparency on whether these sites are statistically significant DEGs and *clsy3*-dependent siRNA clusters is also required.

b. For TEs, the example in Fig. 5I at the 5S-rDNA repeat looks good, but for the MITE the siRNA change is not significant and its not clear what region of the BS data corresponds to the siRNA site. As siRNA levels are globally reduced at MITEs, Gypsy and retroelements (Fig. 5H) additional loci should be assessed. In addition, the change in methylation at the line element shown in Fig. S5D is also very modest.

c. For siren loci, DNA methylation is only assessed for one locus (Fig. 6G) and while reductions are clear in the *poliv*-kd, the changes in the *clsy3*-kd are modest and it would be helpful to see a quantification. This suggest CLSY3 is not the only factor controlling siren loci, and further draw into question the assessment that CLSY4 is not involved. More loci should be assessed to determine the effect of *clsy3* mutants on siren methylation levels.

d. For imprinted genes, changes in methylation between the embryo and endosperm are shown and reductions in siRNAs are shown for the *clsy3*-kd for many genes. However, change in methylation are only assessed at a few loci and only three out of six site showed significant changes.

e. Overall, these results draw into question how much the *clsy3* mutant affects methylation levels. Having a weaker phenotype at the level of methylation vs siRNAs is not unprecedented given siRNAs are often in excess and there are redundancies within the CLSY family in other organisms and well as redundancies between the RdDM and other methylation pathways. However, if the phenotype is weak this should be stated more clearly to promote further exploration into the mechanisms controlling RdDM in the

endosperm. Ideally, genome-wide BS sequencing experiments would be conducted in parallel with the *clsy3*-kd and *poliv*-kd lines to assess the contribution of CLSY3 to RdDM in the endosperm on a global scale.

Minor comments:

1. In Fig S1C one of the gene is labeled OsNF-NC11, but in the text and Fig S1B is labeled OsNF-YC11. Please correct whichever label is in error.
2. In Fig. S1D one gene is named OsRRP, but in the text only a gene ID is given "Os09g0537700". Are these the same gene? Please clarify.
3. The text states "We observed that OsCLSY3 and OsCLSY1 are majorly expressed in endosperm and embryos, respectively. However, unlike Arabidopsis, OsCLSY4 is expressed ubiquitously (Fig.1F, Supplemental Fig. S2A and B). However, in Fig. 1F CLSY1 is not "majorly" expressed in the embryo as it is higher in the young endosperm and anther, for example. Also, CLSY4 shows variation in its expression. Please clarify the text to better reflect the data presented.
4. The text states "Around 3607 transcripts showed embryo-preferred expression, while transcripts that expressed highly in endosperm were around 3686 (Supplemental dataset S1 and Supplemental dataset S2). But some of the genes in these tables have no names and are listed as a "-". What does that mean? Are these unannotated genes? Also, since the text (lines 171-174, lines 181-182, etc.) highlights some embryo and endosperm preferred genes by their names rather than gene IDs, perhaps adding an extra column to datasets S1 and S2 with gene names would be helpful.
5. In Fig. 1B there are 3,475 EN-preferred genes, but table S1 has 3686 genes. Please clarify why the numbers don't match.
6. For dataset 3, 175 epigenetic genes are listed, but the text and Fig. 1B only include 160. Please clarify. Also please add a column with the gene names in addition to the gene IDs.
7. For dataset 3, it would be helpful to generate a list of unique gene IDs and then also annotated this list with the gene names.
8. The text, lines 196-197 reference CHR740, but this is not labeled in Fig. 1D, rather several rice genes are labeled as OsDRD1-like. Please use a common nomenclature in the text and figures.
9. Supplemental Table S4 is listed out of order, before Supplemental Tables S2 and S3. Please reorder.
10. The text states "We observed that for OsCLSY3, 86.1% of transcripts came from the maternal genome when compared to OsSHH1 (BiG) which showed 65.3%. As expected in the case for OsARF22 (a PEG), 86.4% of transcripts were from the paternal genome (Fig.1I, Supplemental Fig. S2F). These analyses demonstrated that OsCLSY3 is a maternally expressed imprinted (MEG) in rice." The data matches for CLSY3 and SHH1, but for ARF22 the sequence shown in Fig. 1I is TC(A/G)GT but in Fig. S2F the stacked barplots are for G(A/G)T. Please clarify, shouldn't it be for C(A/G)T?
11. For the traditional BS in Fig. 2D, please reference Table S1 so it is clear the data is from targeted high-throughput sequencing. Also it is not clear if the same promoter region is being shown in Fig. 2E. For comparison, showing the same region would be helpful. Furthermore, it appears that there is also less methylation at the CLSY4 promoter in the endosperm compared to the embryo despite the text implying that the change in methylation is specific to the promoter of CLSY3. Please clarify the aforementioned points.

12. In Fig S3A, the probe region used for the DNA blots is not clear and the expected size for the intact T-DNA is not indicated making his blot hard to interpret. Please clarify. Also, is there DNA blot data for KO#5?
13. In Fig S4 the transgenic plants labeled as OE and KO are not clear. Are these both for CLSY3? Also, please add information for the *clsy4*-kd.
14. In Fig S4C, the probe region used for the DNA blots is not clear and the expected size for the intact T-DNA is not indicated making his blot hard to interpret. Please clarify.
15. The text states "We obtained a total of 7 transgenic plants with double amiRs (clsy3-kd2) (Fig. 4A, Supplemental Fig. S4D) and 8 plants with single amiR (clsy3-kd1) (Fig. 4A)." But figure S4C shows the opposite, 7 lines for kd1 and 8 for kd2. Also, it is not clear how the numbers correlate with each other between the main and supplemental figure. Please clarify.
16. The text states that "Principal component Analysis (PCA) indicated that the identical pool of 23-24 nt endosperm specific sRNAs were downregulated in *clsy3*-kd and *poliv*-kd endosperm tissues (Fig. 5C)." This cannot be determined from a PCA and should instead be determined by an overlap analysis of reduced siRNA clusters.
17. Please define Class I and Class II TEs with regards to Fig 5 at their first mention in the text.
18. Line 382 references Shortstack analysis and datasets S3 and S4, but I think it should be S4 and S5. Please clarify.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this manuscript, the authors identified OsCLSY3 as a functional ortholog of the Arabidopsis CLASSY3 gene. They found that in contrast to Arabidopsis CLSY3, OsCLSY3 is a maternally expressed imprinted gene and the imprinting status depends on the presence of MITE elements in the promoter region. To dissect the function of OsCLSY3 during endosperm development, the authors generated CRISPR-Cas9 based knock-out lines (KO), RNAi based knockdown lines (kd) and overexpression lines (OE). They found that pollen development was drastically affected in KO lines, resulting in severe sterility; seed size was reduced in kd and seed gemination speed was delayed in OE, together suggesting that OsCLSY3 regulates reproductive development. To understand OsCLSY3 function in detail, the authors made transcriptome analysis on kd endosperm. They found a clear reduction of sRNAs, many of which belong to endosperm-specific loci. Reduction of 23-24nt sRNAs was mapped to gene regions as well as different TE families, correlating with altered gene expression. The authors further showed that reduced 23-24 nt sRNAs were also mapped to 9 of 25 maternally imprinted sRNA loci and all 16 paternally imprinted sRNA loci. Finally, the authors foundthat many imprinted genes flanked by imprinted sRNAs loci had increased expression and reduced gene methylation in kd. Taken together, the authors propose that the maternally imprinted OsCLSY3 control sRNA generation in rice endosperm, which mediates methylation of genes and TEs, in turn regulating gene expression and grain size. This manuscript provides interesting and relevant data advancing our understanding of siRNA formation and function in plants and the role of CLASY proteins in this process. In particular the finding that OsCLSY3 is imprinted and specifically expressed in the endosperm is of high relevance, given that sirenRNA production in Brassica was proposed to be determined by the maternal seed coat. Furthermore, the functional relevance of sirenRNA production for pollen and seed development is relevant and interesting. Nevertheless, to fully support the claims of the authors, some additional experiments and analyses are required. In particular the connection between sirenRNAs and DNA methylation is insufficient and should be backed up with whole genome DNA methylation analyses.

Author's reply: Thank you for your encouraging comments and elaborate summary of the presented work. We agree with your observations and suggestions and have strived hard to address each and every question. We have added CLSY3 ChIP-seq data as well as performed whole genome DNA methylation analysis in the revised manuscript and these have clearly strengthened the manuscript.

Major comments:

1. **Reviewer's comment -** Line 374: Based on the PCA analysis one cannot conclude that a similar set of sRNAs is affected in nrpd1 and clsy3 mutants. This should be directly tested by generating Venn diagrams testing the overlap of affected sRNAs.

   Author's reply: Thank you for this suggestion. We agree that PCA analysis is not sufficient to pinpoint sRNA loci that are affected in nrpd1 and clsy3 mutants. To understand this, an ideal way is to compare those loci by Venn diagram, such as the ones reported previously (Zhou, Palanca, and Law 2018; Zhou et al. 2022).
   However, in our study, comparing between these two studies was difficult due to the different growth stages when the tissues were collected. In poliv-kd, the endosperm was collected a week earlier (15 days post anthesis as in GSE180456), when compared to clsy3-kd datasets, where the endosperm was collected at 20 days after anthesis. In our previous study and this study, only 17896 sRNA loci were overlapping between two WT datasets due to differences in sampling (please see the attached figure A panel). Due to this reason, in all the analysis presented in the revised manuscript, we represented poliv-kd data and clsy3-kd data separately with their respective WT controls.
   Nevertheless, we counted sRNAs from CLSY3-dependent sRNA loci (21652) in poliv-kd and clsy3-kd endosperm with their respective WT controls. The data clearly indicates that CLSY3-dependent sRNA loci were also reduced in poliv-kd (please see the attached figure below Fig B panel). This data was also included in the revised manuscript (Fig. 5C).

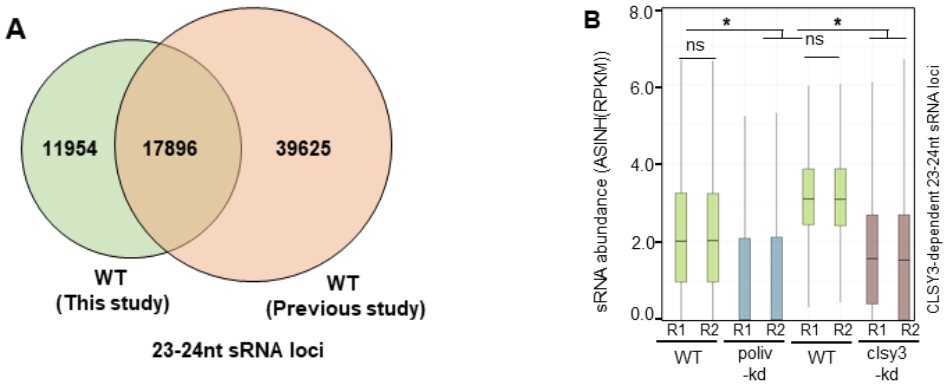


Fig.

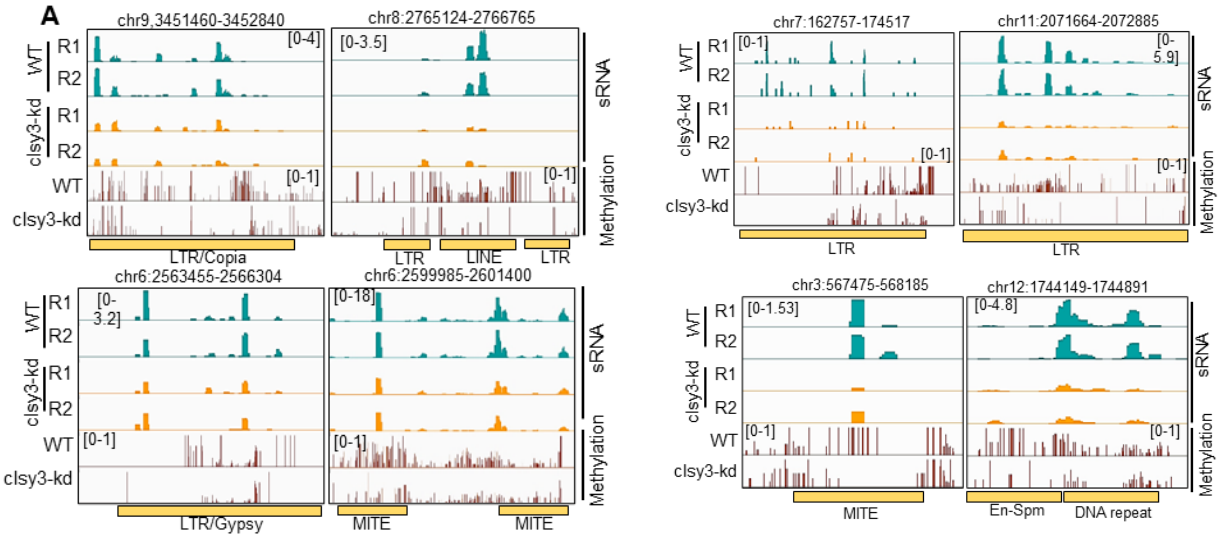
(A) Venn diagram showing overlap between sRNA loci between WT datasets in two studies. (B) Box plot showing presence of CLSY3 dependent sRNA loci in poliv-kd endosperm-derived datasets.

2. **Reviewer's comment -** L392: The selection of loci tested for changes in DNA methylation is unclear. The authors should generate genome-wide bisulfite

data and test DNA methylation at all loci targeted by CLSY3-dependent sRNAs.

Author's reply: Thank you for this suggestion. We have now performed the genome-wide DNA methylation analysis to check the effect of CLSY3 in endosperm-specific DNA methylation. As observed previously, level of DNA methylation in rice endosperm is at much lower level than other tissues such as panicle and leaf.

As expected, we found many of the CLSY3-dependent sRNA loci, TEs and repeat regions showed reduced DNA methylation in clsy3-kd (please see the A panel below). However, we also observed that in several CLSY3-dependent sRNA loci showed redistribution and sometimes gain in DNA methylation (please see B panel). The gain in DNA methylation in CLSY single mutants is a well-known observation (Yang et al. 2018), something that is also mentioned by Reviewer 3. Interestingly, DNA hypermethylation was observed only in CHH context in CLSY3-dependent sRNA loci in clsy3-kd lines (please see C and D panels).
Detailed data on this is presented in the revised manuscript (Fig 5I-5J and Fig S8A -E).



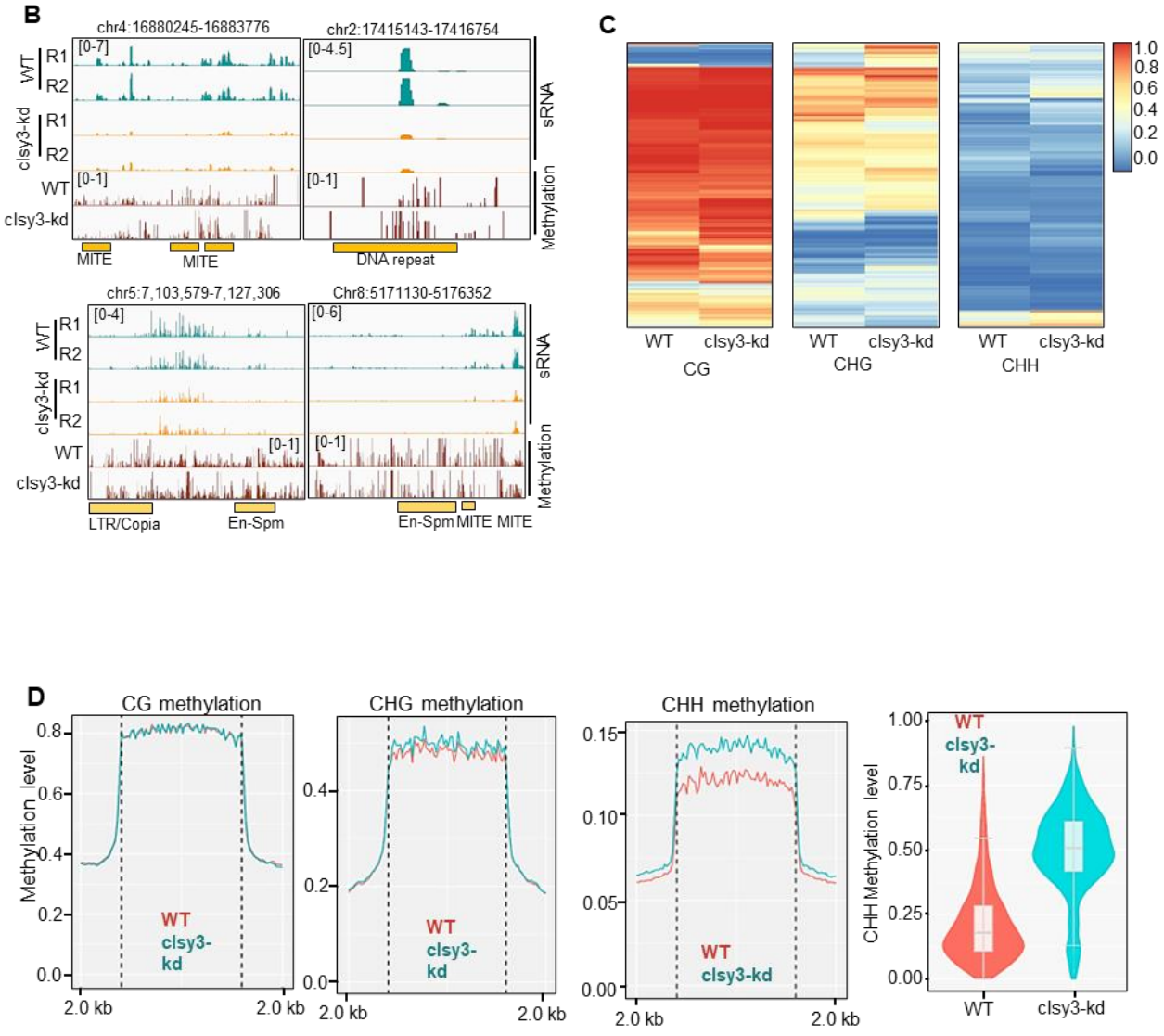


Fig.

(A) IGV screenshots showing sRNA and CHH methylation status of hypomethylated loci in WT and clsy3-kd EN.
(B) IGV screenshots showing sRNA and CHH methylation status of re-distributed/hypermethylated loci in WT and clsy3-kd EN.
(C) Heatmap showing DNA methylation status in WT and clsy3-kd EN in CG, CHG and CHH contexts.
(D) Metaplots showing DNA methylation in CG, CHG and CHH contexts, in WT and clsy3-kd EN and violin plot showing methylation in CHH context.

3. **<u>Reviewer's comment -</u>** L411: To test the ability of sirenRNAs to guide trans methylation, the authors should look at DNA methylation in CLSY3kd endosperm and specifically look at loci that can be targeted by sirenRNAs *in trans*.
Author's reply: Thanks for this suggestion. Recent papers have demonstrated that sirenRNAs are capable of guiding trans-methylation. We

have analysed the CLSY3 dependent sRNAs in a similar way as shown previously (Burgess et al. 2022). The 23-24 nt Shortstack-aligned reads at siren loci were extracted from the bam files using BEDtools intersect. The reads were converted into fasta by samtools and realigned to rice genome (IRGSP genome) with and without masking TEs and siren loci using bowtie with 1, 2 or 3 mismatches (-V 1, -V 2, -V 3). The reads which intersected with different rice genes were found by bedtools multicov. This clearly suggested trans-methylation possibilities. The numbers and details of the analysis are listed below (Panels A to C). Due to hypomethylation state of endosperm, transmethylation capacity by siren sRNAs still need further extensive checking. In this study, we did not proceed it further, as this is a deviation from the main findings.

A

Number of mismatches	After cis-loci masking
1	743
2	2383
3	13853

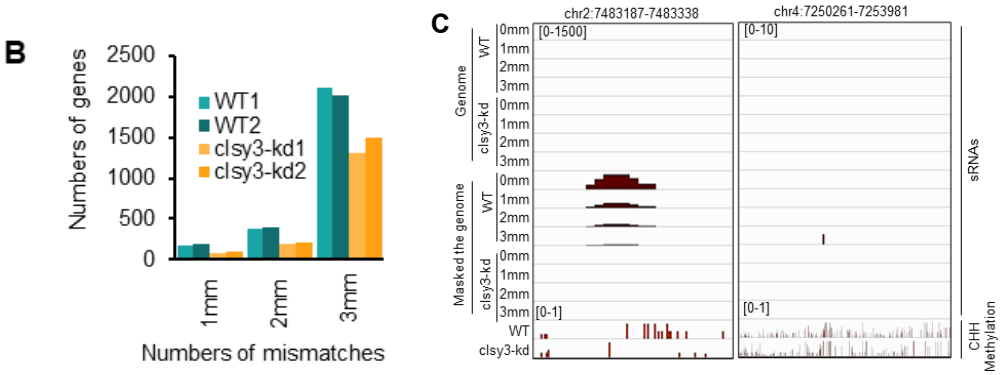


Fig.

(A) Table showing alignment of siren RNA reads after masking the cis-regulatory regions (Right column).

(B) Bar plot showing siren sRNAs aligned with number of genes in WT and clsy3-kd EN after masking the cis-regulatory regions.

(C) IGV screen shots showing deficient trans-methylation at siren sRNA aligned regions in clsy3-kd in different genomic locations.

4. **<u>Reviewer's comment -</u>** L428: Based on this data no conclusion can be drawn regarding antagonistic regulation of sRNAs and gene expression. This would require a correlation plot showing reduction of sRNAs in kd lines and corresponding expression differences.

Author's reply: Thank you for raising this issue. To understand the antagonistic regulation of sRNAs and gene expression, we considered 2kb promoter, 2kb terminator of all annotated genes and overlapped them with CLSY3 dependent sRNA loci. The sRNA and RNA counts were calculated from the overlapped sRNA loci and genes, respectively. Those were plotted as a correlation plot (please see the attached figure below). We find far more genes being upregulated upon kd (when compared with WT, second quadrant) that are associated with the downregulated sRNAs. Using this analysis, we found significant antagonistic regulation between sRNAs and gene expression, confirming the roles of CLSY3 in regulating gene expression through sRNAs. This data is included in the revised manuscript (Fig. 7I). The normalized values of CLSY3-dependent sRNAs and mRNAs are included in the revised script (Supplemental dataset S11).

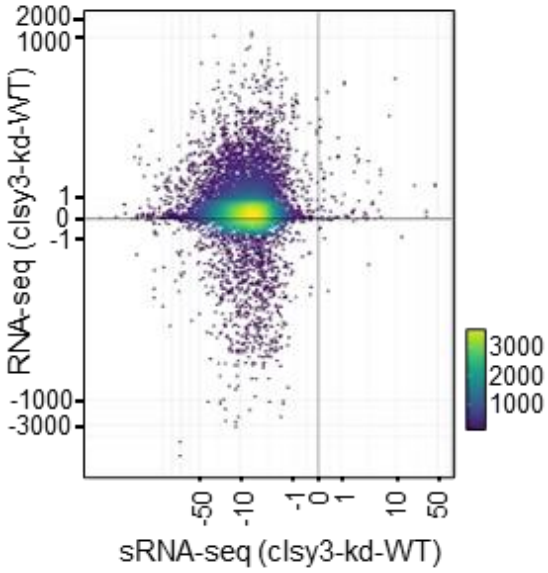


Fig. Correlation plot showing between sRNAs and RNA-seq reads in clsy3-kd EN.

5. <u>Reviewer's comment -</u> It would be interesting to test whether reduced expression of paternally expressed sRNAs in kd lines would also be observed in the cross ko or kd lines x wt.

Author's reply: Thank you for this suggestion. However, due to the following reasons, we are unable to perform those crosses in this study:

1. In Arabidopsis, mutant lines used for such studies are stable (T-DNA mutants) but in rice, we used CRISPR-cas9 based method to generate KO. The drawback is that, if lines are not T-DNA free, they should not be used for crossing experiments. If T-DNA with CRISPR-cas9 and guide RNA persists in the plant genome, it will target the introduced WT allele after the cross. In rice, one single copy insertion line segregated into 3:1 ratio offsprings (Transgenic: Non-transgenic). Ideally, we should be using these non-transgenic lines for

crossing. Unfortunately, the KO plants were sterile and could not be propagated. We got very few seeds (abnormal shaped with poor germination) in KO#1 but the plant was a 3 copy T-DNA insertion line. Hence, the seed numbers were not sufficient for removing 3 copy T-DNA to perform crossing.

2. Similar difficulties also exist with knockdown lines. In case of clsy3-kd, artificial miRNA will target newly introduce paternal and maternal allele equally. Due to that knockdown lines are inappropriate for crossing.
   We hope the Reviewer agrees with this challenge.

6. <u>Reviewer's comment -</u> The authors show in Fig 1F,1G and S2B that OsCLSY3 was expressed in reproductive organs but not in vegetative tissues. However, in Fig 2B, OsCLSY3 expression was detected in young seedlings without AZA treatment; in Fig 2J, MITE siRNAs were weakly present in vegetative leaves, flag leaves and endosperm, suggesting that OsCLSY3 was expressed in vegetative leaves and flag leaves in addition to the endosperm. Similarly, Fig 2G shows that the OsCLSY3 promoter was active in young leaves even without AZA treatment. If OsCLSY3 is not expressed in vegetative tissues, the authors need to explain potential contradictions of data shown in Fig 1 and Fig 2.

Author's reply: Thank you for pointing this concern. The Fig 1F,1G and S2B that indicated very high expression of *OsCLSY3* in reproductive stages were verified from several independent studies and the genome-wide datasets used for the studies. In case of Fig 2B, though *CLSY3* is showing limited expression, the expression level is extremely low, and the increase in expression seen after AZA is at much lower levels when compared to reproductive tissues. Hardly any organ-specific or tissue-specific gene shows zero expression in another tissue/organ. 'Negligible expression' would be the right term and we have used this term in the revised script.

In Fig 2J, we agree with the observation made by the Reviewer. However, the sRNAs in vegetative and flag leaf are higher levels than endosperm. In vegetative and flag leaf, the MITE region methylation was very high when compared to endosperm. The expression of CLSY3 in endosperm was much higher than in leaf tissues. Since RdDM establishes the DNA methylation and other epigenetic marks during early growth stages, low levels of sRNAs might be already sufficient to maintain these marks.

In Fig 2G, we found silencing of OsCLSY3 promoter which was introduced as a transgene was not as much silenced as endogenous OsCLSY3 promoter. Due to this reason, in all GUS assays, the GUS expression was observed but in same tissue the native *OsCLSY3* expression was extremely low.

To investigate these clearly, we have checked the *CLSY3, GUS and FIE1* expression by RT-qPCR analysis in young leaf tissues in P:CLSY3-GUS transgenic lines. The analysis showed that, similar to *FIE1, OsCLSY3* expression also low, but *GUS* expression was 3.5-10-fold higher than *FIE1* and CLSY3 in multiple independent transgenic lines (please see the attached figure below). Also, in mature leaves GUS expression is much lower than seedling leaves (please see the attached figure below).

All these results collectively and conclusively suggest that the transgenic CLSY3 promoter showed higher transcriptional activity than endogenous CLSY3 promoter. Clearly, CLSY3 has negligible expression in vegetative tissues.

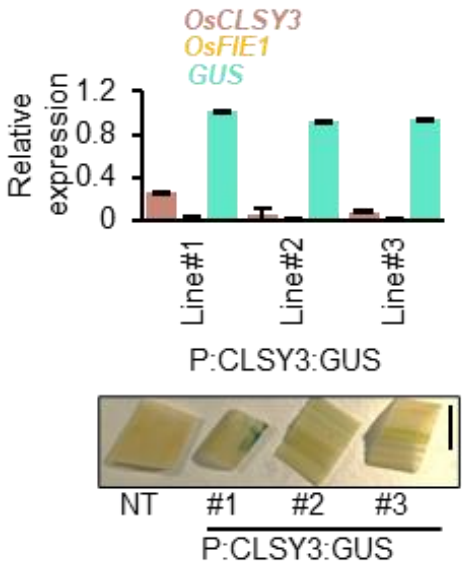


Fig. RT-qPCR analysis of *OsCLSY3*, *OsFIE1* and *GUS* in leaf. *OsActin* served as internal control. Error bar-Standard Error (SE).

7. Reviewer's comment - clsy3-KO showed severe pollen defects, indicating its relevance for pollen development. Since OsCLSY3 is a MEG in the endosperm, the function of OsCLSY3 in the endosperm can could be investigated by crossing of clsy3-KO (mother) with wild type (father). The authors should do this cross and check seed phenotypes at different developmental stages.

Author's reply: This is an excellent suggestion. As discussed earlier, due to the nature of the transgenes generated in this study, and sterility issues, this experiment could not be performed. Please see our related reply above.

8. Reviewer's comment - In clsy3-kd mutants, seed length, width and weight were reduced. The author should cross clsy3-kd with wild type and check parental effects of OsCLSY3 on seed phenotypes.

Author's reply: Again, this is a good suggestion and can be a new direction. The problem to perform this crossing was already discussed in the 5th point above.

9. Reviewer's comment - The authors performed transcriptome analysis on endosperm of 20-25 DAP. However, only mature grain length, width, weight and chalkiness were measured in clsy3-kd and OsCLSY3-OE, which cannot clearly manifest OsCLSY3 function. The authors should check whether early endosperm development is affected in these lines with histological methods.

Author's reply: Thank you for this suggestion. To check the early endosperm development related phenotypes, we performed histochemistry and electron microscopy in the 10-15 DAP endosperm tissues (Panel A). In this analysis, we observed severe defects in *KO (*clsy3) derived endosperm (please see the attached B panel). On the other hand, we also observed that clsy3-kd and OE derived endosperm tissues have defects in cellularization timing. These abnormalities were also confirmed using electron microscopy images (Panel C and D). The data added into revised manuscript (supplemental Fig. S6A-D).

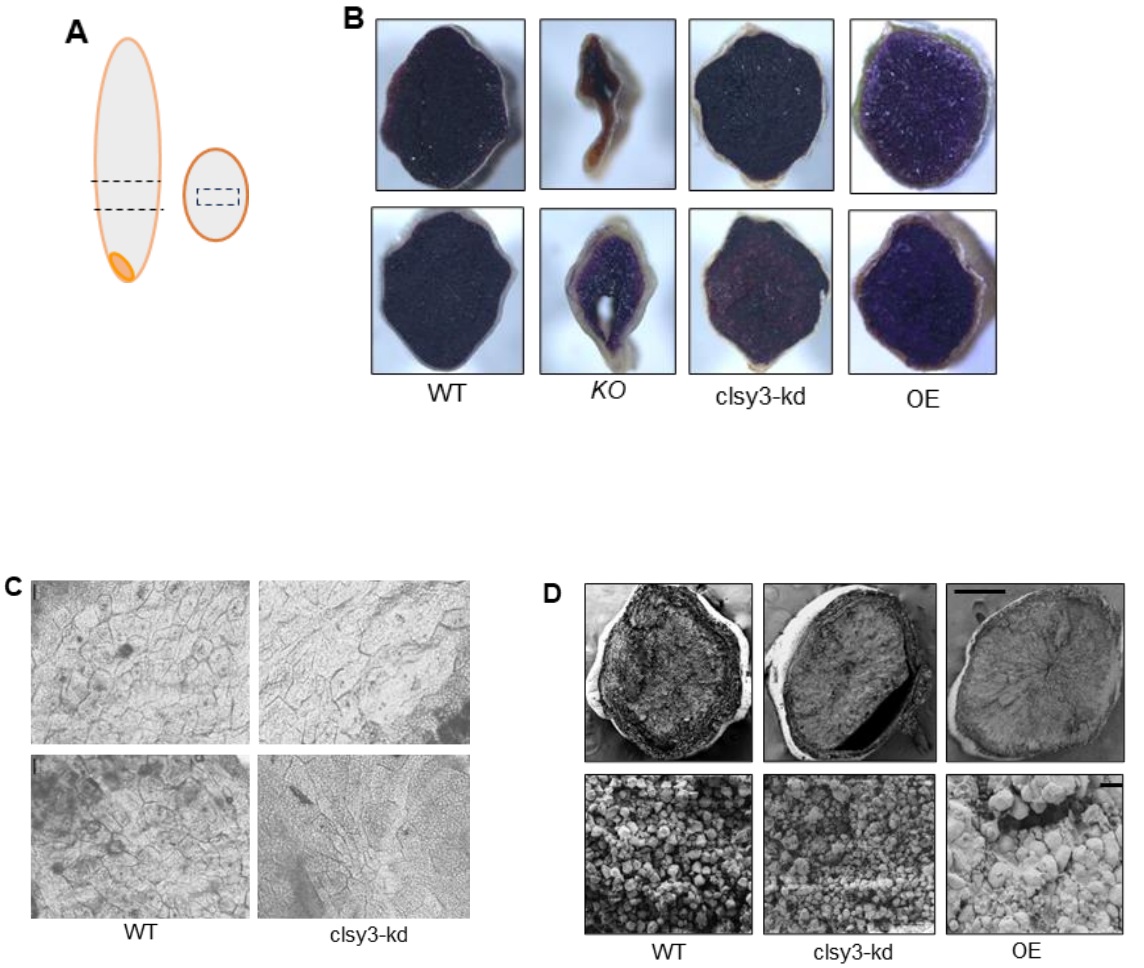


Fig.

(A) Schematic showing sections used for imaging in endosperm. (B) Images showing manually dissected endosperms (15 DAP). (C) Images showing 60 µm cross-section of endosperm (10 DAP). (D) Morphology of endosperm cross section (15 DAP) under SEM microscope (SB-200 µm and 10 µm).

10. Reviewer's comment - Data shown in Figures 1 and 2 are not really adding much to the main story of this manuscript. That OsCLSY3 is imprinted has been previously shown and that MITEs are associated with imprinting is also not new. I

suggest to shorten this part of the manuscript and move data to the supplement.
Author's reply: Thank you for this suggestion. Here, we tried to emphasis on how a tissue-specific regulator of RdDM is regulated in a tissue-specific manner by RdDM pathway itself. Previously, published reports indicated that DNA methylation and demethylation pathway have a cross talk via a "methylstat" mechanism as seen in ROS1 promoter (Williams et al. 2015). There are many direct and indirect evidences that indicated interconnected and interdependent regulatory mechanisms between epigenetic pathways (Martins and Law 2023). However, very few regulators have been studied in detail. In rice, finding such a regulatory mechanism is quite new and interesting. Hence, we thought this aspect makes the presented work more complete.

Minor Comments:

11. <u>Reviewer's comment -</u> Since OsCLSY3 regulates pollen development, does OsCLSY3-OE plants have normal fertility?

Author's reply: Thank you for asking this question. The fertility in CLSY3 OE plants was also not normal. We counted numbers of filled grains per panicle in OE panicles and found that filled grains per panicle were also significantly less in OE plants as shown in revised manuscript (Fig 4L). In OE lines (T0), multiple lines were completely sterile (panel A-B). Also, we observed partial pollen defects in OE plants (Panels C-D).

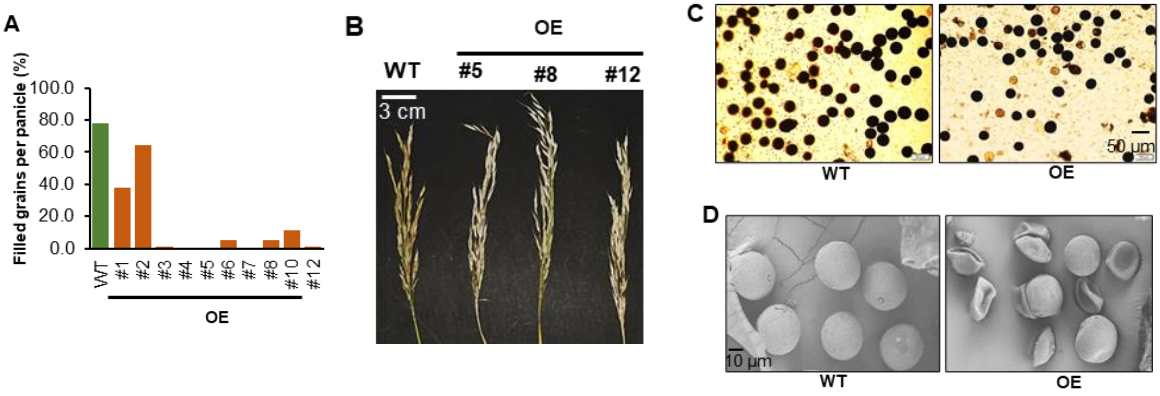


Fig.

(A) Bar plots showing grain filling in OE plants. (B) Image showing panicles of the OE plants. (C) Pollen viability assay of OE (T0) used for the assay, Scale bar (SB)-50µm. (D) Pollen morphology of OE under the electron microscope (SB)-10µm.

12. <u>Reviewer's comment -</u> In Fig S2F, "PB-1" should be "PB1".

Author's reply: Thank you. We have corrected the mistake.

13. Reviewer's comment - In fig S3G, the figures are incorrectly labelled.

Author's reply: Apologies for this mistake. We have corrected this in the revised manuscript.

14. Reviewer's comment - Line 49-52: In many flowering plants, the central cell is a diploid cell (the nuclei fuse before fertilization). However, in some other species, like waterlily, the central cell has only one haploid nucleus, which results in a diploid endosperm after double fertilization. Please rephrase this part correctly.

Author's reply: Thank you for this suggestion. We corrected the text accordingly.

15. Reviewer's comment - Line 87: CHH methylation can also be established by CMT2.

Author's reply: Thank you. We have included this detail.

16. Reviewer's comment - poliv mutant should be replaced by nrpd1 mutant.

Author's reply: We used the term nrpd1 mutant in the revised manuscript.

17. Reviewer's comment - Line 395: should be hypermethylated.

Author's reply: Thank you. We have corrected this mistake.

18. Reviewer's comment - Line77: AtNRPD1 is not an imprinted gene. It is a paternal biased gene (PMID: 35389984; PMID: 24994762).

Author's reply: Thank you for this comment. We changed the text as per this suggestion, although there are conflicting reports.

19. Reviewer's comment - Line 103-107: Although it has been shown that 24 nt sRNAs are maternally biased in young seeds of Arabidopsis, it has also been shown that the paternal genome contributes substantially to Arabidopsis endosperm sRNAs (DOI: 10.1016/j.celrep.2017.11.078=. . Please phrase more carefully.

Author's reply: Thank you for this suggestion. We rephrased the revised text.

20. Reviewer's comment - Line 139: "Mutation in OsCLSY3 negatively affected endosperm formation". The endosperm can form normally, but the development of the endosperm seems affected. So "Mutation in OsCLSY3 negatively affected endosperm development" is more precise.

Author's reply: Thank you for this suggestion. We have modified the sentence.

21. Reviewer's comment - Line 144: "paternally imprinted OsCLSY3" should be "maternally imprinted OsCLSY3".

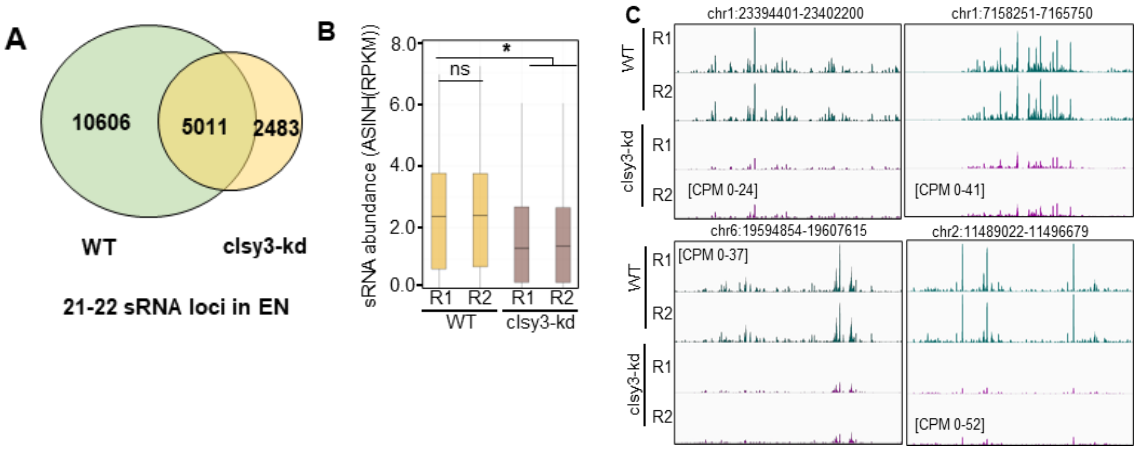
Author's reply: Thank you. We apologise for this error. We have corrected the sentence.

22. Reviewer's comment - Line 233: "OsCLSY3 is a maternally expressed imprinted (MEG) in rice" should be "OsCLSY3 is a maternally expressed imprinted gene (MEG) in rice".

Author's reply: Thank you for spotting this. We have corrected the sentence in the revised manuscript.

23. Reviewer's comment - The authors showed that 21-22 nt sRNAs mapped to miRNA encoding loci were not significantly reduced in Fig 5G. Since previous work revealed that RdDM pathway components affect generation of different sized sRNAs form the same loci ("RNA Pol IV induces antagonistic parent-of-origin effects on Arabidopsis endosperm, PMID: 35389984" and "Polymerase IV plays a crucial role in pollen development in Capsella, PMID:31988265"), it would be relevant to check whether reduced 21-22 nt sRNAs also occur at loci losing 23-24 nt sRNAs. 24. The authors should refer to sirenRNAs and siren loci when referring to endosperm-specific sRNAs/loci.

Author's reply: Thank you for the suggestion. We have found that 21-22 nt sRNAs were also reduced in the clsy3-kd endosperm (please see the attached figure panel A). The 21-22 nt sRNAs were specifically reduced in the siren loci (panels B and C). We observed that all categories of sRNAs were reduced and corresponding transcripts were upregulation in kd lines. The sRNAs of all size classes in control loci were unchanged (please see panel D below). As expected, 21-22 nt sRNAs were also reduced in specific selected loci where we also found reduction in 23-24nt sRNAs (please see panel E). These are included in the revised manuscript (Supplemental Fig. S7E)



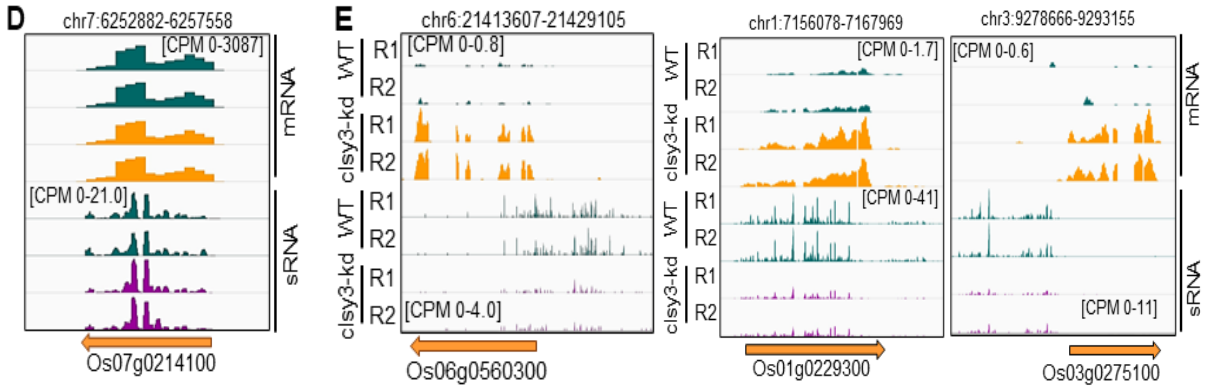


Fig.

Venn diagram showing 21-22 nt sRNA loci across WT and clsy3-kd EN. (B) Boxplot showing status of 21-22 nt sRNAs in siren loci. (C) IGV screenshots showing expression of 21-22 nt sRNAs in the selected siren loci. (D) IGV screen shot showing 21-22 nt sRNA expression in control locus. (E) IGV screenshots showing expression of three selected DEGs and levels of adjacent 21-22 nt sRNA loci in clsy3-kd EN.

We thank the Reviewer for all the comments and suggestions. We have performed all possible experiments including CLSY3 ChIP-seq to identify how this regulator can act on specific sequences. ChIP-seq data also showed how rice CLSY3 has different binding preferences than *Arabidopsis* CLSYs (7115 sequence-specific binding sites, sequence motifs, for example, Fig.6 and Supplemental Fig. S10 in the revised manuscript). All these, along with whole genome DNA methylation analysis has greatly improved the manuscript. Many thanks to the Reviewer for suggesting these modifications.

Reviewer #2 (Remarks to the Author):

In this manuscript, one chromatin remodeling protein, OsCLSY3 was investigated in epigenetic views. They conclude that 1) OsCLSY3 is maternally expressed imprinted gene in endosperm; 2) RdDM controls the DNA methylation on the OsCLSY3 promoter and thus its transcription; 3) OsCLSY3 regulates seed development. The results might be real if the concerns below be fixed.

Author's reply: Thank you for your encouraging comments and for summarizing the key points presented in this work. We agree with your observations and suggestions and addressed each and every comment.

Major comments:

1. **<u>Reviewer's comment:</u>**

One main result is that OsCLSY3 is one MEGs. This important conclusion is only derived from one cross of Whiteponni (WP) x PB1. What is the result of reciprocal cross? For imprinted genes in rice, a lot of crosses have been published, for example, Nipponbare with 9311 (PLoS Genetics, 2011, 7:e1002125); Longtefu with 02428 (New Phytologist, 2017, 216:373-387), Liuqianxin, Rongfeng and Wufeng (Plant Physiology, 2018, 177:1754-1771); Nipponbare, Kitaake, 93-11 and IR64 (PNAS, 2021, 118:e2104445118). The expression pattern of OsCLSY3 in the above crosses will provide more solid conclusion.

Author's reply: Thank you for this suggestion. We checked all the imprinted gene lists presented in those papers. In PLoS Genetics 2011 paper, *OsCLSY3* was listed as a MEG (Nipponbare with 9311). In Plant Physiology 2018 paper also it is described as a MEG in Rongfeng and Liuqianxin and Wufeng-A and Yu6-A genotypes, in both reciprocal crosses. However, in New Phytologist 2017 paper (Longtefu with 02428) and in PNAS 2021 paper, they have not listed OsCLSY3 as a MEG. To capture imprinting status of a gene, we depend on SNPs in parent varieties. Also, for those imprinted genes, there will be a loss in the biased expression pattern during development. It is also possible that some imprinted genes are specific to certain parental lines. In our experiments (crosses and detailed sequencing of the progeny in the right stage of endosperm development), we could clearly deduce that OsCLSY3 is indeed an imprinted MEG.

2. **Reviewer's comment:**

Utilization of transgene with promoter driven GUS to examine its expression pattern is often wrong-directed. The promoter of endogenous OsCLSY3 contained specific DNA methylation and various histone modifications, how to make sure these epigenetic marks are precisely established de novo in transgenes plants? The transcriptional levels of OsCLSY3 is very low in leaves (Fig1G), for OsCLSY3::GUS line, however, the GUS signal is quite high before AZA treatment (Fig2G). The inconsistency of OsCLSY3 expression for endogenous gene and transgene undermines the conclusion. To examine the effect of TE on the transcription of OsCLSY3, the authors also use transgene of OsCLSY3::GUS. The drawback also exists. Why not delete TE in endogenous OsCLSY3?

Author's reply: Thank you for this concern. It is true that expression of GUS gene driven by CLSY3 promoter transcribed more than the endogenous *OsCLSY3* gene. As reviewer pointed out, other epigenetic marks like histone modifications might play roles in silencing the *OsCLSY3* in the vegetative stages. In the transgene promoter, all the epigenetic marks not established as high as the endogenous copy and this well-established phenomenon does not alter the conclusions made in the paper.

RT-qPCR in in the young leaves of GUS transgenic plants showed GUS expression is 3.5-10.0 folds higher than the OsCLSY3 in transgenic lines (please see the attached figure).

The deletion of MITEs from the *OsCLSY3* promoter by CRISPR-cas9 technology is a good suggestion. However, in our experiments, MITEs are the most abundant TEs in rice genome and we could not delete MITEs. We also considered presence of important cis-regulatory elements vicinity of those MITEs TEs. Altering a single nucleotide, addition or deletion is sufficient to KO a gene, but in case of deletion of repetitive regions is not established even in Arabidopsis promoters. Due to all those points we used reporter-based assays. Though this is a well-established scheme, it has its own deficiencies as discussed above.
However, we did generate a transgenic line in which GUS gene is driven by CLSY3 promoter having a single MITE (Tourist MITE). The plant showed more GUS expression when compared to WT CLSY3 promoter with two MITEs. This data included is in revised manuscript (Supplemental Figure S2G-H).

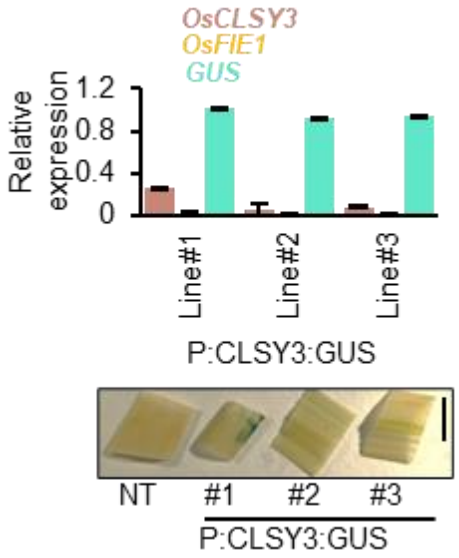


Fig.
RT-qPCR analysis of *OsCLSY3*, *OsFIE1* and *GUS* in leaf. *OsActin* served as internal control. Error bar-Standard Error (SE).

3. **Reviewer's comment:**
The NA methylation on the promoter of OsCLSY3 was reduced in rdr2, pol iv, pol v nrpd/e2, and clsy4 (Fig 3). Beside in pol iv panicle, the transcriptional levels of OsCLSY3 in various tissues of nrpe1, nrpd/e2, rdr2, clsy4 could be easily examined in published data and in this study. The data are important to confirm that the DNA methylation via RdDM in its promoter control the transcripts in various tissues.

Author's reply: Thank you for the suggestion. To check if the reduction of sRNA and DNA methylation had any impact on *OsCLSY3* expression, we analysed the RNA-seq data presented in previously published papers (nrpd1, rdr2, drm2 and poliv-kd (Wang et al. 2022; Hu et al. 2022; Hari Sundar G et al. 2023; L. Xu et al. 2020). We found that *nrpd1* shoot base*, rdr2* seedling *OsCLSY3* level is unchanged but in the poliv-kd panicle and in *drm2* mature leaf tissues, the expression of *OsCLSY3* was more (please see the attached A panel). We also checked expression of *OsCLSY4* as control and found it has slight changes (please see the attached panel B). We confirmed higher level of *OsCLSY3* expression in poliv-kd anther and in 5 days old clsy4-kd seedlings (please see C panel). The data has been incorporated in the revised manuscript (Supplemental Fig. S3A-E).

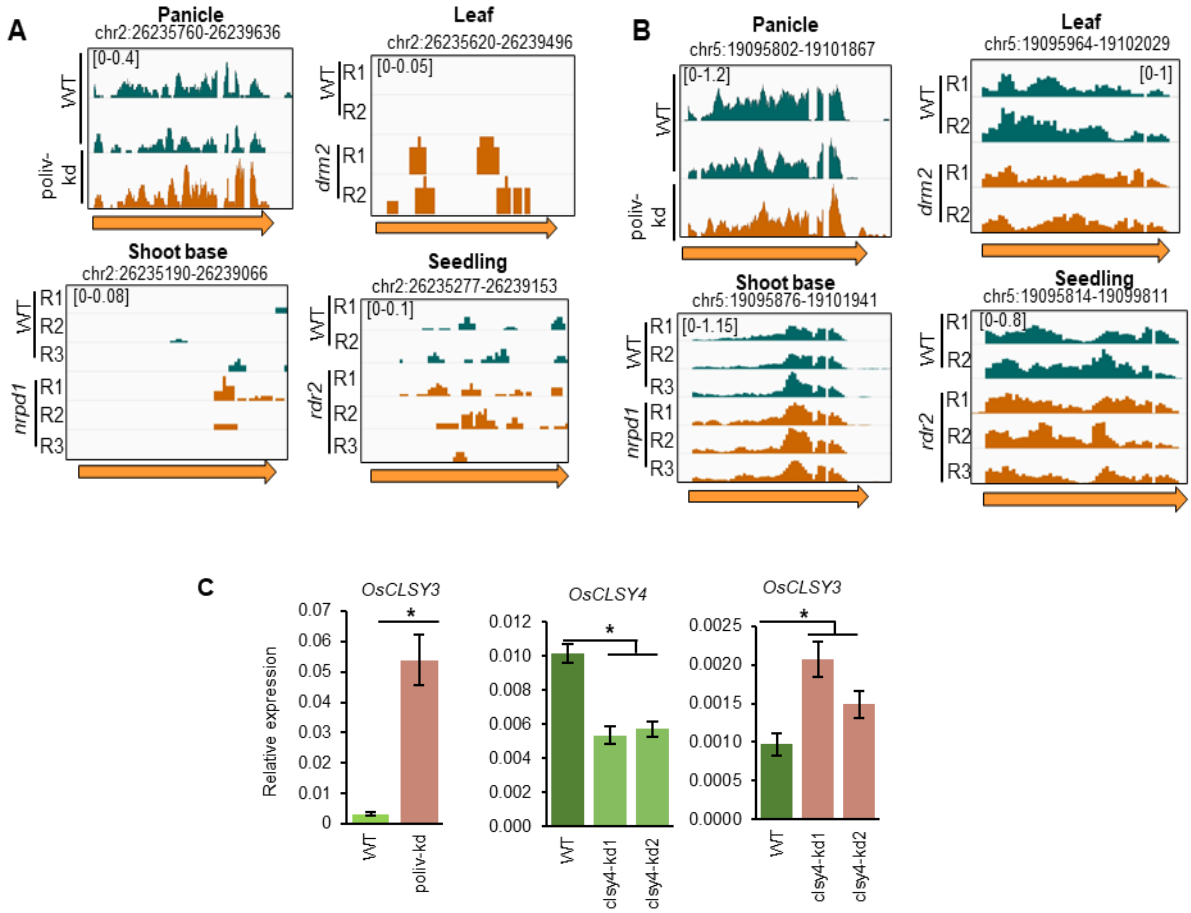


Fig.

(A) IGV screenshots showing expression of *OsCLSY3* in different RdDM pathway mutants.
(B) IGV screenshots showing expression of *OsCLSY4* different RdDM pathway mutants.
(C) RT-qPCR is showing expression *OsCLSY3* in poliv-kd anther (Left) and clsy4-kd 5 days old seedling (Right). OsActin served as internal control. Error bar-Standard Error (SE).

4. **Reviewer's comment:**

The off-target events of CRISPR/Cas9 often occur especially for the homologous genes. It is necessary to examine the sequence of OsCLSY4 and OsCLSY1 in clsy3-ko mutants. The transcriptional levels of OsCLSY4 and OsCLSY1 in clsy3-kd should be examined to make sure the effect of artificial microRNA only decreases the transcriptional levels of OsCLSY3 but not OsCLSY1 and OsCLSY4.

Author's reply: Thank you for this suggestion. Here, we used artificial miRNAs to knock down the gene. The amiR strategy with our sequence-specific modifications very specific in silencing, unlike antisense-RNA based knock down (Ossowski, Schwab, and Weigel 2008; Warthmann et al. 2008; Narjala et al. 2020). We found that both the amiRs employed here targeted two different exons of *OsCLSY3* as expected, but not those of *OsCLSY4* and *OsCLSY1* (please see the attached A panel). We also found transcript levels of other *CLSYs* are not altered in clsy3-kd endosperm derived transcriptome (Panel B). This was also verified with RT-qPCR in clsy3-kd panicle tissues (Panel C).

We found that the gRNA sequence which used to knock out the *OsCLSY3* showed single hit in rice genome when blasted to RAP-DB (Panel D). We amplified ~5kb genomic fragments of *OsCLSY1* and *OsCLSY4* gene from clsy3-KO plants and performed deep sequencing. We did not find modification in other CLSY genes in *clsy3-KO* (Panel E), indicating that the *KO* achieved here is very specific without any off targets.

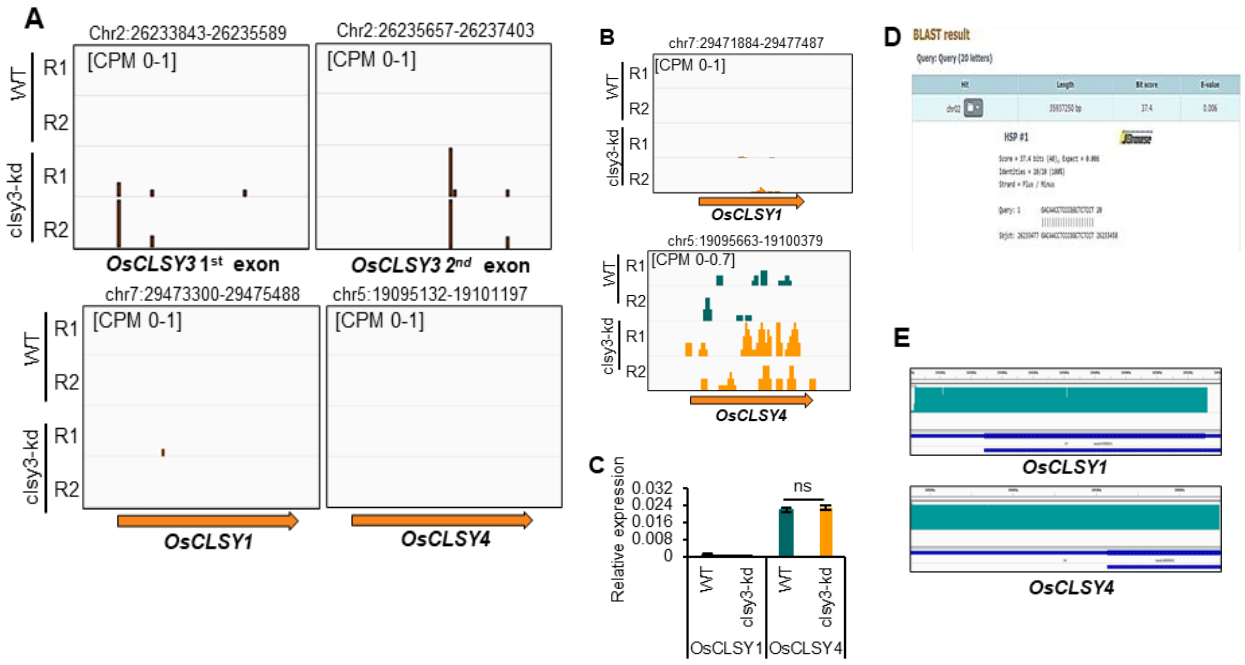


Fig.
(A) IGV screenshots showing alignment of amiRs in different OsCLSYs in EN. (B) IGV screenshots showing expression of *OsCLSY1* and *OsCLSY4* in clsy3-kd EN. (C) RT-qPCR is showing expression Other *OsCLSYs* in clsy3-kd panicle. *OsActin* served as internal control. Error bar-Standard Error (SE). (D) Image showing alignment of CLY3 guide RNA in rice genome. (E) IGV screenshots showing OsCLSY1 and OsCLSY4 genomic loci in clsy3-ko leaf.

5. **<u>Reviewer's comment:</u>**
The phenotypic analysis was conducted only in one line of transgenic plants for OsCLSY3::GUS (Fig. 2F-G), clsy3-kd (Fig. 3E-F, H-M). The results from one line were very vulnerable. At least three independent lines are needed.

Author's reply: Thank you for this comment. Apologies for this confusion. For the OsCLSY3: GUS, we used different lines but the number of replicates used were not mentioned. In all the GUS assays with AZA, we had 5 plants each for the analysis. In phenotyping, we used two transgenic lines OE (#10 and #2). For clsy3-kd, all seed related phenotypes, one single amiR line and one double amiR line were separately taken. To reduce the ambiguity, we wrote OE and clsy3-kd and did not mention the line numbers in detail, and this we have incorporated in the revised script.

Minor suggestions:

<u>Reviewer's comment:</u>

In line 76-77, "AtFIS2, AtMEA, AtVIM5, AtNRPD" should be "FIS2, MEA, VIM5 and NRPD1".

Author's reply: Thank you. We changed the sentence accordingly.

Reviewer's comment:

In line 19, "RNA directed DNA methylation" should be "RNA-directed DNA methylation".
Author's reply: Thank you for pointing out this mistake. We have corrected the sentence.

Reviewer's comment:

The accumulation of 24-nt siRNAs, the methylome in other tissue of clsy3 mutant should be examined.
Author's reply: Our study majorly focusses on endosperm because OsCLSY3 majorly expresses in this tissue. We have not studied other tissues where expression of *OsCLSY4* or *OsCLSY1* predominate. We observed germination, pollen and other reproductive development related defects in *OsCLSY3* transgenic lines. It might have some other roles through RdDM pathway and this detailed study is beyond endosperm-centric study presented here.

Reviewer #3 (Remarks to the Author):
In the current work, entitled "Upstream regulator of genomic imprinting in rice endosperm is a small RNA associated chromatin remodeler", Pal and colleagues identify CLSY3 as an imprinted gene in rice and proceed to characterize how it is imprinted and what affects it has on siRNAs, DNA methylation, gene expression, and rice development. Their work represents several key advances as the data presented clearly demonstrate that CLSY3 is a maternally expressed gene that is regulated by RdDM-mediated methylation at two TEs located in its promoter. Furthermore, clsy3 mutants were found to have reduced siRNA levels in endosperm and they displayed severe fertility defects demonstrating a critical role for CLSY3 in rice reproductive development. However, based on the data presented, it is difficult to assess (1) the role of CLSY4 in regulating CLSY3 expression (2) the strength of the clsy3 mutant relative to poliv, (3) the affects of clsy3 on DNA methylation, and (4) the associations between losses of siRNAs and the expression of genes related to imprinting and/or the observed developmental defects. As detailed below, clarifying these connections will provide additional support to the authors claims and raise the impact of their findings.

Author's reply: Thank you for your encouraging comments and elaborate summary of the work. Your assiduous review pointing all the critical points, suggestions and mistakes has been of immense help to us. We completely agree with your observations and suggestions. We have strived hard to address each and every point that you have raised.

1. The role of CLSY4 in regulating CLSY3 expression and in controlling siren siRNAs remains unclear.

a. Reviewer's comment - The text mentions the generation of clsy4-kd lines with an amiR strategy, but there is no data presented to show these lines

have lower CLSY4 levels. They show the amiR is expressed in Fig. 3F but not that CLSY4 transcript levels or protein are affected. Thus, it is not clear if the reduced siRNAs at mites in these lines is due to reductions in CLSY4.

Author's reply:

a. Thank you for pointing this out. We have now confirmed the levels of amiR expression by sRNA northern and observed a clear reduction in *OsCLSY4* transcript level in the leaf tissues in clsy4-kd lines (Please see the panels attached below). Please note, the phenotype of the clsy4-kd plant that we generated is quite similar to the one that got published while this paper was under revision(D. Xu et al. 2023). In our clsy4-kd lines, we see reduced plant growth and developmental abnormalities similar to the reported research (please see the attached panel A-B). Clearly, these data suggest that the observed reduction in MITE sRNAs is mostly due to reduced CLSY4 levels. Xu et al (2023) also showed a global reduction in MITE-derived sRNAs in seedlings. The data incorporated into revised manuscript (Fig.3F and Supplemental Fig. S3C-E)

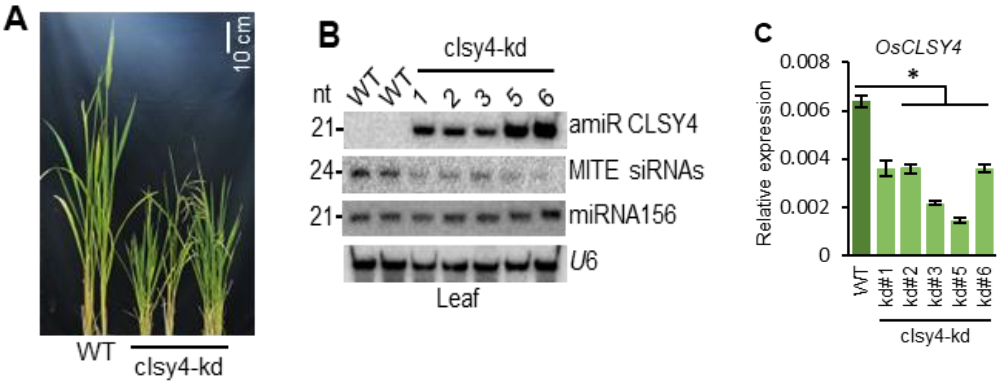


Fig.

(A) Image showing phenotypes of clsy4-kd transgenic lines. (B) sRNA NB showing expression of amiR (C) Bar plot showing expression of *OsCLSY4* on clsy4-kd lines.

b. **Reviewer's comment -** The text states "Using bisulfite sequencing (BS-PCR), we found a reduction of DNA methylation at the OsCLSY3 promoter in leaves of clsy4-kd, which indicated that OsCLSY4 controls expression of OsCLSY3 via RdDM (Fig. 3F)." However, the expression of CLSY3 in the clsy4-kd line was not assessed. To demonstrated a role for CLSY4 in regulating CLSY3 expression RT-qPCR experiments should be conducted in the clsy4-kd lines (once they are vetted as mentioned in "a").

Author's reply: Thank you for this comment. We found *CLSY3* expression to be increased in clsy4-kd 5-day-old seedlings (Please see the figure attached below). This data is now added in the revised manuscript (Supplemental Fig. S3E).

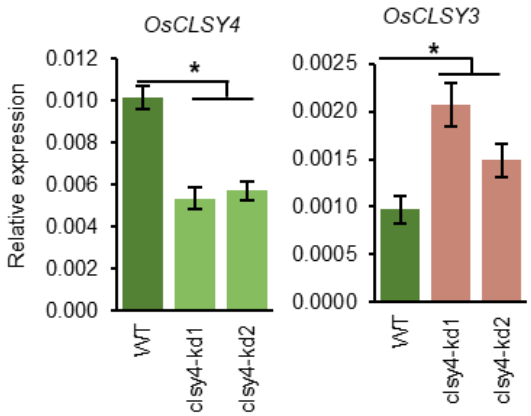


Fig.

Bar plots showing expression of OsCLSY4 and OsCLSY3 in 5 days old seedling. OsActin serve as internal control, Error bar-Standard Error (SE).

c. **Reviewer's comment -** In Fig. 6F the clsy4-kd lines are used again and based on siRNA blots it was suggested that CLSY4 doesn't affect siren loci. However, the quality of the blot is a bit low and it is difficult to assess from this single experiment the role of CLSY4 at siren loci. At a minimum this blot should be repeated. However, including blots at additional siren loci or conducting smRNA-seq experiments in clsy4-kd lines would be advised if the authors want to claim CLSY4 is not involved in siren siRNA production.

Author's reply: As per the suggestion of the reviewer, we tried an additional siren locus (siren 2) which also showed a reduction in accumulation in clsy3-kd, confirming that the reduction in siren loci in clsy-3 kd is a more general phenomenon (this detail is added in the revised manuscript). However, we do believe with limited data, OsCLSY4 also plays a role in siren loci and in endosperm development. Previous reports also discuss such a redundancy in function among rice CLSY proteins (Yang et al. 2018; D. Xu et al. 2023). This data is incorporated in the revised manuscript (Fig. 7F).

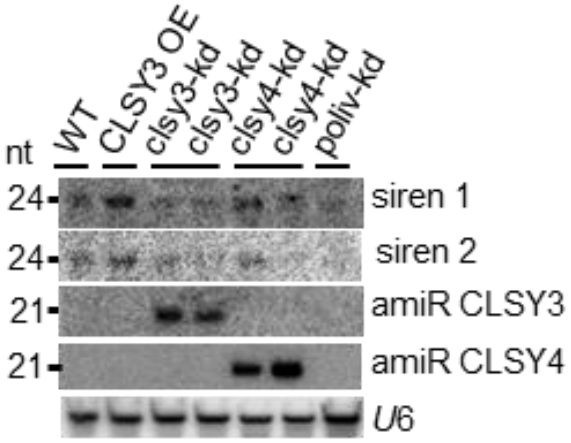


Fig.

sRNA NB showing expression of siren sRNAs in two selected siren loci in different RdDM pathway related transgenic lines.

2. Reviewer's comment - Some clarifications on the role of CLSY3 in siRNA production are required assess its contributions to the endosperm small RNA landscape.

a. Reviewer's comment - It's not clear what parameters were used to call clsy3-dependent siRNAs. The methods mention using DESseq2, but the fold change and FDR cutoffs were not specified. Please add this information to the methods.

Author's reply: Thank you for pointing out the ambiguity between text and the supplemental table. We agree with the concern and corrected it in the revised manuscript. We also added these details in the method section.

Apologies for the confusion. In revised manuscript, we have clearly added the missing information. Here, after Shortstack analysis, we used 'bedtools' intersect to find sRNA loci which are lost in clsy3-kd endosperm tissues. We termed those loci as "CLSY3-depenedent sRNA loci".

b. Reviewer's comment - In Fig. 5D is hard to tell the global decrease in siRNAs from these tracks. With a 70% decrease in siRNA producing loci there should be a way to make this more clear.

Author's reply: In Fig. 5D, we had shown across different chromosomes, how 23-24 nt sRNAs are reduced in clsy3-kd endosperm. We have now added 23-24 nt sRNA loci which were identified through Shortstack analysis followed by 'bedtools' intersect in WT and clsy3-kd endosperm to clearly show this reduction. The result is provided in the revised manuscript in the form of a

venn diagram (Fig 5F and please see below). In text, our statements probably overstated CLSY3 dependency in the previous version. We have corrected this to reduce ambiguity in the revised manuscript.

Fig.

Venn diagram showing CLSY3-dependent 23-24 nt sRNA loci in clsy3-kd EN.

c. Reviewer's comment - Fig. 5F and this text are confusing. "We observed around 70% of sRNA loci lost sRNAs in clsy3-kd compared to WT which were further called CLSY3-dependent sRNA loci (Fig. 5F and Supplemental dataset S6)." Fig. 5F shows the overlap of shortstack clusters showing 8,903 of the clusters called in the clsy3-kd were also called in the WT. However, the 70% comes from dataset S6 showing 21,653 of the 29,850 WT cluster are downregulated in clsy3. Please clarify the text/figure to make this point more clear.

Author's reply: As we discussed in above, in Fig. 5F, after Shortstack analysis followed by bedtools intersect, we got a total of 29850 (23-24 nt) sRNA loci in WT endosperm. Using the same cutoff and threshold values, we found 10865 (23-24 nt sRNA loci) in clsy3-kd samples. In revised manuscript text, we have modified the text to include these new details and corrections.

d. Reviewer's comment - Regarding the siRNA levels at siren loci, the data in Fig. 6A-F and Fig. S6 clearly show effects in both clsy3-kd and poliv-kd lines. However, several questions remain that can be addressed with additional data analysis.

i. Reviewer's comment - While in the majority of siren loci are reduced in the clsy3-kd lines based on the z-score heatmap (Fig. 6c) it is hard to know how robust the changes are. Please indicate what fraction of these loci were identified as/overlap with clsy3-depenend clusters from the earlier analysis. Depending on the numbers, it might also be useful to visualize the siren loci in subgroups rather than all together.

Author's reply: Thank you for pointing out the ambiguity between text and the supplemental table. We agree with the concern and corrected it in the revised manuscript. Thank you for these excellent suggestions.

i. To visualized the siren loci in a better way, we sub-grouped all 801 siren loci mentioned in the previous study (Rodrigues et al. 2021) and performed analysis. The category-1 contains 316 siren loci which are absolutely CLSY3 dependent, are log_2(-1 to -3) fold downregulated in clsy3-kd when compared to WT (RPKM value). The category-2 contains 464 siren loci that might not be exclusively CLSY3 dependent, as they showed both downregulation and somewhat an upregulation under low cutoff conditions. The category-3 contains 17 siren loci which were upregulated. We have plotted them as box plots and heatmaps in the revised manuscript. This subgrouping suggested by the reviewer clearly helped us to identify the impact of CLSY3 in generating siren loci (Please see the attached Fig. below). This data is incorporated in the revised manuscript (Fig S12 A-B).

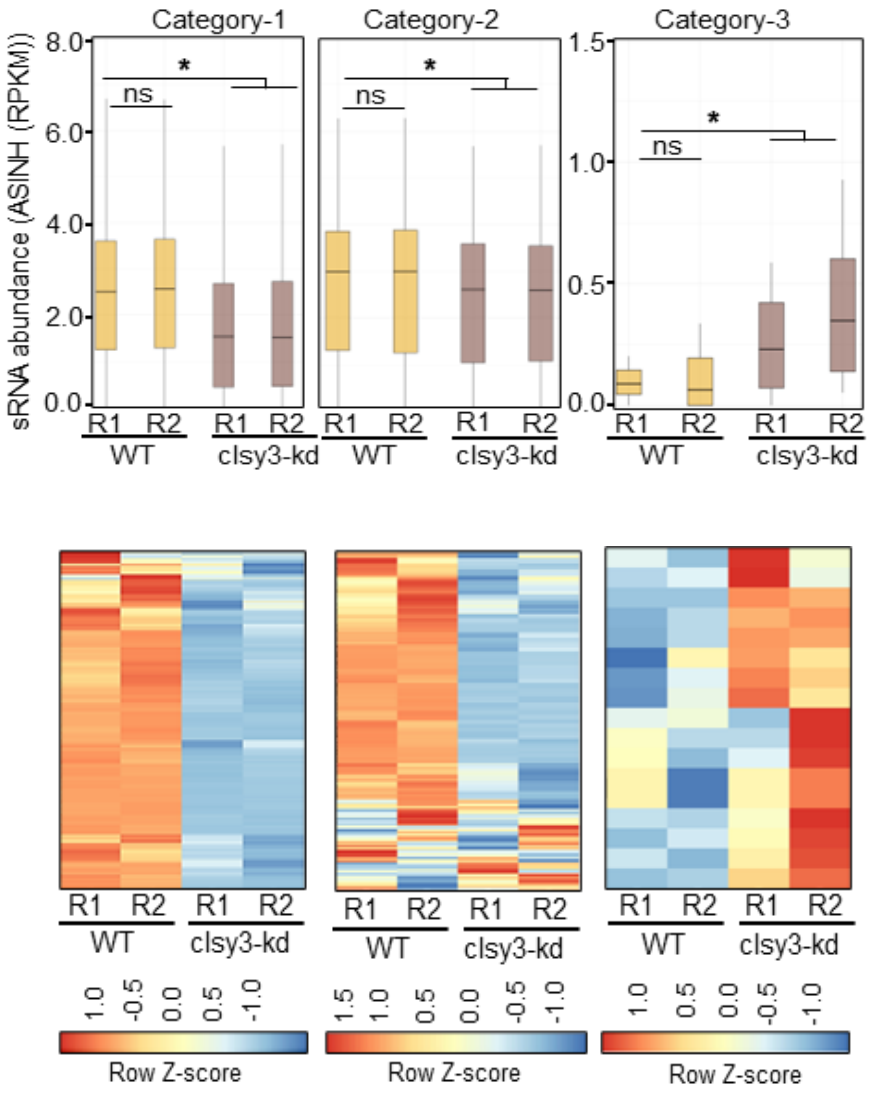


Fig.

Boxplots and heatmaps showing expression of siren sRNAs in the clsy3-kd EN. (Set1-316, Set2-464, Set3-17). [Sets made according to log_2 (clsy3-kd/WT) fold change. Set1 < log_2(-1), Set2 log_2 (-1) < > log_2 (+1), log_2 (+1) < Set3]

ii. Reviewer's comment - For the siren loci data are also shown for the poliv-kd, but always as separated z-score plots or IGV track with different scales than used for the clsy3-kd. Thus, its hard to assess how strong the clsy3-kd line is compared to the poliv-kd line. Rather than separating these data, please plot them together to allow for such comparisons.

Author's reply: Thank you, this is a very good suggestion. We tried to do such an analysis. However, since the tissues used for nrpd1 and clsy3 analysis are not exactly same, we got different number of loci even in WT samples. We were unable to cross compare between clsy3-kd and poliv-kd data in our manuscript and plotted them with their respective WT datasets (please see our reply to Reviewer 1 as well). We did not add the clsy3-kd and poliv-kd in same z-score or IGV tracks with same scales because of already explained reasons. However, in revised manuscript, we incorporated the normalized RPKM values of all siren loci (WT, poliv-kd and clsy3-kd) in detailed (Supplemental dataset S8).

Reviewer's comment -The connections between reduced siRNA levels and altered gene expression are difficult to follow, making it hard to assess the impact of CLSY3 on gene regulation. This is true for siren adjacent genes, imprinted genes and other genes associated with seed/reproductive traits.

a. Reviewer's comment - The main issue is that it is not clear if the comparisons are restricted to statistically significantly reduced siRNA clusters and DE genes in the clsy3-kd lines or rather selected genes of interest showing trends in siRNA and gene expression levels. Z-scores are a nice way to represent data with a wide spread in expression levels, as is often the case for siRNA expression levels, but with just two genotypes included this can also make subtle differences look significant. Understanding the connections between siRNAs and gene expression is further convoluted by the mixed usage of gene names and gene IDs between heatmaps and screenshots making it hard to know what to compare. Please clarify these issues throughout and include gene lists for the siren adjacent genes, seed

development and yield related genes, maternally and paternally expressed siRNA loci etc. Some specific examples are highlighted below.

a. Author's reply: Thank you for bringing up these concerns that have helped us to refine the analysis further.

We designated DEGs in the previous version of the manuscript, that are $log_2$1.5-fold upregulated and downregulated genes (p-value<0.05) in DEseq2 analysis. In the revised manuscript, for siren loci adjacent genes, imprinted genes and other genes associated with seed/reproductive traits, we counted the transcript value with 'bedtools multicov' followed by normalization and plotted the data. We observed a large number of them overlapping with DEGs (nearly 50 siren adjacent genes, 123 imprinted genes and 12 seed/reproductive traits related). Please note that due to this stringent criterion, the genes which were changed significantly but did not cross the DEseq2 threshold value were excluded from the list. We have also added a table listing these genes with their normalized expression values (RPKM) in the revised manuscript (Supplemental dataset S12).

b. Reviewer's comment - Line 427 mentions about 1000 loci with reduced siRNAs and increased expression in clsy3-kd mutants. What are these loci and are they clsy3-dependent siRNA clusters from the DESeq2 analysis? The example, in Fig 6J it doesn't look like siRNAs at the LTR Gypsy are reduced very much. If that is one of the better examples then the changes in gene expression don't appear to be well correlated with changes in siRNA levels.

b. Author's reply: Thank you for pointing this out. We overlapped CLSY3-dependent sRNA loci with all annotated rice genes adjacent to them (with 2kb promoter and terminator). We clearly see an inverse correlation in a large number of CLSY3-dependent sRNA loci and adjacent mRNA expression. Please see the figure attached below. This data is now incorporated in the revisited manuscript (Fig. 7I). The data used for making the plot also incorporated into revised manuscript (Supplemental dataset S11) in details. We agree with the Reviewer that the example in Fig. 6J in the previous version was not an ideal example. We incorporated a better example in the revised manuscript (Fig. 7J).

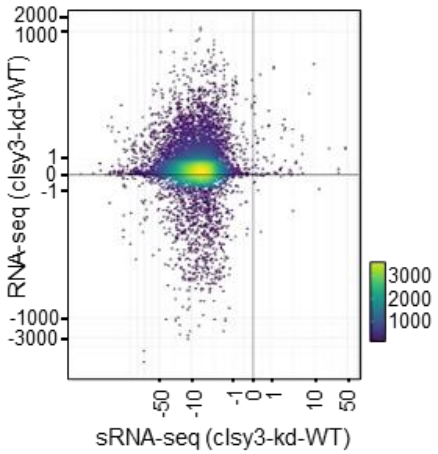


Fig.
Correlation plot showing relation between sRNAs and RNA in clsy3-kd EN.

c. Reviewer's comment - Its not clear which if any of the loci shown in Fig. 6L are statistically significantly downregulated in the clsy3-kd lines. Please indicate which of these genes are part of the DEGs from Fig. 6H.

c. Author's reply: Thank you for this comment. We have now provided the normalized expression values (RPKM) in WT and clsy3-kd of those genes to show their detailed change of the expression in Supplemental dataset S12 in revised manuscript. We also overlapped the listed genes from Fig. 6L with DEGs by bedtools intersect. We found 47 siren adjacent genes and 12 seed development/yield related genes overlapping with the genes listed in Fig. 6H (in the previous version mentioned by the Reviewer).

d.Reviewer's comment -In addition to the heatmap and boxplots in Fig 7A and B, it would be good to note that 74 upregulated and 39 downregulated imprinted genes are changed enough in expression to be captured in the endosperm clsy3-kd DEGs in Fig 6H while the other are not. For B and C given there are up and downregulated genes and some are significant and others not perhaps a breakdown of these classes for the boxpots would be more informative.

d. **Author's reply:** Thank you for this suggestion. In the previous version of the manuscript, we found among 635 published imprinted genes, 124 imprinted genes overlapping with DEGs. We have now separated the imprinted genes as per the Reviewer's suggestion. We find that, among 264 upregulated imprinted genes, 76 were the part of the DEGs. The plots are attached bellow. Please see the revised manuscript Fig. S14A.

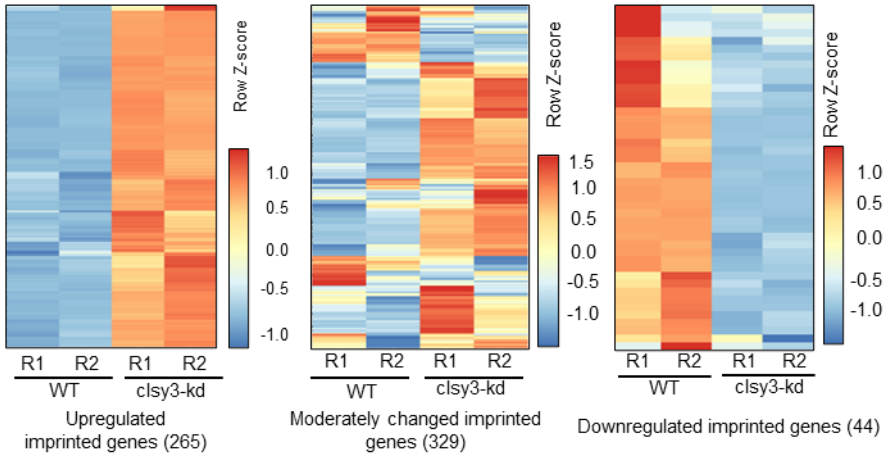


Fig.
Heatmaps showing expression of imprinted genes in clsy3-kd EN (N-635)

e. Reviewer's comment - For Fig 7D it is not clear how the 15 MEG siRNA loci and 16 PEG siRNA loci were defined or how they related to the siRNA clusters identified from the Shortstack analysis. Are these MEG and PEG loci included in the 21,653 clsy3-dep clusters? Like for the effects at siren loci, having the poliv-kd data visualized together, on the same scale, with the clsy3- kd data instead of in the supplement would be helpful to determine the relative strength of clsy3 versus poliv.

e. Author's reply: Thank you for this suggestion. The 15 maternally expressed and 16 paternally expressed sRNA loci were taken from published studies (Rodrigues et al. 2013, 2021; Chen et al. 2018). We have counted abundance of 23-24 nt sRNA from those previously published datasets. We have added the normalized RPKM values of those loci in clsy3-kd and poliv-kd samples along with their coordinates in Supplemental dataset S14 in the revised manuscript.

f. Reviewer's comment - In lines 463-465 the text states "However, most of the upregulated maternally expressed sRNA loci in clsy3-kd were downregulated in poliv-kd, probably due to a compensatory effect likely involving other CLSYs (Supplemental Fig. S8A and B)." However, as Fig. S8A only shows data for the poliv-kd and in Fig. S8B siRNA levels are not clearly higher in the clsy3-kd at either of the maternally expressed siRNA loci shown, it is not clear why these figures are referenced. Please clarify. Are the heatmaps for Fig. S8A and Fig. 7D in the same order? If so, please mention this and also indicate which genes are shown in the screen shots.

Author's reply: Thank you so much for pointing this out this mistake. We apologise for this ambiguity. We have corrected the sentence in the revised manuscript. In previous version, Fig. S8A and Fig. 7D were not exactly in the same order. The normalized values with were used for generated the

heatmaps provided in revised manuscript Supplemental dataset S14 in details.

The imprinted genes in previous manuscript that were used for generating IGV screenshots were also listed with their normalized RPKM values and IDs in the revised version (Supplemental dataset S15).

g. Reviewer's comment - Are the imprinted genes in Fig 7E part of the 74 imprinted DEG from the mRNAseq? And are some of the genes from 7E shown in 7F? If so, please use a common naming system to make this more clear, if not, why were the genes in Fig. 7F chosen?

g. **Author's reply:** Yes, indeed 5 imprinted genes are among the 20 DEGs. We also observed that this number goes up to 9 if we relax the cutoff. The genes shown in Fig. 7F are a subset of the genes shown in 7E. We apologies for this confusion. As per the suggestion, we have added the names of the gene with their RAPDB IDs in the revised manuscript Supplemental dataset S15.

4. **Reviewer's comment -** Throughout the paper the connections between losses of siRNAs in the clsy3-kd line and changes in DNA methylation are weak.

a. Reviewer's comment - As for the links to expression, more transparency in why specific loci were selected to show changes in methylation between the endosperm and embryo and/or for targeted for BS sequencing would be helpful. Furthermore, transparency on whether these sites are statistically significant DEGs and clsy3-dependent siRNA clusters is also required.

a. Author's reply: The endosperm is globally hypomethylated than other tissues. However, we found some regions in endosperm which are hypermethylated in CHH context compared to embryo after analyzing a publicly available whole-genome DNA methylation dataset (Rodrigues et al. 2021). Surprisingly, many of those regions overlapped with CLSY3-dependent sRNA loci and the expression of adjacent genes were altered. We hypothesised that CLSY3 might regulating DNA methylation of those regions. We selected those regions for further BS-PCR and Chop-PCR analysis as mentioned in the previous version of the manuscript. This part is also retained in the revised script as an additional proof.

b. Reviewer's comment - For TEs, the example in Fig. 5I at the 5S-rDNA repeat looks good, but for the MITE the siRNA change is not significant and its not clear what region of the BS data corresponds to the siRNA site. As siRNA levels are globally reduced at MITEs, Gypsy and retroelements (Fig. 5H) additional loci should be assessed. In addition, the change in methylation at the line element shown in Fig. S5D is also very modest.

b. Author's reply: Thank you for this comment. As per the suggestions of reviewers, we have performed the whole genome bisulfite sequencing in clsy3-kd endosperm to understand the global DNA methylation status. In the

revised manuscript, we assessed methylation of multiple MITEs, Gypsy and retroelements. This data is now incorporated in the revised manuscript (Fig. 5I and 5J).

LINE element mentioned were used as a negative control because, LINE and CACTA elements did not show much overlap with CLSY3-dependent sRNAs. This detail in incorporated in the revised manuscript text.

c. Reviewer's comment - For siren loci, DNA methylation is only assessed for one locus (Fig. 6G) and while reductions are clear in the poliv-kd, the changes in the clsy3-kd are modest and it would be helpful to see a quantification. This suggest CLSY3 is not the only factor controlling siren loci, and further draw into question the assessment that CLSY4 is not involved. More loci should be assessed to determine the effect of clsy3 mutants on siren methylation levels.

c. Author's reply: Thank you for pointing out this concern. The PolIV is the key player for generating 23-24-nt sRNAs in RdDM. However, there is a redundancy between CLSYs as shown previously (Yang et al. 2018; D. Xu et al. 2023). Due to these reasons, in poliv-kd the selected siren locus reduction of DNA methylation is clearer than clsy3-kd (Previous manuscript Fig.6G).

As pointed by the Reviewer, after analysing the whole genome methylome and categorizing the siren loci as per sRNA expression, we also observed a possible redundancy between CLSY proteins. The CLSY3 does not seem to be the only factor controlling the siren loci. We quantified total DNA methylation in all siren loci in rice but did not observe any global change in DNA methylation level (Please see the below figure panel A). However, we found in many siren loci, the DNA methylation was decreased as expected. We also observed redistribution and sometimes increase in DNA methylation (CHH context) in clsy3-kd (Please see the below figure panel B). The data is incorporated in the revised manuscript Fig. 7E and Supplemental Fig. S12C-D.

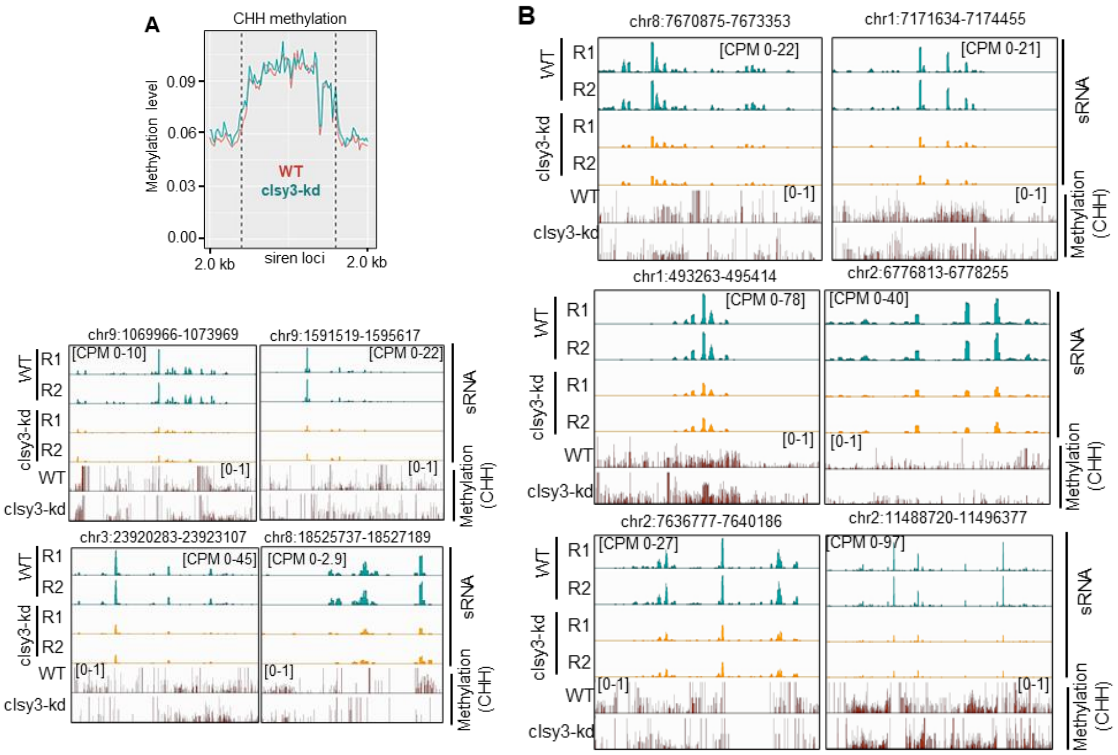


Fig.

(A) Metaplot showing global CHH methylation status of all rice siren loci in clsy3-kd endosperm.
(B) IGV screen shots showing sRNA and CHH methylation level in clsy3-kd endosperm.

d. <u>Reviewer's comment -</u> For imprinted genes, changes in methylation between the embryo and endosperm are shown and reductions in siRNAs are shown for the clsy3-kd for many genes. However, change in methylation are only assessed at a few loci and only three out of six site showed significant changes.

d. Author's reply: Thank you for this concern. In our previous manuscript, we showed changes in methylation between the embryo and endosperm for 6 imprinted genes and 1 DEG named OsTAR1 (Original manuscript Fig S9). The basis of selecting those genes, was not clearly conveyed in previous manuscript. We observed around 124 imprinted genes are significantly changed in the clsy3-kd transcriptome. However, we could not find CLSY3-dependent sRNA loci near to the most of the imprinted genes. Mechanism wise, it is hard to explain how CLSY3 or RdDM pathway regulate them directly. Previous studies found presence of imprinted sRNA loci near to many imprinted genes (Rodrigues et al. 2013; Yuan et al. 2017; Chen et al. 2018; Rodrigues et al. 2021). We selected 20 genes which have imprinted sRNA loci in their vicinity. We found 9 of them were significantly changed in terms of expression in RNA-seq and imprinted sRNAs level also showed

decrease in clsy3-kd endosperm. Among these 9, the 6 genes we observed had more CHH methylation in endosperm compared to embryo clearly. Due to this reason, in our previous manuscript, we used only 6 selected imprinted genes for DNA methylation analysis by targeted bisulfite and chop-PCR. Along with whole genome DNA methylome analysis, we also included the DNA methylation status analysis using targeted bisulfite and chop-PCR (Fig. 8F-H, Supplemental Figure S15 and S16).

e. Reviewer's comment - Overall, these results draw into question how much the clsy3 mutant affects methylation levels. Having a weaker phenotype at the level of methylation vs siRNAs is not unprecedented given siRNAs are often in excess and there are redundancies within the CLSY family in other organisms and well as redundancies between the RdDM and other methylation pathways. However, if the phenotype is weak this should be stated more clearly to promote further exploration into the mechanisms controlling RdDM in the endosperm. Ideally, genome-wide BS sequencing experiments would be conducted in parallel with the clsy3-kd and poliv-kd lines to assess the contribution of CLSY3 to RdDM in the endosperm on a global scale.

e. Author's reply: Thank you for this suggestion. The BS-PCR and chop PCR for selective loci cannot show the clear picture of DNA methylation change in the clsy3-kd endosperm due to their limitations. To address this, we have performed the whole genome BS sequencing. We observed as expected many CLSY3-dependent sRNA loci have decreased DNA methylation in clsy3-kd endosperm (please see the figure below panel A). We found also many CLSY3 dependent sRNA loci gained DNA methylation as previously observed in *Arabidopsis* and rice single clsy mutants indicating redundancy (Yang et al. 2018; D. Xu et al. 2023) (please see the figure below panel B). After analysis the whole genome methylome, we completely agree with the Reviewer that there is a redundancy between CLSYs in rice endosperm. This need to further extensive study with different combinations of CLSYs mutants in rice. This data is incorporated in the revised manuscript.

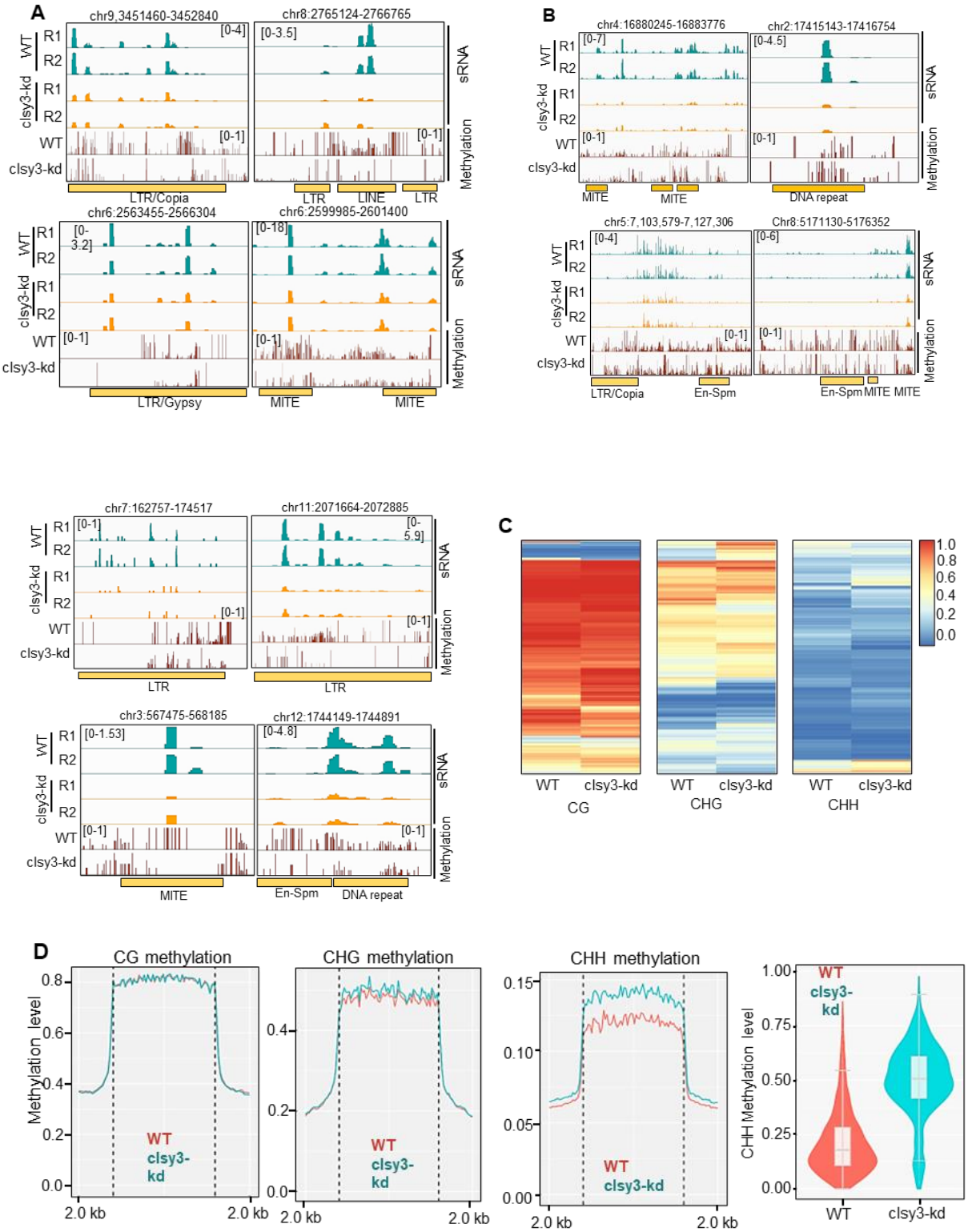

A
chr9,3451460-3452840
chr8:2765124-2766765
chr6:2563455-2566304
chr6:2599985-2601400
WT
clsy3-kd
R1
R2
sRNA
Methylation
LTR/Copia
LTR
LINE
LTR/Gypsy
MITE
B
chr4:16880245-16883776
chr2:17415143-17416754
chr5:7,103,579-7,127,306
Chr8:5171130-5176352
DNA repeat
En-Spm
C
CG
CHG
CHH
D
CG methylation
CHG methylation
CHH methylation
Methylation level
CHH Methylation level
2.0 kb

Fig.

(A) IGV screenshots showing sRNA and CHH methylation status of hypo methylated loci in WT and clsy3-kd EN.
(B) IGV screenshots showing sRNA and CHH methylation status of hyper methylated loci in WT and clsy3-kd EN.
(C) Heatmap showing DNA methylation status in WT and clsy3-kd EN in CG, CHG and CHH context.
(D) Metaplots showing DNA methylation CG, CHG and CHH contexts in WT and clsy3-kd EN and violin plot showing methylation in CHH context.

Minor comments:

1. **Reviewer's comment -** In Fig S1C one of the gene is labeled OsNF-NC11, but in the text and Fig S1B is labeled OsNF-YC11. Please correct whichever label is in error.
   Author's reply: Thank you for pointing out this mistake. We have corrected it in the revised script.
2. **Reviewer's comment -** In Fig. S1D one gene is named OsRRP, but in the text only a gene ID is given "Os09g0537700". Are these the same gene? Please clarify.
   Author's reply: Apologies for this mistake. We have corrected this in the revised text.
3. **Reviewer's comment -** The text states "We observed that OsCLSY3 and OsCLSY1 are majorly expressed in endosperm and embryos, respectively. However, unlike Arabidopsis, OsCLSY4 is expressed ubiquitously (Fig.1F, Supplemental Fig. S2A and B). However, in Fig. 1F CLSY1 is not "majorly" expressed in the embryo as it is higher in the young endosperm and anther, for example. Also, CLSY4 shows variation in its expression. Please clarify the text to better reflect the data presented.
   Author's reply: Thank you for raising this concern. We have checked the expression of those genes much more carefully. The OsCLSY4 is expressed ubiquitously as also documented in a recent paper (D. Xu et al. 2023). The paper showed expression of *OsCLSY1* in stamen (stage S12). However, we were unable to detect *OsCLSY1* expression in our anther RNAseq datasets (Hari Sundar G et al. 2023). It is possible that we did not observe its expression due to sampling stage differences or due the variations in expression between *indica* and *japonica* subspecies. The expression of OsCLSYs were also checked in unfertilized ovule and anther (1 day before anthesis) *OsCLSY3* and *OsCLSY4* are well expressed in both these tissues but not *OsCLSY1*.

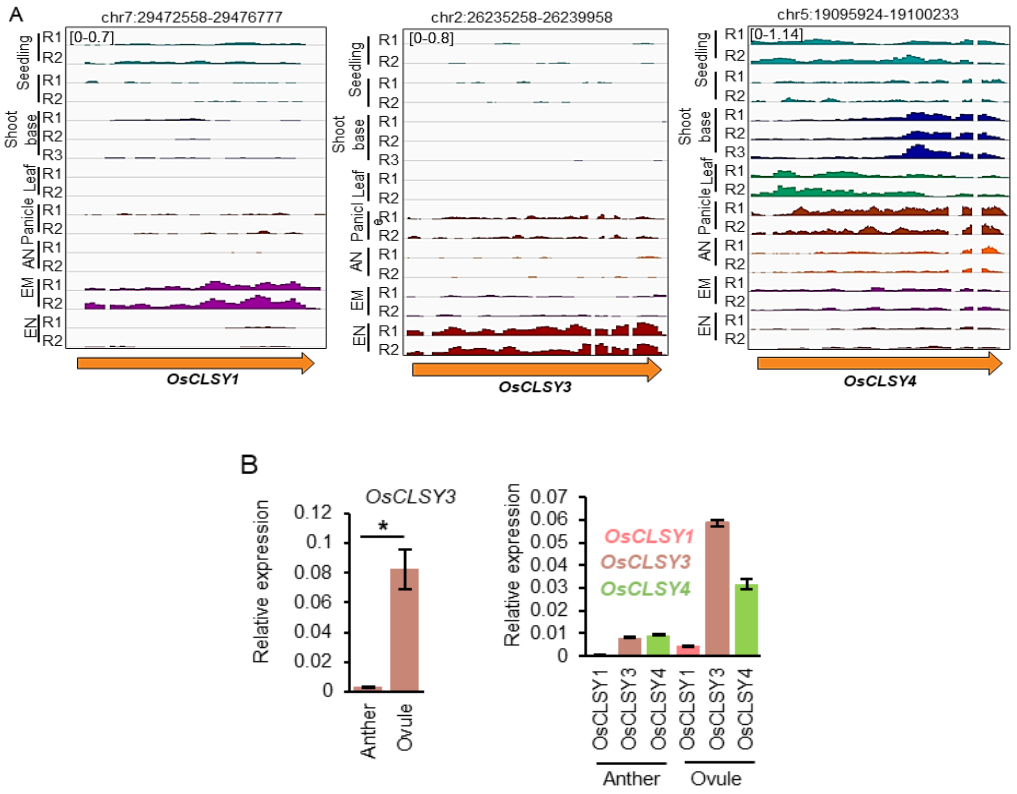


Fig.

(A) IGV screenshots showing expression of rice CLSYs in different tissues (AN-Anther, EM-Embryo, EN-Endosperm).

(B) Bar plots showing expression of *OsCLSYs* in Anther and unfertilized ovule. *OsActin* served as internal control. Error bar-Standard Error (SE).

4. <u>Reviewer's comment -</u> The text states "Around 3607 transcripts showed embryo-preferred expression, while transcripts that expressed highly in endosperm were around 3686 (Supplemental dataset S1 and Supplemental dataset S2). But some of the genes in these tables have no names and are listed as a "-" What does that mean? Are these unannotated genes? Also, since the text (lines 171-174, lines 181-182, etc.) highlights some embryo and endosperm preferred genes by their names rather than gene IDs, perhaps adding an extra column to datasets S1 and S2 with gene names would be helpful.

Author's reply: Thank you for this suggestion. The tables have no names and listed as a **"-"** are unannotated genes or transcripts. As per the suggestion, we have added names of all the genes. Since many genes in rice are not well annotated and carry similar names, some will not have names.

5. <u>Reviewer's comment -</u> In Fig. 1B there are 3,475 EN-preferred genes, but table S1 has 3686 genes. Please clarify why the numbers don't match.

Author's reply: Thank you. This ambiguity was due to genes without names. When we generated Venn diagram, the tool only considered unique gene IDs and the rest were not included. This detail is now mentioned in the revised script.

6. Reviewer's comment - For dataset 3, 175 epigenetic genes are listed, but the text and Fig. 1B only include 160. Please clarify. Also please add a column with the gene names in addition to the gene IDs**.**

Author's reply: Thank you. We have corrected this mistake in the revised list. We added a column with gene names with IDs in the revised manuscript.

8. Reviewer's comment - The text, lines 196-197 reference CHR740, but this is not labeled in Fig. 1D, rather several rice genes are labeled as OsDRD1-like. Please use a common nomenclature in the text and figures.
Author's reply: Thank you for the suggestion. We added these details in the revised phylogenetic tree.

9. Reviewer's comment - Supplemental Table S4 is listed out of order, before Supplemental Tables S2 and S3. Please reorder.

Author's reply: Thank you. We have corrected this in the revised manuscript.

10. Reviewer's comment - The text states "We observed that for OsCLSY3, 86.1% of transcripts came from the maternal genome when compared to OsSHH1 (BiG) which showed 65.3%. As expected in the case for OsARF22 (a PEG), 86.4% of transcripts were from the paternal genome (Fig.1I, Supplemental Fig. S2F). These analyses demonstrated that OsCLSY3 is a maternally expressed imprinted (MEG) in rice." The data matches for CLSY3 and SHH1, but for ARF22 the sequence shown in Fig. 1I is TC(A/G) GT but in Fig. S2F the stacked barplots are for G(A/G) T. Please clarify, shouldn't it be for C(A/G) T?

Author's reply: Thank you very much for pointing out this mistake. We have corrected it in the revised figure.

11. Reviewer's comment - For the traditional BS in Fig. 2D, please reference Table S1 so it is clear the data is from targeted high-throughput sequencing. Also, it is not clear if the same promoter region is being shown in Fig. 2E. For comparison, showing the same region would be helpful. Furthermore, it appears that there is also less methylation at the CLSY4 promoter in the endosperm compared to the embryo despite the text implying that the change in methylation is specific to the promoter of CLSY3. Please clarify the aforementioned points.

Author's reply: Thank you for this suggestion. We added the reference in the revised manuscript. For comparing promoter methylation in *OsCLSY3* and *OsCLSY4* 2 kb upstream regions were taken (equal lengths). As per the suggestion, the targeted BS sequenced regions are now marked. Please note that the levels of DNA methylation in endosperm is low when compared to embryo and other tissues. The DNA methylation at OsCLSY4 promoter is very low which decreases in endosperm as expected.

12. Reviewer's comment - In Fig S3A, the probe region used for the DNA blots is not clear and the expected size for the intact T-DNA is not indicated making his blot hard to interpret. Please clarify. Also, is there DNA blot data for KO#5?

Author's reply: The length between HindIII site and the left border (LB) is 2.4 kb, and the intact T-DNA has to be longer than this. We have marked the region properly in the revised manuscript. We have not added the KO#5 line in this blot as the DNA was of poor quality.

13. Reviewer's comment - In Fig S4 the transgenic plants labelled as OE and KO are not clear. Are these both for CLSY3? Also, please add information for the clsy4-kd.

Author's reply: Thank you for this suggestion. We have added the information and corrected those issues in labelling OE and KO lines.

14. Reviewer's comment - In Fig S4C, the probe region used for the DNA blots is not clear and the expected size for the intact T-DNA is not indicated making his blot hard to interpret. Please clarify.

Author's reply: Thank you for this point. We have indicated the probed region in linear vector maps with an arrow. The length between EcoRI and LB is 2.4 kb. Copies less than 2.4 kb will be considered as truncated T-DNA.

15. Reviewer's comment - The text states "We obtained a total of 7 transgenic plants with double amiRs (clsy3-kd2) (Fig. 4A, Supplemental Fig. S4D) and 8 plants with single amiR (clsy3-kd1) (Fig. 4A)." But figure S4C shows the opposite, 7 lines for kd1 and 8 for kd2. Also, it is not clear how the numbers correlate with each other between the main and supplemental figure. Please clarify.

Author's reply: Thank you for pointing out the mistake in the text and images. Apologies for this mistake. We have corrected these in the revised manuscript.

16. Reviewer's comment - The text states that "Principal component Analysis (PCA) indicated that the identical pool of 23-24 nt endosperm specific sRNAs were downregulated in clsy3-kd and poliv-kd endosperm tissues (Fig. 5C)." This cannot be determined from a PCA and should instead be determined by an overlap analysis of reduced siRNA clusters.

Author's reply: Thank you for this suggestion. We agree with you that the PCA is not the proper way to claim this conclusion. We have changed the text in the revised version.

17. Reviewer's comment - Please define Class I and Class II TEs with regards to Fig 5 at their first mention in the text.
Author's reply: Thank you. We have incorporated this detail.

18. Reviewer's comment - Line 382 references Shortstack analysis and datasets S3 and S4, but I think it should be S4 and S5. Please clarify.

Author's reply: Thank you very much for pointing out this mistake. We have corrected this in revised manuscript.

Thank you very much for all the suggestions and these have clearly helped us to improve the manuscript. We hope the revised manuscript addresses all your comments.

References:

Burgess, Diane, Hiu Tung Chow, Jeffrey W. Grover, Michael Freeling, and Rebecca A. Mosher. 2022. “Ovule SiRNAs Methylate Protein-Coding Genes in Trans.” *The Plant Cell* 34 (10): 3647–64.

Chen, C., T. Li, S. [.]. Zhu, Z. Liu, Z. Shi, X. Zheng, R. Chen, et al. 2018. “Characterization of Imprinted Genes in Rice Reveals Conservation of Regulation and Imprinting with Other Plant Species.” *Plant Physiology* 177 (4): 1754–71.

Hari Sundar G, Vivek, Chenna Swetha, Debjani Basu, Kannan Pachamuthu, Steffi Raju, Tania Chakraborty, Rebecca A. Mosher, and P. V. Shivaprasad. 2023. “Plant Polymerase IV Sensitizes Chromatin through Histone Modifications to Preclude Spread of Silencing into Protein-Coding Domains.” *Genome Research*, June. https://doi.org/10.1101/gr.277353.122.

Hu, Daoheng, Yiming Yu, Chun Wang, Yanping Long, Yue Liu, Li Feng, Dongdong Lu, et al. 2022. “Erratum for: Multiplex CRISPR-Cas9 Editing of DNA Methyltransferases in Rice Uncovers a Class of Non-CG Methylation Specific for GC-Rich Regions.” *The Plant Cell* 34 (4): 1416.

Martins, Laura M., and Julie A. Law. 2023. “Moving Targets: Mechanisms Regulating SiRNA Production and DNA Methylation during Plant Development.” *Current Opinion in Plant Biology* 75 (August): 102435.

Narjala, Anushree, Ashwin Nair, Varsha Tirumalai, G. Vivek Hari Sundar, and Padubidri V. Shivaprasad. 2020. “A Conserved Sequence Signature Is Essential for Robust Plant MiRNA Biogenesis.” *Nucleic Acids Research* 48 (6): 3103–18.

Ossowski, Stephan, Rebecca Schwab, and Detlef Weigel. 2008. “Gene Silencing in Plants Using Artificial MicroRNAs and Other Small RNAs.” *The Plant Journal: For Cell and Molecular Biology* 53 (4): 674–90.

Rodrigues, Jessica A., Ping-Hung Hsieh, Deling Ruan, Toshiro Nishimura, Manoj K. Sharma, Rita Sharma, Xinyi Ye, et al. 2021. “Divergence among Rice Cultivars Reveals Roles for Transposition and Epimutation in Ongoing Evolution of Genomic Imprinting.” *Proceedings of the National Academy of Sciences of the United States of America* 118 (29): e2104445118.

Rodrigues, Jessica A., Randy Ruan, Toshiro Nishimura, Manoj K. Sharma, Rita Sharma, Pamela C. Ronald, Robert L. Fischer, and Daniel Zilberman. 2013. “Imprinted Expression of Genes and Small RNA Is Associated with Localized Hypomethylation of the Maternal Genome in Rice Endosperm.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (19): 7934–39.

Wang, Lili, Kezhi Zheng, Longjun Zeng, Dachao Xu, Tianxin Zhu, Yumeng Yin, Huadong Zhan, Yufeng Wu, and Dong-Lei Yang. 2022. “Reinforcement of

CHH Methylation through RNA-Directed DNA Methylation Ensures Sexual Reproduction in Rice." *Plant Physiology* 188 (2): 1189–1209.
Warthmann, Norman, Hao Chen, Stephan Ossowski, Detlef Weigel, and Philippe Hervé. 2008. "Highly Specific Gene Silencing by Artificial MiRNAs in Rice." *PloS One* 3 (3): e1829.
Williams, Ben P., Daniela Pignatta, Steven Henikoff, and Mary Gehring. 2015. "Methylation-Sensitive Expression of a DNA Demethylase Gene Serves as an Epigenetic Rheostat." *PLoS Genetics* 11 (3): e1005142.
Xu, Dachao, Longjun Zeng, Lili Wang, and Dong-Lei Yang. 2023. "Rice Requires a Chromatin Remodeler for Polymerase IV-Small Interfering RNA Production and Genomic Immunity." *Plant Physiology*, November. https://doi.org/10.1093/plphys/kiad624.
Xu, Le, Kun Yuan, Meng Yuan, Xiangbing Meng, Min Chen, Jianguo Wu, Jiayang Li, and Yijun Qi. 2020. "Regulation of Rice Tillering by RNA-Directed DNA Methylation at Miniature Inverted-Repeat Transposable Elements." *Molecular Plant* 13 (6): 851–63.
Yang, Dong-Lei, Guiping Zhang, Lili Wang, Jingwen Li, Dachao Xu, Cuiru Di, Kai Tang, et al. 2018. "Four Putative SWI2/SNF2 Chromatin Remodelers Have Dual Roles in Regulating DNA Methylation in Arabidopsis." *Cell Discovery* 4 (1): 55.
Yuan, Jingya, Sushu Chen, Wu Jiao, Longfei Wang, Limei Wang, Wenxue Ye, Jie Lu, et al. 2017. "Both Maternally and Paternally Imprinted Genes Regulate Seed Development in Rice." *The New Phytologist* 216 (2): 373–87.
Zhou, Ming, Ceyda Coruh, Guanghui Xu, Laura M. Martins, Clara Bourbousse, Alice Lambolez, and Julie A. Law. 2022. "The CLASSY Family Controls Tissue-Specific DNA Methylation Patterns in Arabidopsis." *Nature Communications* 13 (1): 244.
Zhou, Ming, Ana Marie S. Palanca, and Julie A. Law. 2018. "Locus-Specific Control of the de Novo DNA Methylation Pathway in Arabidopsis by the CLASSY Family." *Nature Genetics* 50 (6): 865–73.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors made efforts to address my comments and added relevant data that improve this manuscript. I think the findings reported are exciting and even though several questions remain, it is an important contribution to our current knowledge. I have a couple of comments that however do not require more experimental work.

1. Line 434: Based on the genome-wide DNA methylation data, there are apparently many more loci that gain rather than lose DNA methylation. It would be important to show how many loci lose and how many gain DNA methylation.

2. L514: To conclude that there is a negative correlation, the authors would need to calculate the correlation coefficient.

3. Line 340: Based on the provides pictures in Figure S4H no conclusion on the endosperm can be drawn. Please rephrase.

4. Line 517ff: If only few of the upregulated genes lose DNA methylation, the statement “This observation suggested that the genes were likely regulated by CLSY3-dependent DNA methylation, specific to endosperm.” is not correct and should be toned down. Based on all data the connection between CLASSY3-dependent siRNAs and DNA methylation remains unclear. This should be stated to avoid the erroneous impression that there is a clear correlation.

5. L563ff: Since the authors generated whole genome bs data, they should use this data for all analyses rather than semi-quantitative chop PCRs or locus-specific bs analyses. They should quantify CHH methylation for all imprinted genes and show the quantified data rather than screenshots of the genome browser.

6. L577: Based on the data I would revise the conclusion that some (not many) imprinted genes are regulated by OsCLSY3 through sRNA-directed DNA methylation.

7. Regarding comment 10; I remain convinced that Figures 1 and 2 do not add much to the novelty of this manuscript, but rather make it unnecessary lengthy.

8. L51: Instead of adding the waterlily example, I suggest to rather write “typically diploid central cell”.

9. L243: remove the “Surprising”. With all what we know how MEGs are regulated, finding a MITE in its promoter is not surprising.

10. Change poliv to nrpd1

11. Figure S13A: The gene identifiers of the MADS-box genes should be included.

12. L560: Instead of “However” the right phrase would rather be “Out of those, 9 genes…”

13. L563: should read “was largely unchanged”

Reviewer #2 (Remarks to the Author):

One very related paper was published on Plant Physiology (kiad 624, 2023, November 22) about one CLSY family member, OsCLSY4 named in this study. There are a lot of contrast results. The authors might carefully reexamine the experimental results with the published data.

Major Concerns

1. Mutation in FEM2/CHR742 caused multiple developmental defects including failure of seed development. The rice plants with simultaneous mutation in CHR740/OsCLSY3/FEL1 and CHR722/FEL2, however, were normal for growth and development including reproductive development. In this manuscript, mutation in CHR740/OsCLSY3/FEL1 resulted in sterility because of pollen and endosperm development. There is no expression for OsCLSY3 in anther (Figure 1d). Why osclsy3 have pollen developmental defects? Given osclsy4 was generated in this study, what is development and DNA methylation phenotype of osclsy4?

2. In this manuscript, the authors claimed that DNA methylation at MITE control the expression of OsCLSY3. Beside the MITE, there is one long transposon in the promoter of OsCLSY3/FEL1. In addition, the dense DNA methylation on this long transposon (Figure 3b, Plant Physiology) might be important for regulating it expression. The OsCLSY3::GUS transgenic plants showed signal in anther (Figure 1h), which is contrast to no expression in anther for endogenous gene in Figure 1c. The inconsistency in expression between transgenic and endogenous OsCLSY3 suggest short promoter used in this study is not enough to reflect it real expression pattern.

3. OsCLSY3 is maternally expressed imprinted gene in endosperm where OsCLSY3 controlled siRNA production and imprinted genes. To confirm this main conclusion, the reciprocal cross between OsCLSY3-kd and WT will provide the maternal and paternal effect of OsCLSY3 on the production of siRNAs, and thus the expression of imprinted genes.

4. To examine siRNA level and gene transcription in endosperm, the dissection of endosperm often has seed coat contamination. The authors need to exclude this possibility.

Reviewer #3 (Remarks to the Author):

In the revised manuscript the authors have addressed many of my previous concerns and have added experiments and analyses that better support their main claims. Based on these changes, I remain enthusiastic about the significance of the work presented. However increased transparency on which genes are likely regulated directly by CLSY3 via the RdDM pathway would better define the scope of CLSY3's role in regulating sRNAs and controlling gene expression and imprinting during rice reproduction. As detailed below, such evidence can be provided by additional analysis of already available data and will increase the impact of the presented work.

Major comments:

1. The identification and reporting of the *clsy3*-dependent sRNA clusters remains unclear. The text states "To identify CLSY3-dependent sRNA loci, we quantified number of loci present in WT and clsy3-kd endosperm by ShortStack analysis (Supplemental dataset S4 and S5). We observed that around 70% of sRNA loci lost 23-24nt sRNAs in *clsy3*-kd, when compared to WT in a bedtools based analysis and these were termed CLSY3-dependent sRNA loci (Fig.5F and Supplemental dataset S6)." The cluster numbers in dataset S4 and S5 match the Venn diagram, which indicates there are 20,947 clusters lost in the *clsy3*-kd and 1,962 clusters gained for a total of 22,909 *clsy3*-dependent sRNA loci. However, dataset S6 lists 21,653 clusters. Please clarify. Were some adjacent clusters merged?

2. The effects of the clsy3-kd on DNA methylation patterns are not well explored leaving it unclear how much this CLSY contributes to epigenetic regulation in the endosperm.

- Please identify hyper and hypo DMRs by comparing the BS-seq data from the WT and *clsy3*-kd lines and correlate these regions with the sRNA clusters that are gained and lost in the *clsy3*-kd line. In the screen shots shown in Fig. 5I, for example, reductions in methylation look clear and it would be helpful to know how common such occurrences are throughout the genome.
- Please also add DMR tracks to all the screen shots where DNA methylation from the BS-seq data is presented.
- In Fig. S8D and E it is not clear what regions were used to make the metaplots and violin plot. Are these over the full set of *clsy3*-dep sRNA regions, both those that gain and lose sRNAs? If so, breaking these into those that gain and lose sRNAs and/or gain and lose methylation at DMRs might reveal the subset of regions where CLSY3 plays the largest role in regulating methylation levels. As presented, it looks like CLSY3 blocks hyper methylation on a much larger scale than it promotes methylation. However, this could be a large effect at a small number of loci that masks smaller reductions in methylation over a larger number of loci.
- Related to the point above, the authors mention that there is also hypermethylation in the *clsy3*-kd line and equate this with observations from *clsy3* mutants in Arabidopsis. However, the examples shown in Fig. 5J and Fig. S8B show hyper methylation at loci that lose sRNAs in the *clsy3*-kd line which is not consistent with redundancy with other CLSY members as proposed (lines 441-442). Correlating losses and gains in sRNAs separately

with DNA methylation levels and/or DMRs will more clearly demonstrate whether or not sRNA changes are linked to the observed CHH hypermethylation on a global scale.

- Given the strong reductions in sRNAs at siren loci, it's surprising there is no decrease in CHH methylation across all the siren loci in the *clsy3*-kd line (Fig. S12C). If the metaplot is split into the same three categories as the boxplots and heatmaps can a decrease be observed at least in the category 1 subset? If not, addressing this disconnect in the discussion would be helpful.

3. The claim that other histone marks are altered in the *clsy3*-kd line is poorly supported. From the immunostaining data in Fig. S9 it is difficult to conclude a redistribution of histone marks without a more quantitative assessment and genetic experimentation. As this is not central to the rest of the work, it should be removed.

4. The claim that the CLSY3 ChIP peaks correlate well with *clsy3*-dependent sRNA clusters is poorly supported. In Fig. 6F, two examples are shown with CLSY3 ChIP at loci with reduced sRNAs in the *clsy3*-kd and the text mentioned this is consistent across CLSY3 ChIP peaks. Please show a figure that demonstrates the correlation on a global scale or alter the conclusion as needed.

5. Based on the data presented in the figures and tables, it's difficult to assess what gene expression changes are statistically significant (i.e. are part of the DEG set from Fig. 7G) and also subject to epigenetic regulation via CLSY3 (i.e. are near *clsy3*-depenent siRNA clusters and DMRs identified in the *clsy3*-kd BS-seq data).

- Lines 517-519 states "Few genes which were upregulated in *clsy3*-kd, overlapped with CLSY3-dependent sRNA loci and their DNA methylation levels were also reduced as expected (Fig. 7J)." Please indicate the number of genes in this category and mark them in the supplemental table.
- Related to the above comment it remains hard to know from the figures and tables which siren adjacent genes, seed development genes, and imprinted genes are (1) DEGs, (2) near clsy3-dependent sRNA loci and (3) show reduced methylation in the clsy3-kd line. In the response to reviewer letter it is stated that "We observed a large number of them overlapping with DEGs (nearly 50 siren adjacent genes, 123 imprinted genes and 12 seed/reproductive traits related)". However, these numbers were not included in the revised manuscript and are not annotated in the supplementary tables. I suggest adding columns to the dataset S12, S13 and S14 tables to indicate which genes are a (1) DEGs in the *clsy3*-kd line, (2) are near *clsy3*-dependent sRNA loci and (3) show reduced methylation in the *clsy3*-kd line.
    - While expression values were added to several tables, without the associated p-values it remains unclear if the changes are statistically significant.

Minor comments:

Main Text

- In the main text (lines 354-55) states "We obtained a total of 7 transgenic plants with double amiRs (*clsy3*-kd2) (Fig. 4A, Supplemental Fig. S5D) and 8 plants with single amiR (*clsy3*-kd1) (Fig. 4A)." As there are some lines in these figures that are not KD lines, I

think adding the line #'s would be helpful. "We obtained a total of 7 transgenic plants with double amiRs (*clsy3*-kd2) (Fig. 4A and Supplemental Fig. S5D, lines 1 and 6-10) and 8 plants with single amiR (*clsy3*-kd1) (Fig. 4A, lines 1-8)."
- The line 457 reference to S4B should be S5B.
- Lines 4880490 state "However, unlike clsy3-kd, majority of the siren loci were downregulated in poliv-kd indicating that OsCLSYs regulate specific siren loci (Supplemental Fig. S11C-D and Supplemental Fig. S12A-B)." However the boxplots in these figures look very similar so its not clear how this data supports the majority of loci being downregulated in the poliv-kd but not the clsy3-kd. Please reword for clarity.

Figures

- 1B, the numbers in the Venn diagram don't match the genes lists in the supplementary tables. Table S2 has 3607 genes (3459 annotated) but Fig. 1B has 3476 genes (Blue circle). The text mentions 159 chromatin regulators and 635 imprinted genes, but Fig. 1B has 160 epigenetic genes (orange circle) and 668 imprinted genes (yellow circle). Please clarify.
- 1D, please add CLSY designations for the Zm and Os genes in addition to their CHR# and gene ID #.
- 1D, the Gene ID mentioned in the text (Os02g0650800) is absent. Other genes don't match either, for example, in the text CHR722 is Os07g0692600 and CHR742 is Os05g0392400. Please cross check all the genes between the text and the tree. Also, as AtCLSY3 and AtCLSY4 group together, what is the evidence that Os02g0650800 is the rice CLSY3 and not CLSY4?
- 1C, the legend is missing that EM=embryo.
- 5D, the yellow color makes the *clsy3-kd* sRNA track hard to see. Please change to a darker color.
- 5I and J, as dataset S6 is already in bed format, please add a track to these screen shots that shows the regions corresponding to *clsy3*-dependent sRNA clusters.
- 6B, is IP a ChIP with non-transgenic material as a neg control? Please clarify what this sample corresponds to in the figure and methods.
- 6B, why is there only one blue and one green line. Is this the average of the repeats? If so, please show them independently and in matching colors to the labels.
- B6, what is "center" referring to? It doesn't seem to be the center of the peaks based on the shape of the enriched region and in the legend genomic features are mentioned. Please clarify.
- 6D, what is H3? A histone 3 chIP?
- 6D, 6F and others, please add (1) tracks for the IP controls to all the ChIP screen shot figures to allow the enrichment over background to be assessed (2) tracks showing the regions corresponding to *clsy3*-dependent sRNA clusters, and (3) tracks showing the CLSY3-ChIP peaks.
- S4A, please add that VC=vector control and explain the experiment. Was an empty vector transformed into the plant or is this a digestion of the plasmid run on the gel as a probe control?

- S4A, please show the Hph-1kb probe location on the map and use a consistent name for this drug resistance element (HygR vs hgh).
- S4A, please add to the map a distal HindIII site of variable distance depending on the genomic insertion site. Or mention this explicitly in the text so it is clear that variable sizes are expected on the southern blot.
- S4H, please specify what is being shown on the image on the right.
- S5A, see comments for S4A.
- S5D, mark line 3 in brown.
- S7G/S8A/S8B, as dataset S6 is already in bed format, please add a track to these screen shots that shows the regions corresponding to clsy3-dependent sRNA clusters.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors made efforts to address my comments and added relevant data that improve this manuscript. I think the findings reported are exciting and even though several questions remain, it is an important contribution to our current knowledge. I have a couple of comments that however do not require more experimental work.

Author's reply: Thank you for your encouraging comments and very useful suggestions.

1. **Reviewer's comment -** Line 434: Based on the genome-wide DNA methylation data, there are apparently many more loci that gain rather than lose DNA methylation. It would be important to show how many loci lose and how many gain DNA methylation.
   Author's reply: Thank you for pointing out this missing information. We found around 21653 CLSY3-dependent sRNA loci (lost sRNAs) and 2620 CLSY3-dependent loci that gained sRNAs (Fig. 4F). As per the suggestion from Reviewer 3, we now analysed DNA methylation levels in CLSY3 dependent sRNA loci. Among the 13259 CLSY3-dependent loci that lost sRNAs (excluding loci that have zero-CHH DNA methylation levels), DNA methylation increased in 7257 loci and decreased in 4953 loci. Among the 1800 CLSY3-dependent sRNA-gained loci, 898 loci showed increased DNA methylation while 723 loci showed decrease in DNA methylation. We plotted the DNA methylation levels at these loci as box plots (please see the figure attached below). We have incorporated these details in the revised manuscript (Supplemental Fig. S9F).

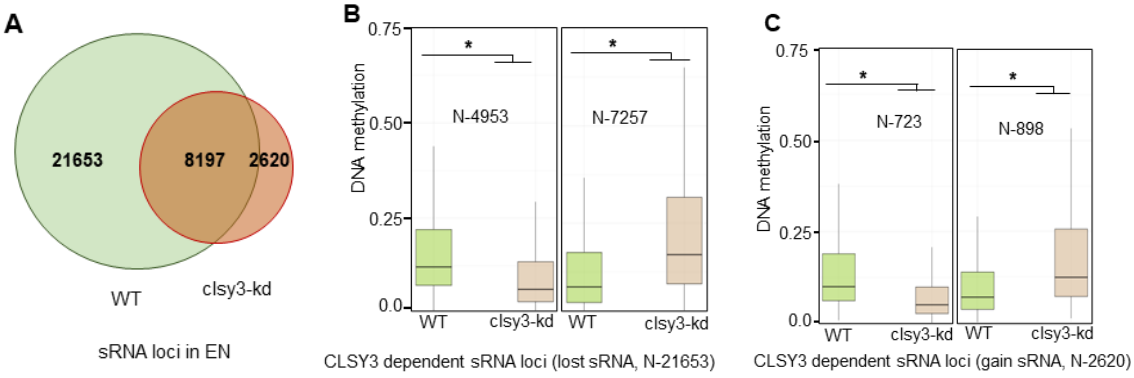


Fig.

(A) Venn diagram showing sRNA status of CLSY3 dependent sRNA loci.

(B) Box plots showing DNA methylation status of CLSY3 dependent sRNA loci (lost sRNA).

(C) Box plots showing DNA methylation status of CLSY3 dependent sRNA loci (gain).

2. **Reviewer's comment -** L514: To conclude that there is a negative correlation, the authors would need to calculate the correlation coefficient.
   Author's reply: Thank you for this suggestion. We have calculated corelation coefficient between sRNA (kd-WT) and RNA (kd-WT) values. We found Pearson's correlation coefficient as -0.96. The value shown here is a strong negative correlation between sRNAs and RNA expression in clsy3-kd endosperm. We also calculated F-value for sRNA and RNA-seq expression which showed $F_{1,12687} = 5.06$, P-value < 0.05. This information is included in the revised script.
3. **Reviewer's comment -** Line 340: Based on the provides pictures in Figure S4H no conclusion on the endosperm can be drawn. Please rephrase.
   Author's reply: Thank you for this comment. We rephrased the sentence.
4. **Reviewer's comment -** Line 517ff: If only few of the upregulated genes lose DNA methylation, the statement "This observation suggested that the genes were likely regulated by CLSY3-dependent DNA methylation, specific to endosperm." is not correct and should be toned down. Based on all data the connection between CLASSY3-dependent siRNAs and DNA methylation remains unclear. This should be stated to avoid the erroneous impression that there is a clear correlation.
   Author's reply: Thank you for pointing out this. We agree. We have toned down those claims in the revised script.
5. **Reviewer's comment -** L563ff: Since the authors generated whole genome bs data, they should use this data for all analyses rather than semi-quantitative chop PCRs or locus-specific bs analyses. They should quantify CHH methylation for all imprinted genes and show the quantified data rather than screenshots of the genome browser.
   Author's reply: Thank you for this helpful suggestion. In our revised manuscript, in supplementary figures, we also included DNA methylation tracks for the chop qPCR tested loci (Fig.7G, Supplemental Fig. S15B and C). We did a quantification of CHH methylation of all published rice imprinted genes (N-635). However, if all of them are taken together, CHH methylation was unchanged between clsy3-kd and WT endosperm (Supplemental Fig. S14D), while 9 loci clearly showed hypomethylation in clsy3-kd (Fig. 7H and Supplemental Fig. S15D).

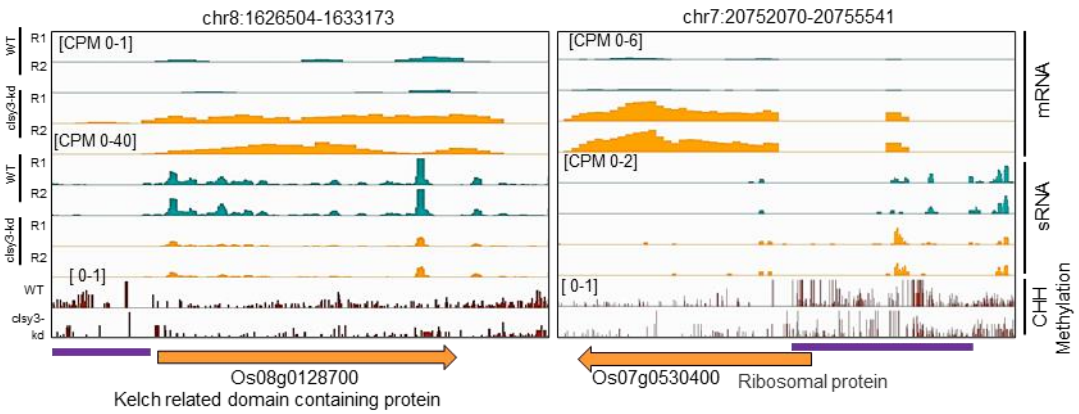


Fig.

IGV screen shots showing sRNA, RNA and CHH methylation status of two selected imprinted genes in WT and clsy3-kd EN.

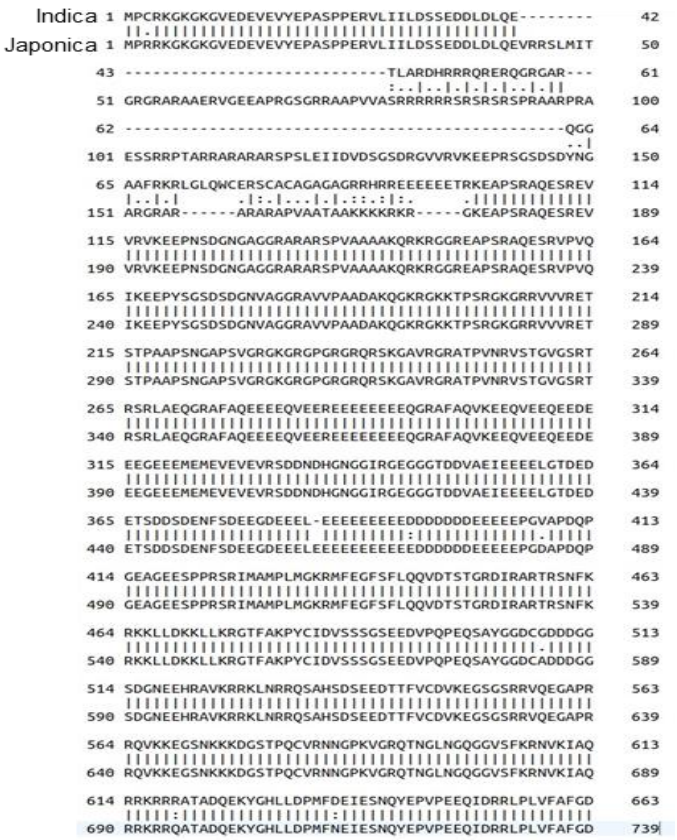
6. **Reviewer's comment -** L577: Based on the data I would revise the conclusion that some (not many) imprinted genes are regulated by OsCLSY3 through sRNA-directed DNA methylation.
   Author's reply: We agree. Although there were many imprinted genes mis-expressed in clsy3-kd endosperm, only some imprinted genes were directly regulated by CLSY3 dependent sRNA-directed DNA methylation. We have modified the sentence accordingly.
7. **Reviewer's comment -** Regarding comment 10; I remain convinced that Figures 1 and 2 do not add much to the novelty of this manuscript, but rather make it unnecessary lengthy.
   Author's reply: Thank you for this suggestion. As per the suggestion, we moved Figure 2 to the supplementary section. We felt Figure 1 is key and must be in the main figure section.
8. **Reviewer's comment -** L51: Instead of adding the waterlily example, I suggest to rather write "typically diploid central cell".
   Author's reply: Thank you for this suggestion. We have changed this sentence in the revised manuscript.
9. **Reviewer's comment -** L243: remove the "Surprising". With all what we know how MEGs are regulated, finding a MITE in its promoter is not surprising.
   Author's reply: Thank you for this suggestion. We corrected the sentence accordingly.
10. **Reviewer's comment -** Change poliv to nrpd1
    Author's reply: Thank you for this suggestion. We have incorporated this in the revised manuscript.
11. **Reviewer's comment -** Figure S13A: The gene identifiers of the MADS-box genes should be included.
    Author's reply: Thank you. Since it is not possible to provide the gene IDs in the heatmap in Fig. S13A, we had provided in column 4 in Supplemental dataset S13.
12. **Reviewer's comment -** L560: Instead of "However" the right phrase would rather be "Out of those, 9 genes..."
    Author's reply: Thank you for this suggestion. We have rephrased the sentence.
13. **Reviewer's comment -** L563: should read "was largely unchanged"
    Author's reply: Thank you for the correction. We rephrased the sentence in the revised script.
    Thank you for all your suggestions and corrections. These have definitely improved manuscript.

Reviewer #2 (Remarks to the Author):

One very related paper was published on Plant Physiology (kiad 624, 2023, November 22) about one CLSY family member, OsCLSY4 named in this study. There are a lot of contrast results. The authors might carefully reexamine the experimental results with the published data.

Author's reply: Thank you for pointing two contrasting points in the study that the Reviewer has cited (paper was cited and discussed by us in the previous version of revised manuscript, please note, this publication came out while our paper was under review). We do find these differences interesting, and we have pointed out how several factors might contribute to these differences:

1. In Figure 3A (Plant Physiology paper) showed absence of expression of any CLSYs in *japonica* rice endosperm. Please note that this was an RT-qPCR analysis and this will not rule out low expression of genes while publicly available RNA-seq datasets also showed adequate expression of CLSY3 and CLSY4 in endosperm (Supplemental Fig. S2B). Please also note that it is not easy to extract good quality RNA from polysaccharide-rich endosperm. In fact, we used a modified method to get RNA (please see methods in previous and revised version of manuscript). We also failed to see expression of CLSY3 in *indica* rice while using the unaltered Trizol method. There is a clear difference between *indica* and *japonica* rice in many traits including habitat, growth pattern, anatomy, seed phenotypes and gene expression in several key genes. Since we have not worked with *japonica* rice, we are unable to exclude the possibility of variations in CLSY expression between these two subspecies.
2. In Figure 4I and 4J (plant Physiology paper), authors showed seed setting rate of CHR740/OsCLSY3/FEL1 knockout was unchanged compare to control. Here again, we are unable to exclude the possibility of differences between *japonica* and *indica* rice. In fact, we saw clear changes in CLSY3 amino acid sequences between *japonica* and *indica* lines in the N terminal end (more than 50% variation in the first 200 amino acids). Overall *indica* CLSY3 was only 90.4% similar to *japonica* CLSY3. Most importantly, this variable region contains Intrinsically disordered region (IDR, a domain implicated in phase separation), and in our KO lines, a peptide of 190 amino acids could be translated and this might code for IDR domain. The *japonica* KO reported in the Plant Phys paper can code for only 120 amino acids. Please note that in our knockdown (kd) lines, we did not see sterility issues or differences in vegetative tissues (similar to Plant Phys paper). Only difference we saw were in endosperm size (seed size), a phenotype that was not counted in Plant Phys paper. Please also note that our molecular analysis was performed in kd lines.



OsCLSY3

Major Concerns

Reviewer's comment - Mutation in FEM2/CHR742 caused multiple developmental defects including failure of seed development. The rice plants with simultaneous mutation in CHR740/OsCLSY3/FEL1 and CHR722/FEL2, however, were normal for growth and development including reproductive development. In this manuscript, mutation in CHR740/OsCLSY3/FEL1 resulted in sterility because of pollen and endosperm development. There is no expression for OsCLSY3 in anther (Figure 1d). Why osclsy3 have pollen developmental defects? Given osclsy4 was generated in this study, what is development and DNA methylation phenotype of osclsy4?

Author's reply: Thank you for mentioning this. Please see our detailed reply above. In Figure 1d, a Z-score heatmap which is relative quantification with reference to other tissues. For absolute quantification, RT-qPCR would give a better clarity. We observed that OsCLSY3 is not expressed in early anther developmental stages but it expressed in later stages of anther development. The expression of OsCLSYs were also checked in unfertilized ovule and anther (1 day before anthesis) by RT-qPCR. *OsCLSY3* and *OsCLSY4* are well expressed in both these tissues. The previous published paper (Xu et al. 2023) also documented that *japonica OsCLSY3* is expressed in S12 stage of the stamen. Please see Fig. A attached below, and this shows extent of CLSY3 expression in *indica* anther tissues.

In our study, we also generated osclsy4-kd plants, phenotypes of which matched phenotypes as showed in Xu et al. 2023. In fem2 (*clsy4*) mutant plants seed setting rate, height and tiller numbers were affected. We also saw all those major phenotypes in our clsy4-kd plants. Please see Fig. B and C attached below reinforcing these points.
We did not perform the whole genome Bisulfite sequencing in clsy4-kd plants. Xu et al 2023 reported that they found CHH hypo-DMRs in fem2 mutants which were mainly associated with miniature inverted repeat TEs (MITEs). In our study, we also showed DNA methylation and sRNAs reduced in selected MITEs which are also present in the CLSY3 promoter (Fig. 2F). The clsy4-kd plant which generated in this study showed same development and DNA methylation related phenotypes (MITEs at CLSY3 promoter) as shown in Xu et al (Fig.2F and supplemental Fig. S4C). Since our paper is not about CLSY4, performing whole genome DNA methylation analysis in our cls4-kd lines is beyond the scope of the study.

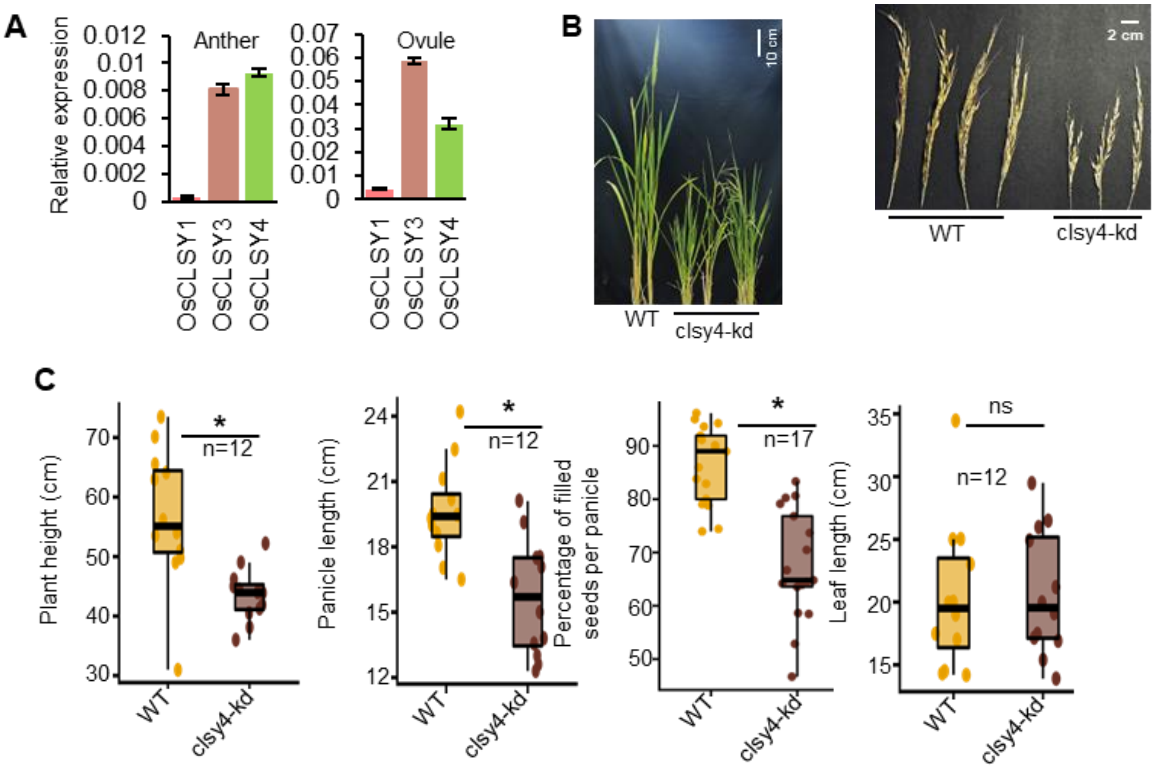


Fig.

(A) Bar plots showing relative expression of *OsCLSY3* in anther and ovule (1 day before pollination).
(B) Images showing phenotypes of clsy4-kd plants.
(C) Box plots showing plant height, panicle length, percentage of filled seeds per panicle and leaf length of the clsy4-kd plants compare to WT.

Reviewer's comment - In this manuscript, the authors claimed that DNA methylation at MITE control the expression of OsCLSY3. Beside the MITE, there is one long transposon in the promoter of OsCLSY3/FEL1. In addition,

the dense DNA methylation on this long transposon (Figure 3b, Plant Physiology) might be important for regulating it expression. The OsCLSY3::GUS transgenic plants showed signal in anther (Figure 1h), which is contrast to no expression in anther for endogenous gene in Figure 1c. The inconsistency in expression between transgenic and endogenous OsCLSY3 suggest short promoter used in this study is not enough to reflect it real expression pattern.

Author's reply: This is indeed interesting. We also observed the presence of a long LTR TE in the promoter of *OsCLSY3*/FEL1. However, this long transposon was present more than 2 Kb upstream of the *OsCLSY3*/FEL1 Transcription start site (TSS). The methylation on this TE might influence the *OsCLSY3* expression but the long LTR is far away from the conventional promoter. More importantly, we did not observe reduced methylation in this region unlike what we saw easily in MITES in endosperm when compared to embryo tissue. Please see the Fig. A attached bellow.

Thank you for pointing out the issue with expression of *OsCLSY3* in anther. We have explained this in detail above. Briefly, our data shows that OsCLSY3 is not expressed in early anther stages (Fig. 1C) but expressed in later stages. The GUS expression also verified this observation. The CLSY3 P: GUS not expressed in the early stages of the anther with its expression restricted at the ovule. Please see the attached figure bellow (Fig. B below).

We agree with you that expression of transgenic and endogenous *OsCLSY3* have some difference that we explained in the previous version of the revised script. Promoter length might contribute to the expression difference between transgenic and endogenous *OsCLSY3*, however, since our hypothesis was to find the role of those MITE TEs in context of *OsCLSY3* expression, we did not explore taking multiple lengths of promoters for promoter analysis. The tissue-specific methylation of MITE TEs is clearly inversely correlated with the gene expression.

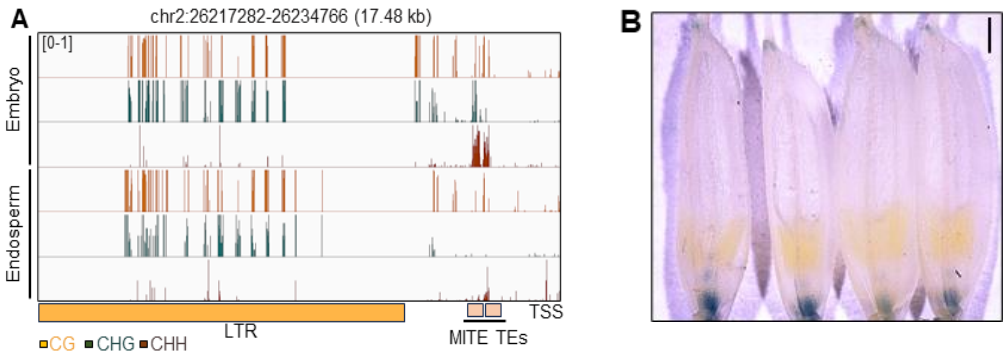


Fig.

(A) IGV screenshot showing DNA methylation change in embryo and endosperm at ~17 kb *OsCLSY3* promoter region.
(B) Images showing GUS expression in the rice spikelets (before emerged) (SB- 2mm).

Reviewer's comment - OsCLSY3 is maternally expressed imprinted gene in endosperm where OsCLSY3 controlled siRNA production and imprinted genes. To confirm this main conclusion, the reciprocal cross between OsCLSY3-kd and WT will provide the maternal and paternal effect of OsCLSY3 on the production of siRNAs, and thus the expression of imprinted genes.

Author's reply: Thank you for this suggestion. In this study, we observed mis expression sRNAs, imprinted genes, seed development genes in clsy3-kd plants endosperm. Our claim is that adequate levels of *OsCLSY3* in endosperm is needed for proper expression of sRNAs and genes. We agree with you that it would be a wonderful addition if we can also resolve the maternal and paternal effects of *OsCLSY3* further. However, due to technical difficulties such as generation of amiR-resistant CLSY3 complementation in clsy3-kd lines, we cannot perform the reciprocal cross in this study with this set of transgenic plants. In case of clsy3-kd, artificial miRNA will target newly introduced paternal and maternal allele equally. Due to that knockdown lines are inappropriate for reciprocal-crossing. For the reciprocal cross, *KO* lines would be better but due to seed viability issues, we are unable to use it for this experiment. We hope the Reviewer agrees with this challenge addressing which is beyond the goals of the current study. We hope to study this aspect further in the coming years.

Reviewer's comment - To examine siRNA level and gene transcription in endosperm, the dissection of endosperm often has seed coat contamination. The authors need to exclude this possibility.

Author's reply: Thank you for pointing out this point. In our study, we took the all the preventions to avoid the seed coat contaminations in our endosperm isolation. The seed coat is green in colour in rice in the stage of seed that were selected. We can clearly separate the seed coat tissue from the endosperm as shown in Fig. A attached below. We also verified and found that previously studied green tissue-specific marker genes such as *RIBONUCLEASE 4* (Os09g0537700), *OsPLT1*(Os03g0197100), *OsAKR2* (Os07g0142900) (Li et al. 2018) were not expressed in our endosperm transcriptomes. Please see the Fig. B and Fig. C attached below. All the figures collectively suggested that our dissection of endosperm is proper and free from the seed coat related contaminations.

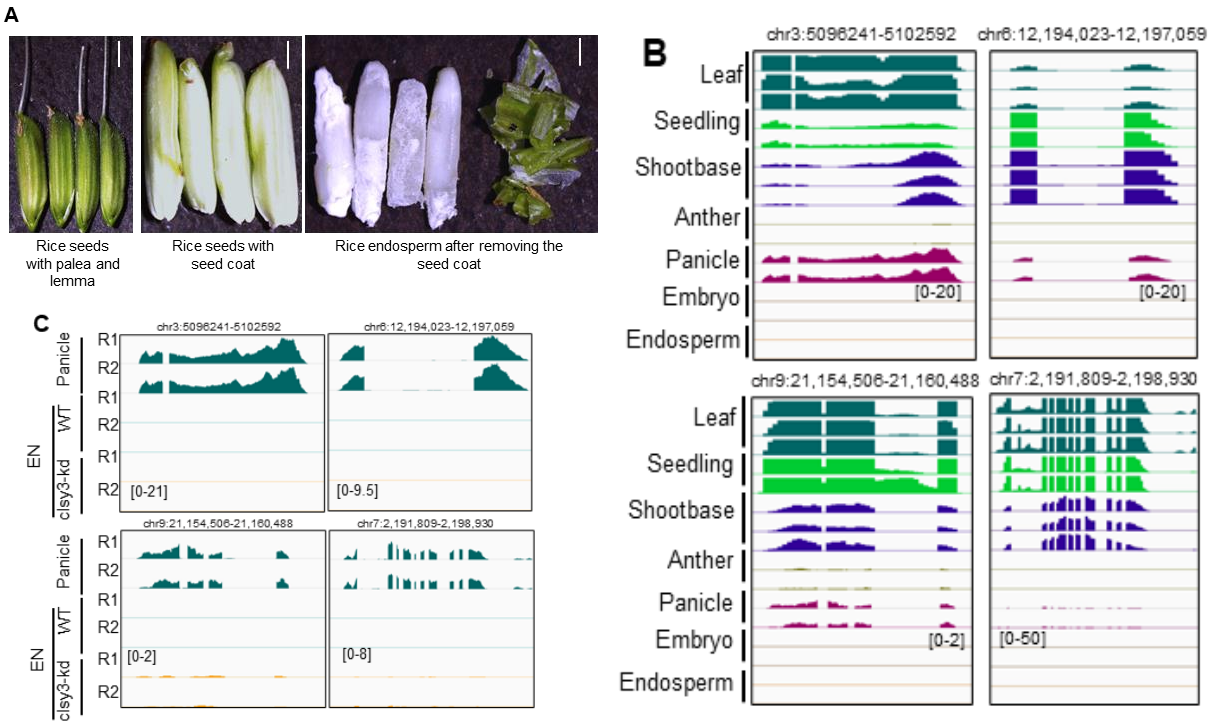


Fig.

(A) Image showing endosperm isolation process in this study (SB-2 mm).
(B) IGV screenshots showing expression of green-tissue specific genes in the different rice tissues.
(C) IGV showing green-tissue specific genes in panicle (WT) and endosperm tissues (WT, clsy3-kd).

We thank the Reviewer for all the suggestions and corrections. We found them very useful and constructive.

Reviewer #3 (Remarks to the Author):

In the revised manuscript the authors have addressed many of my previous concerns and have added experiments and analyses that better support their main claims. Based on these changes, I remain enthusiastic about the significance of the work presented. However increased transparency on which genes are likely regulated directly by CLSY3 via the RdDM pathway would better define the scope of CLSY3's role in regulating sRNAs and controlling gene expression and imprinting during rice reproduction. As detailed below, such evidence can be provided by additional analysis of already available data and will increase the impact of the presented work.

Author's reply: Thank you for your encouraging comments, critical suggestions and detailed review of the manuscript. We have performed newer analysis to enhance the quality of the manuscript.

Major comments:

Reviewer's comment - The identification and reporting of the clsy3-dependent sRNA clusters remains unclear.
The text states "To identify CLSY3-dependent sRNA loci, we quantified number of loci present in WT and clsy3-kd endosperm by ShortStack analysis (Supplemental dataset S4 and S5). We observed that around 70% of sRNA loci lost 23-24nt sRNAs in clsy3-kd, when compared to WT in a bedtools based analysis and these were termed CLSY3-dependent sRNA loci (Fig.5F and Supplemental dataset S6)." The cluster numbers in dataset S4 and S5 match the Venn diagram, which indicates there are 20,947 clusters lost in the clsy3-kd and 1,962 clusters gained for a total of 22,909 clsy3-dependent sRNA loci. However, dataset S6 lists 21,653 clusters. Please clarify. Were some adjacent clusters merged?

Author's reply: Thank you for pointing out this. The issue with such analysis is lack of a tool that looks across data to define and compare loci. Most of the existing tools define region of interest with their own specific parameters, and hence while overlaying another information, the overlapping regions as well as subdivided regions confuse with different numbers. Here, some adjacent clusters are merged. In our previous version of the manuscript, we used bedtools intersect -wa, -wb and -v commands to calculate the number of intersecting loci. The bedtools intersect is sensitive to order of the inputs and presence of overlapping loci in two sets that are being compared. To alleviate these deviations, we have used 'intervene', that merges the double overlaps between the bed files and gives a merged number which is biologically meaningful. In our revised manuscript, we generated venn diagram with "Intervene Shiny app" (Khan and Mathelier 2017) to generate venn diagram to alleviate this mention deviations. The new venn diagram indicated that there are 21653 clusters lost in the clsy3-kd and 2620 clusters gained sRNAs in clsy3-kd endosperm tissues. In our current version of the manuscript, we considered the 24273 clusters as CLSY3-dependent sRNA loci. We apologise for creating this confusion in the previous version of the manuscript. We have updated the Supplemental dataset S6 and Fig. 4F.

Reviewer's comment - The effects of the clsy3-kd on DNA methylation patterns are not well explored leaving it unclear how much this CLSY contributes to epigenetic regulation in the endosperm.

a. Please identify hyper and hypo DMRs by comparing the BS-seq data from the WT and clsy3-kd lines and correlate these regions with the sRNA clusters that are gained and lost in the clsy3-kd line. In the screen shots shown in Fig. 5I, for example, reductions in methylation look clear and it would be helpful to know how common such occurrences are throughout the genome.

Author's reply: Thank you for this suggestion. As previously mentioned, we found 21653 loci which lost sRNA and 2620 loci that gained sRNA in clsy3-kd endosperm tissues. We also analysed DNA methylation status in those loci. Among these, DNA methylation was clearly decreased in 4953 loci, while it increased in 7257 loci. Rest of the loci did not show difference in DNA methylation, or the changes were marginal due to insufficient

methylatable/methylated sites. The data incorporated into the revised manuscript (Supplemental Fig. S9F-H and Supplemental dataset S7).

b. Please also add DMR tracks to all the screen shots where DNA methylation from the BSseq data is presented.

Author's reply: Thank you for this suggestion. Reviewer 1 suggested us to use whole genome bisulfite data in the place of targeted BS-PCR data. We have removed those BS-PCR data in our revised manuscript.

c. In Fig. S8D and E it is not clear what regions were used to make the metaplots and violin plot. Are these over the full set of clsy3-dep sRNA regions, both those that gain and lose sRNAs? If so, breaking these into those that gain and lose sRNAs and/or gain and lose methylation at DMRs might reveal the subset of regions where CLSY3 plays the largest role in regulating methylation levels. As presented, it looks like CLSY3 blocks hyper methylation on a much larger scale than it promotes methylation. However, this could be a large effect at a small number of loci that masks smaller reductions in methylation over a larger number of loci.

Author's reply: Thank you for this excellent suggestion. Among the 13259 CLSY3-dependent loci that lost sRNAs, DNA methylation increased in 7257 loci and decreased in 4953 loci. Among the 1800 CLSY3-dependent sRNA-gained loci, 898 loci showed increased DNA methylation while 723 loci showed decrease in DNA methylation. We plotted those loci as box plots in the revised script separately. Please see the below figure (A and B). We have incorporated these details in the revised manuscript (Supplemental Fig. S9F and H).

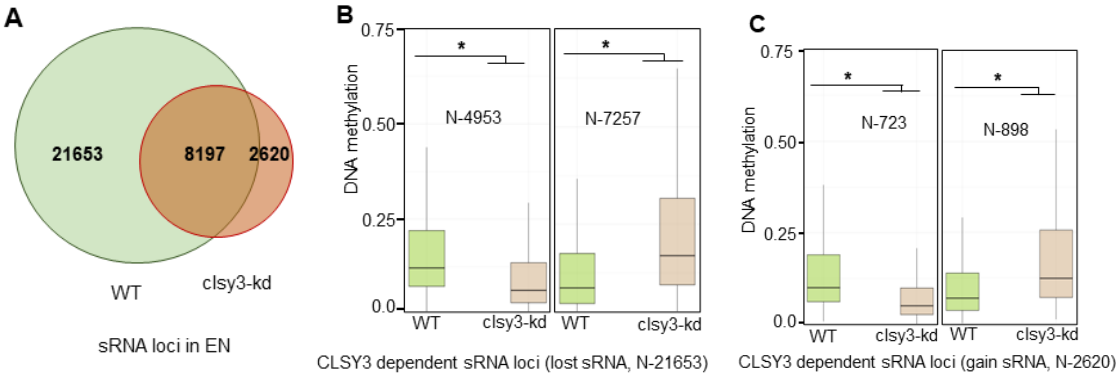


Fig.

(A) Venn diagram showing sRNA status of CLSY3 dependent sRNA loci.

(B) Box plots showing DNA methylation status of CLSY3 dependent sRNA loci (lost sRNA).

(C) Box plots showing DNA methylation status of CLSY3 dependent sRNA loci (gain).

d. Related to the point above, the authors mention that there is also hypermethylation in the clsy3-kd line and equate this with observations from clsy3 mutants in Arabidopsis. However, the examples shown in Fig. 5J and Fig. S8B show hyper methylation at loci that lose sRNAs in the clsy3-kd line which is not consistent with redundancy with other CLSY members as proposed (lines 441-442). Correlating losses and gains in sRNAs separately with DNA methylation levels and/or DMRs will more clearly demonstrate whether or not sRNA changes are linked to the observed CHH hypermethylation on a global scale.

Author's reply: Thank you for this comment. As mentioned, we observed that among many CLSY3 dependent sRNA loci where sRNA levels decreased, there was hypermethylation. Also, among the loci that gained sRNAs in clsy3-kd, some did not gain DNA methylation. Since hypermethylation was recorded in single mutants in *Arabidopsis*, we thought it is important to connect to the existing literature. Yang et al. also suggested demethylation in loci that are maintained by CLSYs or involvement of other sRNA-independent mechanisms in maintaining DNA methylation (Yang et al. 2018). In Yang et al. it is not clear what are the levels of sRNAs in hypermethylated loci. We have now provided number of loci gaining and losing DNA methylation in the revised script.

e. Given the strong reductions in sRNAs at siren loci, it's surprising there is no decrease in CHH methylation across all the siren loci in the clsy3-kd line (Fig. S12C). If the metaplot is split into the same three categories as the boxplots and heatmaps can a decrease be observed at least in the category 1 subset? If not, addressing this disconnect in the discussion would be helpful.

Author's reply: We agree. Thank you for this helpful suggestion. After removing low confident loci (excluding loci that have zero-CHH DNA methylation levels), out of 722 siren loci, DNA methylation was decreased in 248 loci while it increased in 317 loci. There were 157 loci in which DNA methylation did not change.

As per the suggestion from the Reviewer, we have analysed category-wise methylation status of siren loci. In category-1, among 316 siren loci, DNA methylation increased in 134 loci and decreased in 94 loci. In the category-2, DNA methylation decreased in 146 loci and it increased in 176 loci. All these points are included in the revised discussion and figure (Supplemental Fig. S12D).

Reviewer's comment - The claim that other histone marks are altered in the clsy3-kd line is poorly supported. From the immunostaining data in Fig. S9 it is difficult to conclude a redistribution of histone marks without a more quantitative assessment and genetic experimentation. As this is not central to the rest of the work, it should be removed.

Author's reply: Thank you for this suggestion. We agree with this. We have removed the immunostaining data in the revised manuscript.

Reviewer's comment - The claim that the CLSY3 ChIP peaks correlate well with clsy3-dependent sRNA clusters is poorly supported. In Fig. 6F, two examples are shown with CLSY3 ChIP at loci with reduced sRNAs in the clsy3-kd and the text mentioned this is consistent across CLSY3 ChIP peaks. Please show a figure that demonstrates the correlation on a global scale or alter the conclusion as needed.

Author's reply: Thank you for this suggestion. In our global analysis, we extended ChIP-seq peaks to 3 kb on either side and overlapped with CLSY3 dependent sRNA loci (that lost sRNAs). We found that, among 7115 ChIP peaks, 1398 peaks overlapped with CLSY3 dependent sRNA loci. We have included this detail in the revised manuscript. We also found that 23-24nt sRNAs were decreased in the 1398 extended ChIPseq peak-containing regions. This figure is now included in the revised manuscript (Fig.5F and G).

Reviewer's comment - Based on the data presented in the figures and tables, it's difficult to assess what gene expression changes are statistically significant (i.e. are part of the DEG set from Fig. 7G) and also subject to epigenetic regulation via CLSY3 (i.e. are near clsy3-depenent siRNA clusters and DMRs identified in the clsy3-kd BS-seq data).

a. Lines 517-519 states "Few genes which were upregulated in clsy3-kd, overlapped with CLSY3-dependent sRNA loci and their DNA methylation levels were also reduced as expected (Fig. 7J)." Please indicate the number of genes in this category and mark them in the supplemental table.

b. Related to the above comment it remains hard to know from the figures and tables which siren adjacent genes, seed development genes, and imprinted genes are (1) DEGs, (2) near clsy3-dependent sRNA loci and (3) show reduced methylation in the clsy3-kd line. In the response to reviewer letter it is stated that "We observed a large number of them overlapping with DEGs (nearly 50 siren adjacent genes, 123 imprinted genes and 12 seed/reproductive traits related)". However, these numbers were not included in the revised manuscript and are not annotated in the supplementary tables. I suggest adding columns to the dataset S12, S13 and S14 tables to indicate which genes are a (1) DEGs in the clsy3-kd line, (2) are near clsy3-dependent sRNA loci and (3) show reduced methylation in the clsy3-kd line.

While expression values were added to several tables, without the associated p-values it remains unclear if the changes are statistically significant.

Author's reply: Thank you for this comment. We have included all the details of the DEGs and statistics in the revised manuscript. As commented by Reviewer 1, we kept this detail intact in the revised manuscript. However, we have added list of adjacent genes (24 genes, marked in Supplemental dataset S13) wherever there is clear correlation with sRNAs, DNA methylation and DEGs.

Minor comments:

Reviewer's comment - In the main text (lines 354-55) states "We obtained a total of 7 transgenic plants with double amiRs (clsy3-kd2) (Fig. 4A, Supplemental Fig. S5D) and 8 plants with single amiR (clsy3-kd1) (Fig. 4A)." As there are some lines in these figures that are not KD lines, I think adding the line #'s would be helpful. "We obtained a total of 7 transgenic plants with double amiRs (clsy3-kd2) (Fig. 4A and Supplemental Fig. S5D, lines 1 and 6-10) and 8 plants with single amiR (clsy3-kd1) (Fig. 4A, lines 1-8)."

Author's reply: Thank you for these suggestions. We included these suggestions in the revised manuscript.

Reviewer's comment - The line 457 reference to S4B should be S5B.

Author's reply: Thank you. Apologies for this mistake.

Reviewer's comment- Lines 4880490 state "However, unlike clsy3-kd, majority of the siren loci were downregulated in poliv-kd indicating that OsCLSYs regulate specific siren loci (Supplemental Fig. S11C-D and Supplemental Fig. S12A-B)." However, the boxplots in these figures look very similar so its not clear how this data supports the majority of loci being downregulated in the poliv-kd but not the clsy3-kd. Please reword for clarity.

Author's reply: Thank you for pointing out this mistake. We have modified the sentence.

Figures

- 1B, the numbers in the Venn diagram don't match the genes lists in the supplementary tables. Table S2 has 3607 genes (3459 annotated) but Fig. 1B has 3476 genes (Blue circle). The text mentions 159 chromatin regulators and 635 imprinted genes, but Fig. 1B has 160 epigenetic genes (orange circle) and 668 imprinted genes (yellow circle). Please clarify.

  Author's reply: Thank you pointing out these errors. We have corrected the sentence with corrected numbers. This happened due to errors in IDs (from RABDB and MSU databases) that we have corrected.

- 1D, please add CLSY designations for the Zm and Os genes in addition to their CHR# and gene ID #.

  Author's reply: Thank you. We have included this detail.

- 1D, the Gene ID mentioned in the text (Os02g0650800) is absent. Other genes don't match either, for example, in the text CHR722 is Os07g0692600 and CHR742 is Os05g0392400. Please cross check all the genes between the text and the tree. Also, as AtCLSY3 and AtCLSY4 group together, what is the evidence that Os02g0650800 is the rice CLSY3 and not CLSY4?

Author's reply: Thank you. This is due to differences in RAPD ID and MSU ID. We have now included both IDs in the revised manuscript text and Supplemental table S2. Os02g0650800 is closest to AtCLSY3 as mentioned in text. The Os02g0650800 had 29.3% amino acid identity, and 44.5% amino acid similarity with AtCLSY3. The Os02g0650800 had 28.8% amino acid identity, and 41.8% amino acid similarity with AtCLSY4.

- 1C, the legend is missing that EM=embryo.

  Author's reply: Thank you. EM=embryo was already defined in 1A legend.

- 5D, the yellow color makes the clsy3-kd sRNA track hard to see. Please change to a darker color

  Author's reply: Thank you. We have changed the color.

- 5I and J, as dataset S6 is already in bed format, please add a track to these screen shots that shows the regions corresponding to clsy3-dependent sRNA clusters.

  Author's reply: Thank you. We have included this detail in IGV screenshots.

- 6B, is IP a ChIP with non-transgenic material as a neg control? Please clarify what this sample corresponds to in the figure and methods.

  Author's reply: Negative control here is sheared DNA from wild type non-transgenic line. We have included this detail in the legend.

- 6B, why is there only one blue and one green line. Is this the average of the repeats? If so, please show them independently and in matching colors to the labels.

  Author's reply: Thank you. Two replicates are identical and have merged. This detail has been mentioned in the revised script.

- B6, what is "center" referring to? It doesn't seem to be the center of the peaks based on the shape of the enriched region and in the legend genomic features are mentioned. Please clarify.

  Author's reply: Thank you so much. There was a mistake in reference point and apologies for that. We have corrected the mistake.

- 6D, what is H3? A histone 3 chIP?

  Author's reply: Thank you. Yes, it was H3. Now we have removed this track and used only input for better clarity.

- 6D, 6F and others, please add (1) tracks for the IP controls to all the ChIP screen shot figures to allow the enrichment over background to be assessed (2) tracks showing the regions corresponding to clsy3-dependent sRNA clusters, and (3) tracks showing the CLSY3-ChIP peaks.

  Author's reply: Thank you. We have included this detail in Fig. 6F. However, since the objective in Fig.6D was only to show how ChIP with two different tags resulted in identical binding, we did not include the sRNA track. In Fig.6F, we included all three tracks.

- S4A, please add that VC=vector control and explain the experiment. Was an empty vector transformed into the plant or is this a digestion of the plasmid run on the gel as a probe control?

  Author's reply: Thank you. This is a transgenic plant with just vector transformed without any insert. We have included this detail.

- S4A, please show the Hph-1kb probe location on the map and use a consistent name for this drug resistance element (HygR vs hgh).

  Author's reply: Thank you. We have included this detail.

- S4A, please add to the map a distal HindIII site of variable distance depending on the genomic insertion site. Or mention this explicitly in the text so it is clear that variable sizes are expected on the southern blot.

  Author's reply: Thank you. We have included this detail in the revised figure.

- S4H, please specify what is being shown on the image on the right.

  Author's reply: Thank you. We have included a better image showing endosperm abnormalities in *clsy3-KO*.

- S5A, see comments for S4A.

  Author's reply: Thank you. We have included this additional detail.

- S5D, mark line 3 in brown.

  Author's reply: Thank you so much. We have included this additional detail.

- S7G/S8A/S8B, as dataset S6 is already in bed format, please add a track to these screen shots that shows the regions corresponding to clsy3-dependent sRNA clusters.

Author's reply: Thank you. We have included sRNA tracks in these panels.

Thank you for these important suggestions and corrections. We have incorporated all these changes in the revised script. Thank you very much, this is very helpful.

We thank all the three Reviewers for carefully going through the manuscript and for offering excellent suggestions and comments. We also noted that all three Reviewers spent lot of time patiently to enhance the quality of this manuscript. We are indebted to them for their efforts.

References:

Khan, Aziz, and Anthony Mathelier. 2017. "Intervene: A Tool for Intersection and Visualization of Multiple Gene or Genomic Region Sets." *BMC Bioinformatics* 18 (1). https://doi.org/10.1186/s12859-017-1708-7.

Li, Hao, Juan Li, Rongfang Xu, Ruiying Qin, Fengshun Song, Li Li, Pengcheng Wei, and Jianbo Yang. 2018. "Isolation of Five Rice Nonendosperm Tissue-Expressed Promoters and Evaluation of Their Activities in Transgenic Rice." *Plant Biotechnology Journal* 16 (6): 1138–47.

Xu, Dachao, Longjun Zeng, Lili Wang, and Dong-Lei Yang. 2023. "Rice Requires a Chromatin Remodeler for Polymerase IV-Small Interfering RNA Production and Genomic Immunity." *Plant Physiology*, November. https://doi.org/10.1093/plphys/kiad624.

Yang, Dong-Lei, Guiping Zhang, Lili Wang, Jingwen Li, Dachao Xu, Cuiru Di, Kai Tang, et al. 2018. "Four Putative SWI2/SNF2 Chromatin Remodelers Have Dual Roles in Regulating DNA Methylation in Arabidopsis." *Cell Discovery* 4 (1): 55.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have addressed my concerns and further improved the manuscript. I have two minor suggestions, aiming to put the data into context:

L438ff: I do not really agree on the “redistribution” of DNA methylation; at least based on the data this is not really convincing. The data show that some CLSY3-dependent loci are hypermethylated in the kd lines, which is in line with previous data (Yang et al., 2018: doi: 10.1038/s41421-018-0056-8) and I suggest to refer to this publication.

L501ff: The rather limited effect of CLASSY3 on DNA methylation at siren loci is similar to the limited effect in Arabidopsis (Zhou et al., 2022, doi.org/10.1038/s41467-021-27690-x), which the authors could refer to.

Reviewer #2 (Remarks to the Author):

One CLSY type of chromatin remodeling gene, OsCLSY3 was investigated in this work. It specifically affected the endosperm development as the authors concluded. There are three CLSYs in rice genome. The evidence the authors provided were not sufficient to lead the conclusion. The manuscript was not well written.

The imprinted genes were downregulated or upregulated in Osclsy3-kd endosperm just as other DEGs. Since the ratio of maternally- vs. paternally- expressed imprinted genes was not examined, the authors claimed that OsCLSY3 control expression of imprinted genes, which is confusing. The authors should clarify that the transcript levels but not the expression pattern of imprinted was changed.

The author claimed that OsCLSY3 is paternally imprinted gene, which lack evidence. The data of paternal and maternal ratio should be presented.

The FEM2/OsCLSY4 kd rice was created (Figure S4C-S4D). What is the seed setting rate, seed size and endosperm phenotypes of those OsCLSY4-kd mutant? The siRNA levels, DNA methylation and gene transcript on genome scale of OsCLSY4-kd mutant?

Compare the developmental phenotypes and epigenome of OsCLSY3-kd and OsCLSY4-kd is vital to understand their tissue-specific expression and function diversification in indica variety, even they are revealed and compared in Nipponbare, a japonica variety (Plant Physiology, 2024, 194:2149-2164).

The work of FEM2/OsCLSY4 and FEL1/OsCLSY3 published (Plant Physiology, 2024, 194:2149-2164) should be cited in the Abstract, Introduction and Discussion. The different or same results in the two works should be compared and discussed. Especially, distinct difference below should be well cited and discussed, which is critical for the community to notice the contrast results between this work and published data in Plant Physiology.

1) In Nipponbare, a japonica variety, simultaneously knocked out FEL1/OsCLSY3 and FEL2/OsCLSY1 has no effect on growth and development, and on DNA methylation.

2) The transcripts levels of FEL1/OsCLSY3 and FEL2/OsCLSY1 in panicles, stamen and various tissues are substantially lower than that of FEM2/OsCLSY4 (Figure 1A, Plant Physiology).

3) In various developmental stages of seed, FEM2 has quite high expression level than FEL1/OsCLYS3 and FEL2/OsCLSY1 in japonica (Wases et al., 2017; Figure S5 of Plant Physiology).

The obvious difference might be attributed to japonica and indica genetic background as the authors claimed in Response, which should be presented and discussed in the revised manuscript.

All “WT” in Figures and manuscript should be changed into PB1 to avoid misleading of understanding. In Supplementary Figure S4, I guess the “WT” is not PB1, the variety name of various mutant should be replaced with “WT”.

Gene name was desultory throughout the manuscript.

1) As RMR1 and RML in maize, FEM2, FEL1, FEL2 should be added in Figure 1D to avoid confusion in gene name for readers.

2) In Figures and manuscript, CLSY4 should be changed into FEM2/OsCLSY4 (Xu et al., 2024, Plant Physiology), RDR2 should be FEM1/OsRDR2 (Wang et al., 2022, Plant Physiology), POLIV and POLV should be Pol IV and Pol V, DRM2, NRPD1 and AGO4 should be OsDRM2, OsNRPD1 and OsAGO4, CLSY3 should be FEL1/OsCLSY3.

3) There is no gene name for Os08g0289400 in Figure 1D, which is strange.

4) The CLSY1-4 in Arabidopsis was wrongly named as AtCLSY1-AtCLSY4 in this manuscript. In many place, OsCLSY3 and OsCLSY4 were written as CLSY3 and CLSY4, like in Figure 3, 4, 5, 6, 7 and Supplementary Figure S3, S4, S6.

5) The CLSY homologs in maize were named as RMR1, RML1, RML2, and RML3 in Figure 1D. In Figure S2C, however, they were named as ZmCLSY1, ZmCLSY3, ZmCLSY4.

The seed size of OsCLSY3-kd was reduced than WT as shown in Figure 3D, 3F. In Figure 3M, however, they were comparable to WT.

Reviewer #3 (Remarks to the Author):

In their revised manuscript the authors addressed my previous comments. Listed below are minor comments that arose during the last revision.

1. Given the gains and losses of siRNAs and methylation in the clsy3 mutant the authors should avoid claiming effects at specific loci are “direct” effects/targets of CLSY3 unless they demonstrate these loci are those that are directly bound by CLSY3 from their ChIP data.

2. For Figure S9F and H, the text mentions a lower # of loci being assessed after “excluding loci that have zero-CHH methylation levels) but the figures still are marked with N-21653 (lost sRNA) and N-2620 (gain sRNA). Please label the figures with the actual number of loci included in the plots.

3. For Fig. S4H please included WT samples for comparison and add a more detailed description in the text/legend describing the nature of the defects observed.

4. Fig. 5B, legend states Replicates are almost merged. Please clarify, is the data merged computationally, or are the replicates just very similar and thus the lines are overlapping?

5. DMRs tracks were only added for Figs. S15D, 6J, and 7H.

a. Please add DMR track to the screen shots shown in Figs. 6E, 7G, S9A, S9B, S12E, S15B, S15C, S16A, and S16B. Without these added it’s difficult to assess if the changes in methylation shown are statistically significant. If none of these regions contain DMRs, the authors should alter their conclusions accordingly.

b. Please also add a table for the locations of the hyper and hypoDMRs.

6. Thank you for clarifying the overlap analysis for the Shortstack sRNA clusters in the wt and clsy3 mutants. This is now clear and the numbers in the figure and table match for the WT sample but for the clsy3-kd the table has 10,865 clusters but the venn diagram only shows 10,817. Please correct this discrepancy.

7. For Fig. S12D, the number of hypo and hyper siren loci shown, N-248 and N-317 does not match the totals described in the text: “In category-1, among 316 siren loci, DNA methylation increased in 134 loci and decreased in 94 loci. In the category-2, DNA methylation decreased in 146 loci and it increased in 176 loci.” For hypo 94+146=240 and for hyper 134+176=310. Please clarify.

8. As Fig 4F only shows three screen shots, this statement is too broad: “These CLSY3-bound regions were adjacent to, or overlapped with CLSY3-dependent sRNA loci (Fig. 5F).” The authors should either show this statement is generally true across all CLSY3-bound regions, or alter their conclusions to fit the data presented.

9. Related to the CLSY3 ChIP, the authors found “Among 7115 CLSY3-bound peaks, 1398 peaks clearly overlapped with CLSY3-dependent sRNA loci.” This overlap seems very low. Is this more than expected by chance across all siRNA regions regardless of their dependency on CLSY3? Statistics would be help for interpreting this data.

10. Lines 489-493, I think the S11 and S12 references are swapped. Please double check.

11. Lines 529-531 state “This observation suggested that significant number of the DEGs in clsy3-kd lines were directly regulated by CLSY3-dependent DNA methylation specific to endosperm.” However, this is in reference to the correlation between sRNA levels and mRNA levels, so its not clear how this relates to CLSY3-dependent DNA methylation.

12. Lines 537-539 states “The adjacent genes of around 1258 located next to CLSY3-dependent sRNA loci were mis-expressed in clsy3-kd (Supplemental Fig. S13A)”. Is the 1258 the number of adjacent genes or the number of CLSY3-dependent sRNA loci?

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have addressed my concerns and further improved the manuscript. I have two minor suggestions, aiming to put the data into context:

Author's reply: Thank you very much for going through and for your encouraging comments. We have addressed both these suggestions in the revised script.

Reviewer's comment - L438ff: I do not really agree on the "redistribution" of DNA methylation; at least based on the data this is not really convincing. The data show that some CLSY3-dependent loci are hypermethylated in the kd lines, which is in line with previous data (Yang et al., 2018: doi: 10.1038/s41421-0180056-8) and I suggest to refer to this publication.

Author's reply: Thank you for this suggestion. We have changed the text accordingly and cited the suggested literature.

Reviewer's comment - L501ff: The rather limited effect of CLASSY3 on DNA methylation at siren loci is similar to the limited effect in Arabidopsis (Zhou et al., 2022, doi.org/10.1038/s41467-021-27690-x), which the authors could refer to.

Author's reply: Thank you for this suggestion. We have cited the suggested publication.

Reviewer #2 (Remarks to the Author):

One CLSY type of chromatin remodeling gene, OsCLSY3 was investigated in this work. It specifically affected the endosperm development as the authors concluded. There are three CLSYs in rice genome. The evidence the authors provided were not sufficient to lead the conclusion. The manuscript was not well written.

Reviewer's comment: The imprinted genes were downregulated or upregulated in Osclsy3-kd endosperm just as other DEGs. Since the ratio of maternally- vs. paternally- expressed imprinted genes was not examined, the authors claimed that OsCLSY3 control expression of imprinted genes, which is confusing. The authors should clarify that the transcript levels but not the expression pattern of imprinted was changed.

Author's reply: Thank you for your comments. We have provided multiple lines of evidence to show that OsCLSY3 is specifically expressed in endosperm and contributes to its development using detailed genetics, whole genomics, epigenetics and phenotypic analysis. We have made several changes in the language and expression to address the above comments.

Although we claimed that OsCLSY3 controls expression levels of imprinted genes, we did not claim if they are of paternal or maternal origin. In other words, we have not claimed OsCLSY3 regulates imprinting status of imprinted genes. As Reviewer

has kindly pointing out, to avoid the further confusion, we have clearly stated these in the revised script.

In *Arabidopsis*, Pol IV derived sRNAs regulate expression and imprinting status of many imprinted genes (Vu et al. 2013; Erdmann et al. 2017; Kirkbride et al. 2019; Satyaki and Gehring 2022). In this work, we observed CLSY3 dependent sRNAs regulate expression levels of imprinted genes in rice endosperm.

Reviewer's comment: The author claimed that OsCLSY3 is paternally imprinted gene, which lack evidence. The data of paternal and maternal ratio should be presented.

Author's reply: Thank you for the opportunity to clarify. The data is clearly presented in Fig.1I that also the relevant controls, to show that OsCLSY3 is maternally expressed. Relevant controls here include a well-known paternally expressed gene (*OsARF22*) (Luo et al. 2011; Chen et al. 2018). as well as bi-allelic gene that expresses in both parents (*OsSHH1*). The presented data clearly shows *OsCLSY3* expression is more biased towards the maternal side. We only suggest that CLSY3 might be paternally imprinted based on these results. It has been well documented that imprinting status of genes change during the development of endosperm, and that specific types of cells and not the complete endosperm might be showing uniform genomic imprinting (Picard et al. 2021). Please also note that two published studies also showed that *OsCLSY3* is maternally biased, but the gene was not annotated in their lists (Luo et al. 2011; Chen et al. 2018). In our study, the raw data regarding paternal and maternal ratio is presented in Supplemental Figure S2F. The NGS data derived from crossing experiments clearly suggesting maternal expression and suggesting the paternal imprinting of *OsCLSY3* has been shared with wider scientific community with a link in GEO.

Reviewer's comment: The FEM2/OsCLSY4 kd rice was created (Figure S4C-S4D). What is the seed setting rate, seed size and endosperm phenotypes of those OsCLSY4-kd mutant? The siRNA levels, DNA methylation and gene transcript on genome scale of OsCLSY4-kd mutant?

Author's reply: This paper is about function of OsCLSY3 in endosperm. Hence, global siRNA and transcript levels and genome scale DNA methylation of **clsy4-kd** lines are not presented in this paper, as they are beyond the scope of this study. However, we have provided additional phenotypic data of clsy4-kd lines, summary of which is as follows.

The recent paper from Xu et al., 2024 nicely documented phenotypes and global sRNA, transcript and DNA methylation level of fem2/osclsy4 mutant in *japonica* rice 18 days old seedling (Plant Physiology, 2024, 194:2149-2164). We also found many similar phenotypes in our clsy4-kd in PB1 *indica* rice. The phenotypes such as plant height, panicle length, seed setting rate of clsy4-kd plants were similar with fem2 as described in Xu et al 2024 paper. We included all these phenotypic data in the revised version of the manuscript (Supplemental Figure S4F-H).

The only difference in the data that we presented here with Xu et al., 2024 is the following: we detected *OsCLSY3* and *OsCLSY4* expression in *indica* rice

endosperm. We have provided multiple evidence to show that *OsCLSY3* in *indica* line that we used here has endosperm-specific expression (using RT-qPCR and RNA-seq). Using the same analysis, we also show OsCLSY4 has low expression in endosperm. We also found that independent osclsy3-kd, and osclsy4-kd transgenic plants in *indica* background show smaller seeds and endosperm defects. All these data have been included in the revised script with sufficient details in text.

Reviewer's comment: Compare the developmental phenotypes and epigenome of OsCLSY3-kd and OsCLSY4-kd is vital to understand their tissue-specific expression and function diversification in indica variety, even they are revealed and compared in Nipponbare, a japonica variety (Plant Physiology, 2024, 194:2149-2164).

Author's reply: Thank you for this suggestion. The question is indeed very interesting. However, this paper is on understanding function of *OsCLSY3* in rice endosperm, and clearly comparing the developmental phenotypes and epigenome of osclsy3-kd and osclsy4-kd between *indica* and *japonica*, side by side, is beyond the scope of this study. Since rice cultivars show variations in the expression of development and yield-associated genes between them, one has to include multiple *indica* and multiple *japonica* lines derived from different geographical origins to perform such as a study.

We are completely agreeing with Reviewer that Xu et al., 2024 showed role of FEM2/OsCLSY4 in rice. However, we are disagreeing with the point that Xu et al also studied role of FEL1 (*OsCLSY3*), FEL2 (*OsCLSY1*) in epigenome regulation properly. Kindly note that Xu et al., 2024 did not use single mutants of FEL1 and FEL2, but used only double mutants. Comparing our data of osclsy3-kd with the *fel1fel2* double mutant is incorrect and also please note that *fel1fel2* genomic studies have not been undertaken (as per GSE215857). Hence Xu et al., 2024 paper did not reveal and compare the developmental phenotypes and epigenome of osclsy3-kd and osclsy4-kd in Nipponbare, a *japonica* variety, a direct comparison with *indica* lines is not possible.

Reviewer's comment: The work of FEM2/OsCLSY4 and FEL1/OsCLSY3 published (Plant Physiology, 2024, 194:2149-2164) should be cited in the Abstract, Introduction and Discussion. The different or same results in the two works should be compared and discussed. Especially, distinct difference below should be well cited and discussed, which is critical for the community to notice the contrast results between this work and published data in Plant Physiology.

1) In Nipponbare, a japonica variety, simultaneously knocked out FEL1/OsCLSY3 and FEL2/OsCLSY1 has no effect on growth and development, and on DNA methylation.

2) The transcripts levels of FEL1/OsCLSY3 and FEL2/OsCLSY1 in panicles, stamen and various tissues are substantially lower than that of FEM2/OsCLSY4 (Figure 1A, Plant Physiology).

3) In various developmental stages of seed, FEM2 has quite high expression level than FEL1/OsCLYS3 and FEL2/OsCLSY1 in japonica (Wases et al., 2017; Figure S5 of Plant Physiology).

The obvious difference might be attributed to japonica and indica genetic background as the authors claimed in Response, which should be presented and discussed in the revised manuscript.

Author's reply: Thank you for this suggestion. We had already cited the paper Xu et al., 2024 after it got published when our paper was undergoing review. We have also now cited the paper (Plant Physiology, 2024, 194:2149-2164) in the revised manuscript in results and discussion. Please note that our Introduction lists CLSY only as a partner in RdDM without discussing any details of its function, similar to other RdDM members. The difference (only one difference) and similarities (several) between Xu et al., 2024 and our paper are also included in the revised discussion in detail. We did not find Wases et al (2017) reference that the Reviewer has mentioned above.

Reviewer's comment: All "WT" in Figures and manuscript should be changed into PB1 to avoid misleading of understanding. In Supplementary Figure S4, I guess the "WT" is not PB1, the variety name of various mutant should be replaced with "WT".

Gene name was desultory throughout the manuscript.

Author's reply: Thank you for this suggestion. To avoid the confusion, in revised version we used "PB1". The *japonica* origin of WT in the Supplementary Figure S4 has been included in the revised figure and its legend.

1) **Reviewer's comment:** As RMR1 and RML in maize, FEM2, FEL1, FEL2 should be added in Figure 1D to avoid confusion in gene name for readers.

Author's reply: Thank you for this suggestion. We included the FEM2, FEL1, FEL2 names in Figure 1D and Supplemental dataset 2 to avoid the confusion.

2) **Reviewer's comment:** In Figures and manuscript, CLSY4 should be changed into FEM2/OsCLSY4 (Xu et al., 2024, Plant Physiology), RDR2 should be FEM1/OsRDR2 (Wang et al., 2022, Plant Physiology), POLIV and POLV should be Pol IV and Pol V, DRM2, NRPD1 and AGO4 should be OsDRM2, OsNRPD1 and OsAGO4, CLSY3 should be FEL1/OsCLSY3.

Author's reply: Thank you for this suggestion. We have incorporated additional names whenever the genes were named first time. Upon consulting with the Handling Editor, to keep the nomenclature as per the Nat Comm rules, we have used the CLSY name.

3) **Reviewer's comment:** There is no gene name for Os08g0289400 in Figure 1D, which is strange.

Author's reply: In RAPDB, the gene is vaguely named as SNF2-related domain containing protein. We included the name in the revised manuscript, Figure 1D and Supplemental dataset 2.

4) **Reviewer's comment:** The CLSY1-4 in Arabidopsis was wrongly named as AtCLSY1-AtCLSY4 in this manuscript. In many places, OsCLSY3 and OsCLSY4 were written as CLSY3 and CLSY4, like in Figure 3, 4, 5, 6, 7 and Supplementary Figure S3, S4, S6.

Author's reply: Thank you for pointing out this. We have corrected these in the revised manuscript to avoid confusion regarding the species these genes are derived from.

5) **Reviewer's comment:** The CLSY homologs in maize were named as RMR1, RML1, RML2, and RML3 in Figure 1D. In Figure S2C, however, they were named as ZmCLSY1, ZmCLSY3, ZmCLSY4.

Author's reply: Thank you for pointing out the mistakes. We have corrected them in the revised Figure S2C.

Reviewer's comment: The seed size of OsCLSY3-kd was reduced than WT as shown in Figure 3D, 3F. In Figure 3M, however, they were comparable to WT.

Author's reply: Thank you for the opportunity to explain this. For the measurement of seed size in Figure 3D and 3F, completely dry seeds were used. Please note that in Figure 3M however, the seeds were imbibed in water for 4 days. The germination process changed the entire seed morphology. Due to this, it was not possible to compare seed size in Figure 3M.

We thank the Reviewer for all the comments and wonderful suggestions.

Reviewer #3 (Remarks to the Author)

In their revised manuscript the authors addressed my previous comments. Listed below are minor comments that arose during the last revision.

Author's reply: Thank you for the detailed review and helpful suggestions. We have incorporated the suggestions in the revised manuscript-please see below.

1. **Reviewer's comment:** Given the gains and losses of siRNAs and methylation in the clsy3 mutant the authors should avoid claiming effects at specific loci are "direct" effects/targets of CLSY3 unless they demonstrate these loci are those that are directly bound by CLSY3 from their ChIP data.
   Author's reply: Thank you for pointing out this. We have toned down the claims in the revised version.

2. **Reviewer's comment:** For Figure S9F and H, the text mentions a lower # of loci being assessed after "excluding loci that have zero-CHH methylation levels) but the figures still are marked with N-21653 (lost sRNA) and N-2620 (gain sRNA). Please label the figures with the actual number of loci included in the plots.
   Author's reply: Thank you for this suggestion. We have incorporated the number of the loci in the revised figure.
   For the N-21653 (loci that lost sRNA), after excluding loci that have zero-CHH methylation levels, we got 13259 loci. Among those, 7257 loci showed hypermethylation, 4953 loci showed hypomethylation and in1049 loci, DNA methylation was unchanged. For the N-2620 (gained sRNA loci), after excluding loci that have zero-CHH methylation levels, we got 1800 loci. Among those, 898 loci showed hypermethylation, 723 loci showed hypomethylation and in 179 loci, DNA methylation was unchanged.

3. **Reviewer's comment:** For Fig. S4H please included WT samples for comparison and add a more detailed description in the text/legend describing the nature of the defects observed.
   Author's reply: Thank you. We have incorporated the WT samples of same stage with *KO* for better comparison. We also added nature of the defects in the revised text for better clarity.

4. **Reviewer's comment:** Fig. 5B, legend states Replicates are almost merged. Please clarify, is the data merged computationally, or are the replicates just very similar and thus the lines are overlapping?

   Author's reply: The replicates were very similar (Fig. 5B) due to that it is hard to visualize them separately in the figure and lines were overlapping.

5. **Reviewer's comment:** DMRs tracks were only added for Figs. S15D, 6J, and 7H.
   Author's reply: We thought these are the only ones that need extra tracks as per the Reviewer's comment and added DMR tracks. The coloured bars are regions with changes in DNA methylation between samples as mentioned in the legend.

   a. **Reviewer's comment:** Please add DMR track to the screen shots shown in Figs. 6E, 7G, S9A, S9B, S12E, S15B, S15C, S16A, and S16B. Without these added it's difficult to assess if the changes in methylation shown are statistically significant. If none of these regions contain DMRs, the authors should alter their conclusions accordingly.
      Author's reply: Thank you for this suggestion. We wish to clarify that in Figs. S15D, 6J, and 7H what we showed with violet lines (mention in the text of legend) are easily distinguishable regions with variations in DNA methylation which were located close to CLSY3-dependent sRNA loci. These are not part of DMRs that were identified using DMRcaller. The legend has been modified to bring this clarity.

   b. **Reviewer's comment:** Please also add a table for the locations of the hyper and hypoDMRs.
      Author's reply: The DNA methylation status of the CLSY3 dependent sRNAs (gained and lost), siren loci were already added in Supplemental dataset S7. Now, to bring better clarity, we also added a table which listed hypo- and hyper- DMRs found by DMRcaller. These extra details have been added in Supplemental dataset S7.

6. Thank you for clarifying the overlap analysis for the Shortstack sRNA clusters in the wt and clsy3 mutants. This is now clear and the numbers in the figure and table match for the WT sample but for the clsy3-kd the table has 10,865 clusters but the venn diagram only shows 10,817. Please correct this discrepancy.

Author's reply: Thank you for pointing this. The output of Shortstack sRNA clusters have been provided in supplemental datasets S4, S5 and S6. The venn diagram was generated using online intervene online tool. As pointed out previously, the 48 sRNA loci were merged when the tool was considering the loci. We have included this detail in the method section to remove ambiguity.

7. For Fig. S12D, the number of hypo and hyper siren loci shown, N-248 and N-317 does not match the totals described in the text: "In category-1, among 316 siren loci, DNA methylation increased in 134 loci and decreased in 94 loci. In the category-2, DNA methylation decreased in 146 loci and it increased in 176 loci." For hypo 94+146=240 and for hyper 134+176=310. Please clarify.
   Author's reply: Thank you for this comment. We provide the numbers clearly for each category below.
   Among 797 siren loci, we found 722 siren loci for further analysis (excluding loci that have zero-CHH methylation levels). In those loci, 248 loci were hypomethylated, 317 loci were hypermethylated and in 157 loci, DNA methylation was unchanged.
   In category 1, among 316 siren loci, 94 loci were hypo methylated and 134 loci were hypermethylated. In category 2, among 464 loci, 146 loci were hypo methylated and 175 loci were hypermethylated. In category 3, there are 17 loci. Among them, 7 loci were hypomethylated and 6 loci got hypermethylation.
   Total hypermethylated loci: 134 + 175 + 6 (category-3) = 315 (2 less).
   Total hypomethylated loci: 94 + 146 + 7 (category-3) = 247 (1 less).
   We have checked the numbers carefully and there is no error. We lost 3 loci when the analysis was performed using intervene tool and this is commonly observed while generating venn diagrams. This problem we faced in the CLSY3 dependent sRNA loci count also. This detail is now added in the method section to avoid ambiguity.

8. As Fig 4F only shows three screen shots, this statement is too broad: "These CLSY3-bound regions were adjacent to, or overlapped with CLSY3-dependent sRNA loci (Fig. 5F)." The authors should either show this statement is generally true across all CLSY3-bound regions, or alter their conclusions to fit the data presented.
   Author's reply: Thank you. This broad statement was however followed by specific details. Since the statement was very general as pointed out, we have altered the statement in the revised manuscript.

9. Related to the CLSY3 ChIP, the authors found "Among 7115 CLSY3-bound peaks, 1398 peaks clearly overlapped with CLSY3-dependent sRNA loci." This overlap seems very low. Is this more than expected by chance across all siRNA regions regardless of their dependency on CLSY3? Statistics would be help for interpreting this data.
   Author's reply: Thank you for these suggestions. We have performed the hypergeometric test and found the overlap is statistically not significant. Please note that, as mentioned in the manuscript, ChIP was performed in panicle tissue

before anthesis whereas sRNAs and methylation data are from 20 days old endosperm. Please note that targeting modes to induce DNA methylation and sRNA production did not overlap between tapetum and ovule as observed in Arabidopsis CLSY3 bound regions, indicating further studies are require the mechanistic basis for this observation (Long et al. 2021; Zhou et al. 2022; Martins and Law 2023).

10. Lines 489-493, I think the S11 and S12 references are swapped. Please double check.
Author's reply: Thank you for pointing this mistake. We have corrected it in the revised manuscript.

11. Lines 529-531 state "This observation suggested that significant number of the DEGs in clsy3kd lines were directly regulated by CLSY3-dependent DNA methylation specific to endosperm." However, this is in reference to the correlation between sRNA levels and mRNA levels, so its not clear how this relates to CLSY3-dependent DNA methylation.
Author's reply: Thank you for pointing out this. The sentence was indeed ambiguous. We have changed the sentence in the revised version.

12. Lines 537-539 states "The adjacent genes of around 1258 located next to CLSY3-dependent sRNA loci were mis-expressed in clsy3-kd (Supplemental Fig. S13A)". Is the 1258 the number of adjacent genes or the number of CLSY3-dependent sRNA loci?
Author's reply: The 1258 numbers are significantly mis-expressed genes in clsy3-kd endosperm which are located 2 kb adjacent to the CLSY3-dependent sRNA loci.

References

Chen, C., T. Li, S. [.]. Zhu, Z. Liu, Z. Shi, X. Zheng, R. Chen, et al. 2018. "Characterization of Imprinted Genes in Rice Reveals Conservation of Regulation and Imprinting with Other Plant Species." *Plant Physiology* 177 (4): 1754–71.

Erdmann, Robert M., Prasad R. V. Satyaki, Maja Klosinska, and Mary Gehring. 2017. "A Small RNA Pathway Mediates Allelic Dosage in Endosperm." *Cell Reports* 21 (12): 3364–72.

Kirkbride, Ryan C., Jie Lu, Changqing Zhang, Rebecca A. Mosher, David C. Baulcombe, and Z. Jeffrey Chen. 2019. "Maternal Small RNAs Mediate Spatial-Temporal Regulation of Gene Expression, Imprinting, and Seed Development in Arabidopsis." *Proceedings of the National Academy of Sciences of the United States of America* 116 (7): 2761–66.

Long, Jincheng, James Walker, Wenjing She, Billy Aldridge, Hongbo Gao, Samuel Deans, Martin Vickers, and Xiaoqi Feng. 2021. "Nurse Cell--Derived Small RNAs Define Paternal Epigenetic Inheritance in Arabidopsis." *Science (New York, N.Y.)* 373 (6550): eabh0556.

Luo, Ming, Jennifer M. Taylor, Andrew Spriggs, Hongyu Zhang, Xianjun Wu, Scott Russell, Mohan Singh, and Anna Koltunow. 2011. "A Genome-Wide Survey of

Imprinted Genes in Rice Seeds Reveals Imprinting Primarily Occurs in the Endosperm." *PLoS Genetics* 7 (6): e1002125.
Martins, Laura M., and Julie A. Law. 2023. "Moving Targets: Mechanisms Regulating SiRNA Production and DNA Methylation during Plant Development." *Current Opinion in Plant Biology* 75 (August): 102435.
Picard, Colette L., Rebecca A. Povilus, Ben P. Williams, and Mary Gehring. 2021. "Transcriptional and Imprinting Complexity in Arabidopsis Seeds at Single-Nucleus Resolution." *Nature Plants* 7 (6): 730–38.
Satyaki, Prasad R. V., and Mary Gehring. 2022. "RNA Pol IV Induces Antagonistic Parent-of-Origin Effects on Arabidopsis Endosperm." *PLoS Biology* 20 (4): e3001602.
Vu, Thiet Minh, Miyuki Nakamura, Joseph P. Calarco, Daichi Susaki, Pei Qi Lim, Tetsu Kinoshita, Tetsuya Higashiyama, Robert A. Martienssen, and Frédéric Berger. 2013. "RNA-Directed DNA Methylation Regulates Parental Genomic Imprinting at Several Loci in Arabidopsis." *Development (Cambridge, England)* 140 (14): 2953–60.
Zhou, Ming, Ceyda Coruh, Guanghui Xu, Laura M. Martins, Clara Bourbousse, Alice Lambolez, and Julie A. Law. 2022. "The CLASSY Family Controls Tissue-Specific DNA Methylation Patterns in Arabidopsis." *Nature Communications* 13 (1): 244.
Zhou, Ming, Ana Marie S. Palanca, and Julie A. Law. 2018. "Locus-Specific Control of the de Novo DNA Methylation Pathway in Arabidopsis by the CLASSY Family." *Nature Genetics* 50 (6): 865–73.