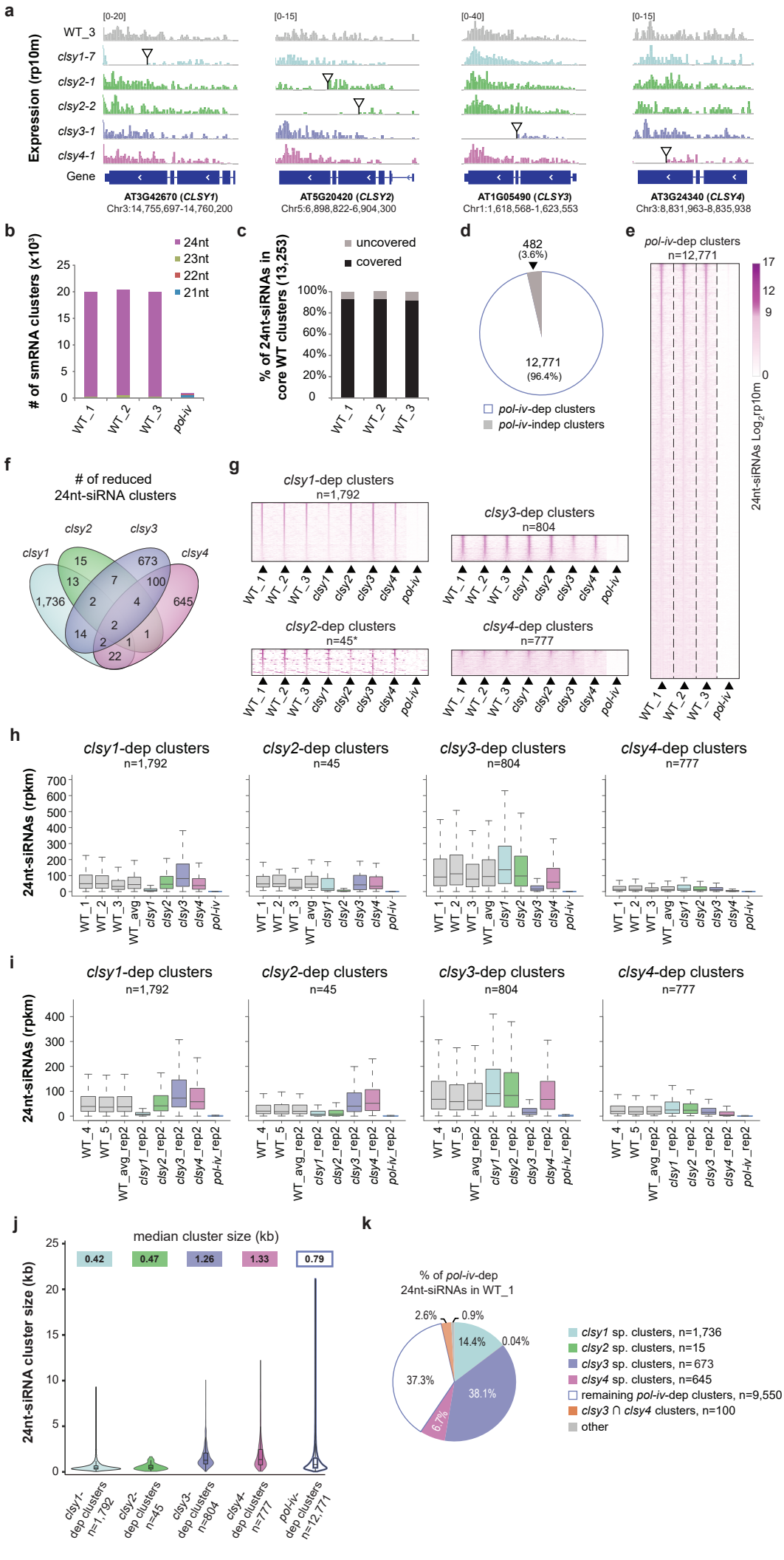
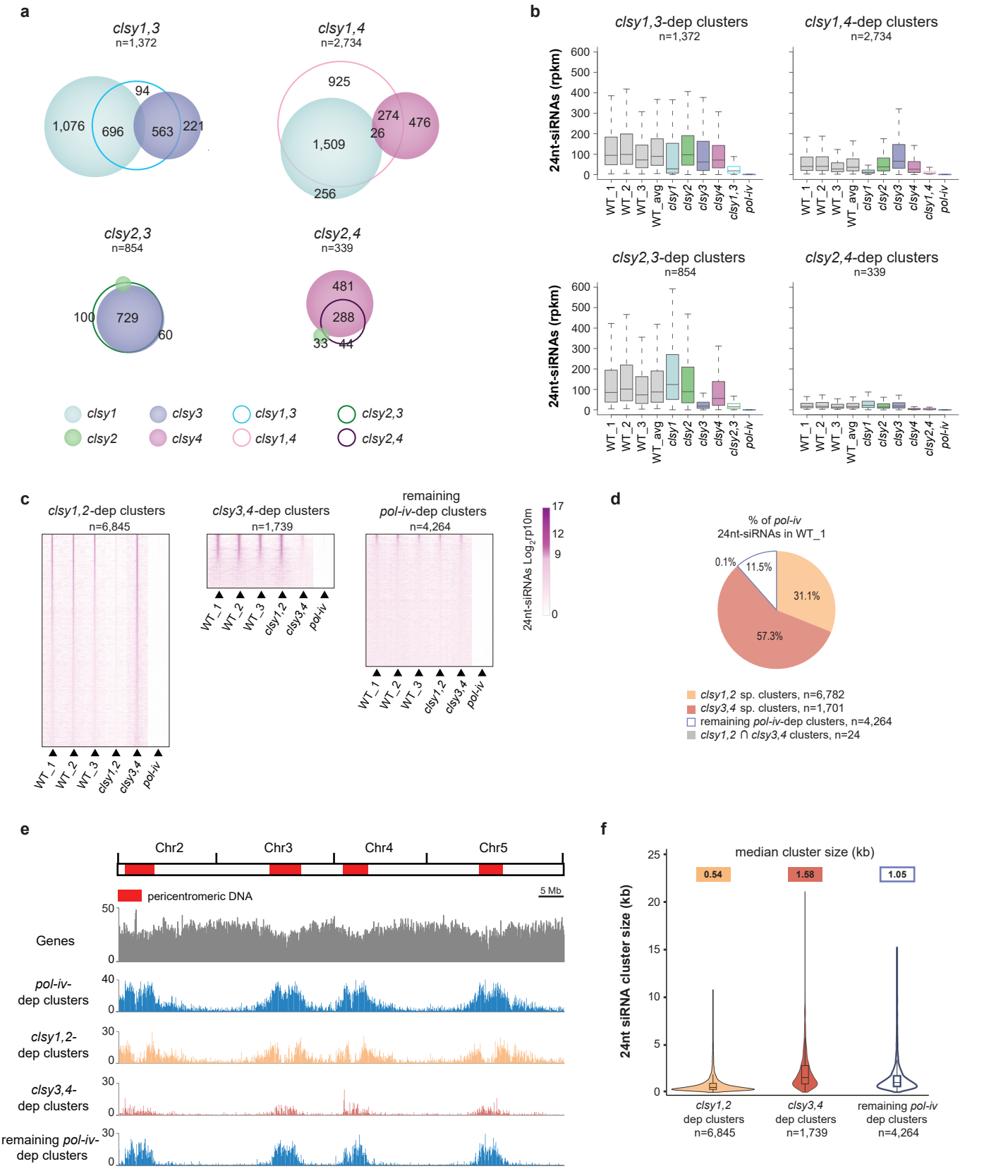
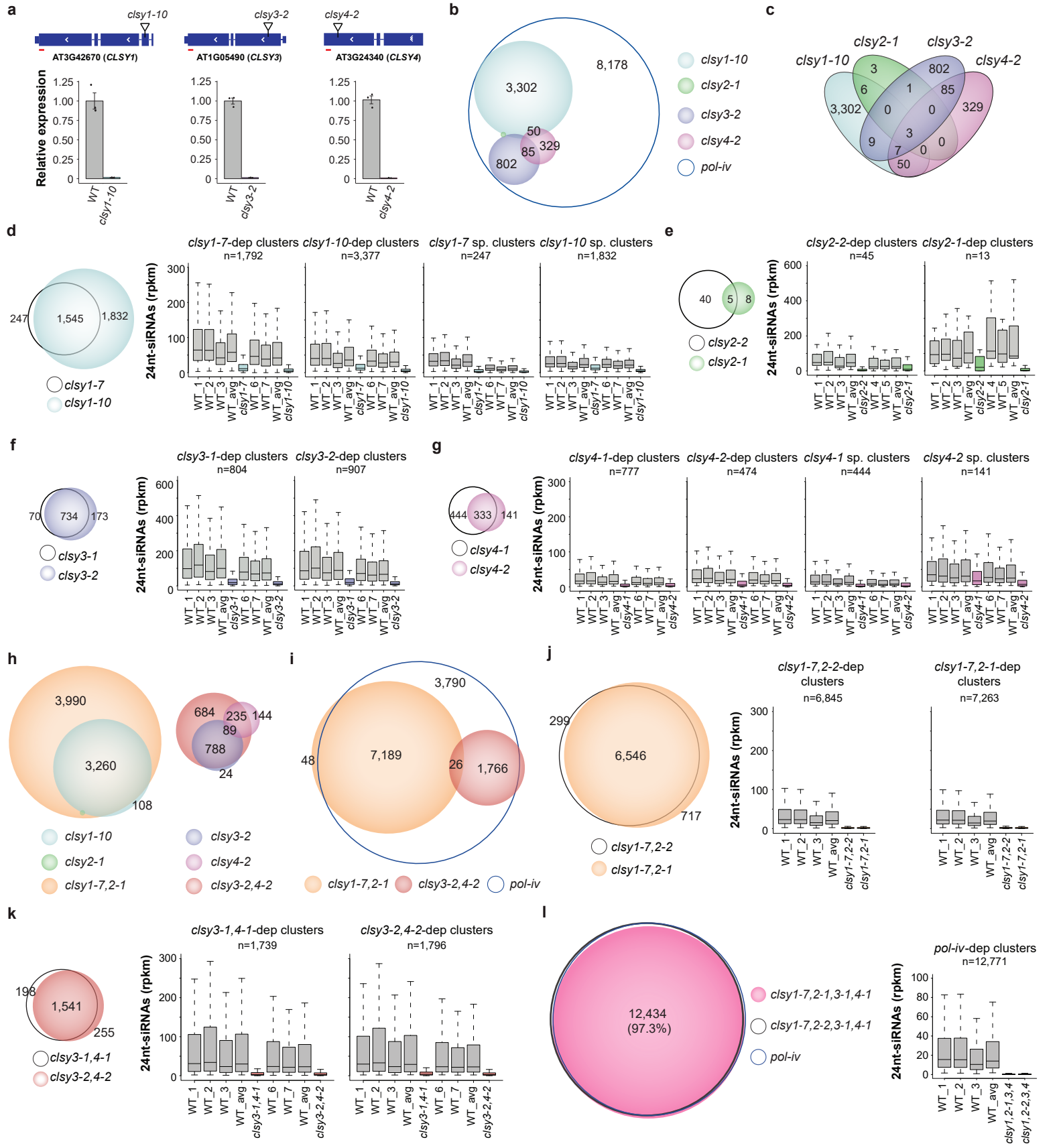



**Supplementary Figure 1. Identification, validation, and characterization of *clsy*-dependent 24nt-siRNA clusters.**

(**a**) Genome browser screen shots depicting the expression levels (reads per 10 million, rp10m) of the four *CLSY* genes in wild-type (WT) plants or the indicated *clsy* mutant alleles. The approximate locations of the T-DNA insertions for each allele are indicated by the hollow black triangles and the scale is indicated in brackets. (**b**) Bar graph showing the number of 21nt-, 22nt-, 23nt-, and 24nt-smRNA clusters identified from small RNA-seq experiments in three WT replicates (WT_1, WT_2, and WT_3) and the *pol-iv* mutant using the ShortStack tool suite. Cluster numbers for all other genotypes can be determined from Supplementary Table 3. (**c**) Bar graph showing the percent of 24nt-siRNAs in the three WT small RNA datasets that map to the core set of clusters (*i.e.* "covered" vs "uncovered"). For details regarding the generation of the core clusters, see Materials and Methods. (**d**) Pie chart showing the percent of core 24nt-siRNA clusters that are *pol-iv*-dependent. (**e**) Heatmap showing the levels of 24nt-siRNAs at *pol-iv*-dependent 24nt-siRNA clusters in the genotypes indicated below. The black triangle marks the midpoint of the 24nt-siRNA clusters and the flanking regions extend +/- 5kb. (**f**) Unscaled Venn diagram showing the overlap between loci with reduced 24nt-siRNA levels in the *clsy* single mutants. (**g**) Heatmaps showing the levels of 24nt-siRNAs at *clsy*-dependent 24nt-siRNA clusters in the genotypes indicated below. Here and elsewhere, the heatmaps are scaled as in **e** such that the Y-axis reflects the number of clusters analyzed, with the exception of the *clsy2*-dependent clusters that are marked by an asterisk (*) to indicate that this heatmap was expanded by 20-fold due to the small sample size. (**h** and **i**) Boxplots showing 24nt-siRNA levels at the indicated *clsy*-dependent clusters in the *clsy* single mutants as compared to *pol-iv*. In **h**, only uniquely-mapped reads from the same data set used in Figure 1 are plotted and in **i**, both unique- and multi-mapping reads from an independent, biological replicate (rep2) are plotted. (**j**) Violin plots showing the size distribution of 24nt-siRNA clusters dependent on the mutants listed below. Median cluster sizes are indicated above each sample. Here, and in all subsequent supplementary figures, the boxplots and violin plots show the interquartile range (IQR) with the median shown as the black line and the whiskers corresponding to 1.5 times the IQR. For the violin plots, the outlines indicate kernel probability density. (**k**) Pie chart showing the percent of *pol-iv*-dependent 24nt-siRNAs present in a representative WT sample, WT_1, that fall into 24nt-siRNA clusters dependent on each *clsy* single mutant or none (*i.e.* "remaining *pol-iv*-dep clusters"). As clusters controlled by *clsy3* and *clsy4* show the largest overlap (∩), these clusters are treated independently while all the other intersecting clusters are grouped together and represented by the "other" category. sp. indicates *clsy*-specific reduced 24nt-siRNA clusters.

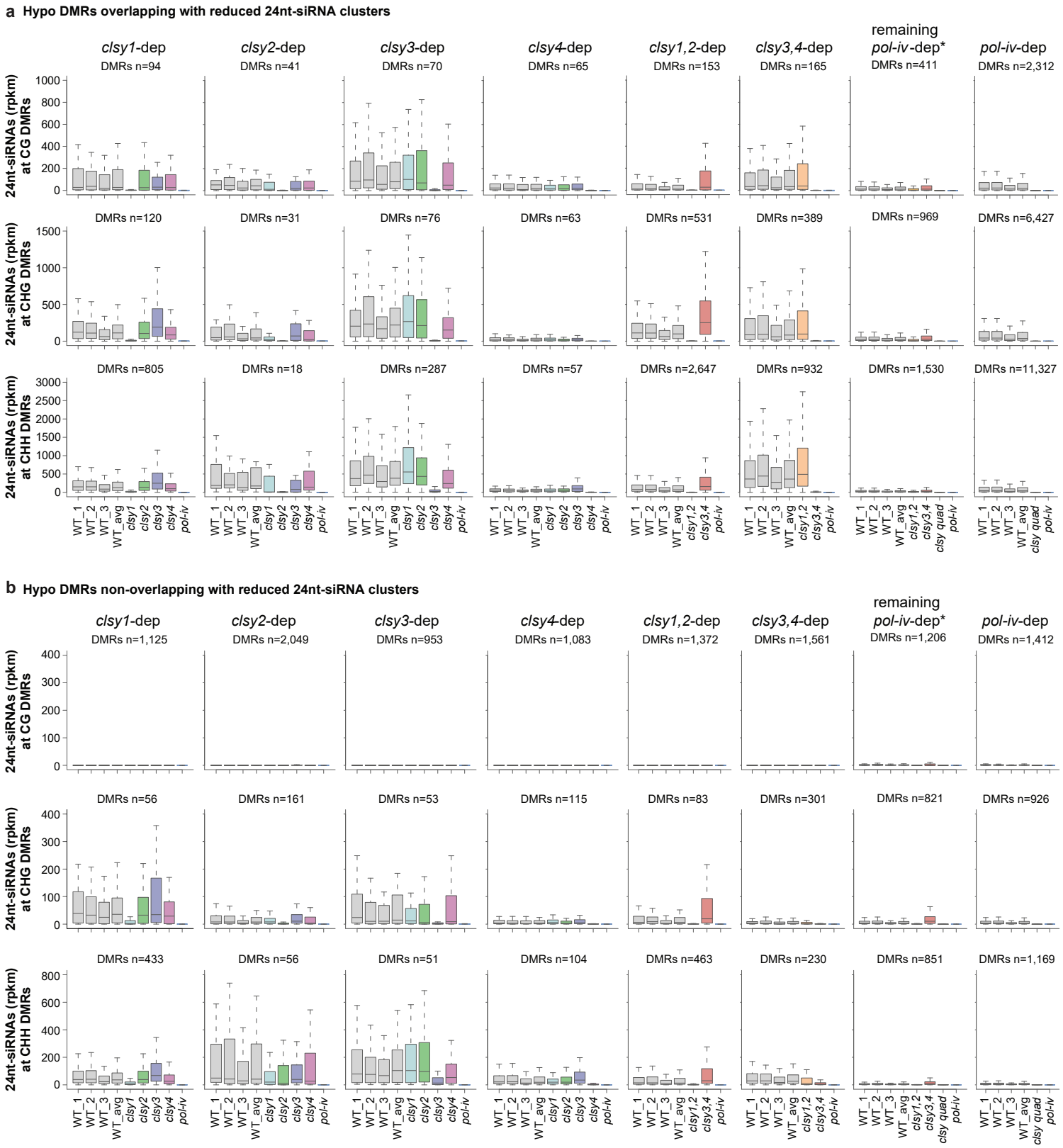



**Supplementary Figure 2. Characterization of 24nt-siRNA clusters dependent on various *clsy* double mutants.**
(**a**) Scaled Venn diagrams showing the relationships between loci with reduced 24nt-siRNA levels in the indicated *clsy* single and double mutants. (**b**) Boxplots showing 24nt-siRNA levels at the indicated *clsy*-dependent clusters in *clsy* single or double mutants as compared to WT controls and *pol-iv*. (**c**) Proportional heatmaps showing the levels of 24nt-siRNAs at *clsy*-dependent 24nt-siRNA clusters in the genotypes indicated below. The black triangles mark the midpoints of the 24nt-siRNA clusters and the flanking regions extend +/- 5kb. (**d**) Pie chart showing the percent of *pol-iv*-dependent 24nt-siRNAs present in a representative WT sample, WT_1, that fall into 24nt-siRNA clusters dependent on each *clsy* double mutant or neither (*i.e.* "remaining *pol-iv*-dep clusters"). sp. indicates *clsy*-specific reduced 24nt-siRNA clusters. Clusters controlled by both double mutants (∩) were also included. (**e**) Chromosome views of 24nt-siRNA clusters dependent on the genotypes indicated on the left, where the scale is the number of clusters per 100kb bin. Red regions correspond to the pericentromeric DNA[56]. (**f**) Violin plots showing the size distribution of 24nt-siRNA clusters dependent on the mutants listed below. Median cluster sizes are indicated above each sample.

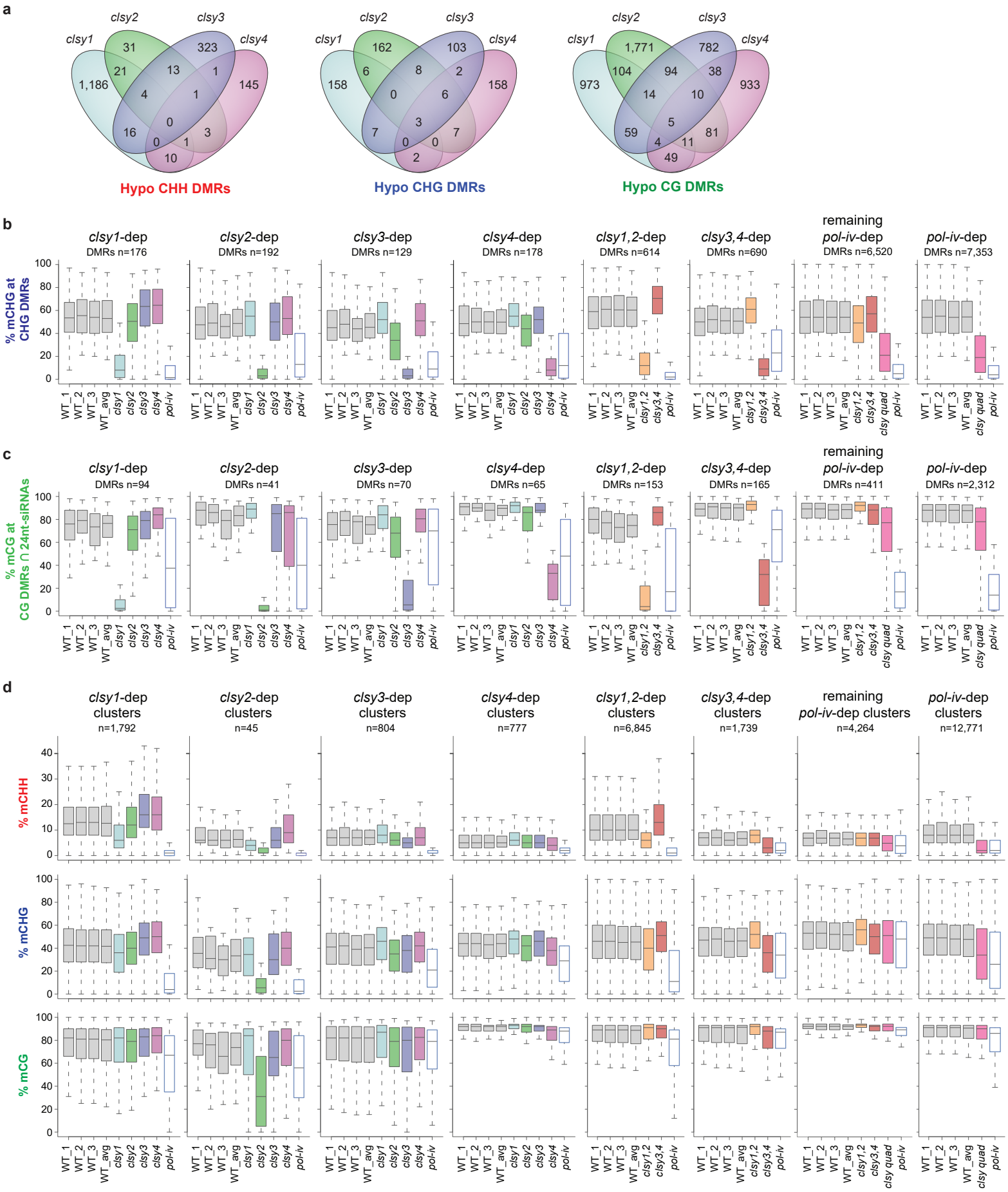



**Supplementary Figure 3. Confirmation of *clsy* mutant phenotypes with alternate mutant alleles.**

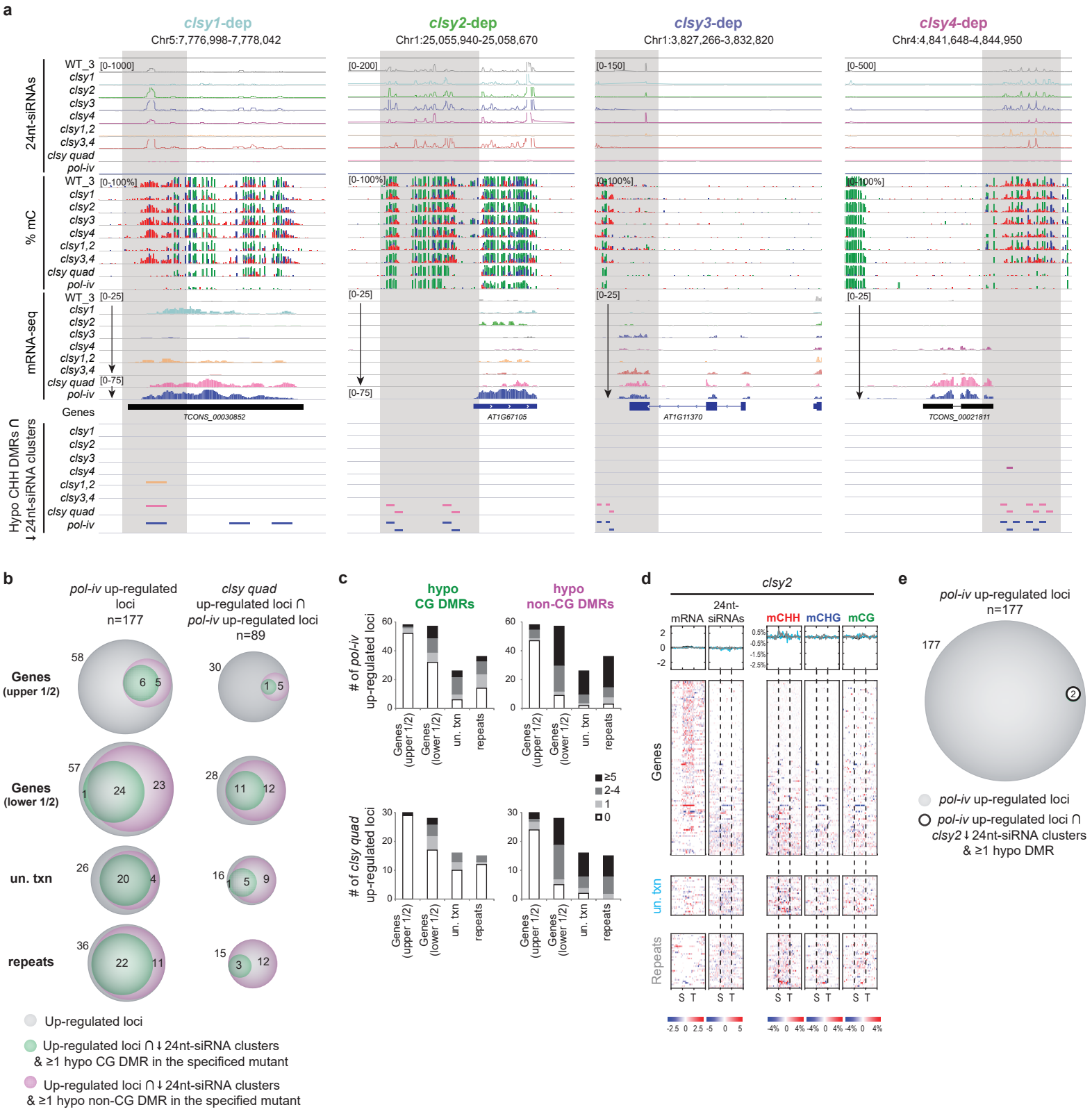
(**a**) Reverse transcriptase quantitative PCR (RT-qPCR) assays showing the expression levels of the *CLSY1*, *CLSY3* and *CLSY4* genes in WT plants or the indicated *clsy* mutant alleles. The samples were normalized relative to *ACTIN2* and the error bars represent the standard error between three technical replicates. In the gene models above, hollow black triangles indicate the locations of the T-DNA insertions for each allele and red lines indicate the regions amplified by PCR. (**b**) Scaled Venn diagrams based on the reduced 24nt-siRNA clusters provided in Supplementary Table 4 showing the relationships between loci with reduced 24nt-siRNA levels in the *clsy* single mutants. For readability, only overlaps >20 are labeled. A small number of overlaps are not shown due to spatial constraints, but the unscaled Venn diagram in **c** shows all the overlaps. (**d-g** and **j-l**) Scaled Venn diagrams (left) and boxplots (right) showing the overlaps in reduced 24nt-siRNA clusters and reductions in 24nt-siRNA levels, respectively, in the indicated genotypes. For the Venn diagrams, the *clsy* alleles characterized in the main text are depicted as black outlined circles while the new alleles are colored. In **e** the diagram is upscaled by 25-fold for readability. For **d** and **g**, where many reduced 24nt-siRNA clusters appear to be specific to just one of the two alleles based on the DESeq2 thresholds (FC≥2; FDR≤0.01), additional boxplots were included to demonstrate that even at these loci, the two alleles show similar trends. (**h** and **i**) Scaled Venn diagrams showing relationships between loci with reduced 24nt-siRNA levels in the *clsy* single, double, and *pol-iv* mutants.



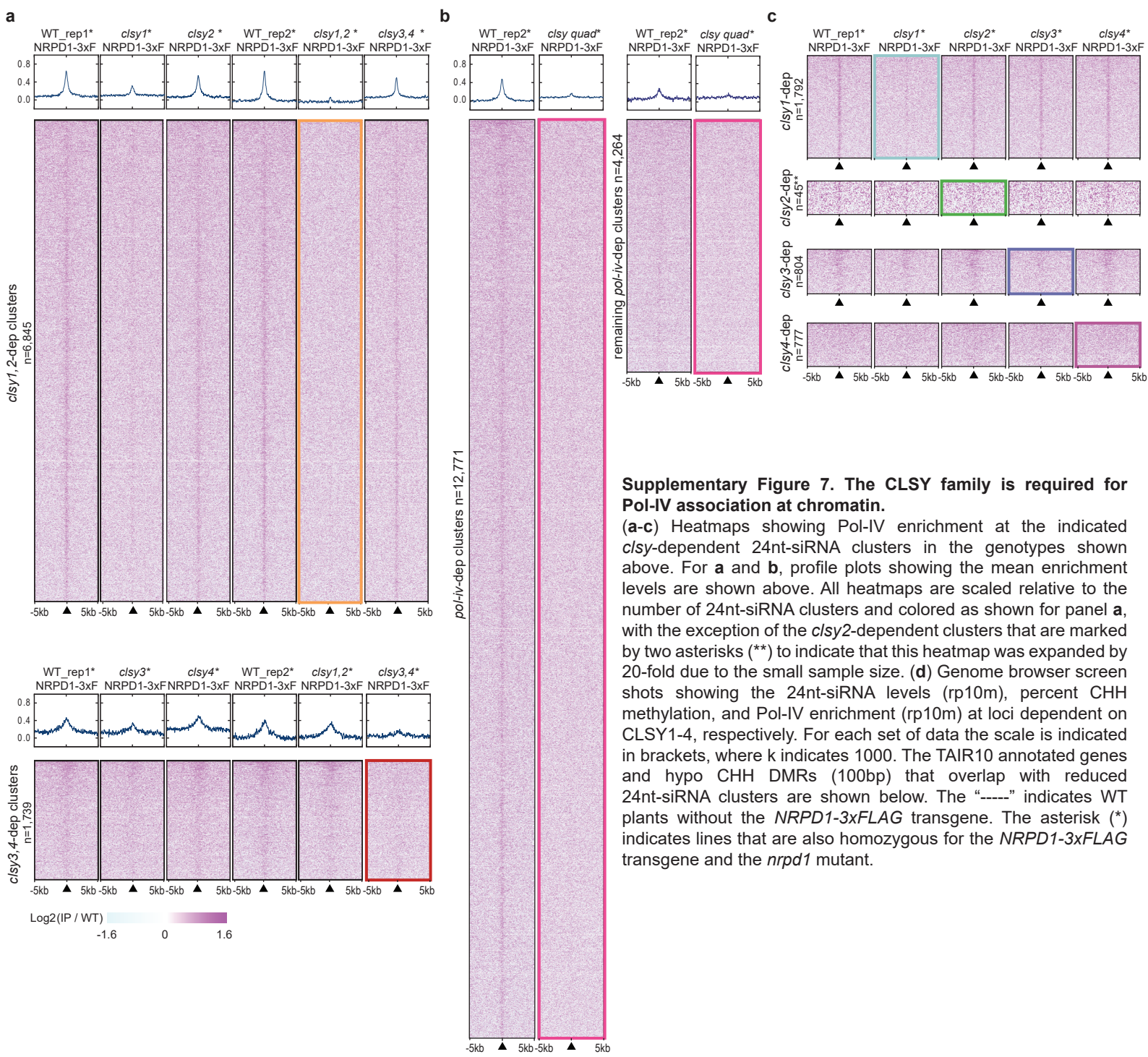

**Supplementary Figure 4. Analysis of 24nt-siRNA levels at DMRs identified in the *clsy* mutants.**
Boxplots showing 24nt-siRNA levels in each *clsy* single, double or quadruple mutant as compared to each other and *pol-iv* at hypo DMRs that either overlap with (**a**), or fail to overlap with (**b**), reduced 24nt-siRNA clusters in the various *clsy* mutants. For **b**, the reductions at these loci were either too weak (*e.g.* ≤2fold reduced) or too small (*e.g.* only 100bp out of a larger 24nt-siRNA cluster) to pass the thresholds used for calling reduced 24nt-siRNA clusters. The asterisk (*) indicates that the overlap between hypo DMRs and reduced 24nt-siRNA clusters for this category of DMRs were determined differently because they do not correspond to a genotype. For the genotype-specific DMRs (*e.g. clsy1*-dep DMRs), the total DMRs are split between those that do, or do not, overlap with reduced 24nt-siRNA clusters identified in the same genetic background. For the *pol-iv*-dependent hypo DMRs not identified in either the *clsy1,2* or *clsy3,4* double mutants ("remaining *pol-iv*-dep DMRs"), the DMRs that overlap with 24nt-siRNA clusters that are reduced in the *pol-iv* mutant but not the *clsy1,2* or *clsy3,4* mutants were then determined to define the "remaining *pol-iv*-dep DMRs" that overlap with the "remaining *pol-iv*-dep reduced 24nt-siRNA clusters". A similar analysis was preformed to identify the "remaining *pol-iv*-dep DMRs" that do not overlap with reduced 24nt-siRNA clusters.




**Supplementary Figure 5. Characterization of hypo DMRs and quantification of DNA methylation levels in *clsy* mutants.**
(**a**) Unscaled Venn diagrams showing the relationships between hypo DMRs in the CHH, CHG, and CG contexts in the *clsy* single mutants. (**b**) Boxplots showing the levels of CHG methylation at the hypo CHG DMRs identified in each *clsy* single, double or quadruple mutant compared to each other, WT controls and *pol-iv*. (**c**) Boxplots showing the levels of CG methylation at the hypo CG DMRs that overlap with the reduced 24nt-siRNA clusters identified in each *clsy* single, double or quadruple mutant compared to each other and *pol-iv*. (**d**) Boxplots showing the levels of CG, CHG and CHH methylation at all the reduced 24nt-siRNA clusters identified in each *clsy* single, double or quadruple mutant as compared to each other, WT controls, and *pol-iv*.

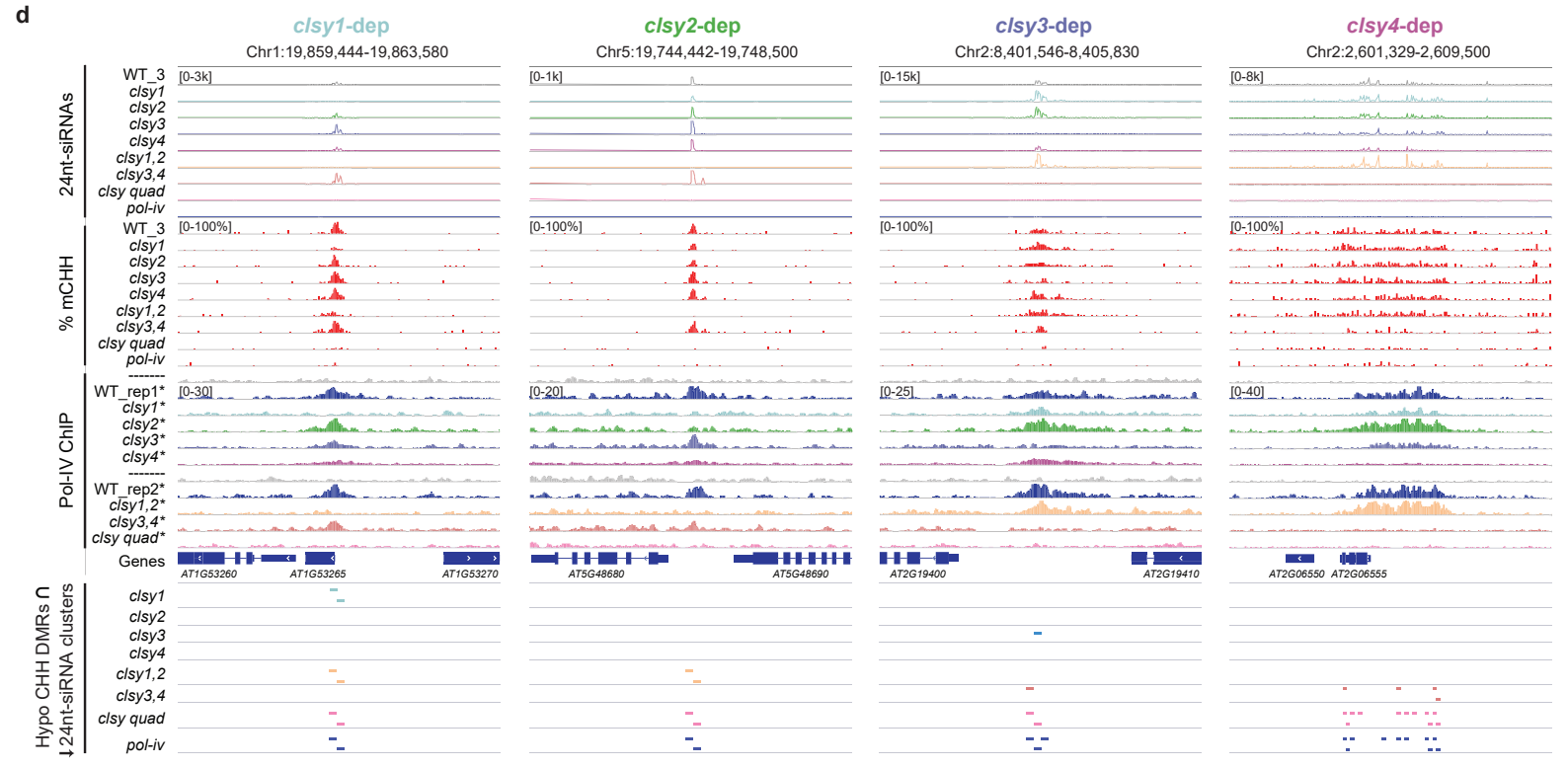


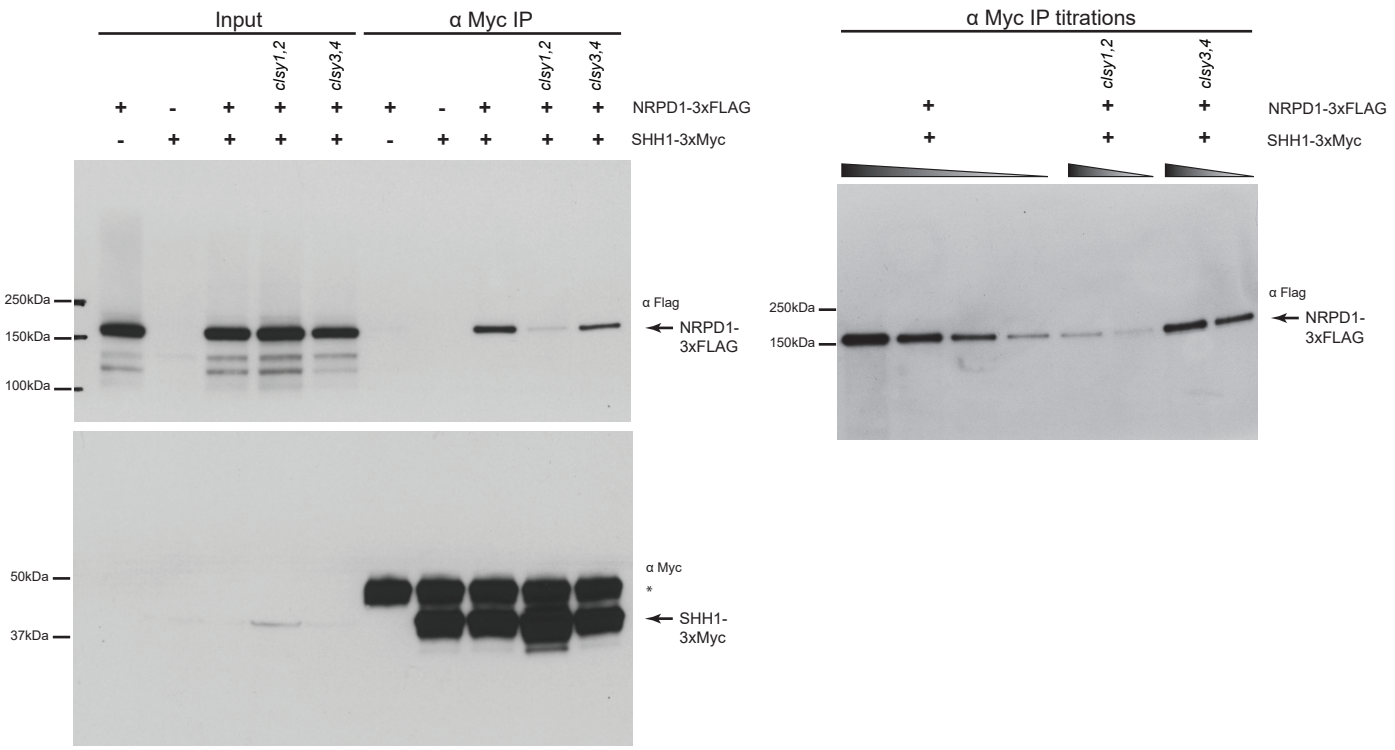

**Supplementary Figure 6. Associations between *clsy*-up-regulated loci, reduced 24nt-siRNA clusters, and DMRs.**
(**a**) Genome browser screen shot showing the levels of 24nt-siRNAs (rp10m), DNA methylation and gene expression (rp10m) at example loci regulated by each CLSY protein. For each set of data, the scale is indicated in brackets, with CG, CHG, and CHH methylation shown in green, blue, and red, respectively. The regions showing the most prominent *clsy*-specific reductions in 24nt-siRNAs and methylation are highlighted in grey. The TAIR10 annotated genes, newly-identified transcripts (TCONS_#), or hypo CHH DMRs (100bp) that overlap with reduced 24nt-siRNA clusters are shown below. (**b**) Scaled Venn diagrams showing the number of up-regulated genes in the *pol-iv* (left) or *clsy* quadruple (right) mutants that overlap with both reduced 24nt-siRNA clusters and ≥1 hypo DMRs in either the CG or non-CG contexts in the *pol-iv* or *clsy* quadruple mutants, respectively. (**c**) Bar plots indicating the number of up-regulated loci shown in **b** that are associated with zero, 1, 2-4, or ≥5 hypo DMRs in the CG or non-CG context that also overlap with reduced 24nt-siRNA clusters. (**d**) Heatmaps and profile plots showing the expression of all the up-regulated TAIR10 genes, unannotated transcripts (un. txn), and TAIR10 repeats shown in Figure 4a, as described in Figure 4b, except in the *clsy2* mutant background. (**e**) Scaled Venn diagram showing the number of up-regulated loci in the *pol-iv* mutant that overlap with both reduced 24nt-siRNA clusters and ≥1 hypo DMR in any sequence context in the *clsy2* mutant. The ∩ symbol indicates the intersection between datasets.




**Supplementary Figure 7. The CLSY family is required for Pol-IV association at chromatin.**

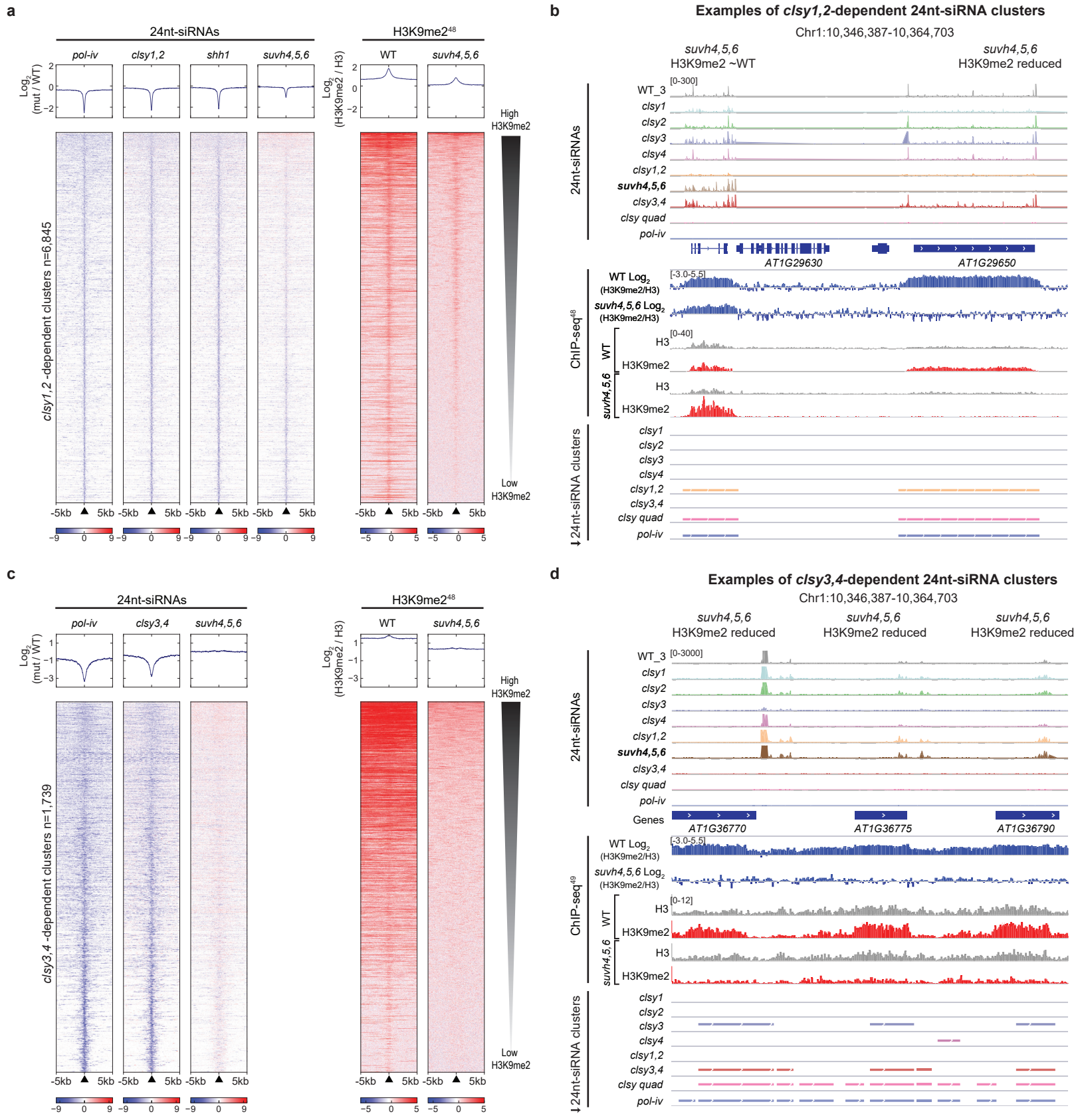
(**a**-**c**) Heatmaps showing Pol-IV enrichment at the indicated *clsy*-dependent 24nt-siRNA clusters in the genotypes shown above. For **a** and **b**, profile plots showing the mean enrichment levels are shown above. All heatmaps are scaled relative to the number of 24nt-siRNA clusters and colored as shown for panel **a**, with the exception of the *clsy2*-dependent clusters that are marked by two asterisks (**) to indicate that this heatmap was expanded by 20-fold due to the small sample size. (**d**) Genome browser screen shots showing the 24nt-siRNA levels (rp10m), percent CHH methylation, and Pol-IV enrichment (rp10m) at loci dependent on CLSY1-4, respectively. For each set of data the scale is indicated in brackets, where k indicates 1000. The TAIR10 annotated genes and hypo CHH DMRs (100bp) that overlap with reduced 24nt-siRNA clusters are shown below. The "-----" indicates WT plants without the *NRPD1-3xFLAG* transgene. The asterisk (*) indicates lines that are also homozygous for the *NRPD1-3xFLAG* transgene and the *nrpd1* mutant.






**Supplementary Figure 8. Co-immunoprecipitation assays showing that CLSY1/2 are required for the association between SHH1 and Pol-IV.**

Uncropped Western blots showing the levels of NRPD1-3xFLAG or SHH1-3xMyc from co-immunoprecipitation (co-IP) experiments in the genetic backgrounds indicated above each lane. For each blot the antibody (α) used is indicated in the upper right corner and the sizes of the protein markers are indicated on the left. An asterisk (*) marks a background band present in the α-Myc IP and the bands corresponding to the NRPD1-3xFLAG and SHH1-3xMyc proteins are marked with arrows. For the IP titrations, the gradient triangles represent a series of 2-fold dilutions starting from undiluted IP samples.




**Supplementary Figure 9. Correlating 24nt-siRNAs and H3K9me2 levels in the *suvh4,5,6* triple mutant at *clsy*-dependent 24nt-siRNA clusters.**

(**a** and **c**) Heatmaps and profile plots showing the levels of 24nt-siRNAs ($Log_2$(mut/WT)) and H3K9 di-methylation ($Log_2$(H3K9me2/H3))[48] in the indicated genotypes at *clsy1,2*-dependent or *clsy3,4*-dependent 24nt-siRNA clusters +/- 5kb, respectively. As indicated by the shaded triangle, the heatmaps for each panel are ordered from high to low based on the mean level of H3K9me2 enrichment in the *suvh4,5,6* triple mutant. (**b** and **d**) Genome browser screen shots showing the levels of 24nt-siRNAs (rp10m) and H3K9me2 modifications at example loci in which H3K9me2 levels remain high or are reduced compared to WT controls. For each set of data, the scale is indicated in brackets, and for the ChIPseq data, both the $Log_2$-normalized tracks (H3K9me2/H3) used to generate the heatmaps in panel a, as well as the individual H3 and H3K9me2 tracks (rp10m) are included. The TAIR10 annotated genes and reduced 24nt-siRNA clusters are shown below.