

---

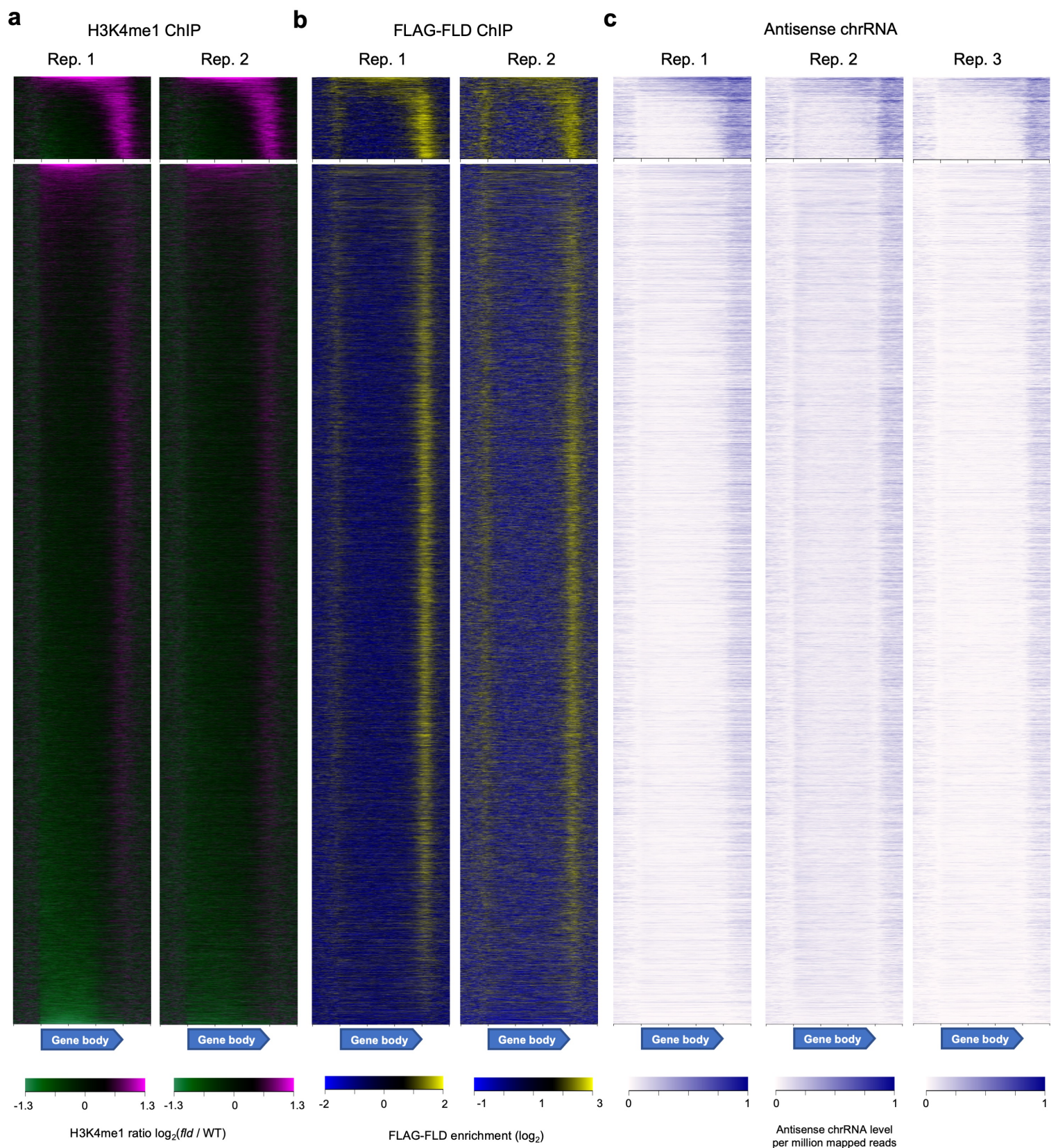
**Supplementary information**

---

**Chromatin-based mechanisms to  
coordinate convergent overlapping  
transcription**

---

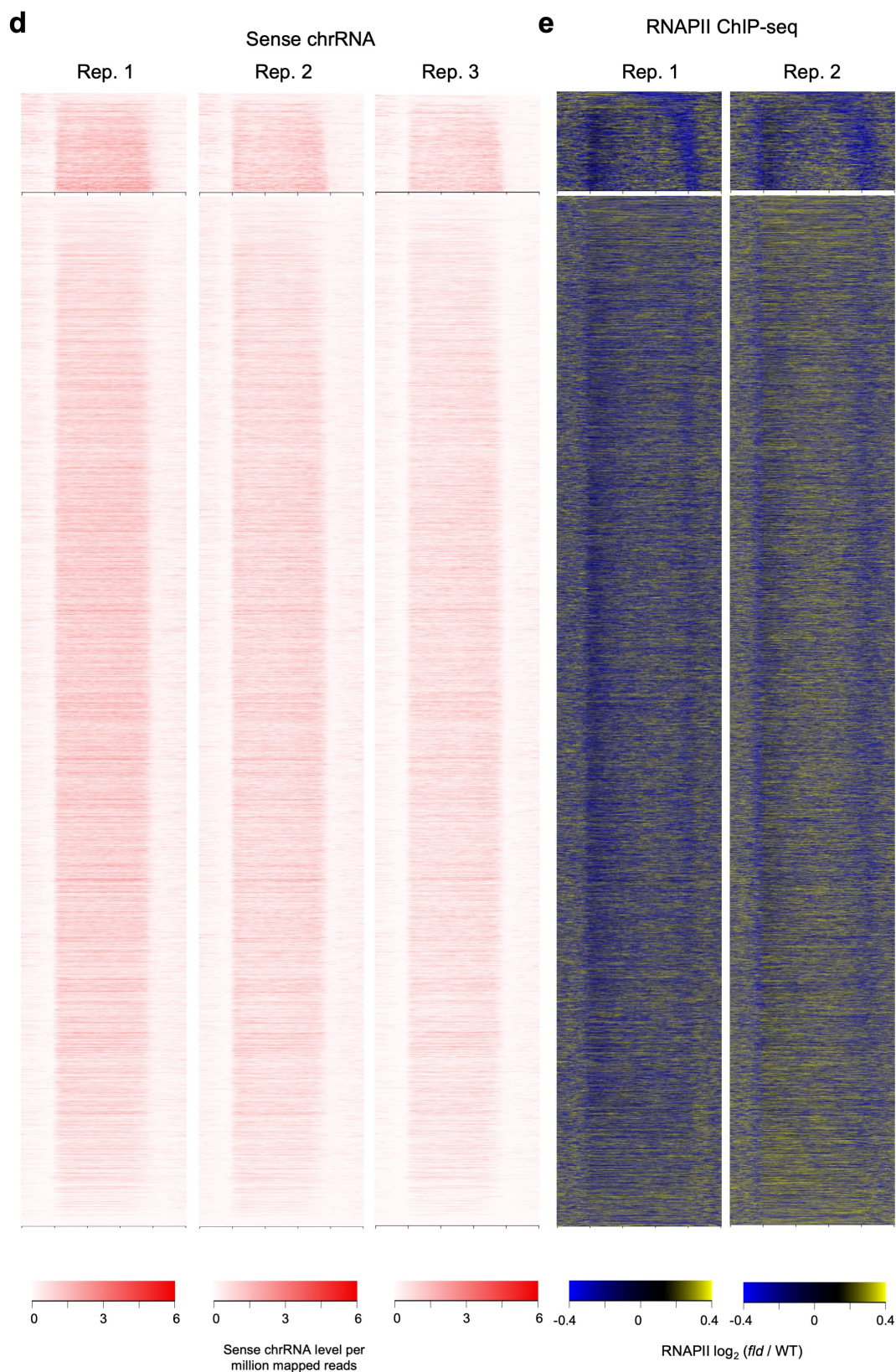
In the format provided by the  
authors and unedited



**Supplementary Fig. 1 Heatmaps for all genes and in biological replicates.**

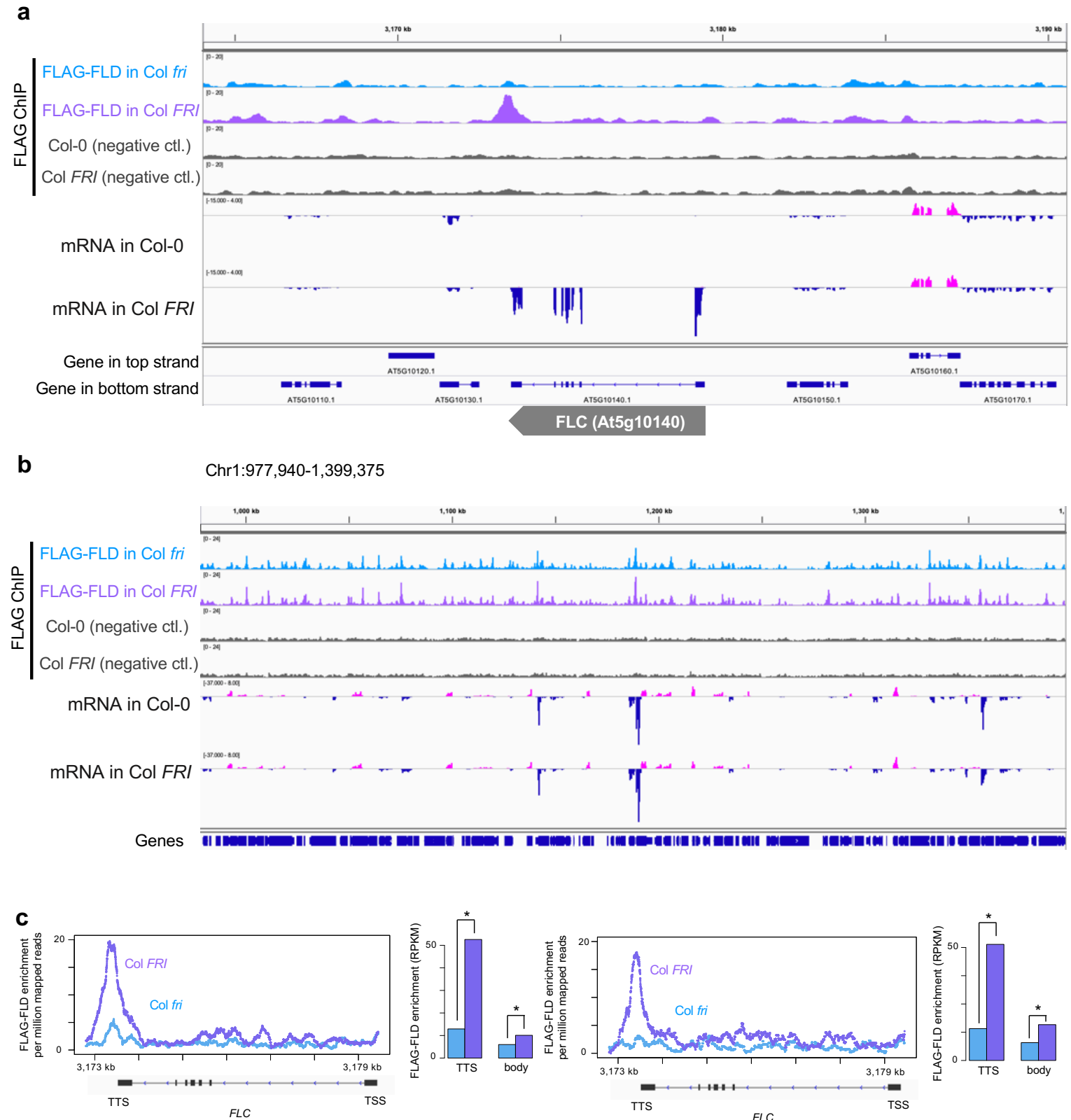
**a**, Heatmaps showing H3K4me1 change in *fld* around all sorted by H3K4me1 change, related to Fig. 1d. In addition to FLD target genes (1,656 genes) shown on top, all other genes are also shown on bottom. Heatmaps for two biological replicates are sorted based on replicate 1. **b**, Heatmap showing FLAG-FLD enrichment, related to Fig. 2a. Genes are divided and sorted in the same way as in **a**. **c**, Heatmap showing antisense chrRNA level in WT, related to Fig. 3b. Genes are divided and sorted in the same way as in **a**.





**Supplementary Fig. 1 (continued)**

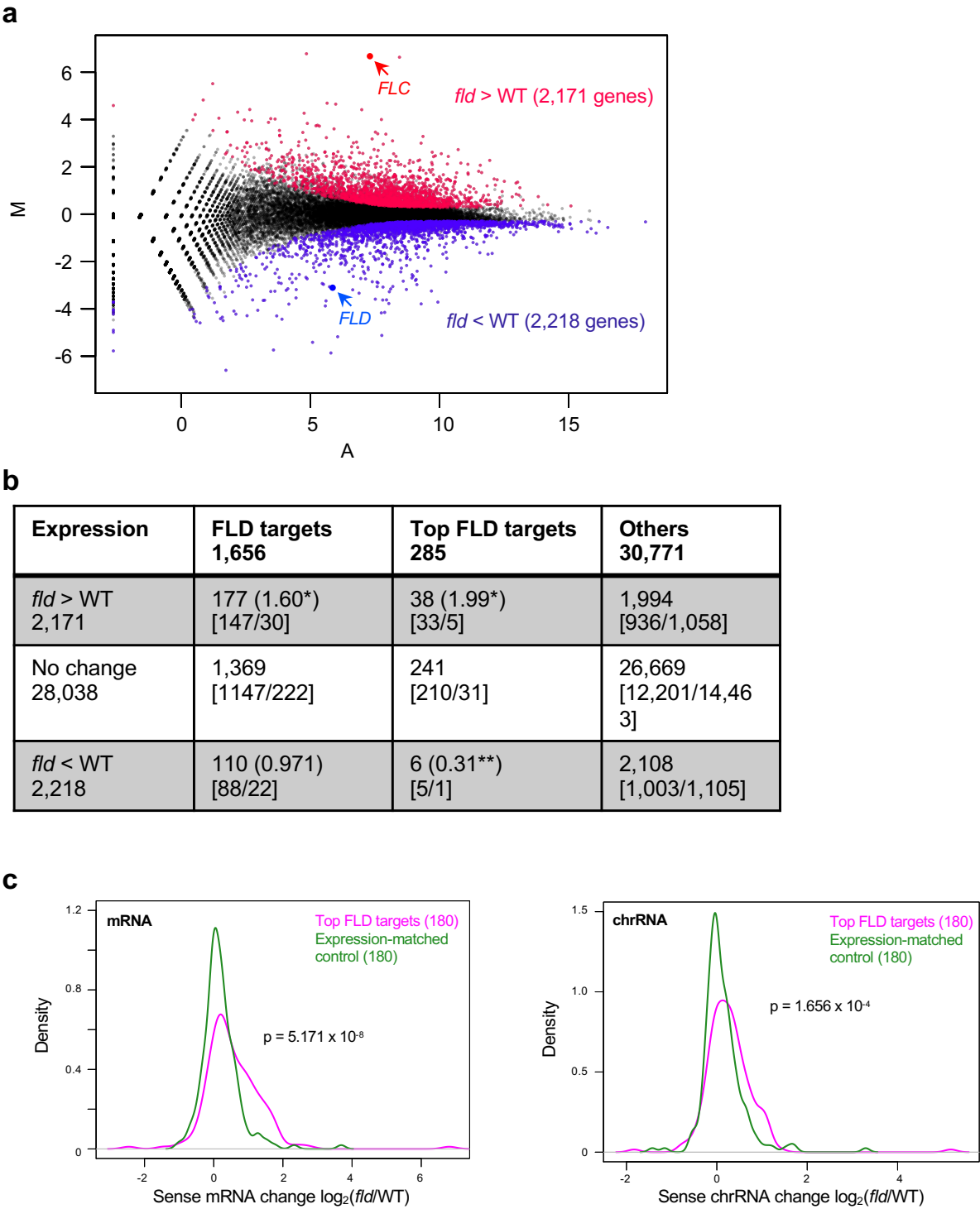
**d**, Heatmap showing antisense chrRNA level in WT. Genes are divided and sorted in the same way as in **a**. **e**, Heatmap showing RNAPII changes between WT and *fld*, related to Fig. 4a and Extended Data Fig. 8a-d. Genes are divided and sorted in the same way as in **a**.



### Supplementary Fig. 2 Effects of FRI-induced transcription on FLD localization.

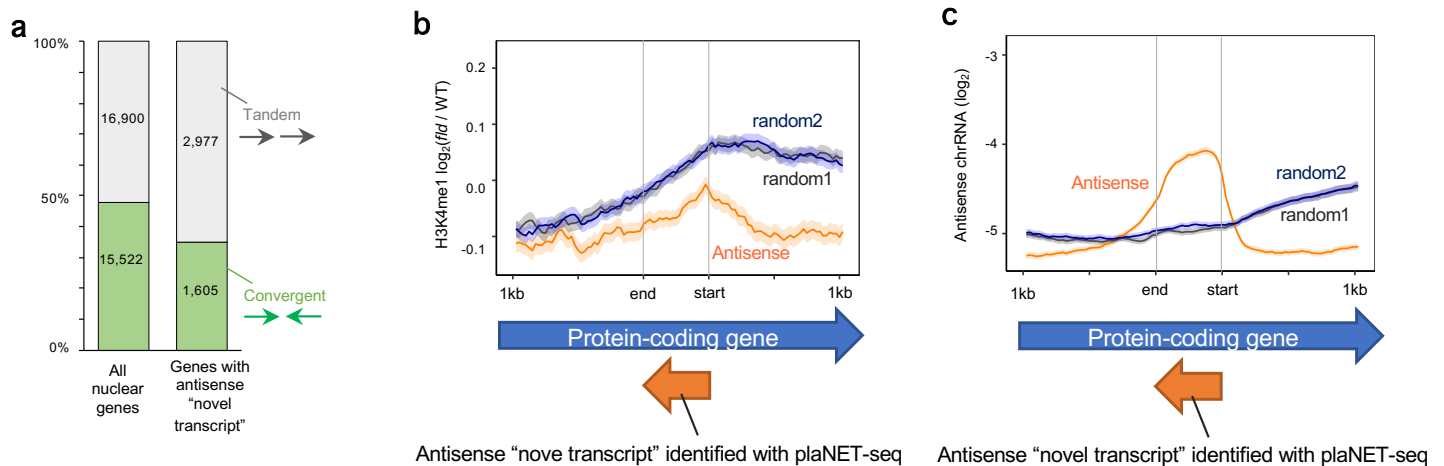
**a, b,** Browser views of FLD localization and mRNA (top strand, magenta; bottom strand, dark blue) comparing the *FLC* nonexpressor, *Col fri*, and the *FLC* expressor, *Col FRI*. FLAG ChIP-seq of nontransgenic *Col-0* plants and *Col FRI* plants were used as negative controls. The pattern around the *FLC* locus (**a**) and the global pattern (**b**) are shown. **c,** Results for biological replicates of the result shown in Fig 2d. For bar graphs, RPKM (read counts per kilobases per million mapped reads) are calculated for 1kb region around TTS and the rest of gene body of *FLC* separately. Asterisks indicate  $p < 2.2 \times 10^{-16}$  for two-sided Fisher's exact test using the numbers of uniquely mapped total reads as sample sizes.





**Supplementary Fig. 3 Effect of *fld* on transcriptome.**

**a**, MA plot comparing mRNA-seq in WT and *fld*. Values from 3 biological replicates are averaged. Significantly (FDR < 0.05) upregulated (2,171 genes) and downregulated (2,218 genes) in *fld* compared to WT were colored in red and purple, respectively. *FLC* and *FLD* genes are highlighted. **b**, Table summarizing effect of *fld* on H3K4me1 and mRNA. The numbers in each cell indicates the number of genes. The numbers in parentheses indicate the enrichment values and asterisks indicate significant overrepresentation tested by hypergeometric test ( $p = 3.35 \times 10^{-10}$ , and  $4.31 \times 10^{-5}$ , respectively). Two asterisks denote significant underrepresentation ( $p = 2.36 \times 10^{-4}$ ). The numbers in brackets indicate the numbers of convergent and tandem genes before and after slashes, respectively. **c**, Kernel density plots for ratios of mRNA (left) and chrRNA (right) between WT and *fld*. Ratios for Top FLD targets are plotted as Fig. 3f and 3g, respectively. “Expression-matched control” genes are selected based on two criteria; the genes other than FLD targets, and shows closest chrRNA level (RPKM) to each Top FLD target. The  $p$ -value is based on two-sided Mann-Whitney-Wilcoxon test comparing two groups of genes.



**Supplementary Fig. 4 Novel transcripts identified with plaNET-seq and its relationship with H3K4me1.**

**a**, Proportion of convergent and tandem genes in all nuclear genes (in TAIR10; 32,422) and protein-coding genes that are associated with "novel transcripts" identified plaNET-seq<sup>20</sup>, in the antisense strand (4,582). Convergent genes are under-represented in that fraction of genes (hypergeometric  $p = 3.89 \times 10^{-80}$ ). **b**, **c**, Averaged profiles of H3K4me1 change between in WT and in *fld* (**b**), and antisense chrRNA (**c**), around "novel transcripts" antisense to protein-coding genes (4,582). The arrangement of protein-coding gene and antisense transcript pair is shown in the bottom. Two randomly selected populations (7,270 and 7,283, respectively) of size-matched regions within protein-coding genes are shown as controls. The ribbon indicates 95% confidence interval for the mean.