nature plants

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41477-017-0100-y

In the format provided by the authors and unedited.

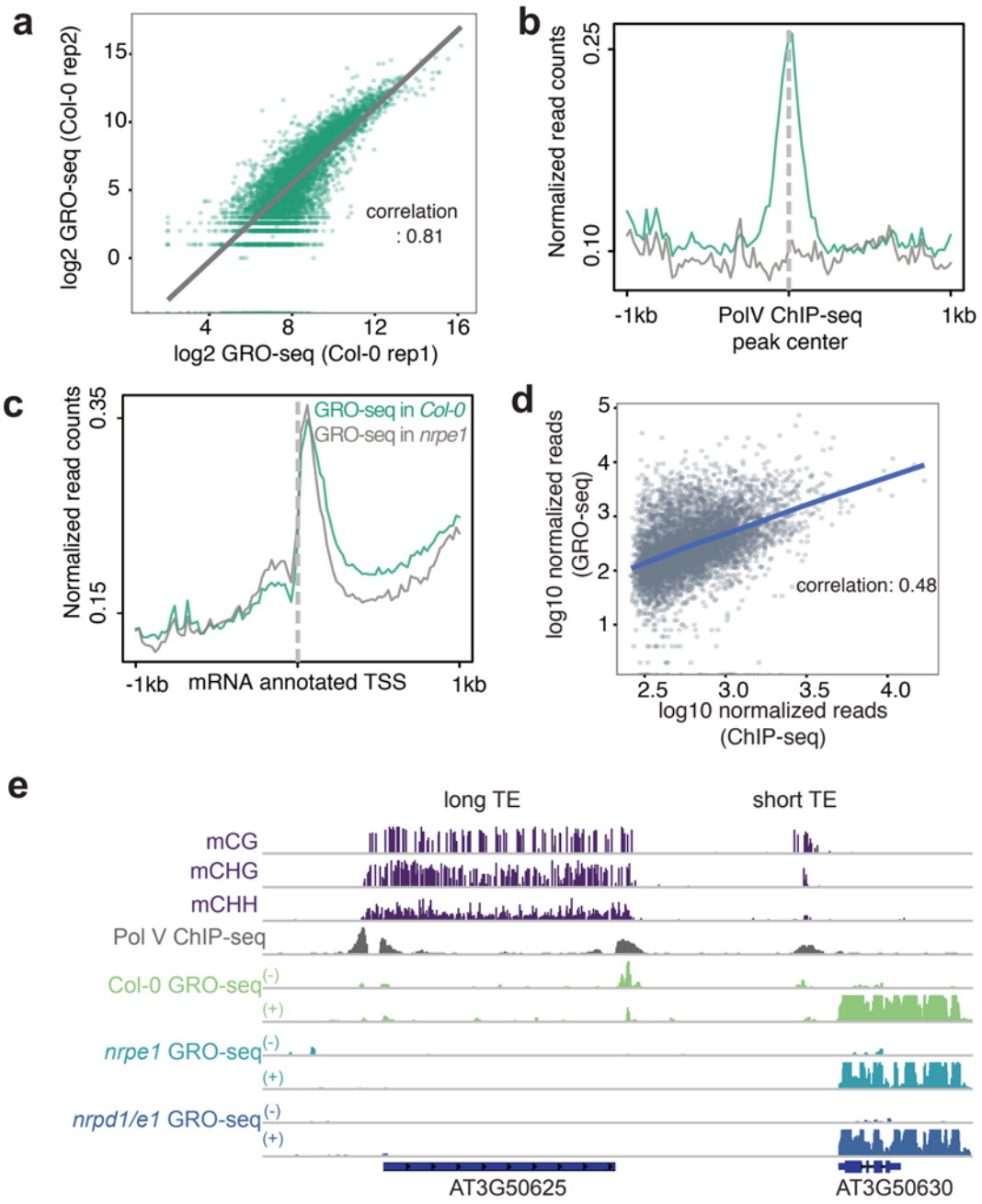
RNA-directed DNA methylation involves co-transcriptional small-RNA-guided slicing of polymerase V transcripts in *Arabidopsis*

Wanlu Liu[1,2], Sascha H. Duttke[3,4,5,6], Jonathan Hetzel[3,4,5], Martin Groth[2], Suhua Feng[2,7], Javier Gallego-Bartolome[2], Zhenhui Zhong[2,8], Hsuan Yu Kuo[2], Zonghua Wang[8], Jixian Zhai[2,9], Joanne Chory[3,4,5] and Steven E. Jacobsen[1,2,7,10]*

[1]Molecular Biology Institute, University of California at Los Angeles, Los Angeles, CA, USA. [2]Department of Molecular, Cell and Developmental Biology, University of California at Los Angeles, Los Angeles, CA, USA. [3]Plant Biology Laboratory, Salk Institute for Biological Studies, La Jolla, CA, USA. [4]Division of Biological Sciences, University of California at San Diego, La Jolla, CA, USA. [5]Howard Hughes Medical Institute, Salk Institute for Biological Studies, La Jolla, CA, USA. [6]Department of Cellular & Molecular Medicine, School of Medicine, University of California at San Diego, La Jolla, CA, USA. [7]Eli & Edythe Broad Center of Regenerative Medicine & Stem Cell Research, University of California at Los Angeles, Los Angeles, CA, USA. [8]State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, College of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou, China. [9]Institute of Plant and Food Science, Department of Biology, Southern University of Science and Technology, Shenzhen, China. [10]Howard Hughes Medical Institute, University of California at Los Angeles, Los Angeles, CA, USA. Wanlu Liu and Sascha H. Duttke contributed equally to this work. *e-mail: jacobsen@ucla.edu

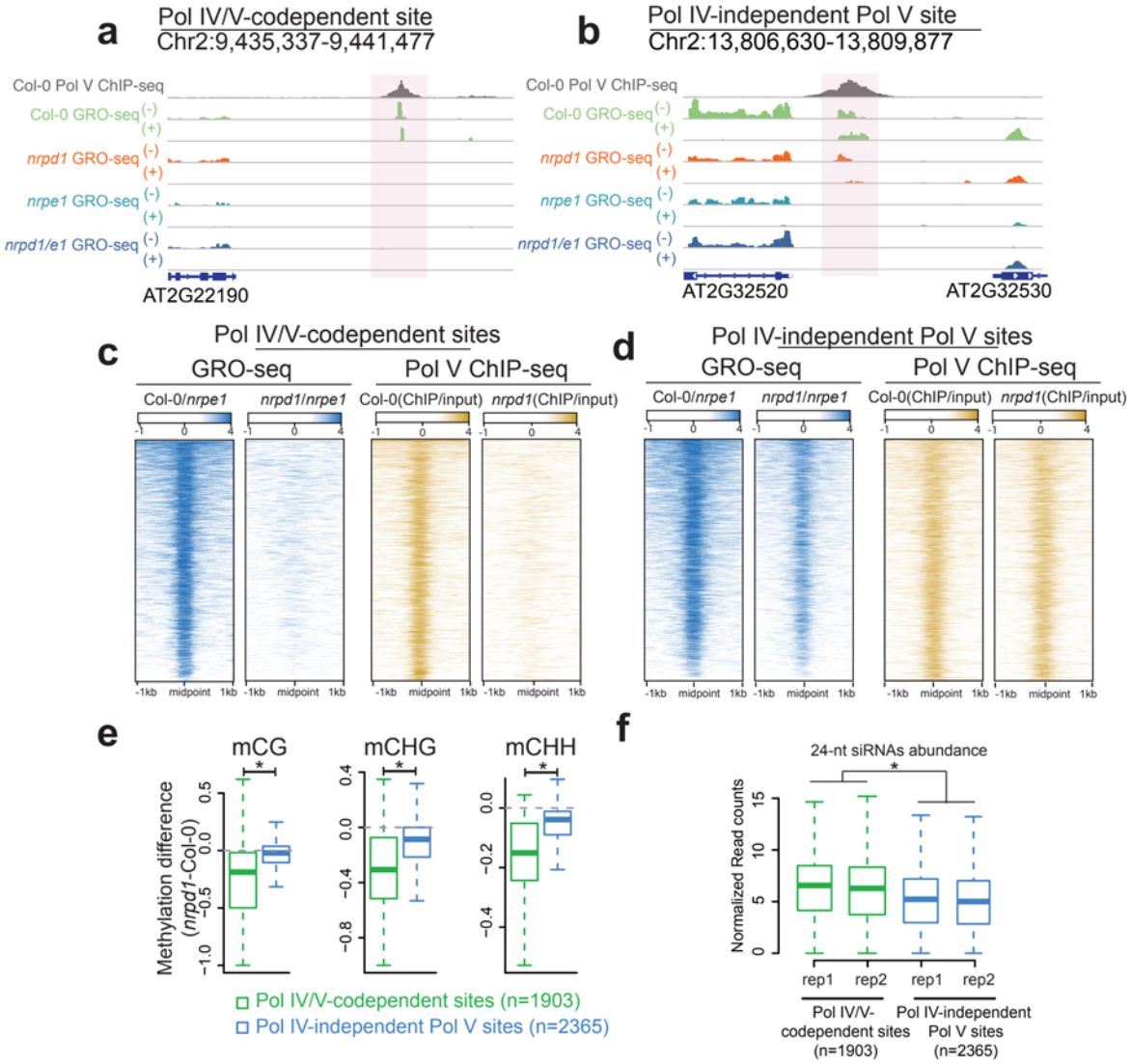
© 2018 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Supplementary Fig.1



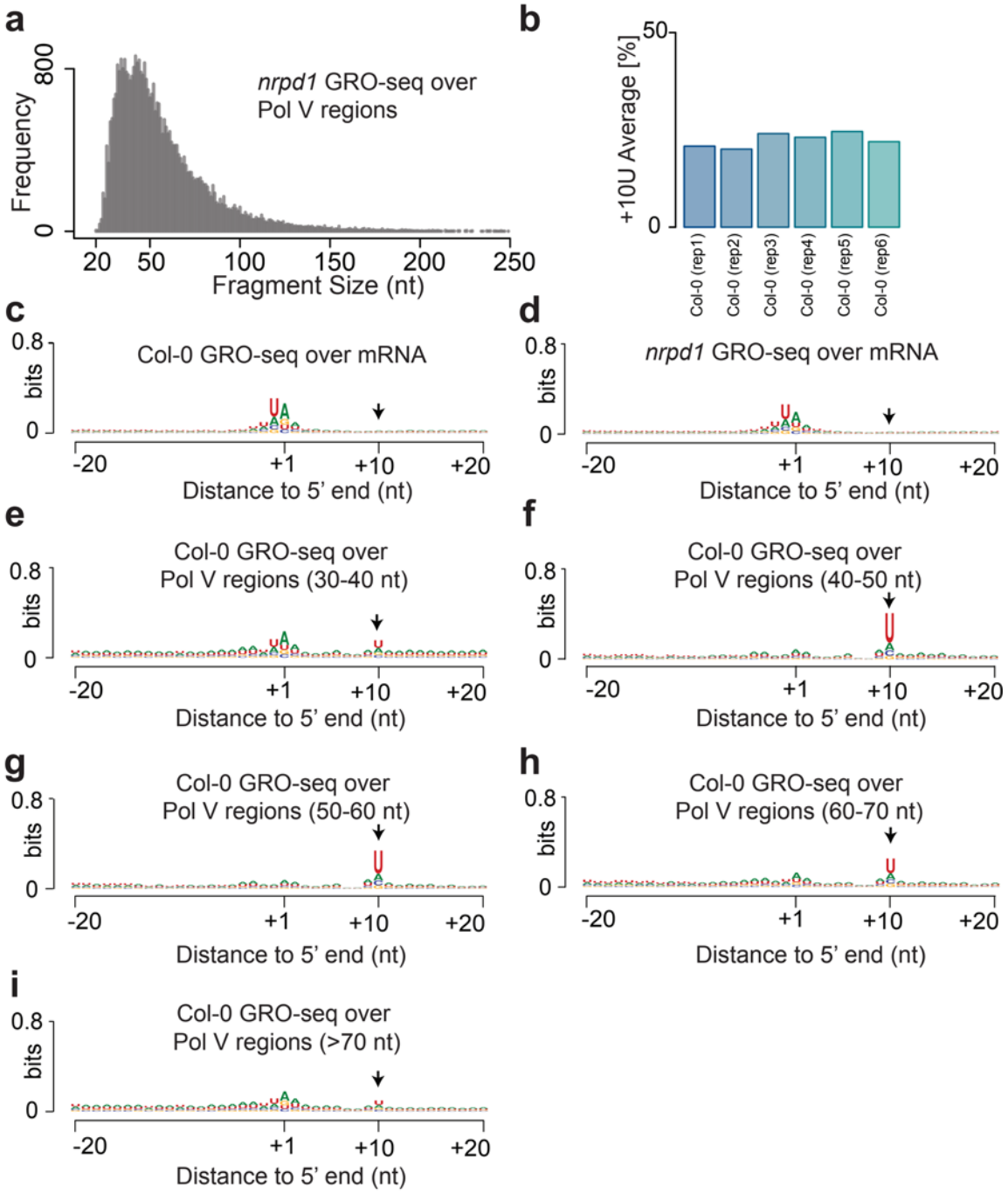
Supplementary Figure 1. Modified GRO-seq is able to capture nascent Pol V-dependent transcripts. **a,** Scatterplot of signals from two independent GRO-seq experiments in Col-0. The Pearson's correlation coefficient is calculated and shown on the plot. **b,** Metaplot showing GRO-seq signals over Pol V-occupied regions in Col-0 and *nrpe1*. **c,** Metaplot showing GRO-seq signals over annotated genes in Col-0 and *nrpe1*. **d,** Scatterplot of normalized signals from Pol V ChIP-seq versus GRO-seq in Col-0. The Pearson's correlation coefficient are calculated and shown on the plot. **e,** Genome browser screenshot for CG, CHG, and CHH methylation in Col-0, Pol V ChIP-seq signals in Col-0, and GRO-seq signals in Col-0, *nrpe1*, and *nrpd1/e1* of a representative long TE and a representative short TE. Plus (+) and Minus (-) indicate the strandness of GRO-seq signal.

Supplementary Fig.2



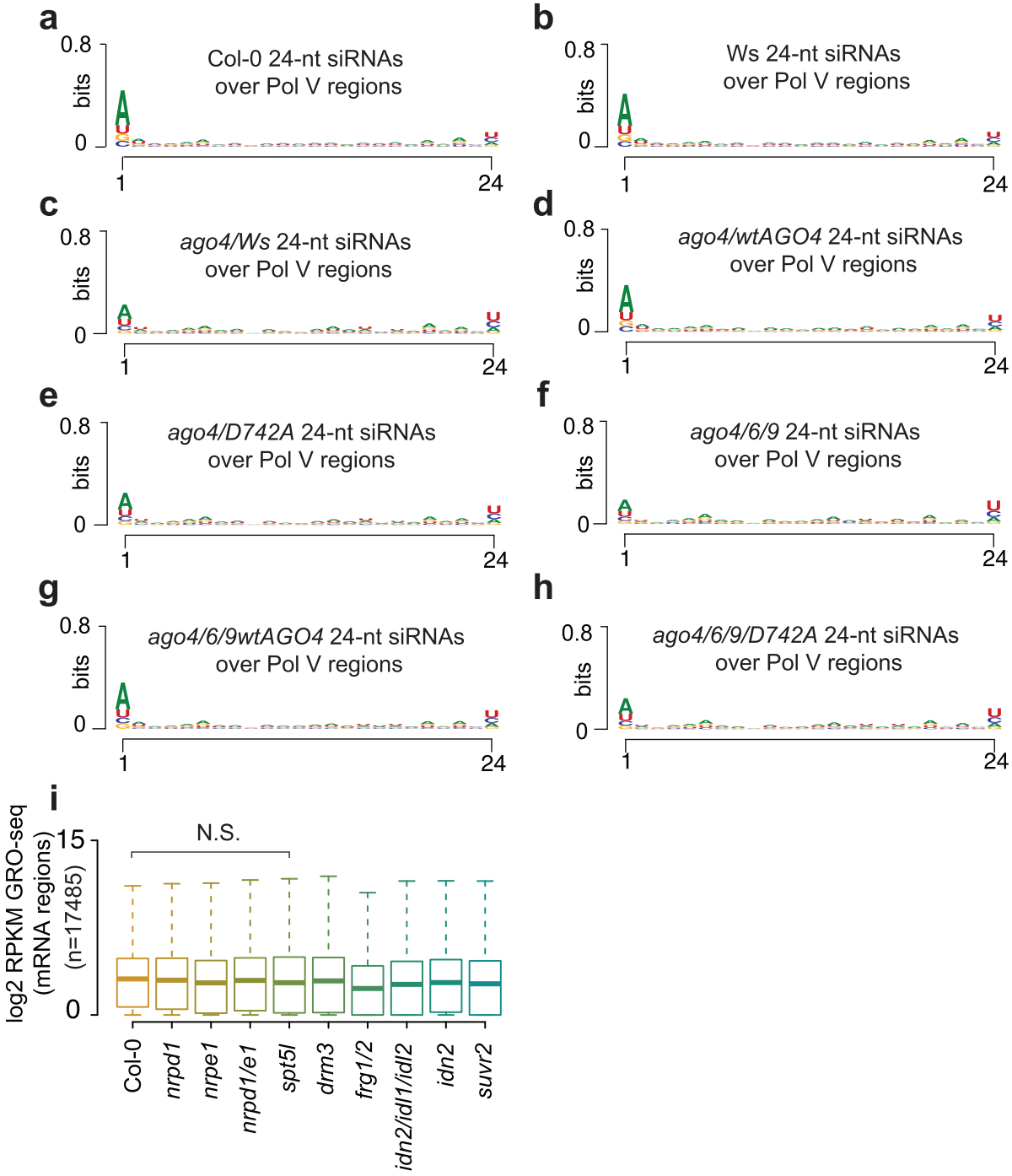
Supplementary Figure 2. Characterization of Pol IV/V-codependent sites and Pol IV-independent Pol V sites. **a,b,** Genome browser screenshot for Pol V ChIP-seq signals in Col-0 and GRO-seq signals in Col-0, *nrpe1*, *nrpd1*, and *nrpd1/e1* of a representative Pol IV/V-codependent site (**a**) and Pol IV-independent Pol V site (**b**) . Plus (+) and Minus (-) indicate the strandness of GRO-seq signal. **c,d,** Heatmap of log2 ratio of GRO-seq in Col-0 vs. *nrpe1*, GRO-seq in *nrpd1* vs. *nrpd1*, Pol V ChIP signals in Col-0, and Pol V ChIP-seq signals in *nrpd1* plotted over Pol IV/V-codependent sites (**c**) and Pol IV-independent Pol V sites (**d**). **e,** Boxplot of CG, CHG, and CHH methylation difference in *nrpd1* vs. Col-0. **p-value < 0.05* (Welch Two Sample t-test). **f,** Normalized 24-nt siRNAs abundance in Col-0 over Pol IV/V-codependent sites and Pol IV-independent Pol V sites. **p-value < 0.05* (Welch Two Sample t-test).

Supplementary Fig.3



Supplementary Figure 3. Pol V transcripts with different lengths are sliced. **a,** Size distribution of nascent transcripts in *nrpd1* over Pol V-dependent regions. Replicates were merged for this plot. **b,** The percentage of U presented over genomic average at position 10 from the 5' ends of nascent transcripts captured with GRO-seq in six biological replicates for Col-0. **c,d,** The relative nucleotide bias of each position in the upstream and downstream 20-nt of nascent RNAs generated from the top 1,000 expressed annotated gene regions in Col-0 (**c**) and nrpd1 (**d**). Replicates were merged for plot (**c-d**). **e-i**, The relative nucleotide bias of each position in the upstream and downstream 20-nt of nascent transcripts of 30- to 40-nt long (**e**), 40- to 50-nt long (**f**), 50- to 60-nt long (**g**), 60- to 70-nt long (**h**) and 70-nt and longer (**i**) captured in Col-0. Replicates were merged for plot (**e-i**).

Supplementary Fig.4



Supplementary Figure 4. 24nt-siRNAs retain strong enrichment of A at position 1 for *ago4*, *ago4/6/9* mutant and *ago4* or *ago4/6/9* mutant expressing wtAGO4 or D742A. **a-h,** The relative nucleotide bias of each position for 24-nt siRNAs over Pol V dependent regions in Col-0 (**a**), Ws (**b**), *ago4/Ws* (**c**), *ago4/wtAGO4* (**d**), *ago4/D742A* (**e**), *ago4/6/9* (**f**), *ago4/6/9/wtAGO4* (**g**) and *ago4/6/9/D742A* (**h**). **i,** Boxplot of normalized GRO-seq signals from top 1,000 expressed annotated gene in Col-0, nrpd1, *nrpe1*, *nrpd1/e1*, *spt5l*, *drm3*, *frg1/2*, *idn2/idl1/idl2*, *idn2*, and *suvr2*. N.S., not significant.