CRISPR targeting of H3K4me3 activates gene expression and unlocks centromeric crossover recombination in Arabidopsis

Jenia Binenbaum[1,*], Vanda Adamkova[1,*], Hannah Fryer[1], Linhao Xu[1], Nicola Gorringe[1], Piotr Włodzimierz[1,3], Robin Burns[1], Ashot Papikian[2,4], Steven E. Jacobsen[2,5], Ian Henderson[1], C. Jake Harris[1,#]

[1] Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK.

[2] Department of Molecular, Cell and Developmental Biology, University of California, Los Angeles, CA, USA.

[3] Current Address: Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland.

[4] Current Address: Plant Molecular and Cellular Biology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, USA.

[5] Howard Hughes Medical Institute, University of California, Los Angeles, CA, USA.

* These authors contributed equally.

[#] For correspondance: cjh92@cam.ac.uk

Abstract:

H3K4me3 is a fundamental and highly conserved chromatin mark across eukaryotes, playing a central role in many genome-related processes, including transcription, maintenance of cell identity, DNA damage repair, and meiotic recombination. However, identifying the causal function of H3K4me3 in these diverse pathways remains a challenge, and we lack the tools to manipulate it for agricultural benefit. Here we use the CRISPR-based SunTag system to direct H3K4me3 methyltransferases in the model plant, *Arabidopsis thaliana*. Targeting of SunTag-SDG2 activates the expression of the endogenous reporter gene, *FWA*. We show that SunTag-SDG2 can be employed to increase pathogen resistance by targeting the H3K4me3-dependent disease resistance gene, *SNC1*. Meiotic crossover recombination rates impose a limit on the speed with which new traits can be transferred to elite crop varieties. We demonstrate that targeting of SunTag-SDG2 to low recombining centromeric regions can significantly stimulate crossover formation. Finally, we reveal that the effect is not specific to SDG2 and is likely dependent on the H3K4me3 mark itself, as the orthogonal mammalian-derived H3K4me3 methyltransferase, PRDM9, produces a similar effect on gene expression with reduced off-target potential. Overall, our study supports an instructive role for H3K4me3 in transcription and meiotic recombination and opens the door to precise modulation of important agricultural traits.

Introduction:

Re-programming of transcriptional landscapes in response to developmental and environmental cues critically underlies a plant's ability to adapt and optimise their performance. The potential to re-wire and fine-tune expression of endogenous genes therefore holds great promise for agricultural improvement. Chromatin modifications associated with gene expression play an essential role in this process (Lloyd and Lister, 2021).

Trimethylation of histone 3 lysine 4 (H3K4me3) is a well-studied and highly conserved chromatin mark. It's enrichment around the transcriptional start site of expressed genes is observed across eukaryotes (Schwaiger et al., 2014), and is thought to assist in transcriptional maintenance and memory of cell identity (Wang and Helin, 2024). Despite the ubiquity of H3K4me3 and its positive correlation with gene expression, whether it plays an instructive role in transcription remains a matter of debate, due in part to the limited evidence of transcriptional impact when H3K4me3 levels are perturbed (Francisco et al., 2023; Howe et al., 2017). In animals, TAF3 directly binds to H3K4me3 and recruits basal transcription factor machinery, providing a plausible mechanism to stimulate transcription (Lauberth et al., 2013; Vermeulen et al., 2007). Recent work in embryonic stem cells shows that H3K4me3 plays a role in RNA Polymerase II pause-release and elongation, rather than transcriptional initiation (H. Wang et al., 2023). In plants, there are no TAF3 homologues and whether similar mechanisms exist to promote transcription downstream of H3K4me3 remains unknown.

H3K4me3 also plays a key role in several other genome-related processes. For instance, meiotic recombination, which allows for genetic exchange between parental chromosomes (Borde et al., 2009). This process is essential for breeders to transfer traits from wild relatives into elite cultivars. However, the positioning of meiotic recombination events is non-uniform across the genome, with gene-rich chromosomal arms having relatively higher rates of recombination, and centromeric repeat-rich regions being more recombination supressed (Rowan et al., 2019; Tock et al., 2021). Loci encoding desirable traits that reside within these low-recombining regions of the genome are largely intractable to selection by traditional breeding approaches.

Further, in mammals, the H3K4me3 SET domain methyltransferase containing protein, PRDM9, helps to define active sites of crossover recombination (Powers et al., 2016). While H3K4me3 levels are positively correlated with recombination rates in plants (Choi et al., 2013; Rowan et al., 2019), whether deposition of H3K4me3 is sufficient to promote crossover formation is entirely unknown.

One way to reveal the causal role of an epigenetic modification is to directly deposit the mark at targeted genomic locations (Dubois and Roudier, 2021; Harris et al., 2023; Policarpi and Hackett, 2021; Villiger et al., 2024; Wang and Yamaguchi, 2024). Epigenome editing tools, such as the CRISPR-based SunTag system (Tanenbaum et al., 2014), have proven to be highly efficient at targeting DNA methylation addition, removal and histone modification (Gallego-Bartolomé et al., 2018; Papikian et al., 2019; Pflueger et al., 2018; Policarpi et al., 2024; M. Wang et al., 2023), although their utility for modulating agronomically useful traits is only just beginning to be realised (Selma and Orzáez, 2021), and is limited by the range of possible marks that can be directed (Lloyd and Lister, 2021). In the model plant *Arabidopsis thaliana*, SDG2 is the main histone methyltransferase responsible for deposition of H3K4me3 (Guo et al., 2010). Here, we describe CRISPR directed H3K4me3 deposition in plants and demonstrate its utility in a range of scenarios. We translationally fuse the SDG2 methyltransferase domain to the scFv single chain antibody, which can be recruited to the SunTag epitope chain connected to dead Cas9 (dCas9). We show that SunTag-SDG2 mediated H3K4me3 deposition can activate the expression of endogenous genes. Further, by targeting a disease resistance gene, we generate plants with enhance resistance capabilities. Next, we show that SunTag-SDG2 can unlock meiotic crossover recombination across centromeric regions. Finally, we incorporate the methyltransferase domain from PRDM9 into SunTag – which was recently shown to instruct transcription in a mammalian context (Policarpi et al., 2024) – showing that it is also highly efficient for transcriptional activation in plants, and displays reduced off-target potential. These results demonstrate the power and versatility of a functional H3K4me3 targeting system for modulation of critical genome templated processes, such as transcriptional enhancement and meiotic recombination.

102 **Results:**

103

104 **SunTag SDG2 mediates transcriptional activation of FWA**

105

106 The SunTag system we chose is composed of three components (Gardiner et al.,
107 2022; Tanenbaum et al., 2014) (Figure 1A): 1) A catalytically deactivated (nuclease
108 deficient) Cas9 (dCas9) translationally fused to a tail with 10x copies of the GCN4
109 epitope, each separated by a 22 amino acid flexible linker. 2) An effector module that
110 encodes a single chain fragment variable (scFv) antibody that recognises GCN4,
111 super folder GFP (sfGFP), and the effector of interest (e.g. enzymatic modifier /
112 recruitment scaffold). 3) *U6* promoter driven CRISPR guide RNA. When the three
113 components are expressed within the same cell, the guide RNA recruits the dCas9-
114 10xGCN4 to the target locus of interest, while the GCN4 single chain antibody
115 containing effector module binds to the dCas9-10xGCN4 epitope tail. The system
116 allows for conformational flexibility and the potential for high stochiometric
117 concentration of the effector at the locus of interest.

118

119 We inserted the C-terminal coding sequence of the H3K4 methyltransferase, SDG2
120 (amino acids 1571-2335), into the SunTag system (Figure 1B, Supp. Fig. S1). As a
121 control, we also generated a version of SDG2 with an amino acid change (Y1903F)
122 predicted to abolish catalytic activity (Supp. Fig. S2) (Du et al., 2014; Guo et al., 2010).
123 Hereafter, we refer to these as SDG2 and dSDG2 (for catalytically deactivated). We
124 targeted the SunTag system to the promoter of the epigenetically repressed *FWA* gene
125 (Soppe et al., 2000), using the previously published CRISPR guide RNA 4, which is
126 complementary to two tandem repeat regions directly upstream of the *FWA*
127 transcription start site (TSS) (Gallego-Bartolomé et al., 2018). These were
128 transformed into the *rdr6* background, as previous results have shown that SunTag is
129 less prone to silencing and achieves higher gene expression (M. Wang et al., 2023).
130 We observed that SunTag:SDG2:FWA_g4 activated *FWA* mRNA expression, while the
131 SunTag:dSDG2:FWA_g4 and no guide (SunTag:SDG2:No_g) RNA controls did not
132 (Fig 1C, Supp. Fig. 1C). Importantly, effector module expression was detected in all
133 constructs, indicating that both guide RNA targeting and catalytic activity are required
134 for *FWA* activation (Sup Fig 3A). To validate the presence of the SunTag and test for
135 the deposition of H3K4me3, we performed ChIP qPCR and observed significant

136 enrichment both for the presence of SunTag and for H3K4me3 accumulation at the
137 target site in these transgenic lines (Figure 1D and Sup Fig 3B), confirming that
138 SunTag-SDG2 binds the *FWA* locus and deposits H3K4me3. To investigate
139 SunTag:SDG2:FWA_g4 binding genome-wide, we performed high-throughput
140 sequencing (anti-HA ChIP-seq), which confirmed the previously reported high
141 specificity of SunTag targeting (Papikian et al., 2019), with the *FWA* target locus as
142 the top binding target (Figure 1E, Supp. Fig. 4). Only 3 conserved binding sites were
143 identified in the SunTag:SDG2:FWA_g4 lines; *FWA* and two off-target sites previously
144 identified as having sequence-similarity to FWA_g4 (Papikian et al., 2019).
145 Furthermore, transcriptome analyses confirmed *FWA* as the highest fold change
146 differentially expressed upregulated gene (Figure 1F). Together, these data indicate
147 that deposition of H3K4me3 by SunTag:SDG2 is sufficient to overcome epigenetic
148 silencing and drive transcriptional activation in a locus-specific manner.

Figure 1

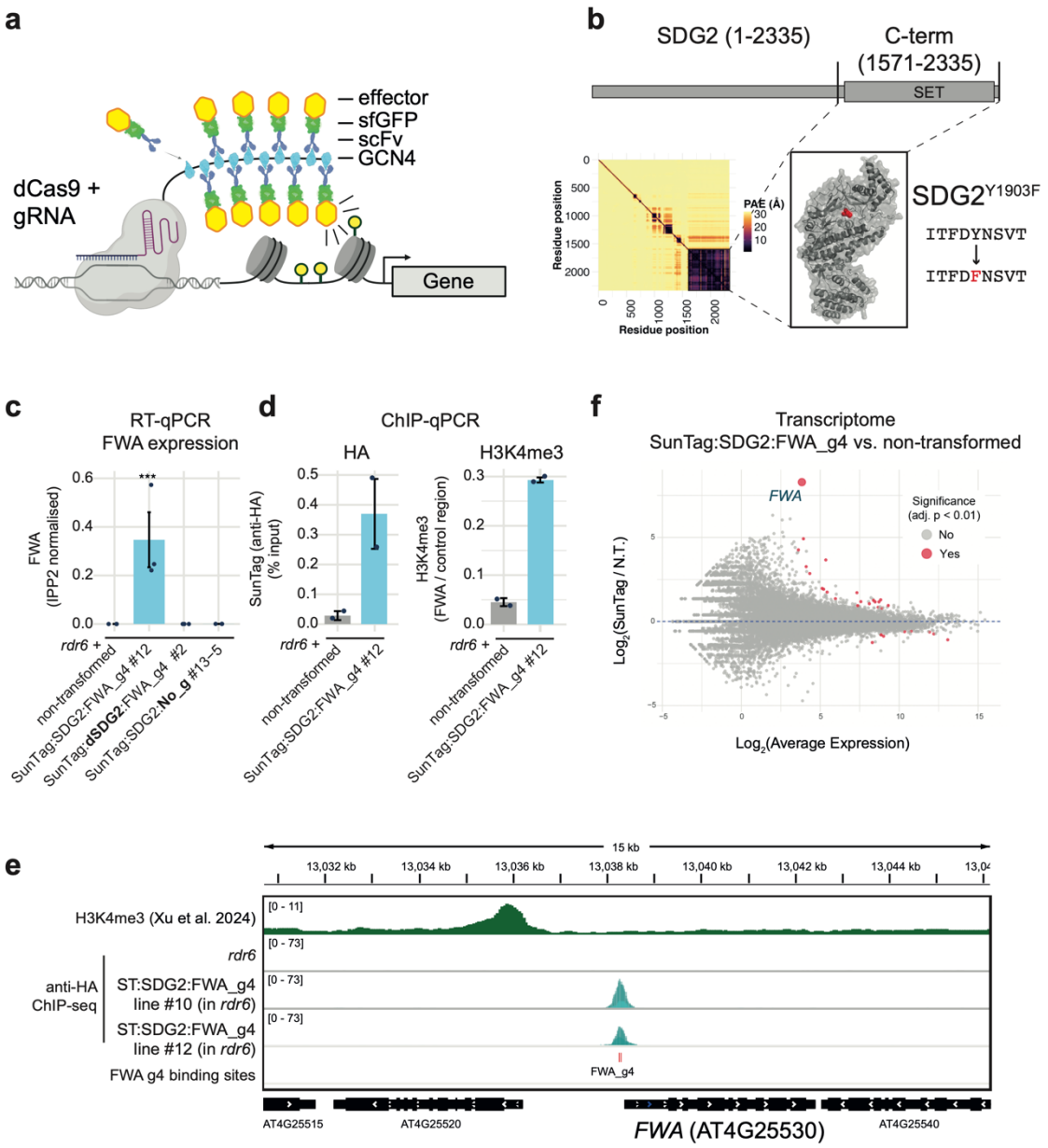


Figure 1: SunTag:SDG2 activates *FWA* mRNA expression.

A) Schematic showing the SunTag system for modulation of chromatin (Gardiner et al., 2022; Papikian et al., 2019). B) Upper panel shows SDG2 with catalytic SET domain (drawn approximately to scale). Lower left depicts predicted alignment error (PAE) plot from AlphaFold3 (Abramson et al., 2024) prediction of the SDG2 coding sequence with a box drawn around the region cloned into SunTag (amino acids 1571-2335). Lower right depicts the AlphaFold3 model for $SDG2_{1571-2335}$, with the position of the amino acid change (Y1903F) for dSDG2 indicated (shown in red). C) RT-qPCR for *FWA* expression from the genotypes indicated. Error bars represent SEM from three biological replicates. *** indicates p-value

159 <0.001 (students T-test). D) ChIP qPCR for the presence of SunTag (left panel) and H3K4me3
160 enrichment (right panel) over the TSS of *FWA*. The control region (AT5G65130) used for
161 normalisation has high H3K4me3. Error bars represent SEM from 2 biological replicates. E)
162 Genome browser image showing the enrichment of two independent SunTag:SDG2:FWA_g4
163 lines (2 biological replicates each, bio replicate tracks are overlaid) at the *FWA* TSS. F) MA
164 plot comparing the transcriptome of SunTag:SDG2:FWA_g4 as compared to the non-
165 transformed control (*rdr6*). Differentially expressed genes are show in in red (FDR < 0.01),
166 with *FWA* labelled and enlarged for visibility.
167
168 ***SNC1* mediated modulation of disease resistance**
169
170 Given that SunTag:SDG2 can modulate chromatin and activate gene expression, we
171 were interested use the tool to modify an important plant trait. Gene expression during
172 pathogen attack is a key factor in determining disease outcome. *SNC1* is disease
173 resistance gene, encoding a nucleotide-binding leucine-rich repeat protein (M. Y.
174 Wang et al., 2023) that sits within the partially epigenetically repressed *RPP5* gene
175 cluster (Yi and Richards, 2009). *SNC1* expression activity is modulated by H3K4me3
176 (Yang et al., 2023), and its level of expression is positively correlated with basal
177 resistance (Yang et al., 2023; Yi and Richards, 2009). Therefore, *SNC1* is an attractive
178 target for proof-of-principle modulation of disease resistance by epigenome
179 engineering.
180
181 We designed a guide RNA to target the promoter region of *SNC1* directly upstream of
182 the TSS and inserted it into our SunTag:SDG2 and SunTag:dSDG2 constructs (Figure
183 2A). As high pathogen resistance is generally associated with dwarfed plants, we
184 inspected the lines for growth and developmental phenotypes. In our *SNC1* targeting
185 SunTag:SDG2:SNC1_g4, but not our SunTag:dSDG2:SNC1_g4 controls, we
186 observed small, dwarfed phenotype with reduced rosette diameter (Figure 2AB).
187 These plant phenotypes resembled those of our positive control *bal* lines, which
188 overexpress *SNC1*, resulting in the small dwarfed plant phenotype and high basal
189 disease resistance (Yi and Richards, 2009). As the phenotype is consistent with
190 upregulation of *SNC1* by SunTag, we confirmed the increased expression and
191 successful targeting of SunTag:SDG2 to the *SNC1* locus (Figure 2CD).
192

193 To assess whether these lines show increased pathogen resistance, we challenged
194 the plants with the generalist plant pathogen, *Pseudomonas syringae* pathovar *tomato*
195 strain DC3000 (*Pst*). The recently described bioluminescent *Pst* (*Pst::LUX*) was used
196 to allow for non-destructive quantification of pathogen colonisation (Furci et al., 2021).
197 We developed a high-throughput assay system, inoculating 7-day-old seedlings with
198 *Pst::LUX* and imaging to quantify the levels of *Pst::LUX* at 24 hour intervals over the
199 course of infection. We challenged a range of control genotypes that are known to be
200 *Pst* hyper-resistant (*bal*), constitutively primed (*edr1*) and susceptible (*NahG*),
201 recapitulating the expected pattern of *Pst::LUX* colonisation dynamics in these
202 backgrounds (Supp. Fig. S5) (Furci et al., 2021). Next, we challenged our *SNC1*
203 targeting lines with *Pst::LUX*, finding that SunTag:SDG2:SNC1_g4 lines displayed
204 hyper resistance that was nearly identical in magnitude to that of the *SNC1*
205 overexpressing positive control, *bal* (Figure 2E). The results demonstrate that
206 epigenome-engineering of a single defense gene, *SNC1*, is sufficient to generate
207 plants with improved disease resistance phenotypes.

Figure 2

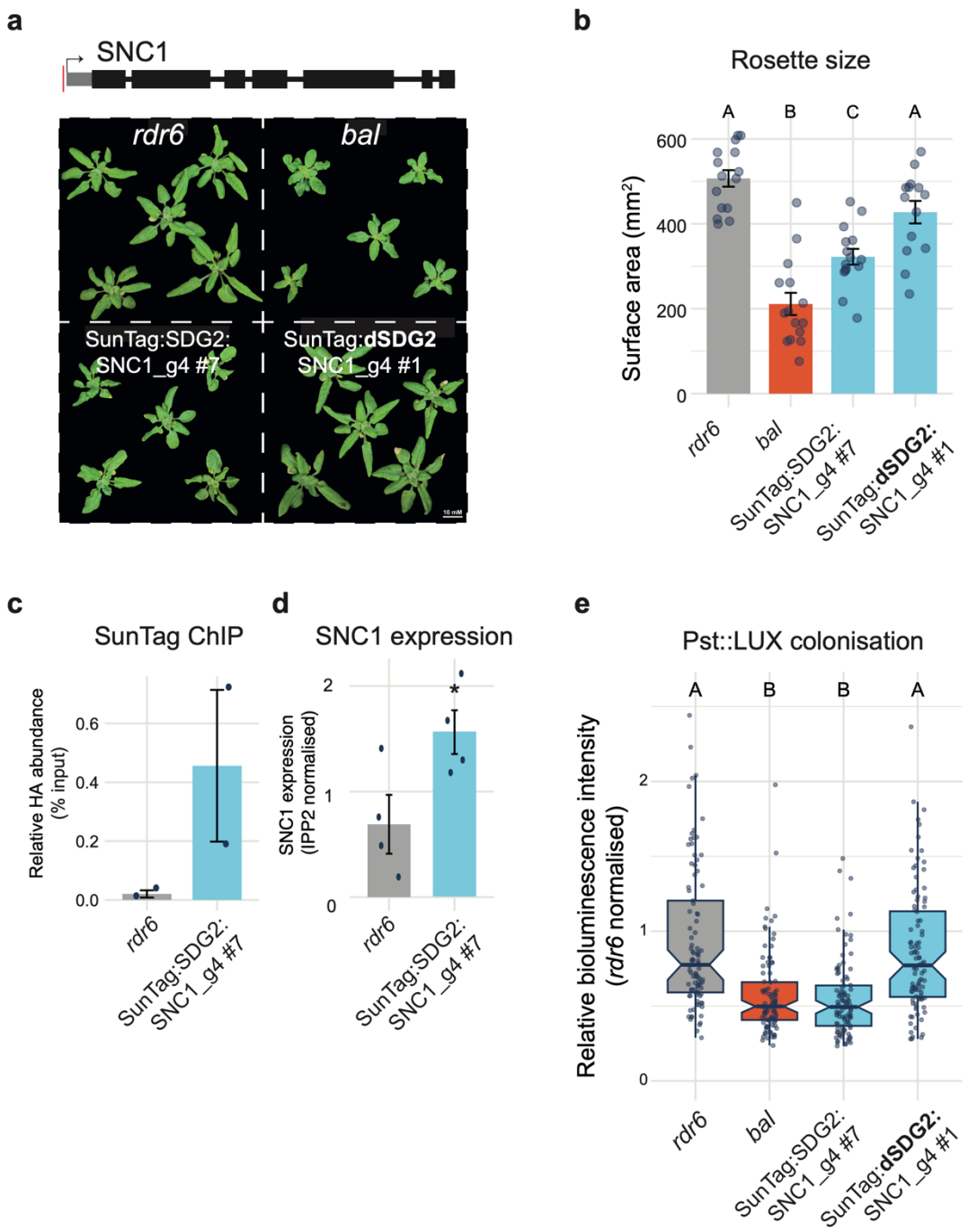


Figure 2: SunTag:SDG2 targeting to *SNC1* enhances resistance to *P. syringae*. A) Upper, schematic of the *SNC1* gene, with the red line indicating the *SNC1* gRNA target site, drawn to scale. Lower, representative images of 3-week old plants from the genotypes indicated. B) Rosette size quantification. Each dot represents an individual plant. Error bars represent SEM. Different letters indicate significant difference by ANOVA with post-hoc Tukey HSD. C) ChIP qPCR for the presence of SunTag at SNC1. D) RT-qPCR for *SNC1* expression. Error bars represent SEM. * indicates p-value < 0.05. E) Pst::LUX assay for colonisation

quantification at 3 days post inoculation. Different letters indicate significant difference by ANOVA with post-hoc Tukey HSD.

Centromeric targeting of SDG2 drives increased meiotic crossover recombination

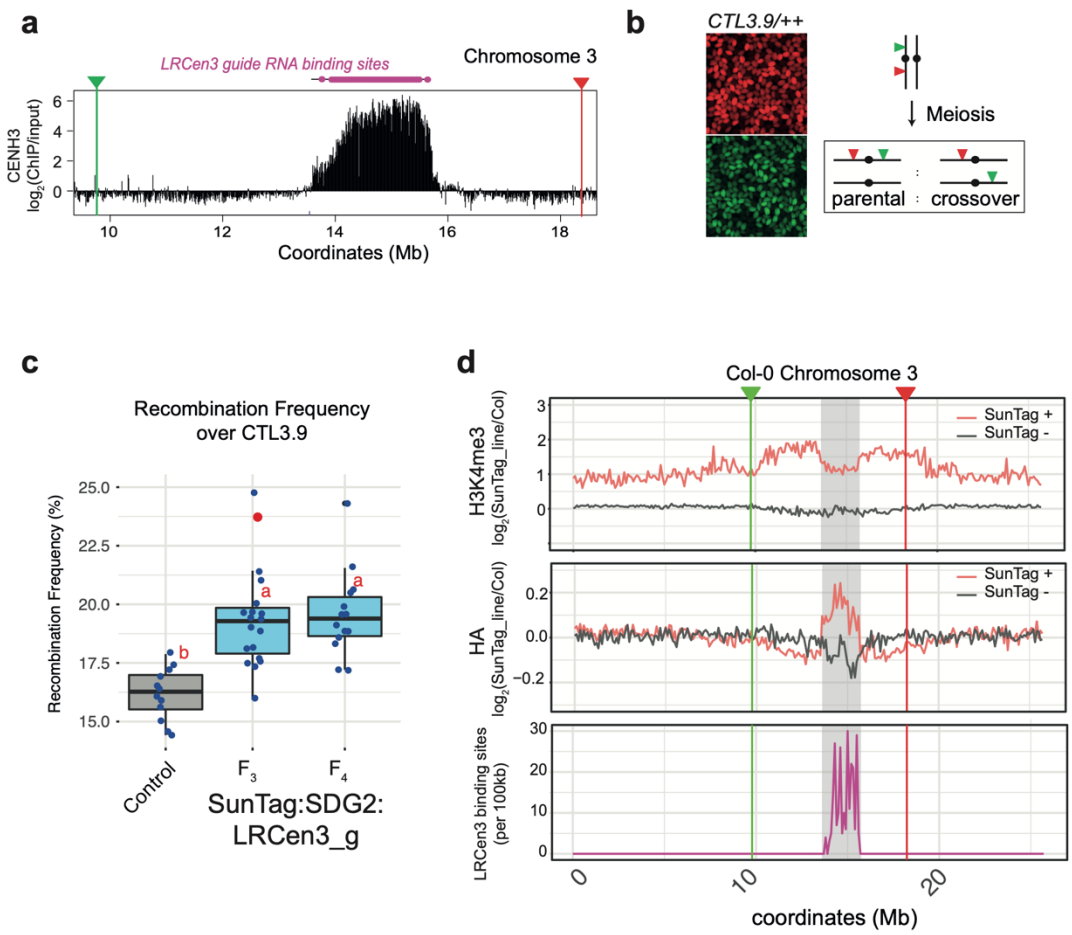
Beyond transcription, H3K4me3 is associated with many other fundamental genome related processes including meiotic recombination. In plants, meiotic crossover recombination rates broadly correlate with epigenetic territories and are non-uniform across the genome (Rowan et al., 2019; Tock et al., 2021). In *A. thaliana*, recombination occurs at a relatively high frequency in euchromatic arms, it is significantly reduced over the more heterochromatic pericentromeric regions, and drops to virtually zero over the centromere (Naish et al., 2021; Yelina et al., 2012). Crossover hotspots are often found to be high at gene promoters where H3K4me3 is enriched (Choi et al., 2013). In previous work, we showed that gene-associated crossover recombination hotspots can be supressed by small RNA directed DNA methylation (Yelina et al., 2015). Therefore, we hypothesised that recombination within the centromeric regions might be increased by SunTag mediated epigenetic activation via deposition of H3K4me3.

The Arabidopsis centromeres are composed of highly repetitive *CEN178* satellite repeat arrays (Naish et al., 2021; Wlodzimierz et al., 2023). We designed a guide RNA that perfectly matches to 250 individual *CEN178* repeats within the centromeric region of chromosome 3 (LRCen3_g), and inserted this guide into the SunTag:SDG2 system. As centromere-proximal crossover recombination events are relatively rare, to quantitatively assess the impact on recombination rate over centromere 3, we crossed (non-transgenic) Col-0 plants with *CTL3.9* (Wu et al., 2015), which is a transgenic a line that encodes red and green fluorescent markers flanking *CEN3*. The red and green T-DNA insertions span a region of approximately 9Mb (positions 9.7Mb and 18.8Mb, respectively, on the Col-CEN assembly) of chromosome 3, while the centromeric satellite arrays are located from 13.6-15.7Mb (Figure 3A). Meiotic recombination events that occur within this region result in progeny inheriting either red or green fluorescent markers, the frequency of which can be measured by automated seed imaging (Kbiri et al., 2022). We transformed *CTL3.9* F_1 double

250 hemizygous plants with the SunTag:SDG2:LRCen3_g construct (Figure 3AB). The F_2
251 plants represented individual transformants of the SunTag (T_1 for SunTag), and so F_3
252 seeds from individual F_2 plants were imaged to measure the rate of recombination, in
253 centiMorgans, over the centromere 3 spanning region of *CTL3.9* (Fernandes et al.,
254 2024).
255
256 Remarkably, the SunTag:SDG2:LRCen3_g containing lines showed significantly
257 elevated crossover recombination within the *CTL3.9* interval, which was not observed
258 in the no guide RNA controls (ANOVA with post-hoc Tukey HSD cutoff <0.05) (Figure
259 3C, Supp. Fig. 6). Some individual transgenic plants displayed a substantial increase
260 in *CTL3.9* crossover recombination rate, over 50% higher as compared the average
261 in non-transgenics, across this megabase-scale centromere-spanning region.
262 Progeny from the most highly recombining line that also showed Mendelian
263 inheritance of the *CTL3.9* T-DNAs was sown to assess whether the effect could persist
264 in the subsequent (F_4) generation. The plants retained a significantly elevated
265 crossover frequency across *CTL3.9* in the next generation, suggesting that the
266 SunTag:SDG2:LRCen3_g effect is relatively stable in the presence of the transgene
267 (Figure 3C).
268
269 To examine the distribution of SunTag:SDG2:LRCen3_g in the high recombination
270 lines (P4), we performed ChIP-seq (anti-HA) comparing progeny from sibling plants
271 that either have (+SunTag) or do not have (-SunTag) the SunTag machinery, as well
272 as non-transgenic (Col-0) controls. Chromosome-wide plots at 100 kb resolution
273 showed that SunTag:SDG2:LRCen3_g is enriched over the entire centromeric region
274 of chromosome 3 (Figure 3D). SunTag was also enriched over the other four *A.*
275 *thaliana* centromeres (Supp. Fig. 7). As these regions are highly repetitive, we
276 reasoned that the LRCen3 guide RNA might additionally bind the other centromeres
277 due to mismatch tolerance. Indeed, while the LRCen3 guide RNA aligns perfectly to
278 250 loci within the centromere of chromosome 3, by allowing a single mismatch the
279 guide is predicted to bind regions in all 5 centromeres, which is consistent with the
280 ChIP-seq data observed (Supp. Fig. 8). Having confirmed that SunTag is enriched
281 within the centromeric regions in these lines, we were next interested to examine
282 levels of H3K4me3. ChIP-seq of H3K4me3 showed that levels were broadly increased
283 throughout the genome (Figure 3D). This increase was stronger in centromeric regions

284 as compared to euchromatic arms, as expected, and was most prominent over the
285 pericentromere. As the markers on *CTL3.9* span the centromere and adjacent
286 pericentromeric regions (Fernandes et al., 2024) (Figure 3D), this is consistent with
287 the significantly elevated crossover recombination rate observed. Overall, the results
288 demonstrate that centromerically targeted SunTag-SDG2 can elevate the meiotic
289 crossover potential of these typically recombination-suppressed regions.

Figure 3



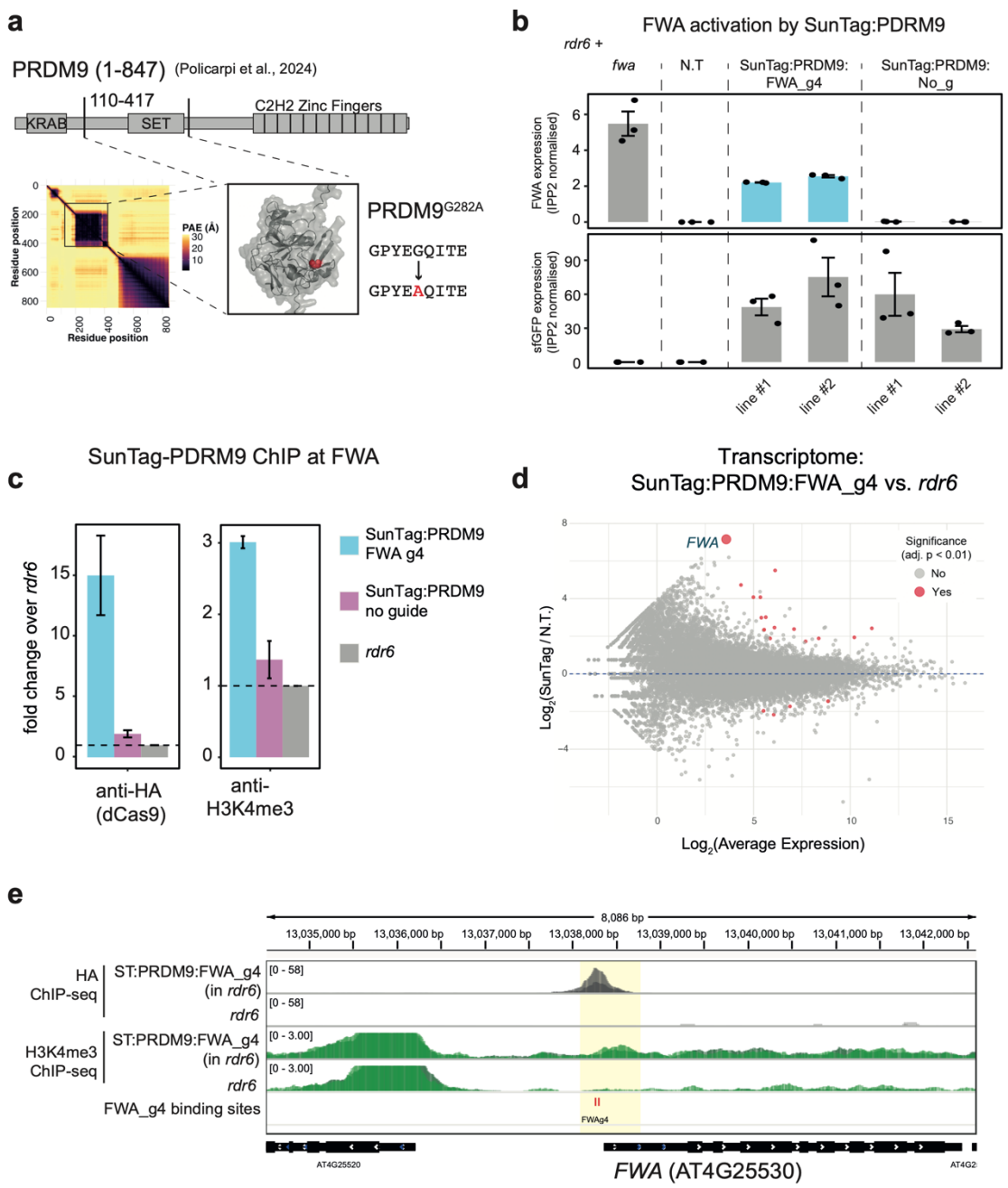
290

291 **Figure 3: Centromeric targeting of SunTag:SDG2 elevates meiotic crossover**
292 **recombination rate.**
293 A) Region of chromosome 3, showing CENH3 ChIP-seq enrichment (Wlodzimierz et al., 2023)
294 with *CTL3.9* (Wu et al., 2015) red / green T-DNA marker positions and LRCen3 guide RNA
295 binding sites indicated. B) Left panel, representative image of red and green fluorescing seeds
296 from a *CTL3.9* double homozygous line. Right panel, schematic of showing how crossovers
297 are identified within *CTL3.9*. C) Boxplot showing crossover recombination frequency over
298 *CTL3.9* in centiMorgans. Control data is from Col-0 (non-transgenic) crossed to *CTL3.9*.

299 Different letters indicate significant difference by ANOVA with post-hoc Tukey HSD. The red
300 dot indicates the seed set line taken to the next generation (F_4). D) Chromosome-wide plots
301 of chromosome 3. Upper and middle panels show enrichment of H3K4me3 and SunTag (anti-
302 HA) by ChIP-seq in sibling lines with / without SunTag (+/-). Enrichment is calculated as log_2
303 fold-change over non-transgenic (Col-0) controls in 100 kb windows. The lower panel depicts
304 an LRCen3 guide RNA binding site density over chromosome 3. *CTL3.9* marker positions are
305 indicated, and centromeric regions are shown in grey. All data are mapped the Col-CEN
306 genome assembly (Naish et al., 2021).
307
308 **Mammalian PRDM9 drives efficient transcriptional activation in Arabidopsis**
309
310 As the enrichment of H3K4me3 in SunTag:SDG2:LRCen3_g was not exclusive to the
311 guide RNA targeted centromeric regions, we reasoned that the SunTag system may
312 afford some level of off-target activity due to non-guide RNA directed ectopic
313 interactions. Consistent with this observation, we also noticed that our SunTag:SDG2
314 no guide RNA controls frequently exhibited pleotropic developmental phenotypes
315 (Supp. Fig. S9). To quantify global levels of H3K4me3, we performed bulk histone
316 westerns, comparing no guide RNA lines to successful *FWA* expression-activating
317 guide RNA containing lines. Importantly, the SDG2 effector itself is expressed to
318 similar levels as in the *FWA* targeting lines (Supp. Fig. S3). The SunTag:SDG2 no
319 guide lines all showed significant (3-4 fold) increases in global levels of H3K4me3, as
320 compared to the *FWA* guide containing lines (Supp. Fig. 10A). This indicates that the
321 presence of the guide RNA may enhance the targeting specificity of the SunTag
322 complex by reducing the frequency of non-specific interactions that result in off-target
323 deposition of H3K4me3. We also noticed that a small number of
324 SunTag:dSDG2:FWA_g4 lines could activate *FWA*, but this effect was limited to lines
325 where the dSDG2 effector was massively overexpressed (Supp. Fig. 10B). This
326 suggests that either dSDG2 retains some residual catalytic activity, or that it can recruit
327 endogenous machinery to initiate transcription.
328
329 In order to reduce off-target activity and the potential for co-recruitment of endogenous
330 complexes, we replaced SDG2 with an orthogonal effector, PRDM9. PRDM9 is a
331 mammalian derived H3K4me3 methyltransferase, and Policarpi et al (2024) recently
332 showed that the catalytic domain could be incorporated into the SunTag system for

333 targeting of H3K4me3 mammalian cells. Therefore, we inserted the PRDM9
334 methyltransferase domain, as well as the catalytically deactivated version, into our
335 SunTag system and targeted *FWA* (Figure 4A). Strikingly, SunTag:PRDM9:FWA_g4
336 was highly effective at activating *FWA* mRNA expression (Figure 4B), (23/32 in the T_1
337 generation, 72%), while none of the dPRDM9 (0/16) and no-guide RNA (0/19) versions
338 were able to do so (Supp. Fig. 11). Importantly, we could detect robust expression of
339 the effector in all constructs (Supp. Fig. 11), and ChIP qPCR confirmed the presence
340 of SunTag and the enrichment of H3K4me3 at *FWA* (Fig 4C). At the transcriptome
341 level, the effect of SunTag:PRDM9:FWA_g4 was highly specific, with only 19 genes
342 identified as differentially expressed (15 up, 4 down, FDR<0.01), and with *FWA* itself
343 having the highest fold change among them (>158 fold change over *rdr6*) (Figure 4D).
344 Unlike the no-guide SDG2 lines, we did not observe any developmental defects in the
345 PRDM9 no-guides, and correspondingly, we observed only minor transcriptional
346 defects when comparing the no-guide lines to non-transformed *rdr6* controls (Supp.
347 Fig. S12). Finally, we performed ChIP-seq in these lines, confirming *FWA* as the most
348 significant and enriched peak for SunTag binding and revealing deposition of
349 H3K4me3 at the transcriptional start site of *FWA* (Fig 4E). Overall, the results
350 demonstrate that PRDM9 is highly effective for deposition of H3K4me3 in plant as well
351 as mammalian SunTag systems. Furthermore, as PRDM9 is not likely to have
352 endogenous partners in the plant nucleus, the results provide strong evidence in
353 favour of a direct role for H3K4me3 in transcriptional stimulation.

Figure 4



354

355

356 **Figure 4: SunTag PRDM9 is sufficient for activation of *FWA*.**

357 A) Upper panel shows PRDM9 with catalytic SET domain (drawn approximately to scale).
358 Lower left depicts predicted alignment error (PAE) plot from AlphaFold3 prediction of the
359 PRDM9 coding sequence with a box drawn around the region cloned into SunTag (amino
360 acids 110-417). Lower right depicts the AlphaFold3 model for $PRDM9_{110\text{-}417}$, with the amino
361 acid change for dPRDM9 indicated (position shown in red). B) RT-qPCR for *FWA* (upper
362 panel) and the effector module (sfGFP, lower panel) for the genotypes indicated. Two

363 independent T_3 lines are used per genotype. Error bars represent SEM. C) ChIP qPCR for the
364 presence of SunTag (left panel) and H3K4me3 enrichment (right panel) over the TSS of *FWA*.
365 Error bars represent SEM from 2 biological replicates. D) MA plot comparing the transcriptome
366 of SunTag:PRDM9:FWA_g4 as compared to the non-transformed control (*rdr6*). Differentially
367 expressed genes are show in in red (FDR < 0.01), with *FWA* labelled and enlarged for visibility.
368 E) Genome browser image showing efficient SunTag targeting and H3K4me3 enrichment at
369 *FWA*. Biological replicate tracks from independent lines are overlaid.
370
371 **Discussion:**
372
373 Here we show that SunTag mediated targeting of H3K4me3 facilitates control of gene
374 expression states, can allow for improved pathogen resistance capabilities, and can
375 drive targeted increases in centromere-proximal recombination. From a fundamental
376 biological perspective, our data provide strong support in favour of an instructive role
377 for H3K4me3 in transcription, as we show that site-specific deposition of H3K4me3
378 results in activation of mRNA expression. As the effect is dependent on the presence
379 of a functional catalytic domain, and is observed using both native and non-native
380 methyltransferases, co-recruitment of endogenous complexes can be likely ruled out
381 as an explanatory factor in driving this transcriptional enhancement. Furthermore, we
382 reveal H3K4me3 targeting as a novel strategy for stimulating crossover formation in
383 low-recombining, repeat-rich regions of the genome. The results open the door to non-
384 GM approaches for precise genome engineering and rational design of agriculturally
385 relevant traits.

386 **Materials & Methods**

387

388 Plant materials and growth conditions:

389 All *Arabidopsis thaliana* lines are in a Col-0 ecotype background. Plants grown on
390 plates were sown on to 1/2x Murashige-Skoog (MS) medium, 1% sucrose and 0.8%
391 agar (pH 5.7), stratified for 2 d at 4 °C in the dark, then transferred to growth chambers
392 (Percival CU-41L4D) at 21 °C under long day light conditions (16 hour light/8 hour
393 dark). Plants grown on soil were grown in F2 soil at 20°C under long day light
394 conditions (16 hour light/8 hour dark). Lines: *rdr6-15* (SAIL_617_H07), and *fwa* lines
395 were previously described (M. Wang et al., 2023). NahG and *edr1* were kindly provided
396 by Prof. Juriaan Ton, previously described in (Furci et al., 2021). *bal* lines were
397 previously described in (Yi and Richards, 2009). The CTL 3.9 line was previously
398 described in (Fernandes et al., 2024; Wu et al., 2015). All transgenic plants were
399 generated by *Agrobacterium tumefaciens* strain AGL1 using floral dip (Clough and
400 Bent, 1998).

401

402 Plasmid construction:

403 To generate the SunTag:SDG2:FWA_g4 construct: the CDS of the catalytic domain of
404 SDG2 (from amino acids 1571-2335) was amplified and coned into into the BsiWI
405 linearised SunTag construct (previously described (Papikian et al., 2019)), using In-
406 Fusion (Takara). To generate dSDG2: the conserved amino acid predicted to abrogate
407 catalytic activity when mutated (Y1903F) was identified using ConSurf (Ashkenazy et
408 al., 2016). Overlapping PCR was used to generate a version of SDG2 containing the
409 desired mutation (see Supplementary Table 1), and the amplified PCR product was
410 cloned into a BsiWI linearized SunTag-SDG2 construct by In-Fusion (Takara). To
411 generate SunTag:PRDM9, the catalytic domain of mouse PRDM9 (110-417aa)
412 (described in Policarpi et al 2024) was synthesised (TWIST Bioscience) and cloned
413 into the SunTag using the same BsiWI linearization and In-Fusion method described
414 above. Subsequently, overlap PCR was used to generate a version of PRDM9 with
415 the G282A mutation (see Supplementary Table 1). The SNC1, LRCen3, and the no
416 guides controls, were cloned using the protocol described in (Ghoshal and Gardiner,
417 2021) (see Supplementary Table 1).

418

419 Chromatin Immunoprecipitation:

420 The chromatin immunoprecipitation experiments were performed as previously
421 described (Harris et al., 2024). 2g of 14-day-old seedlings were collected per biological
422 replicate. For SunTag-SDG2 FWA and SunTag-SDG2 SNC1 input material are pooled
423 T_2 plants. For the SunTag:PRDM9:FWA_g4 in *rdr6*, 2 independent T_3 lines expressing
424 FWA, 2 independent T_3 lines NoG were used alongside *rdr6* controls. 7 µl of anti-HA
425 (3F10, Merck) and 5 µl of anti-H3K4me3 (ab8580, abcam) were added to up to 2 ml
426 of sheared chromatin per sample for pulldown. The ChIP purified DNA was directly
427 used for ChIP-qPCR and ChIP-seq library preparation. Primer information is listed in
428 Suppletory Table 1. Values were normalised to Input ($2^{-(\text{ChIP Ct} - \text{Input Ct})}$*100). ChIP-seq
429 libraries were generated using NuGen Ovation Ultra Low System V2 kits according to
430 manufacturer's instructions and were sequenced on an Illumina NovaSeq X
431 instrument with 150bp paired-end reads.
432
433 Crossover measurement:
434 For the recombination rate experiment, SunTag:SDG2:LRCen3_g/No_g plamids were
435 transformed into F_1 plants derived from a cross between Col-0 and *CTL3.9* (double
436 homozygous for the seed coat fluorescent markers) (Wu et al., 2015). 70-90 individual
437 T_1 transgenic plants were recovered per construct after selection on hygromycin (F_2
438 generation). Seeds derived from individual T_1s (F_3) were imaged and scored for
439 recombination rate using the *SeedScoring* CellProfiler pipeline, previously described
440 (Fernandes et al., 2024; Kbiri et al., 2022) with approximately 1,000 seeds per
441 biological sample. Note, seed sets were pre-screened to remove lines in which either
442 of the fluorescent markers had become homozygous (presence / absence) in the F_2
443 generation). For the F_4 data, seeds derived from the most highly recombining and non-
444 distorted line (P4) were sown (F_3), grown for a generation and seeds from individual
445 plants were collected (F_4) for recombination rate analysis as above, again filtering
446 seed sets from which either of the fluorescent markers had become homozygous
447 (presence / absence) in the previous generation.
448
449 Western blots:
450 Western blot were performed as previously described (Ichino et al., 2021) with minor
451 modifications described below. For detection of the large dCAS9-10xGCN4 and SDG2
452 effector components as shown in Fig S1, single leaf punch (1cm diameter) samples
453 from individual plants were used as input. Tissue was ground in a Tissuelyser II

454 (Retsch) and resuspended in 100 ul of a 60:40 ratio of 2XSDS:8M Urea, mixed, and
455 boiled at 95 degrees C for 5 minutes. Samples were centrifuged at >10,000xg to
456 remove debris and loaded into a 3-8% tris acetate gel using MOPS running buffer.
457 After electrophoresis, protein was transferred from the gel to a methanol activated
458 PVDF membrane in Tobin Buffer (25mM Tris, 192mM Glycine, 20% methanol, 0.035%
459 SDS). Membranes were blocked in 5% milk (w/v) in PBST (Phosphate Buffered Saline,
460 0.1% Tween20) and then incubated with HRP conjugated anti-HA at a 1:3,000 dilution
461 in block. The blot was imaged by ECL detection.
462
463 For the western blots used to detect levels of H3K4me3, 100 mg of 7-day-old seedlings
464 were frozen in liquid nitrogen and ground using a Tissuelyser II (Retsch). Total
465 histones were extracted using the The EpiQuik™ Total Histone Extraction Kit
466 (EpigenTek) according to the manufacturers instructions. 5 µl of sample for H3 and 20
467 µl of sample for H3K4me3 were loaded onto a NuPage 4-12% BIS-Tris gel (Invitrogen)
468 in NuPage MOPS SDS running buffer (Invitrogen). Transfer from gel to PVDF
469 membrane was performed using an iBlot 3 machine with the dedicated transfer stack
470 (Invitrogen). Following transfer, membrane was blocked in 5% milk (w/v) in TBST (Tris
471 Buffered Saline, 0.1% Tween20) and shaken in room temp for 1 hour. Blocking buffer
472 was discarded and membranes were incubated in primary antibody (anti-H3, ab1791
473 and anti-H3K4me3, ab8580, Abcam) in blocking buffer (1:1,000) shaking overnight in
474 4C. Following morning membrane was rinsed in TBST 5 times for 5 minutes in room
475 temp. Secondary antibody was added for 1 hour (Goat anti-mouse HRP 1:10,000 in
476 blocking buffer, ab6789, Abcam). Membrane was washed again 5 times for 5 minutes.
477 Detection was performed using a SuperSignal West Pico chemiluminescent substrate
478 (#34080, Thermo) using an ImageQuant 800 (Amersham).
479
480 <u>Fluorescent imaging:</u>
481 Roots of 7-day-old plants grown in vertical 1/2x MS solid media plates were imaged
482 using a Leica DM6000B epifluorescent microscope.
483
484 <u>Rosette size surface area measurement:</u>
485 3-week-old plants grown on soil were imaged using a Google Pixel 7 mounted on a
486 tripod. Images were then analyzed using Fiji.
487

488 *Pseudomonas syringae* colonisation assay:
489 Infection assays were performed with *P. syringae* pathovar *tomato* strain DC3000
490 (*Pst*) strain expressing a stable chromosomal insertion of the *lux- CDABE* operon from
491 *Photorhabdus luminescens* (*Pst::LUX*) (Fan et al. 2008), using a slightly modified
492 version of a previously published protocol (Furci et al. 2021). Plants were grown on
493 1/2X MS solid media supplemented with 200 µg/ml Timentin in 96-well plates (Falcon)
494 for 7 days. Prior inoculation, *Pst::LUX* were grown in King's B media containing 50
495 µg/ml Kanamycin and 50 µg/ml Rifampicin overnight at 28 °C. Bacteria were pelleted
496 by centrifugation, washed in 10 mM MgSO4 and finally re-suspended to an OD600 of
497 0.2 in 10 mM MgSO4 containing 0.015% v/v Silwet-L77. 7 days-old seedlings were
498 sprayed with bacteria solution and 96-well plates were sealed with parafilm to maintain
499 100% relative humidity. Seedlings were imaged at 1-4 days post infection (dpi). Prior
500 imaging using (ImageQuant 800, Amersham), plates were left in the dark for 2 mins.
501 Bacterial bioluminescence images were acquired with 4 mins exposure time and bright
502 field images were taken using OD settings of the ImageQuant 800. Image-based
503 quantification of bioluminescence was carried out in Fiji. For each well, the well outline
504 was obtained from bright field images and added to ROI Manager. These ROIs were
505 then transposed onto the bioluminescence images. The bioluminescence intensity
506 from infected seedlings was obtained using the Fiji functions Analyse, Measure and
507 Mean value.
508
509 RT-qPCR:
510 RNA was extracted from indicated plant material using TRIzol reagent (Invitrogen) and
511 the Direct-Zol RNA MiniPrep (Zymo) kit, including in-column DNase treatment
512 following the manufacturer's instructions. cDNA was synthesized using SuperScript IV
513 (Invitrogen). qRT-PCRs were performed using a Luna Universal qPCR Master Mix
514 (NEB) and a CFX connect Real-time PCR detection system (Bio-rad). For SunTag-
515 SDG2 targeting FWA and SNC1, RNA was extracted from 14-day-old whole seedlings
516 grown on 1/2x MS plates. SNC1 targeted lines were treated with Pst::LUX 4H prior to
517 harvesting. For SunTag:PRDM9 qRT-PCR, RNA was extracted from 4 weeks-old leaf
518 tissue of all T_1 lines and 2 independent T_3 lines per construct. 450 ng of total RNA was
519 used for cDNA synthesis. Primers are listed in Supplementary Table 1.
520
521 Quant-Seq Library construction:

522 For Suntag-SDG2 QuantSeq, 350 ng of total RNA from ~10 pooled T_2 plants were
523 used as input material. For Suntag:PRDM9 Quant-seq, 350 ng of total RNA from 3 T_1
524 SunTag:PRDM9:FWA_g4 plants expressing FWA, 3 T_1 no guide SunTag:PRDM9
525 lines and *rdr6* and plants was used for library preparation. Libraries were prepared
526 using QuantSeq 3 'mRNA-Seq Library Prep Kit FWD (Lexogen) according to
527 manufacturer's instructions and sequenced on Illumina NovaSeq 6000 PE150
528 instrument.
529
530 Bioinformatic Analysis:
531 ChIP-seq analysis:
532 The ChIP-seq analysis pipeline was performed as previously described with minor
533 modifications. Bowtie2 (version 2.5.0) was used to map the PE150 bp read data in
534 fastq format to the TAIR10 genome (--no-unal), and were converted to bam format
535 using Samtools (version 1.10). Reads were de-duplicated using the samtools fixmate
536 and markdup commands (see https://github.com/C-Jake-Harris/Binenbaum-SunTag-
537 H3K4me3 for more details). Tracks were generated in DeepTools (version 3.5.1) using
538 bamCoverage (--normalizeUsing RPGC, --effectiveGenomeSize 135000000 --binSize
539 10) with multicopy regions blacklisted (--blackListFileName) using the regions
540 identified in (Klasfeld et al., 2022). For analysis of reads mapping to the centromere,
541 a modified version of the previously described pipeline (Wlodzimierz et al., 2023) was
542 used. Briefly, PE150 bp read data in fastq format were mapped to the Col-CEN
543 genome (Naish et al., 2021) using Bowtie2 (version 2.2.5) with (--very-sensitive -k 200
544 --no-unal --no-discordant) and were filtered for primary alignments using samtools
545 view (-F 256 -q 5) prior to deduplication and track generation, as above. Peaks were
546 called using MACS2 (version 2.2.9.1) with default parameters. Correlation plots were
547 generated in DeepTools using MultiBamSummary (--binSize 25) and PlotCorrelation
548 (-c pearson --removeOutliers --plotNumbers -p heatmap).
549
550 QuantSeq Analysis:
551 In accordance with the manufacturer's (Lexogen) instructions, only read (*1.fq.gz)
552 from PE150 bp data was used for downstream analysis. Briefly, reads were trimmed
553 using cutadapt (version 1.18) and were mapped to the genome using STAR (version
554 2.7.10b) to the TAIR10 genome (--quantMode GeneCounts --alignIntronMax 10000 --
555 outSAMmultNmax 20). Gene counts were used as input for DESeq2 analysis in R.

Differentially expressed genes were defined using an adjusted p-value (Bonferroni method) cutoff of <0.01.

gRNA binding analysis:

For the LRCen3 binding and mismatch analysis, the LRCen3 guide RNA sequence (5'-AGGCTTACAAGATTGGGTTG-3'), was mapped to the Col-CEN genome using bowtie2 (version 2.2.5). The following options were used for no mismatches (-f -a --end-to-end --np 0 --score-min L,0,0) and for 1 mismatch (-f -a --end-to-end --score-min L,0,-1). Mapped reads were sorted and converted to bam format using Samtools (version 1.10), and subsequently converted to .bed format using the bamtobed function in bedtools (2.20.1). Bed files were used for downstream analysis in R to generate chromosome-wide density plots over 100kb regions.

Data availability:

The high-throughput sequencing data generated has been deposited to NCBI (GSEXXXXXX). Previously published datasets: CENH3 (Wlodzimierz et al., 2023), H3K4me3 (Xu et al., 2024). Custom scripts used are available on github (https://github.com/C-Jake-Harris/Binenbaum-SunTag-H3K4me3).

Acknowledgements

We would like to thank Jurriaan Ton's group at the University of Sheffield for their generous gift of the *Pst::LUX* strain, and for their support in helping to establish the infection and quantification protocol in the Harris group. **Funding:** C.J.H. is supported by a Royal Society University Research Fellowship (URF\R1\201016) and an ERC Starting Grant (TransPlantMemory), funded by UKRI EPSRC (EP/X025306/1). J.B. is supported by a Marie Curie postdoctoral fellowship, funded by UKRI (EP/Z001749/1). V.A. is supported by a 4-year BBSRC-funded doctoral studentship. I.R.H is supported by an ERC Advanced Grant (EvoPanCen). N.G. is supported by a BBSRC-UKRI doctoral studentship. S.E.J. is an investigator of the Howard Hughes Medical Institute.

Author contributions:

C.J.H., I.R.H. and S.E.J. conceived, designed and supervised the research. C.J.H. performed experiments, NGS bioinformatic analysis, and assembled the manuscript. J.B. performed the majority of the experiments relating to SunTag:SDG2. V.A.

590 performed the experiments relating to SunTag:PRDM9. H.F. performed the majority of
591 experiments and analysis relating to meiotic recombination. L.X. performed the
592 SunTag:SDG2 centromere-targeting ChIP-seq. N.G. provided the Col-0 x *CTL3.9* F_1
593 seeds and guidance on the analysis of recombination frequency. P.W. and R.B.
594 contributed to the design of the centromere targeting guide RNA. A.P. cloned SDG2
595 into SunTag.
596
597 **References**

598 Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore
599 L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung CC, O'Neill M,
600 Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli
601 O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov
602 M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A,
603 Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED,
604 Zielinski M, Žídek A, Bapst V, Kohli P, Jaderberg M, Hassabis D, Jumper JM.
605 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3.
606 *Nature* **630**:493–500. doi:10.1038/s41586-024-07487-w
607 Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016.
608 ConSurf 2016: an improved methodology to estimate and visualize evolutionary
609 conservation in macromolecules. *Nucleic Acids Res* **44**:W344–W350.
610 doi:10.1093/NAR/GKW408
611 Borde V, Robine N, Lin W, Bonfils S, Géli V, Nicolas A. 2009. Histone H3 lysine 4
612 trimethylation marks meiotic recombination initiation sites. *EMBO J* **28**:99–111.
613 doi:10.1038/emboj.2008.257
614 Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, Hardcastle TJ, Ziolkowski
615 PA, Copenhaver GP, Franklin FCH, Mcvean G, Henderson IR. 2013. Arabidopsis
616 meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters.
617 *Nat Genet* **45**:1327–1338. doi:10.1038/ng.2766
618 Clough SJ, Bent AF. 1998. Floral dip: A simplified method for Agrobacterium-mediated
619 transformation of Arabidopsis thaliana. *Plant J* **16**:735–743. doi:10.1046/j.1365-
620 313X.1998.00343.x
621 Du J, Johnson LM, Groth M, Feng S, Hale CJ, Li S, Vashisht AA, Gallego-Bartolome
622 J, Wohlschlegel JA, Patel DJ, Jacobsen SE. 2014. Mechanism of DNA
623 methylation-directed histone methylation by KRYPTONITE. *Mol Cell* **55**:495–504.

624 doi:10.1016/j.molcel.2014.06.009
625 Dubois A, Roudier F. 2021. Deciphering Plant Chromatin Regulation via
626 CRISPR/dCas9-Based Epigenome Engineering. *Epigenomes* **5**:17.
627 doi:10.3390/epigenomes5030017
628 Fernandes JB, Naish M, Lian Q, Burns R, Tock AJ, Rabanal FA, Wlodzimierz P,
629 Habring A, Nicholas RE, Weigel D, Mercier R, Henderson IR. 2024. Structural
630 variation and DNA methylation shape the centromere-proximal meiotic crossover
631 landscape in Arabidopsis. *Genome Biol* **25**:1–31. doi:10.1186/s13059-024-
632 03163-4
633 Francisco M, Id P, Id PS. 2023. Histone methyltransferase activity affects metabolism
634 in human cells independently of transcriptional regulation.
635 doi:10.1371/journal.pbio.3002354
636 Furci L, Pardo DP, Ton J. 2021. A rapid and non - destructive method for spatial –
637 temporal quantification of colonization by Pseudomonas syringae pv . tomato
638 DC3000 in Arabidopsis and tomato. *Plant Methods* 1–8. doi:10.1186/s13007-021-
639 00826-2
640 Gallego-Bartolomé J, Gardiner J, Liu W, Papikian A, Ghoshal B, Kuo HY, Zhao JM-C,
641 Segal DJ, Jacobsen SE. 2018. Targeted DNA demethylation of the *Arabidopsis*
642 genome using the human TET1 catalytic domain. *Proc Natl Acad Sci*
643 **115**:201716945. doi:10.1073/pnas.1716945115
644 Gardiner J, Ghoshal B, Wang M, Jacobsen SE. 2022. CRISPR–Cas-mediated
645 transcriptional control and epi-mutagenesis. *Plant Physiol* **188**:1811–1824.
646 doi:10.1093/plphys/kiac033
647 Ghoshal B, Gardiner J. 2021. CRISPR-dCas9-Based Targeted Manipulation of DNA
648 Methylation in Plants. *Springer Protoc Handbooks* **2**:57–71.
649 doi:https://doi.org/10.1007/978-1-0716-1657-4_5
650 Guo L, Yu Y, Law JA, Zhang X. 2010. Set domain group2 is the major histone H3
651 lysine 4 trimethyltransferase in Arabidopsis. *Proc Natl Acad Sci U S A* **107**:18557–
652 18562. doi:10.1073/pnas.1010478107
653 Harris CJ, Amtmann A, Ton J. 2023. Epigenetic processes in plant stress priming :
654 Open questions and new approaches. *Curr Opin Plant Biol* **75**:102432.
655 doi:10.1016/j.pbi.2023.102432
656 Harris CJ, Zhong Z, Ichino L, Feng S, Jacobsen SE. 2024. H1 restricts euchromatin-
657 associated methylation pathways from heterochromatic encroachment. *Elife* 1–

658 17.
659 Howe FS, Fischl H, Murray SC, Mellor J. 2017. Is H3K4me3 instructive for transcription
660 activation? *BioEssays* **39**:1–12. doi:10.1002/bies.201600095
661 Ichino L, Boone BA, Strauskulage L, Harris CJ, Kaur G, Gladstone MA, Tan M, Feng
662 S, Jami-alahmadi Y, Duttke SH. 2021. MBD5 and MBD6 couple DNA methylation
663 to gene silencing through the J-domain protein SILENZIO. *Science (80-)*
664 **1439**:1434–1439.
665 Kbiri N, Dluzewska J, Henderson IR, Ziolkowski PA. 2022. Quantifying Meiotic
666 Crossover Recombination in Arabidopsis Lines Expressing Fluorescent
667 Reporters in Seeds Using SeedScoring Pipeline for CellProfiler In: Lambing C,
668 editor. Plant Gametogenesis. Methods in Molecular Biology. New York, NY:
669 Springer US. pp. 121–134. doi:10.1007/978-1-0716-2253-7_10
670 Klasfeld S, Roulé T, Wagner D. 2022. Greenscreen: A simple method to remove
671 artifactual signals and enrich for true peaks in genomic datasets including ChIP-
672 seq data. *Plant Cell* **34**:4795–4815. doi:10.1093/plcell/koac282
673 Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG. 2013.
674 H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and
675 selective gene activation. *Cell* **152**:1021–1036. doi:10.1016/j.cell.2013.01.052
676 Lloyd JPB, Lister R. 2021. Epigenome plasticity in plants. *Nat Rev Genet* **0123456789**.
677 doi:10.1038/s41576-021-00407-y
678 Naish M, Alonge M, Wlodzimierz P, Tock AJ, Abramson BW, Schmücker A,
679 Mandáková T, Jamge B, Lambing C, Kuo P, Yelina N, Hartwick N, Colt K, Smith
680 LM, Ton J, Kakutani T, Martienssen RA, Schneeberger K, Lysak MA, Berger F,
681 Bousios A, Michael TP, Schatz MC, Henderson IR. 2021. The genetic and
682 epigenetic landscape of the Arabidopsis centromeres. *Science (80-)* **374**.
683 doi:10.1126/science.abi7489
684 Papikian A, Liu W, Gallego-bartolomé J, Jacobsen SE. 2019. Site-specific
685 manipulation of Arabidopsis loci using CRISPR-Cas9 SunTag systems. *Nat*
686 *Commun* **10**. doi:10.1038/s41467-019-08736-7
687 Pflueger C, Tan D, Swain T, Nguyen T, Pflueger J, Nefzger C, Polo JM, Ford E, Lister
688 R. 2018. A modular dCas9-SunTag DNMT3A epigenome editing system
689 overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A
690 constructs. *Genome Res* **28**:1193–1206. doi:10.1101/gr.233049.117
691 Policarpi C, Hackett JA. 2021. Epigenetic editing : Dissecting chromatin function.

692 *BioEssays* 1–16.
693 Policarpi C, Munafò M, Tsagkris S, Carlini V, Hackett JA. 2024. Systematic epigenome
694 editing captures the context-dependent instructive function of chromatin
695 modifications. *Nat Genet* **56**:1168–1180. doi:10.1038/s41588-024-01706-w
696 Powers NR, Parvanov ED, Baker CL, Walker M, Petkov PM, Paigen K. 2016. The
697 Meiotic Recombination Activator PRDM9 Trimethylates Both H3K36 and H3K4 at
698 Recombination Hotspots In Vivo. *PLoS Genet* **12**:1–24.
699 doi:10.1371/journal.pgen.1006146
700 Rowan BA, Heavens D, Feuerborn TR, Tock AJ, Henderson IR, Weigel D. 2019. An
701 Ultra High-Density Arabidopsis thaliana Crossover. *Genetics* **213**:771–787.
702 Schwaiger M, Schönauer A, Rendeiro AF, Pribitzer C, Schauer A, Gilles AF, Schinko
703 JB, Renfer E, Fredman D, Technau U. 2014. Evolutionary conservation of the
704 eumetazoan gene regulatory landscape. *Genome Res* **24**:639–650.
705 doi:10.1101/gr.162529.113
706 Selma S, Orzáez D. 2021. Perspectives for epigenetic editing in crops. *Transgenic*
707 *Res* **30**:381–400. doi:10.1007/s11248-021-00252-z
708 Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M,
709 Peeters a J. 2000. The late flowering phenotype of fwa mutants is caused by
710 gain-of-function epigenetic alleles of a homeodomain gene. *Mol Cell* **6**:791–802.
711 Tanenbaum ME, Gilbert L a, Qi LS, Weissman JS, Vale RD. 2014. A versatile protein
712 tagging system for signal amplification in single molecule imaging and gene
713 regulation. *Cell* **159**:635–646. doi:10.1016/j.cell.2014.09.039
714 Tock AJ, Holland DM, Jiang W, Osman K, Sanchez-Moran E, Higgins JD, Edwards
715 KJ, Uauy C, Franklin FCH, Henderson IR. 2021. Crossover-active regions of the
716 wheat genome are distinguished by DMC1, the chromosome axis, H3K27me3,
717 and signatures of adaptation. *Genome Res* **31**:1614–1628.
718 doi:10.1101/gr.273672.120
719 Vermeulen M, Mulder KW, Denissov S, Pijnappel WWMP, van Schaik FMA, Varier
720 RA, Baltissen MPA, Stunnenberg HG, Mann M, Timmers HTM. 2007. Selective
721 Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4.
722 *Cell* **131**:58–69. doi:10.1016/j.cell.2007.08.016
723 Villiger L, Joung J, Koblan L, Weissman J, Abudayyeh OO, Gootenberg JS. 2024.
724 CRISPR technologies for genome, epigenome and transcriptome editing. *Nat Rev*
725 *Mol Cell Biol*. doi:10.1038/s41580-023-00697-6

726 Wang H, Fan Z, Shliaha P V., Miele M, Hendrickson RC, Jiang X, Helin K. 2023.
727 H3K4me3 regulates RNA polymerase II promoter-proximal pause-release.
728 *Nature* **615**:339–348. doi:10.1038/s41586-023-05780-8
729 Wang H, Helin K. 2024. Roles of H3K4 methylation in biology and disease. *Trends*
730 *Cell Biol* **xx**:1–14. doi:10.1016/j.tcb.2024.06.001
731 Wang M, Zhong Z, Gallego-bartolomé J, Li Z, Feng S, Kuo HY, Kan RL, Lam H, Richey
732 JC, Tang L, Zhou J, Liu M, Jami-Alahmadi Y, Wohlschlegel JA, Jacobsen SE.
733 2023. A gene silencing screen uncovers diverse tools for targeted gene
734 repression in Arabidopsis . *Nat Plants*. doi:10.1038/s41477-023-01362-8
735 Wang MY, Chen J Bin, Wu R, Guo HL, Chen Y, Li ZJ, Wei LY, Liu C, He SF, Du M
736 Da, Guo Y long, Peng YL, Jones JDG, Weigel D, Huang JH, Zhu WS. 2023. The
737 plant immune receptor SNC1 monitors helper NLRs targeted by a bacterial
738 effector. *Cell Host Microbe* **31**:1792-1803.e7. doi:10.1016/j.chom.2023.10.006
739 Wang X, Yamaguchi N. 2024. Cause or effect: Probing the roles of epigenetics in plant
740 development and environmental responses. *Curr Opin Plant Biol* **81**:102569.
741 doi:10.1016/j.pbi.2024.102569
742 Wlodzimierz P, Rabanal FA, Burns R, Naish M, Primetis E, Scott A, Mandáková T,
743 Gorringe N, Tock AJ, Holland D, Fritschi K, Habring A, Lanz C, Patel C, Schlegel
744 T, Collenberg M, Mielke M, Nordborg M, Roux F, Shirsekar G, Alonso-Blanco C,
745 Lysak MA, Novikova PY, Bousios A, Weigel D, Henderson IR. 2023. Cycles of
746 satellite and transposon evolution in Arabidopsis centromeres. *Nature* **618**:557–
747 565. doi:10.1038/s41586-023-06062-z
748 Wu G, Rossidivito G, Hu T, Berlyand Y, Poethig RS. 2015. Traffic lines: New tools for
749 genetic analysis in Arabidopsis thaliana. *Genetics* **200**:35–45.
750 doi:10.1534/genetics.114.173435
751 Xu L, Wang Y, Li X, Hu Q, Adamkova V, Xu J, Harris CJ, Ausin I. 2024. The H3K4me3
752 binding ALFIN-LIKE proteins recruit SWR1 for gene-body deposition of H2A.Z.
753 *bioRxiv*. doi:https://doi.org/10.1101/2024.09.12.612642
754 Yang L, Wang Z, Hua J. 2023. Multiple chromatin-associated modules regulate
755 expression of an intracellular immune receptor gene in Arabidopsis. *New Phytol*
756 **237**:2284–2297. doi:10.1111/nph.18672
757 Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, Miller N,
758 Drouaud J, Grelon M, Copenhaver GP, Mezard C, Kelly K a, Henderson IR. 2012.
759 Epigenetic remodeling of meiotic crossover frequency in Arabidopsis thaliana

760 DNA methyltransferase mutants. *PLoS Genet* **8**:e1002844.
761 doi:10.1371/journal.pgen.1002844
762 Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B, Henderson IR. 2015. DNA
763 methylation epigenetically silences crossover hot spots and controls
764 chromosomal domains of meiotic recombination in Arabidopsis. *Genes Dev*
765 **29**:2183–2202. doi:10.1101/gad.270876.115
766 Yi H, Richards EJ. 2009. Gene duplication and hypermutation of the pathogen
767 Resistance gene SNC1 in the arabidopsis bal variant. *Genetics* **183**:1227–1234.
768 doi:10.1534/genetics.109.105569
769

770 **Figure Legends**

771

772 **Figure 1: SunTag:SDG2 activates *FWA* mRNA expression.**

773 A) Schematic showing the SunTag system for modulation of chromatin (Gardiner et
774 al., 2022; Papikian et al., 2019). B) Upper panel shows SDG2 with catalytic SET
775 domain (drawn approximately to scale). Lower left depicts predicted alignment error
776 (PAE) plot from AlphaFold3 (Abramson et al., 2024) prediction of the SDG2 coding
777 sequence with a box drawn around the region cloned into SunTag (amino acids 1571-
778 2335). Lower right depicts the AlphaFold3 model for $SDG2_{1571-2335}$, with the position of
779 the amino acid change (Y1903F) for dSDG2 indicated (shown in red). C) RT-qPCR for
780 *FWA* expression from the genotypes indicated. Error bars represent SEM from three
781 biological replicates. *** indicates p-value <0.001 (students T-test). D) ChIP qPCR for
782 the presence of SunTag (left panel) and H3K4me3 enrichment (right panel) over the
783 TSS of *FWA*. The control region (AT5G65130) used for normalisation has high
784 H3K4me3. Error bars represent SEM from 2 biological replicates. E) Genome browser
785 image showing the enrichment of two independent SunTag:SDG2:FWA_g4 lines (2
786 biological replicates each, bio replicate tracks are overlaid) at the *FWA* TSS. F) MA
787 plot comparing the transcriptome of SunTag:SDG2:FWA_g4 as compared to the non-
788 transformed control (*rdr6*). Differentially expressed genes are show in in red (FDR <
789 0.01), with *FWA* labelled and enlarged for visibility.

790

791 **Figure 2: SunTag:SDG2 targeting to *SNC1* enhances resistance to *P. syringae*.**

792 A) Upper, schematic of the *SNC1* gene, with the red line indicating the *SNC1* gRNA
793 target site, drawn to scale. Lower, representative images of 3-week old plants from the
794 genotypes indicated. B) Rosette size quantification. Each dot represents an individual
795 plant. Error bars represent SEM. Different letters indicate significant difference by
796 ANOVA with post-hoc Tukey HSD. C) ChIP qPCR for the presence of SunTag at SNC1.
797 D) RT-qPCR for *SNC1* expression. Error bars represent SEM. * indicates p-value <
798 0.05. E) Pst::LUX assay for colonisation quantification at 3 days post inoculation.
799 Different letters indicate significant difference by ANOVA with post-hoc Tukey HSD.

800

801 **Figure 3: Centromeric targeting of SunTag:SDG2 elevates meiotic crossover**
802 **recombination rate.**

803 A) Region of chromosome 3, showing CENH3 ChIP-seq enrichment (Wlodzimierz et
804 al., 2023) with *CTL3.9* (Wu et al., 2015) red / green T-DNA marker positions and
805 LRCen3 guide RNA binding sites indicated. B) Left panel, representative image of red
806 and green fluorescing seeds from a *CTL3.9* double homozygous line. Right panel,
807 schematic of showing how crossovers are identified within *CTL3.9*. C) Boxplot
808 showing crossover recombination frequency over *CTL3.9* in centiMorgans. Control
809 data is from Col-0 (non-transgenic) crossed to *CTL3.9*. Different letters indicate
810 significant difference by ANOVA with post-hoc Tukey HSD. The red dot indicates the
811 seed set line taken to the next generation (F_4). D) Chromosome-wide plots of
812 chromosome 3. Upper and middle panels show enrichment of H3K4me3 and SunTag
813 (anti-HA) by ChIP-seq in sibling lines with / without SunTag (+/-). Enrichment is
814 calculated as $\log_2$ fold-change over non-transgenic (Col-0) controls in 100 kb windows.
815 The lower panel depicts an LRCen3 guide RNA binding site density over chromosome
816 3. *CTL3.9* marker positions are indicated, and centromeric regions are shown in grey.
817 All data are mapped the Col-CEN genome assembly (Naish et al., 2021).
818
819 **Figure 4: SunTag PRDM9 is sufficient for activation of *FWA*.**
820 A) Upper panel shows PRDM9 with catalytic SET domain (drawn approximately to
821 scale). Lower left depicts predicted alignment error (PAE) plot from AlphaFold3
822 prediction of the PRDM9 coding sequence with a box drawn around the region cloned
823 into SunTag (amino acids 110-417). Lower right depicts the AlphaFold3 model for
824 $PRDM9_{110\text{-}417}$, with the amino acid change for dPRDM9 indicated (position shown in
825 red). B) RT-qPCR for *FWA* (upper panel) and the effector module (sfGFP, lower panel)
826 for the genotypes indicated. Two independent T_3 lines are used per genotype. Error
827 bars represent SEM. C) ChIP qPCR for the presence of SunTag (left panel) and
828 H3K4me3 enrichment (right panel) over the TSS of *FWA*. Error bars represent SEM
829 from 2 biological replicates. D) MA plot comparing the transcriptome of
830 SunTag:PRDM9:FWA_g4 as compared to the non-transformed control (*rdr6*).
831 Differentially expressed genes are show in in red (FDR < 0.01), with *FWA* labelled and
832 enlarged for visibility. E) Genome browser image showing efficient SunTag targeting
833 and H3K4me3 enrichment at *FWA*. Biological replicate tracks from independent lines
834 are overlaid.
835

836 **Supplementary Figure 1:** A) epifluorescence image of rootstocks from 2-week old
837 representative plants showing nuclear localisation of sfGFP from SunTag lines. B)
838 Western blot for the dCas9-10xGCN4 and effector module in SunTag lines showing
839 activation of FWA expression by RT-qPCR show in C).
840
841 **Supplementary Figure 2:** A) Amino acid alignment between SDG2 and two other well
842 characterised A. thaliana hisone methyltransferases, SUVH4 (H3K9me2) and CLF
843 (H3L27me3). The Y at amino acid position 1903 of SDG2 that was mutated to F in the
844 SunTag:dSDG2 constructs corresponds to the Y593F in SUVH4(/KYP) that has
845 previously been shown to abolish methyltransferase catalytic activity (Du et al., 2014)
846 (position indicated in green rectangle). B) CONSURF analysis (Ashkenazy et al.,
847 2016) on the SDG2 predicted SET domain (amino acids 1761-1911) showing the
848 maximum level of conservation for $SDG2_{Y1903}$ (position indicated in green rectangle).
849
850 **Supplementary Figure 3:** A) qRT-PCR for the expression of FWA (upper panel) and
851 sfGFP (lower panel) with individual biological replicates shown. B) ChIP qPCR with
852 two independent FWA targeting lines.
853
854 **Supplementary Figure 4:** A) Venn diagram showing the overlap between the peaks
855 called from HA ChIP-seq of two independent SunTag:SDG2 FWA targeting lines. B)
856 Genome browser images for the overlapping peak regions, the top most significant
857 peak is FWA, while the other two are off-target regions with sequence similarity to the
858 FWA binding site that were previously identified (Papikian et al., 2019).
859
860 **Supplementary Figure 5:** Colonisation dynamics of Pst::LUX in the genotypes
861 indicated. Error bars represent SEM.
862
863 **Supplementary Figure 6:** A) Recombination rate over *CTL3.9* of the F_3 lines (as
864 shown in main Figure 3). Blue arrow (also shown in B) points the P4 seed set taken to
865 the next generation (also see main Figure 3C, red dot). B) Distortion ratios for the red
866 / green T-DNAs from *CTL3.9* from the $F_{3'}$s. C) Recombination rate over *CTL3.9* ($F_{3'}$s)
867 including all scorable individuals including no guide lines. D) Recombination rate over
868 *CTL3.9* ($F_{3'}$s) filtered for individuals with no distortion (i.e. 3:1).
869

870 **Supplementary Figure 7:** A-B) ChIP replicate comparisons by Pearsons correlation
871 clustering using 25bp bin resolution. C-D) ChIPseq data chromosomal-wide plots as
872 shown in Figure 3, over all chromosomes from sibling lines with or without
873 SunTag:SDG2:LRCen3 (+/-). Enrichment is calculated as log2 fold change over non-
874 transgenic (Col-0) controls in 100kb windows. *CTL3.9* marker positions are indicated,
875 and centromeric regions are shown in grey.
876
877 **Supplementary Figure 8:** A) LRCen3 guide RNA binding site density plot with 0
878 mismatches or B) with 1 mismatch tolerance.
879
880 **Supplementary Figure 9:** Representative images from 3-week old plants.
881
882 **Supplementary Figure 10:** A) Upper panel, western blot showing bulk levels of
883 H3K4me3 normalised to H3 in the genotypes indicated. Lower panel, ratio of
884 H3K4me3/H3. B) RT-qPCR for expression of FWA (upper) and sfGFP (middle) and
885 ratio (lower) in the genotypes indicated.
886
887 **Supplementary Figure 11:** A) RT-qPCR of FWA expression (upper) and sfGFP
888 expression (lower) for individual T_1 plants from the genotypes indicated. B) summary
889 table showing the number of FWA activating T_1 lines per construct.
890
891 **Supplementary Figure 12:** A) MA plot for transcriptome of SunTag:SDG2:No_guide
892 versus the non-transformed control (*rdr6*). B) MA plot for transcriptome of
893 SunTag:PDRM9:No_guide versus the non-transformed control (*rdr6*). Differentially
894 expressed genes are depicted in red (FDR < 0.1). Note that FWA is not identified as
895 differentially expressed in either comparison.
896
897 **Supplementary Table 1:** Oligo List.

SunTag cloning	
SDG2_BsiWI_Fw	GAGGTCGGACCGGTCGTACG
SDG2_mutation_Rev	ATTAAAATCGAAAGTTATCTCCTCGCC
SDG2_mutation_Fw	GATTTTAATTCTGTAACTGAGAGTAAG
SDG2_BsiWI_Rev	CGAGATCCTCCTCCCGTACG
Universal F1	GTCATCTATGTTACGAAGCTTTCGTTGAACAACGG
Universal R2	TCGTGCTCCACCATGGTACCAAAAAAAGCACCGACTCGGTGC

pFWA_gRNA4_F2	acggaaagatgtatgggcttGTTTTAGAGCTAGAAATAGC
pFWA_gRNA4_R1	aagcccatacatctttccgtCAATCACTACTTCGACTCTA
pSNC1_gRNA4_F2	ctatgcaagtcttctaagaaGTTTTAGAGCTAGAAATAGC
pSNC1_gRNA4_R1	ttcttagaagacttgcatagCAATCACTACTTCGACTCTA
LRCen3_g_F2	aggcttacaagattgggttgGTTTTAGAGCTAGAAATAGC
LRCen3_g_R1	caacccaatcttgtaagcctCAATCACTACTTCGACTCTA
PRDM9 cloning into SunTag	
RP_PRDM9_G282A_3'	cgagatcctcctcccgtacgtcaaa
FP_PRDM9_G282A_5'	GAGGTCGGACCGGTCGTACGTATC
RP_PRDM9_G282A_middle	CTCTGTGATCTGAGCCTCATAGGGGCCA
FP_PRDM9_G282A_middle	TGGCCCCTATGAGGCTCAGATCACAGAG
ChIP-qPCRs	
High_H3K4me3_AT5G65130_F	CCAGACATTGTTCGTCAAGGACAC
High_H3K4me3_AT5G65130_R	CAACGTTATATCCGACTCAGGAGATG
Low_H3K4me3_AT2G40004up_F	CAGCTACAAGTTCAACTAATCATGC
Low_H3K4me3_AT2G40004up_R	CTTTAATTGCCCGTCATTACGCTAC
pFWA_ChIP_Fw	AATATCAGATCTTGCGCCGCTC
pFWA_ChIP_Rev	GAAAAACAGACAAATCGGGAACC
pSNC1_ChIP_Fw1	ATGGCCTTATCTTGTTAGG
pSNC1_ChIP_Fw2	GAGTAGAAGCTTTCCAAGG
RT-qPCRs	
sfGFP_qPCR_Fw	CAAAGATGACGGGACCTACAA
sfGFP_qPCR_Rev	GTACTCGAGTTTGTGTCCAAGA
FWA_qPCR_Fw	TTAGATCCAAAGGAGTATCAAAG
FWA_qPCR_Rev	CTTTGGTACCAGCGGAGA
SNC1_qPCR_Fw	TTTGAAGAGACATGCAAGGC
SNC1_qPCR_Rev	TCGGCAAGCTCTTCAATCATGG

Figure 1

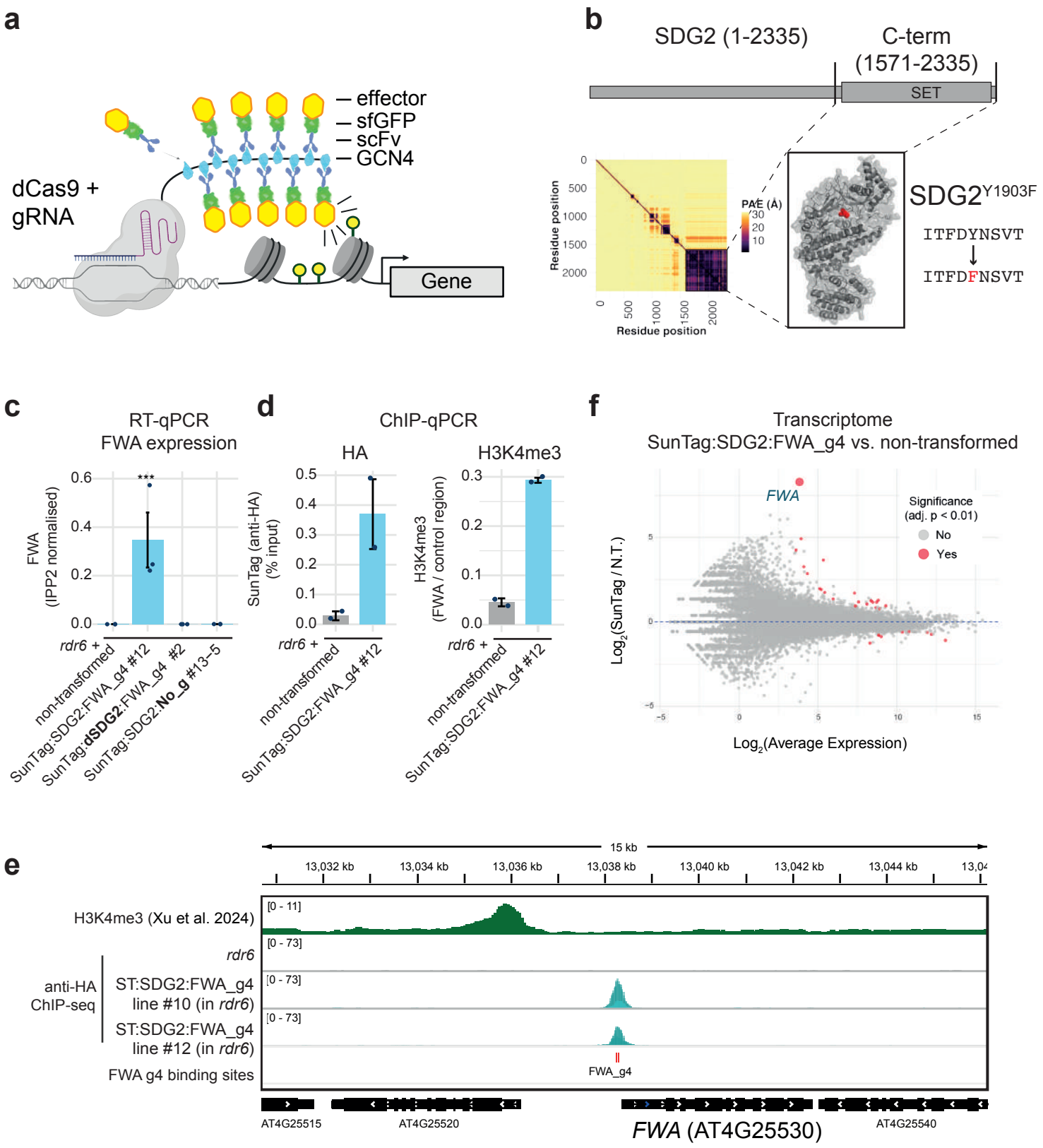

a
effector
sfGFP
scFv
GCN4
dCas9 +
gRNA
Gene
b
SDG2 (1-2335)
C-term
(1571-2335)
SET
Residue position
PAE (Å)
SDG2Y1903F
ITFDYNSVT
ITFDFNSVT
c
RT-qPCR
FWA expression
FWA
(IPP2 normalised)
rdr6 +
non-transformed
SunTag:SDG2:FWA_g4 #12
SunTag:SDG2:FWA_g4 #2
SunTag:dSDG2:FWA_g4
SunTag:SDG2:No_g #13-5
d
ChIP-qPCR
HA
H3K4me3
SunTag (anti-HA)
(% input)
H3K4me3
(FWA / control region)
f
Transcriptome
SunTag:SDG2:FWA_g4 vs. non-transformed
FWA
Significance
(adj. p < 0.01)
No
Yes
Log2(SunTag / N.T.)
Log2(Average Expression)
e
15 kb
13,032 kb
13,034 kb
13,036 kb
13,038 kb
13,040 kb
13,042 kb
13,044 kb
H3K4me3 (Xu et al. 2024)
rdr6
anti-HA
ChIP-seq
ST:SDG2:FWA_g4
line #10 (in rdr6)
ST:SDG2:FWA_g4
line #12 (in rdr6)
FWA g4 binding sites
FWA_g4
AT4G25515
AT4G25520
FWA (AT4G25530)
AT4G25540

Figure 2

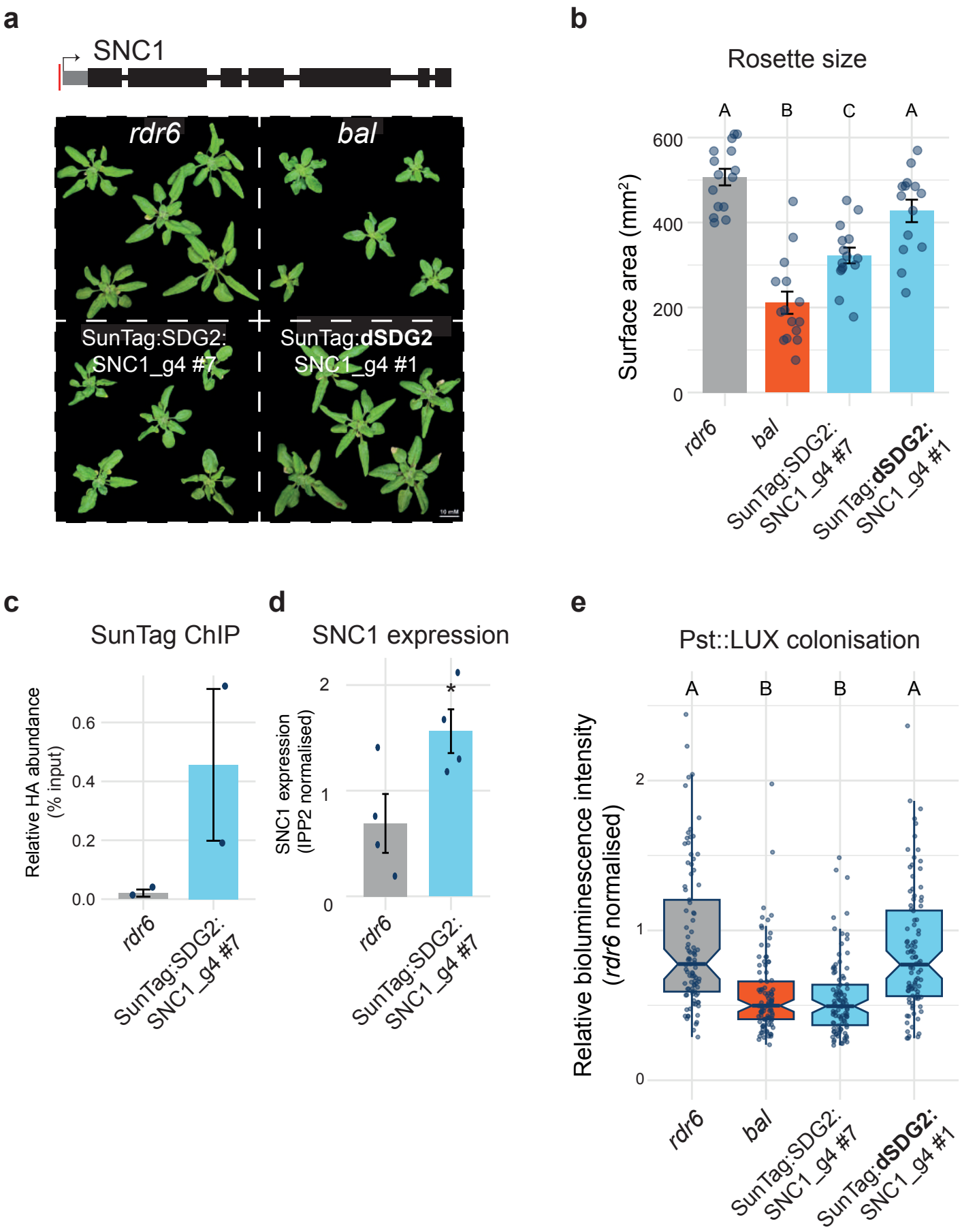

a
SNC1
rdr6
bal
SunTag:SDG2:
SNC1_g4 #7
SunTag:dSDG2
SNC1_g4 #1
b
Rosette size
Surface area (mm²)
A
B
C
A
600
400
200
0
rdr6
bal
SunTag:SDG2:
SNC1_g4 #7
SunTag:dSDG2:
SNC1_g4 #1
c
SunTag ChIP
Relative HA abundance
(% input)
0.6
0.4
0.2
0.0
rdr6
SunTag:SDG2:
SNC1_g4 #7
d
SNC1 expression
SNC1 expression
(IPP2 normalised)
2
1
0
*
rdr6
SunTag:SDG2:
SNC1_g4 #7
e
Pst::LUX colonisation
Relative bioluminescence intensity
(rdr6 normalised)
A
B
B
A
2
1
0
rdr6
bal
SunTag:SDG2:
SNC1_g4 #7
SunTag:dSDG2:
SNC1_g4 #1

Figure 3

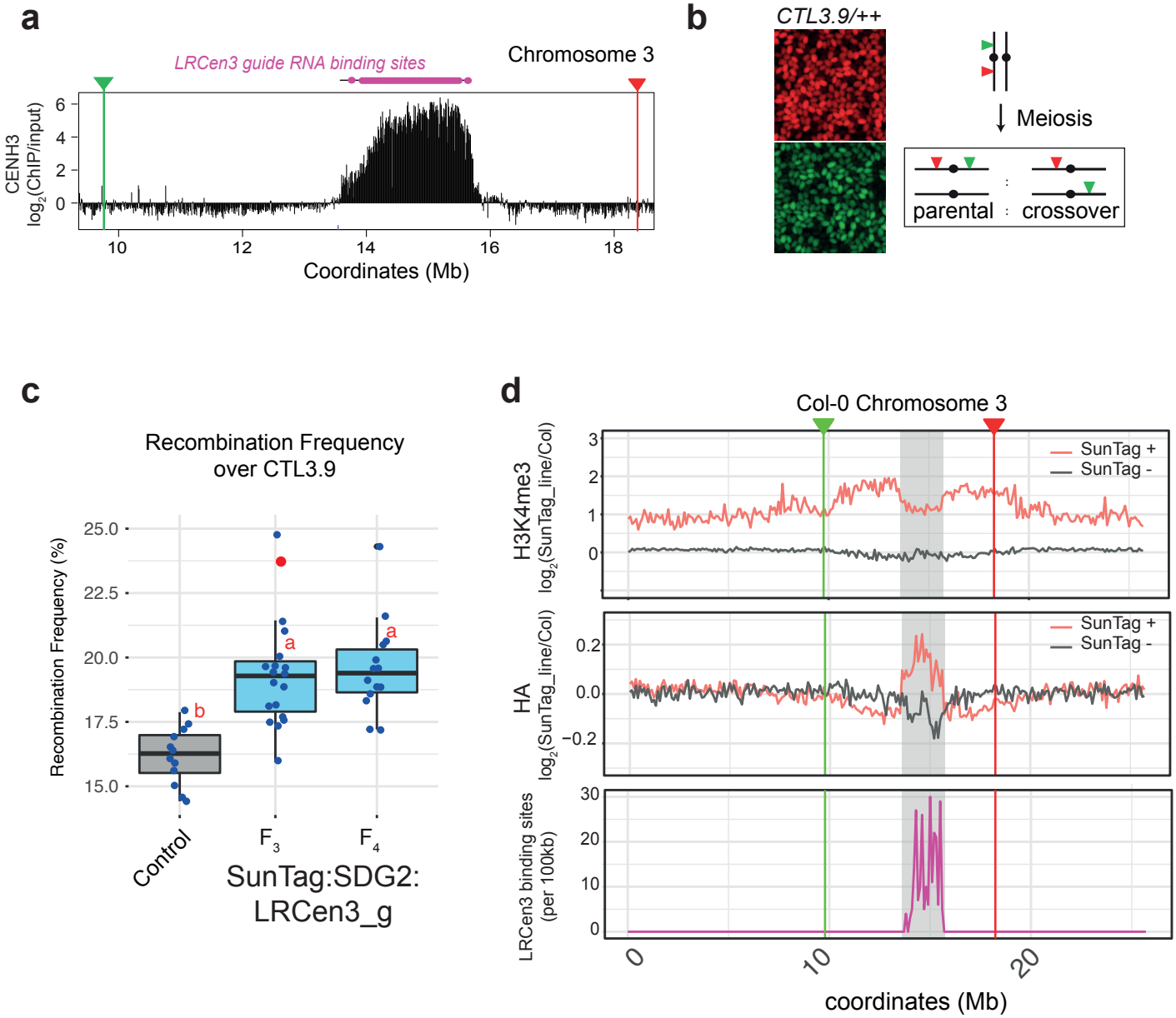
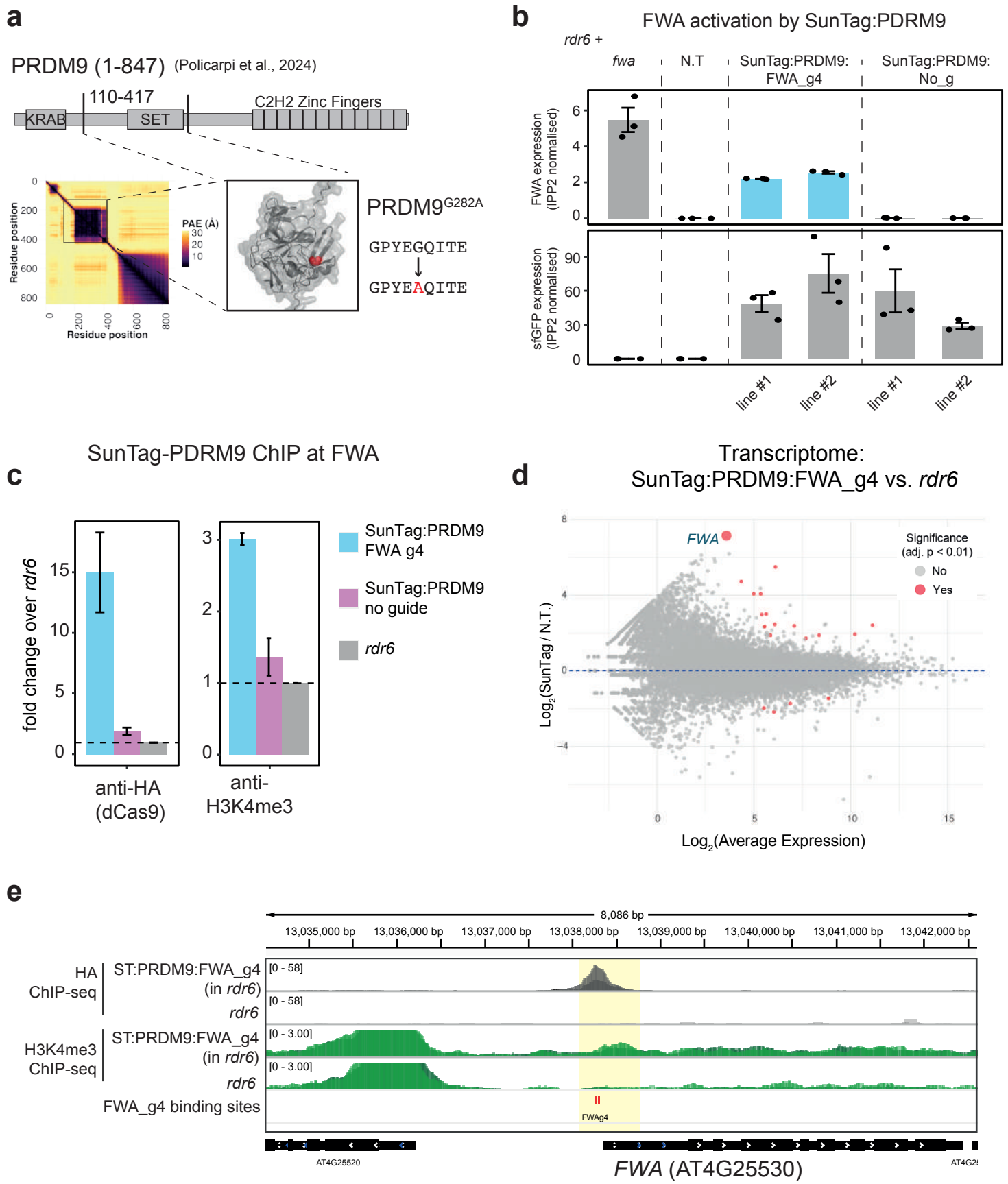
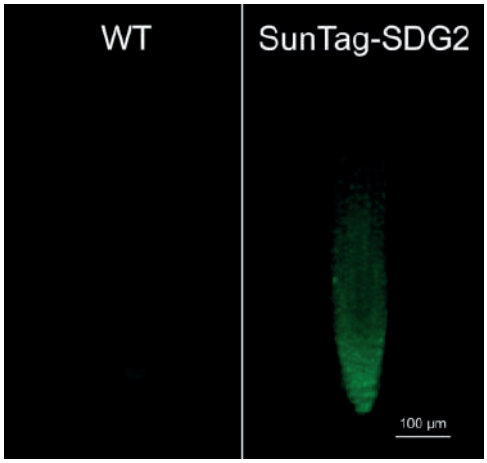

a
LRCen3 guide RNA binding sites
Chromosome 3
CENH3
log2(ChIP/input)
6
4
2
0
10
12
14
16
18
Coordinates (Mb)
b
CTL3.9/++
Meiosis
parental
crossover
c
Recombination Frequency
over CTL3.9
Recombination Frequency (%)
25.0
22.5
20.0
17.5
15.0
b
a
a
Control
F3
F4
SunTag:SDG2:
LRCen3_g
d
Col-0 Chromosome 3
H3K4me3
log2(SunTag_line/Col)
SunTag +
SunTag -
HA
log2(SunTag_line/Col)
LRCen3 binding sites
(per 100kb)
coordinates (Mb)

Figure 4

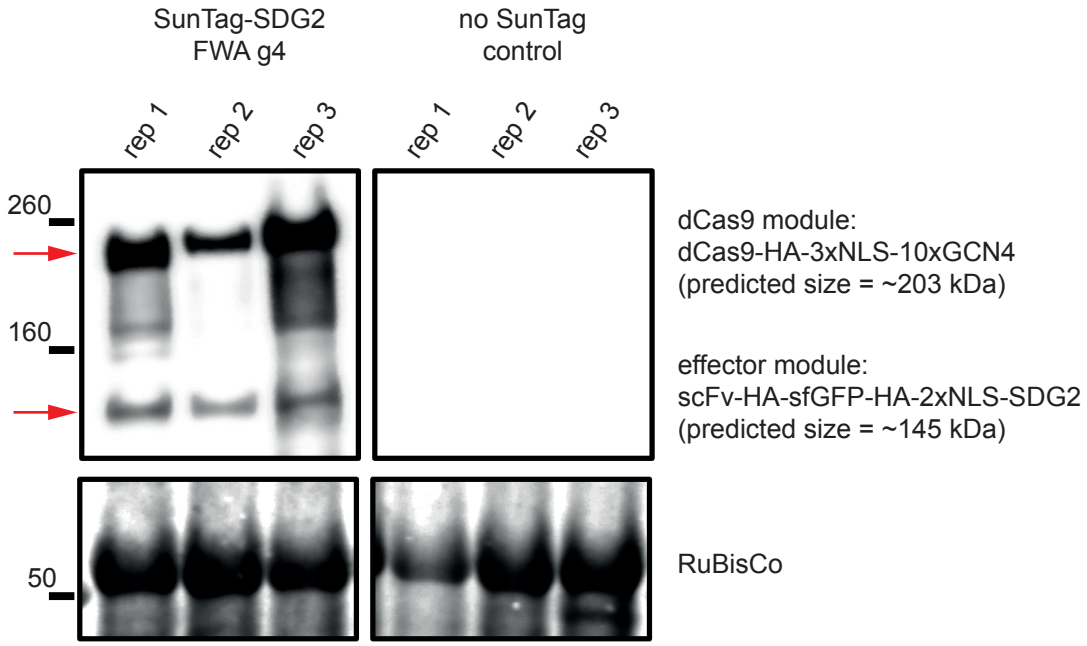

a
PRDM9 (1-847) (Policarpi et al., 2024)
110-417
KRAB
SET
C2H2 Zinc Fingers
Residue position
PAE (Å)
PRDM9G282A
GPYEGQITE
GPYEAQITE
b
FWA activation by SunTag:PDRM9
rdr6 +
fwa
N.T
SunTag:PRDM9: FWA_g4
SunTag:PRDM9: No_g
FWA expression (IPP2 normalised)
sfGFP expression (IPP2 normalised)
line #1
line #2
line #1
line #2
c
SunTag-PDRM9 ChIP at FWA
fold change over rdr6
anti-HA (dCas9)
anti-H3K4me3
SunTag:PRDM9 FWA g4
SunTag:PRDM9 no guide
rdr6
d
Transcriptome: SunTag:PRDM9:FWA_g4 vs. rdr6
FWA
Significance (adj. p < 0.01)
No
Yes
Log2(SunTag / N.T.)
Log2(Average Expression)
e
8,086 bp
13,035,000 bp
13,036,000 bp
13,037,000 bp
13,038,000 bp
13,039,000 bp
13,040,000 bp
13,041,000 bp
13,042,000 bp
HA ChIP-seq
ST:PRDM9:FWA_g4 (in rdr6)
rdr6
H3K4me3 ChIP-seq
ST:PRDM9:FWA_g4 (in rdr6)
rdr6
FWA_g4 binding sites
[0 - 58]
[0 - 3.00]
FWAg4
AT4G25520
FWA (AT4G25530)

S1

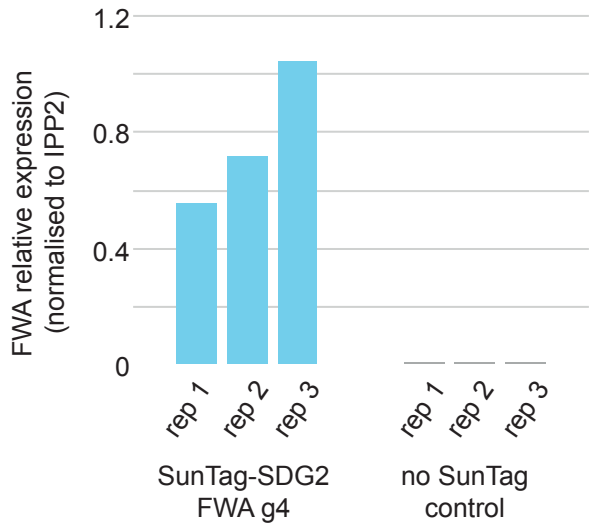
a


WT
SunTag-SDG2
100 µm

b


SunTag-SDG2
FWA g4
no SunTag
control
rep 1
rep 2
rep 3
rep 1
rep 2
rep 3
260
160
50
dCas9 module:
dCas9-HA-3xNLS-10xGCN4
(predicted size = ~203 kDa)
effector module:
scFv-HA-sfGFP-HA-2xNLS-SDG2
(predicted size = ~145 kDa)
RuBisCo

c


FWA relative expression
(normalised to IPP2)
1.2
0.8
0.4
0
rep 1
rep 2
rep 3
rep 1
rep 2
rep 3
SunTag-SDG2
FWA g4
no SunTag
control

S2

a SDG2 alignment over SET domain:

```
SDG2    1858 ANYASRICHSCRPNCEAKVTAVDG----HYQIGIYSVRAIEYGEEITFDYNSVTESKEEY 1913
SUVH4    544 GNFARFINHSCEPNLFVQCVLSSHQDIRLARVVLFAADNISPMQELTYDYGYALDSVHGP  603
CLF      821 GDKLKFANHSPEPNCYAKVIMVAG----DHRVGIFAKERILAGEELFYDYRYEPDRAPAW  876
```

b CONSURF analysis:

The conservation scale:

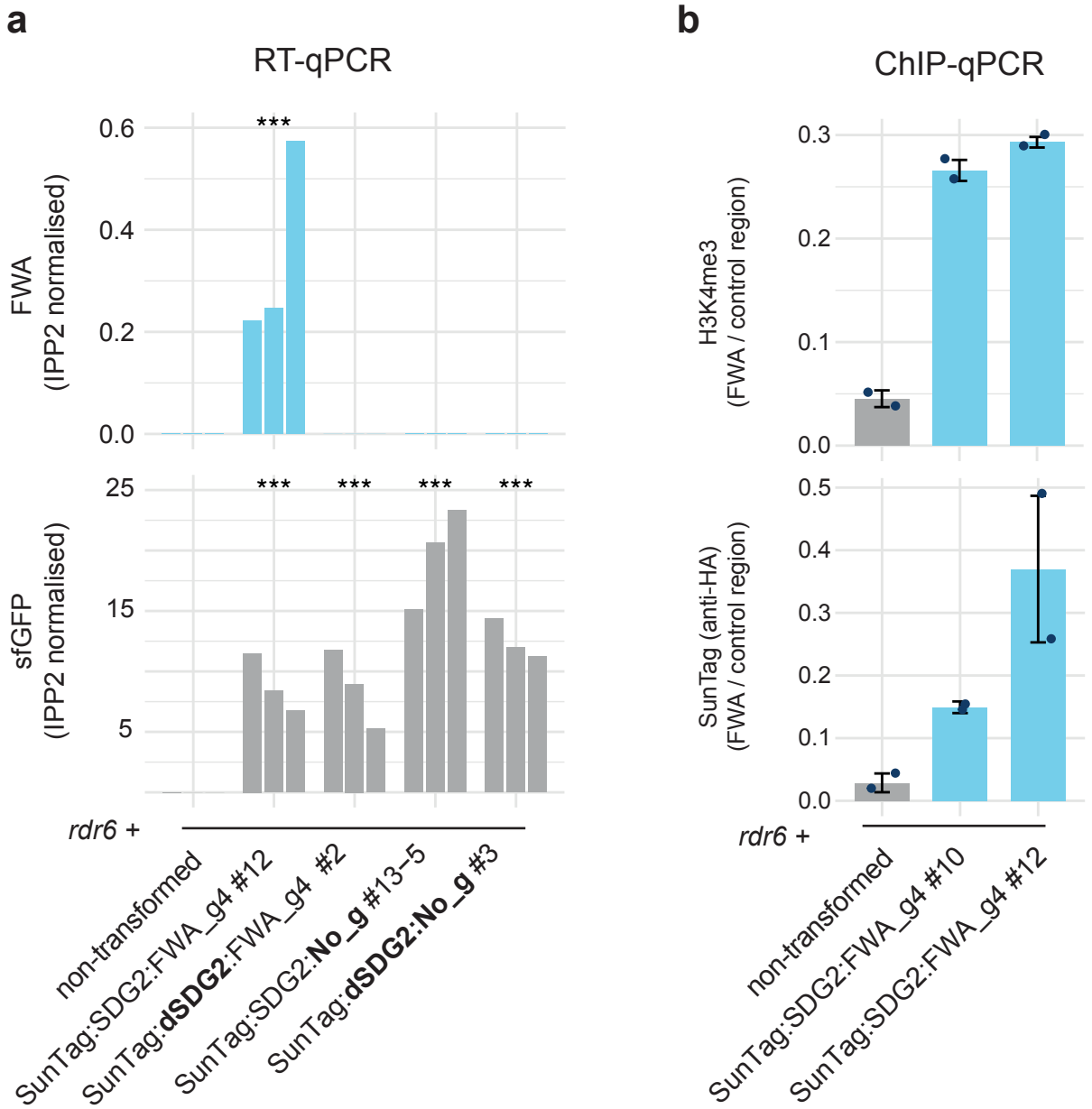
e - An exposed residue according to the neural network algorithm.

b - A buried residue according to the neural network algorithm.

f - A predicted functional residue (highly conserved and exposed).

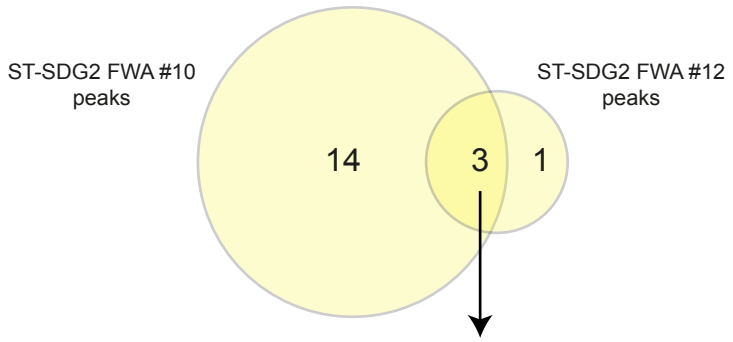
s - A predicted structural residue (highly conserved and buried).

S3


a
RT-qPCR
FWA
(IPP2 normalised)
0.6
0.4
0.2
0.0
***
sfGFP
(IPP2 normalised)
25
15
5
***
***
***
***
rdr6 +
non-transformed
SunTag:SDG2:FWA_g4 #12
SunTag:dSDG2:FWA_g4 #2
SunTag:SDG2:No_g #13–5
SunTag:dSDG2:No_g #3
b
ChIP-qPCR
H3K4me3
(FWA / control region)
0.3
0.2
0.1
0.0
SunTag (anti-HA)
(FWA / control region)
0.5
0.4
0.3
0.2
0.1
0.0
rdr6 +
non-transformed
SunTag:SDG2:FWA_g4 #10
SunTag:SDG2:FWA_g4 #12

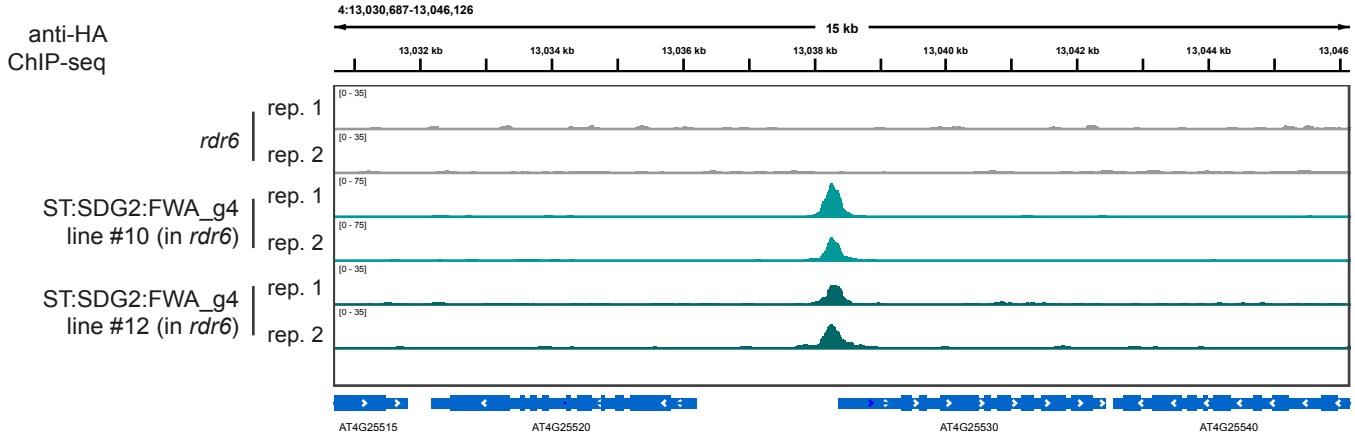
S4

a

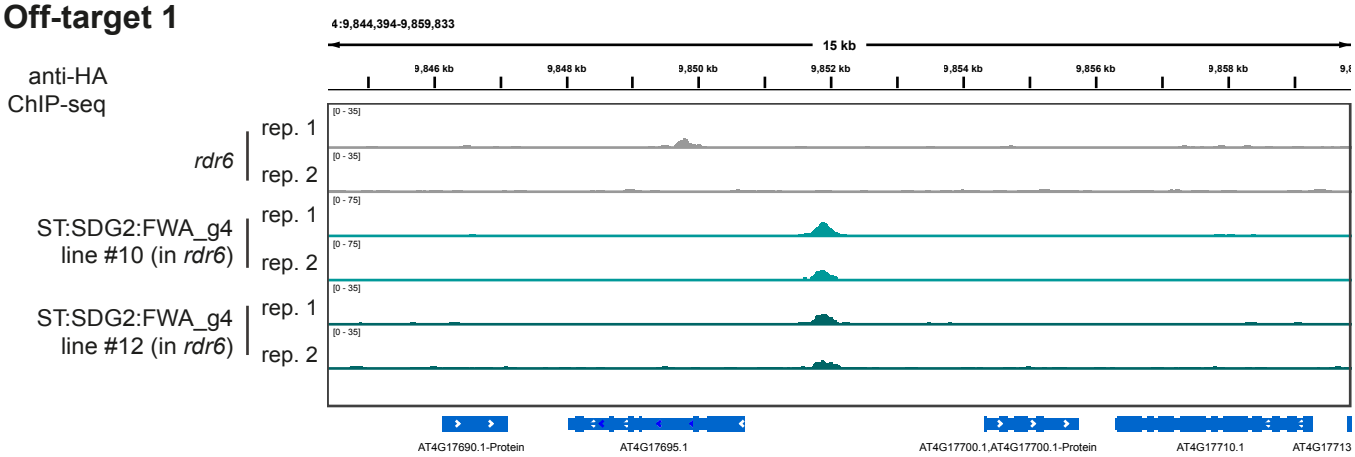


b

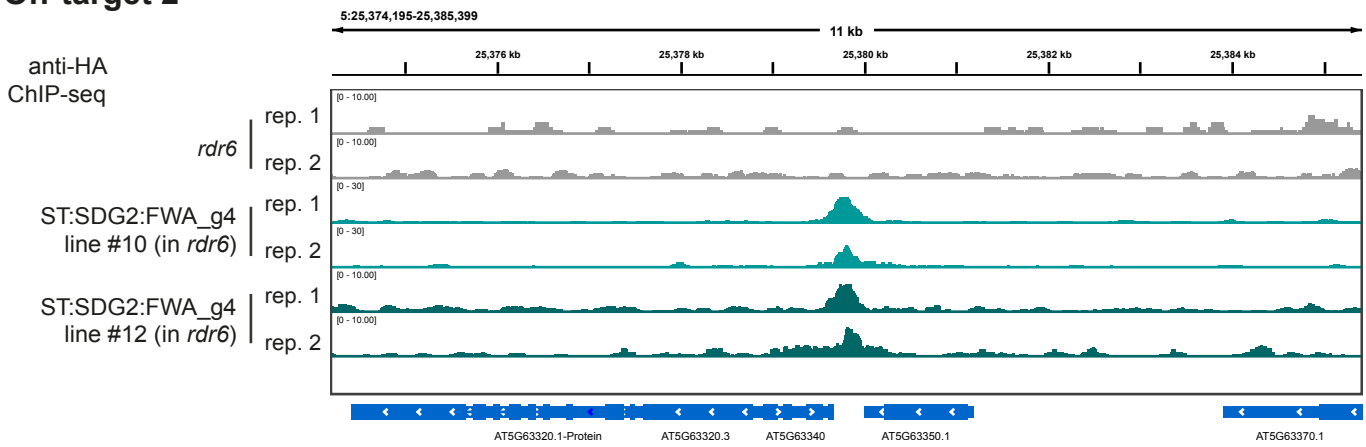
FWA (target)



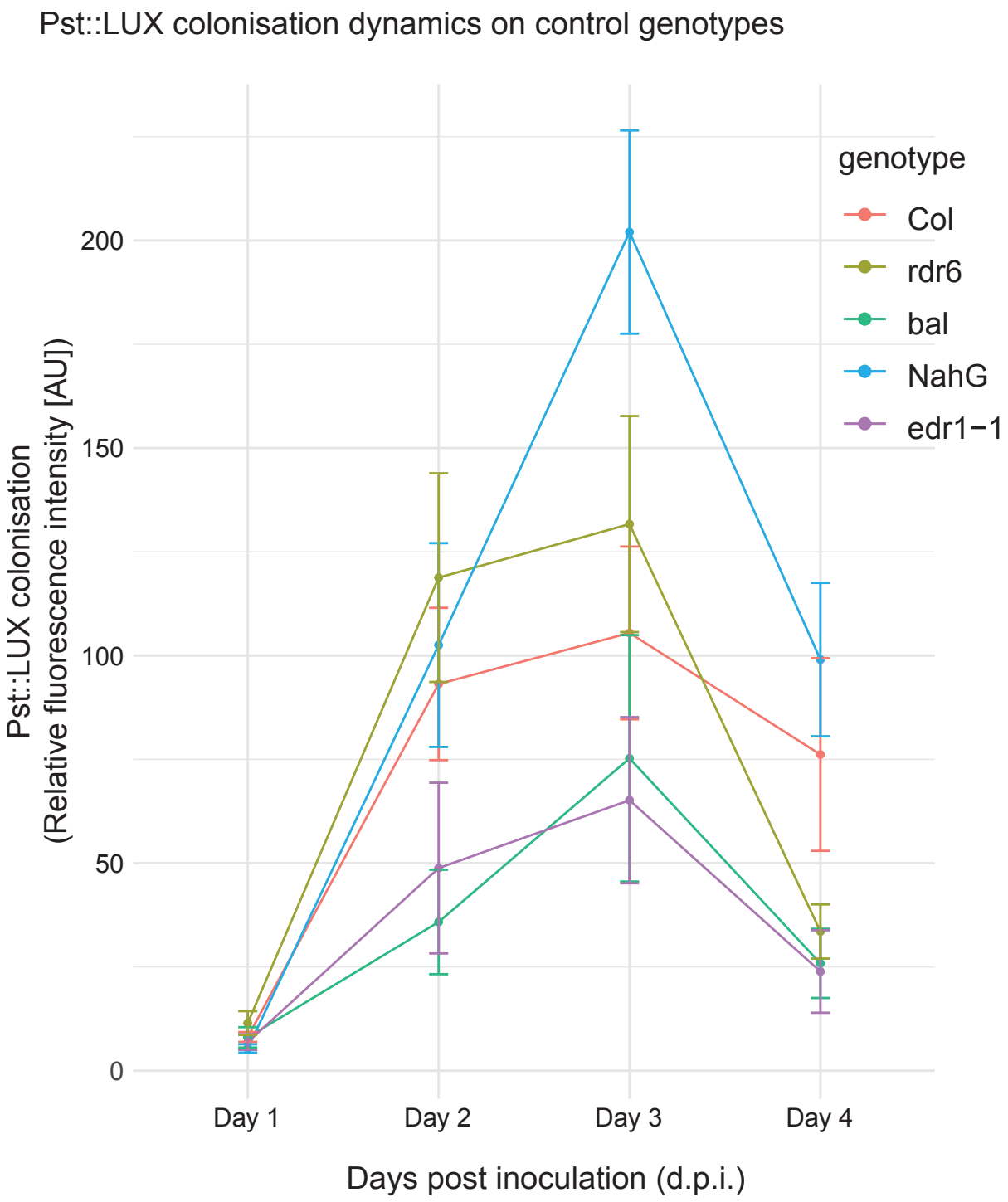
Off-target 1



Off-target 2



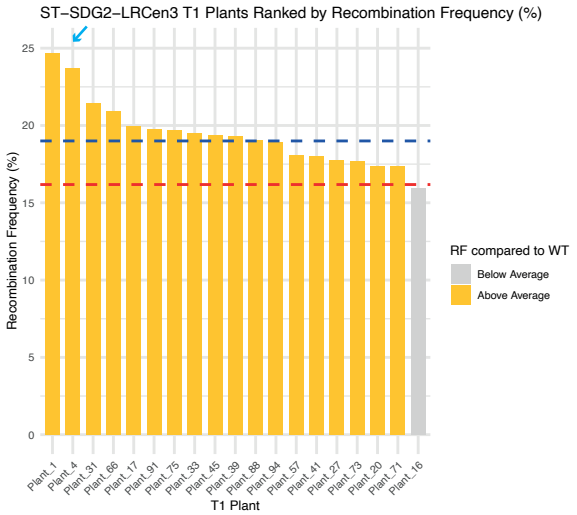
S5


Pst::LUX colonisation dynamics on control genotypes
genotype
Col
rdr6
bal
NahG
edr1−1
Pst::LUX colonisation
(Relative fluorescence intensity [AU])
200
150
100
50
0
Day 1
Day 2
Day 3
Day 4
Days post inoculation (d.p.i.)

S6

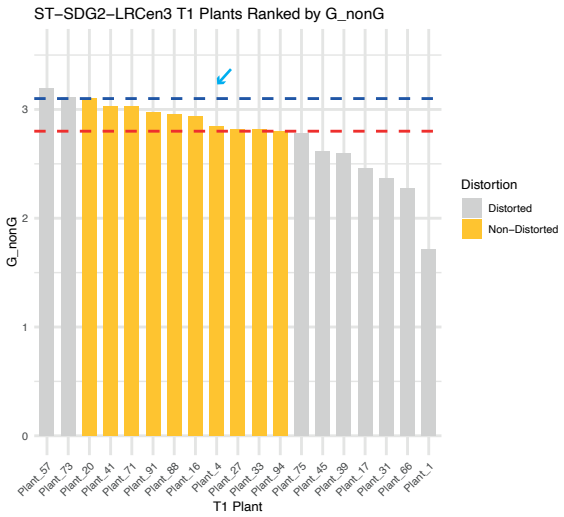
a

Recombination rate of SunTag:SDG2:LRCen3_g independent transformants over CTL3.9

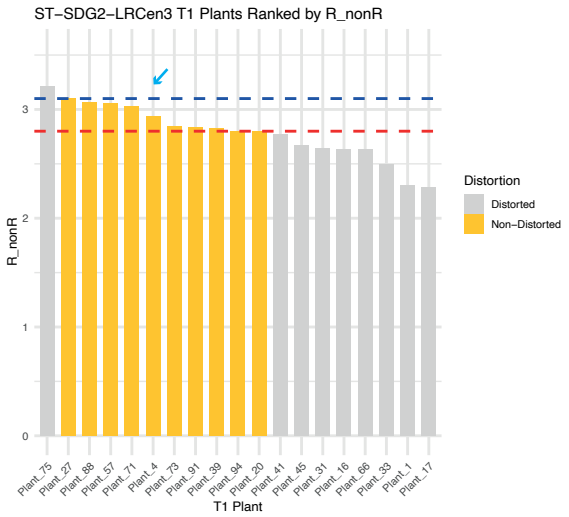


b

Green to non-green distortion



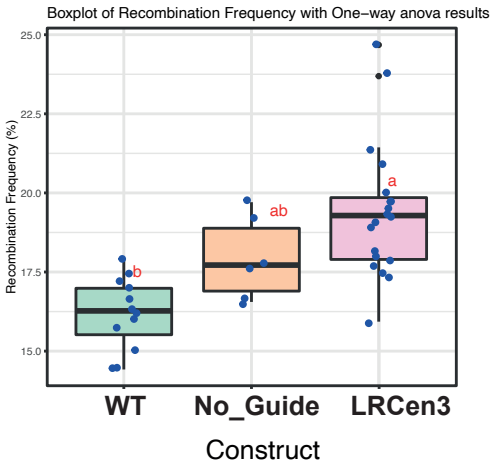
Red to non-red distortion



Recombination frequency over CTL3.9 with different SunTag-SDG2 guide RNAs

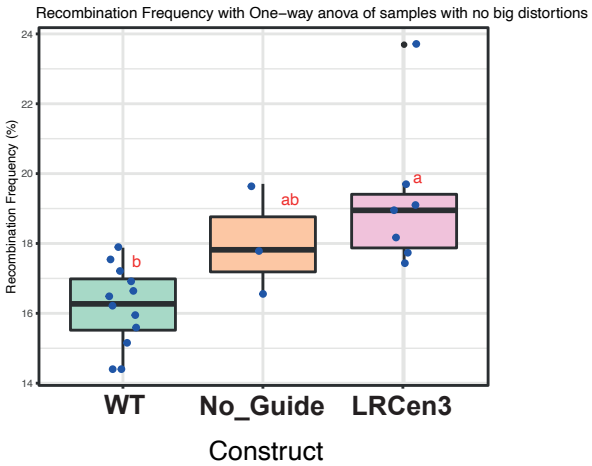
c

all scored individuals



d

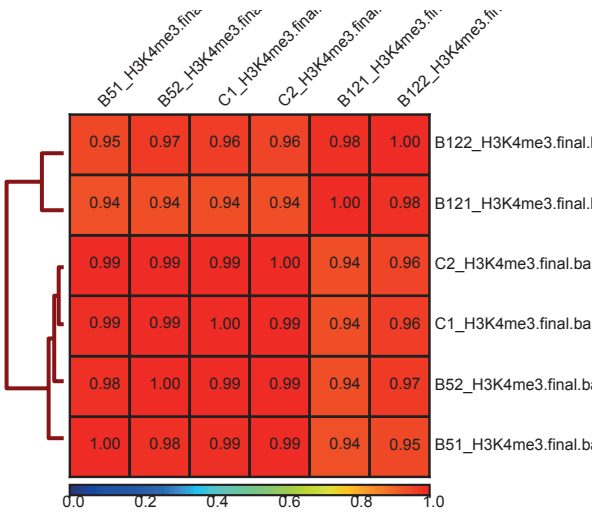
filtered for distortion



S7

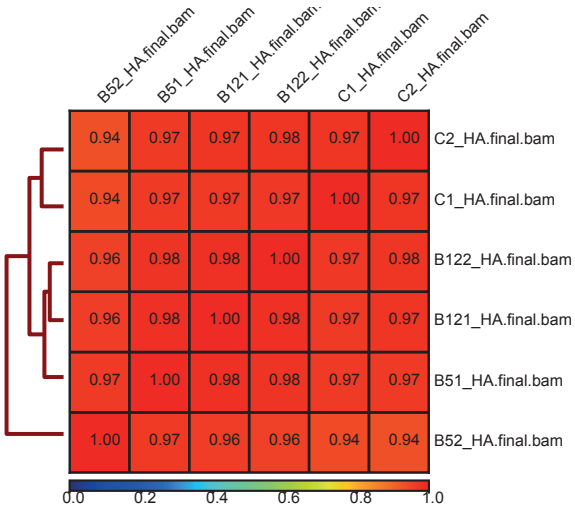
a

anti-H3K4me3 bam comparison



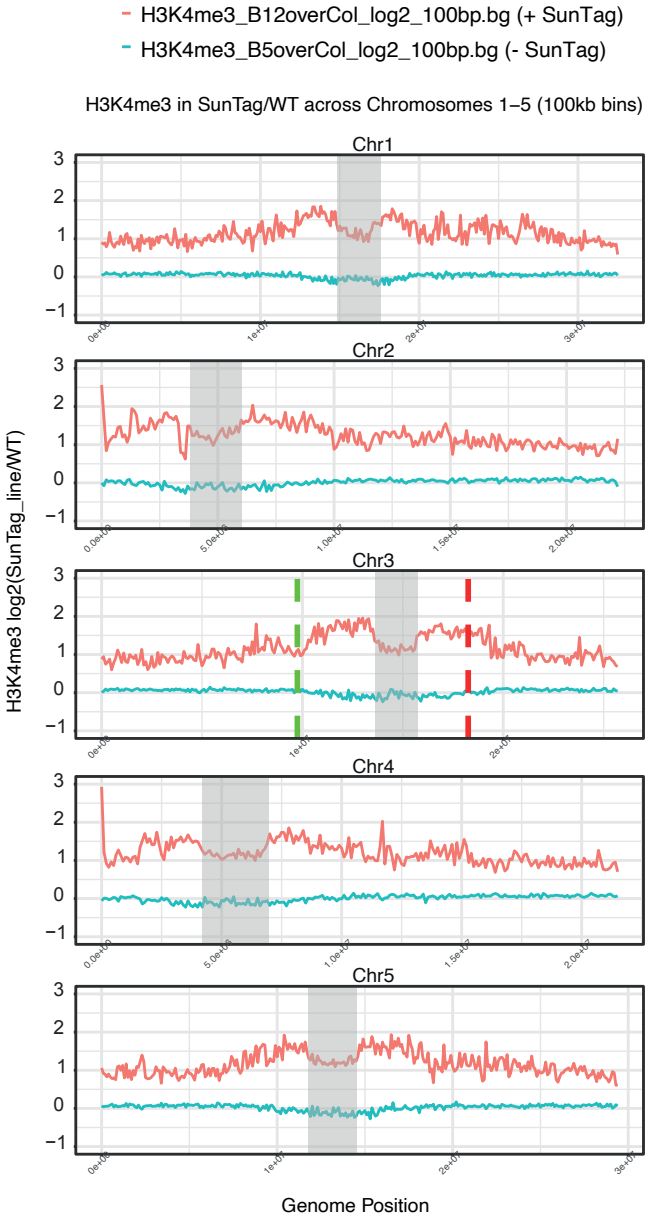
b

anti-HA comparison bam comparison



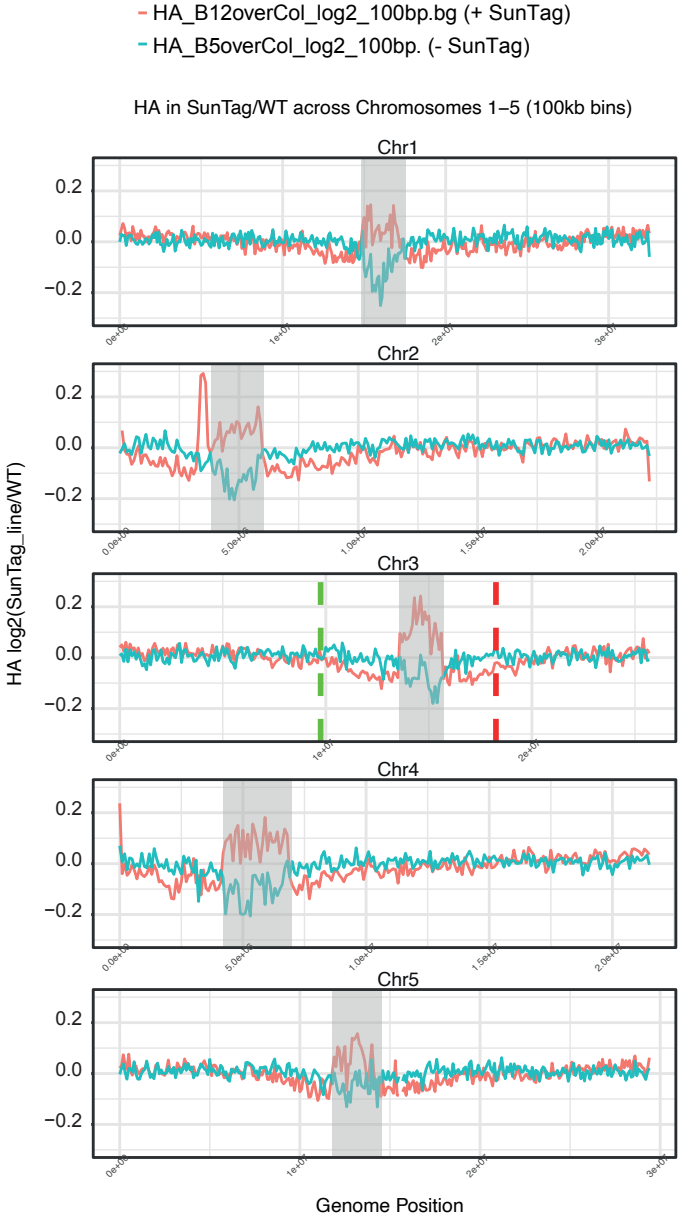
c

$\log_2$(SunTag/Col-0) H3K4me3 over all chromosomes

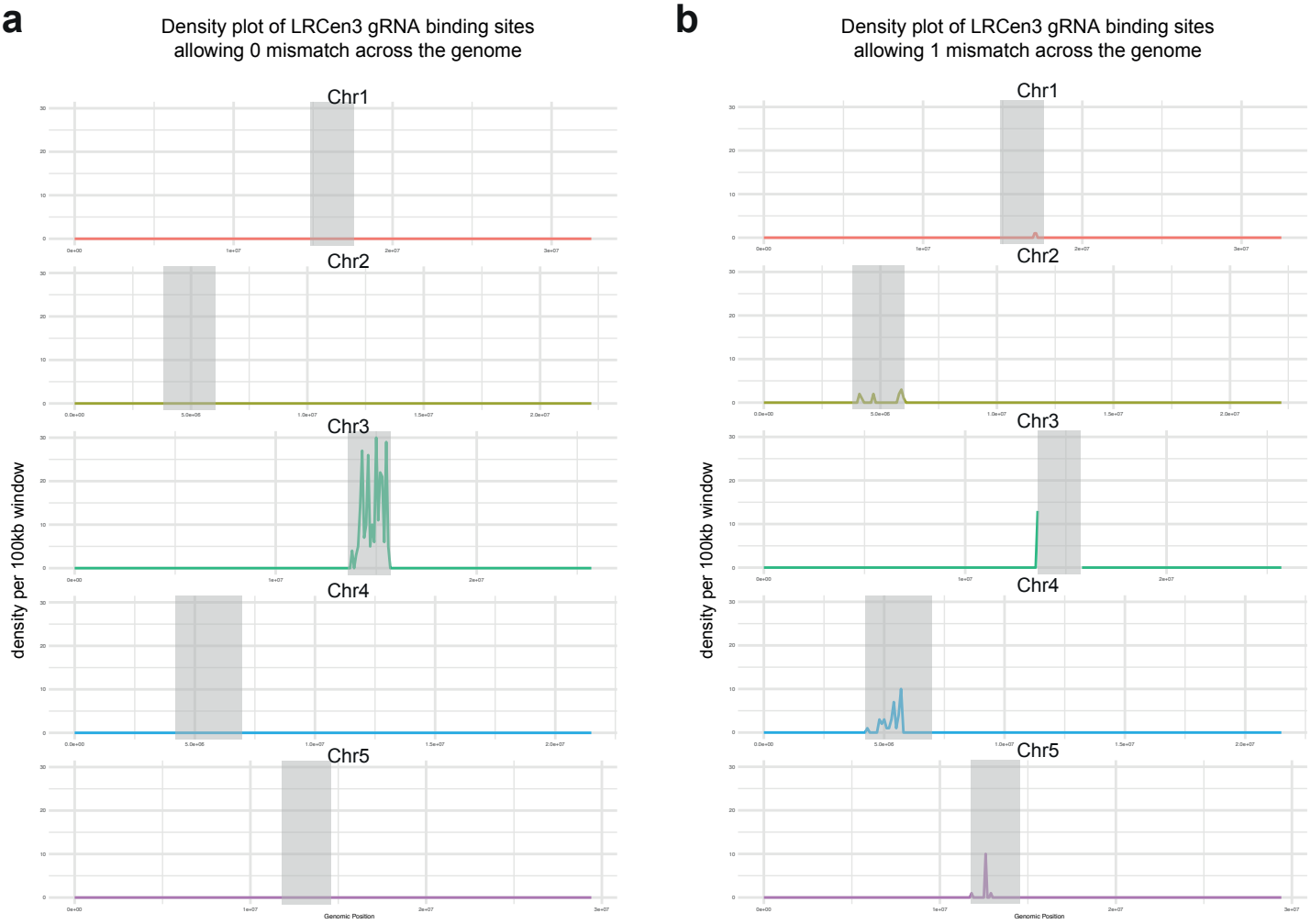


d

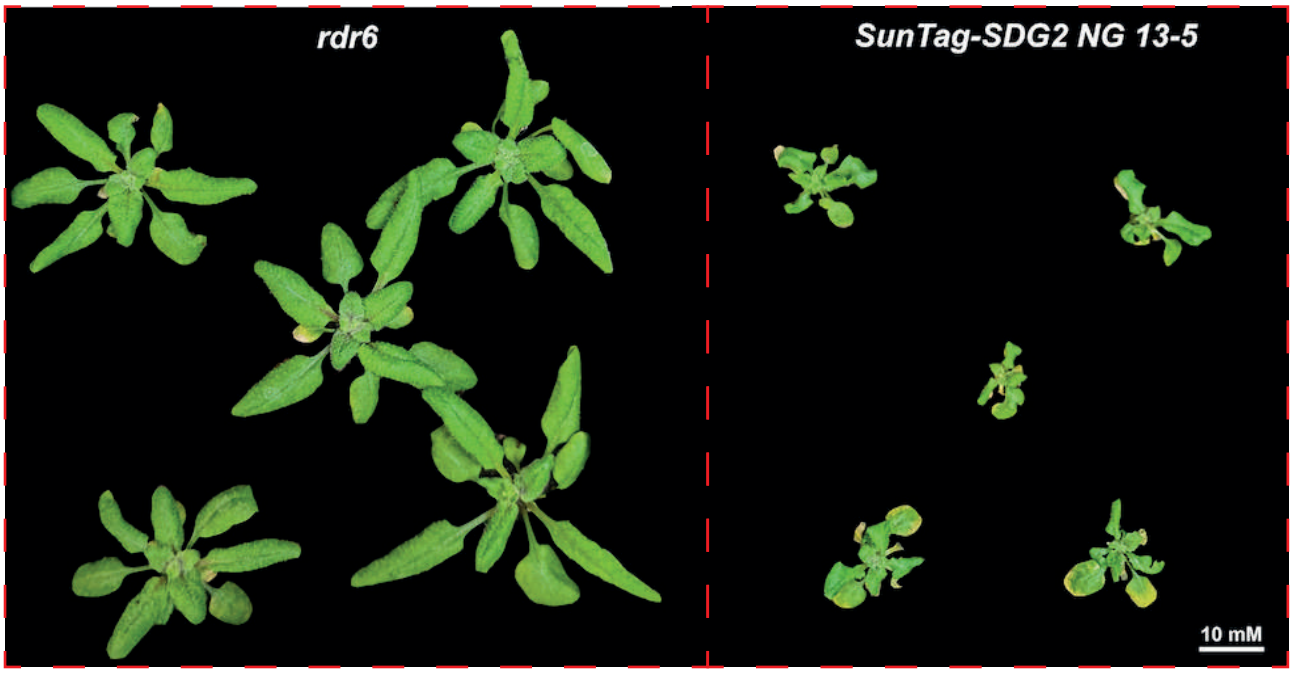
$\log_2$(SunTag/Col-0) HA over all chromosomes



S8


a
Density plot of LRCen3 gRNA binding sites
allowing 0 mismatch across the genome
Chr1
Chr2
Chr3
Chr4
Chr5
density per 100kb window
Genomic Position
b
Density plot of LRCen3 gRNA binding sites
allowing 1 mismatch across the genome
Chr1
Chr2
Chr3
Chr4
Chr5
density per 100kb window
Genomic Position

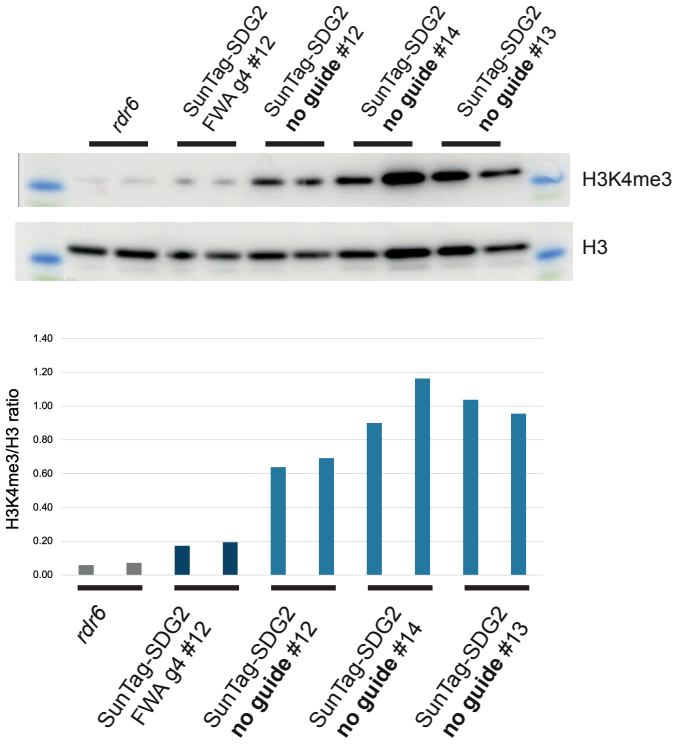
S9



S10

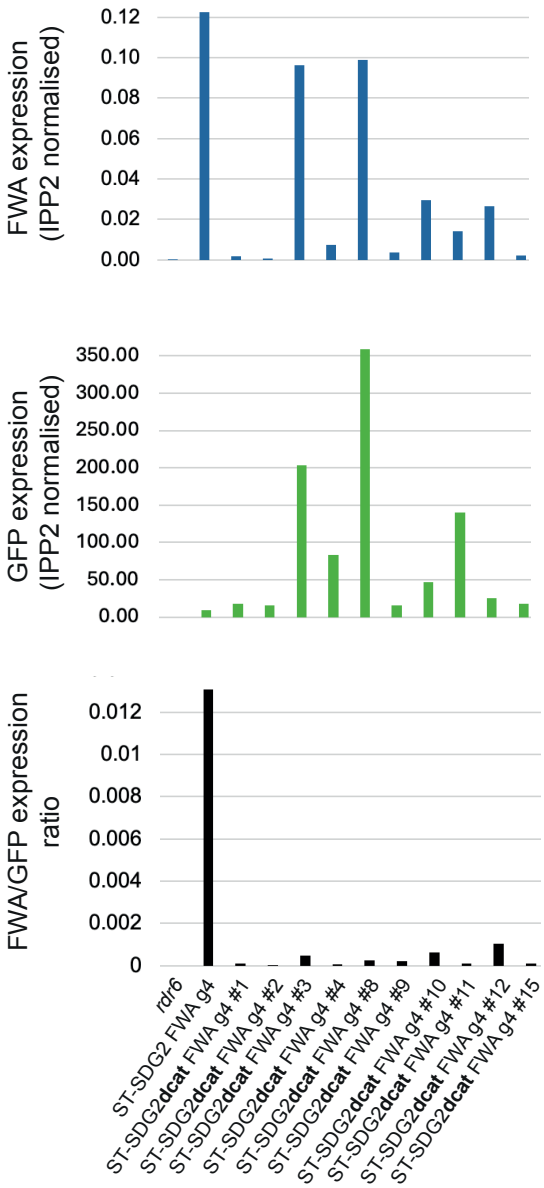
a

SunTag-SDG2 no guide causes bulk increases in H3K4me3



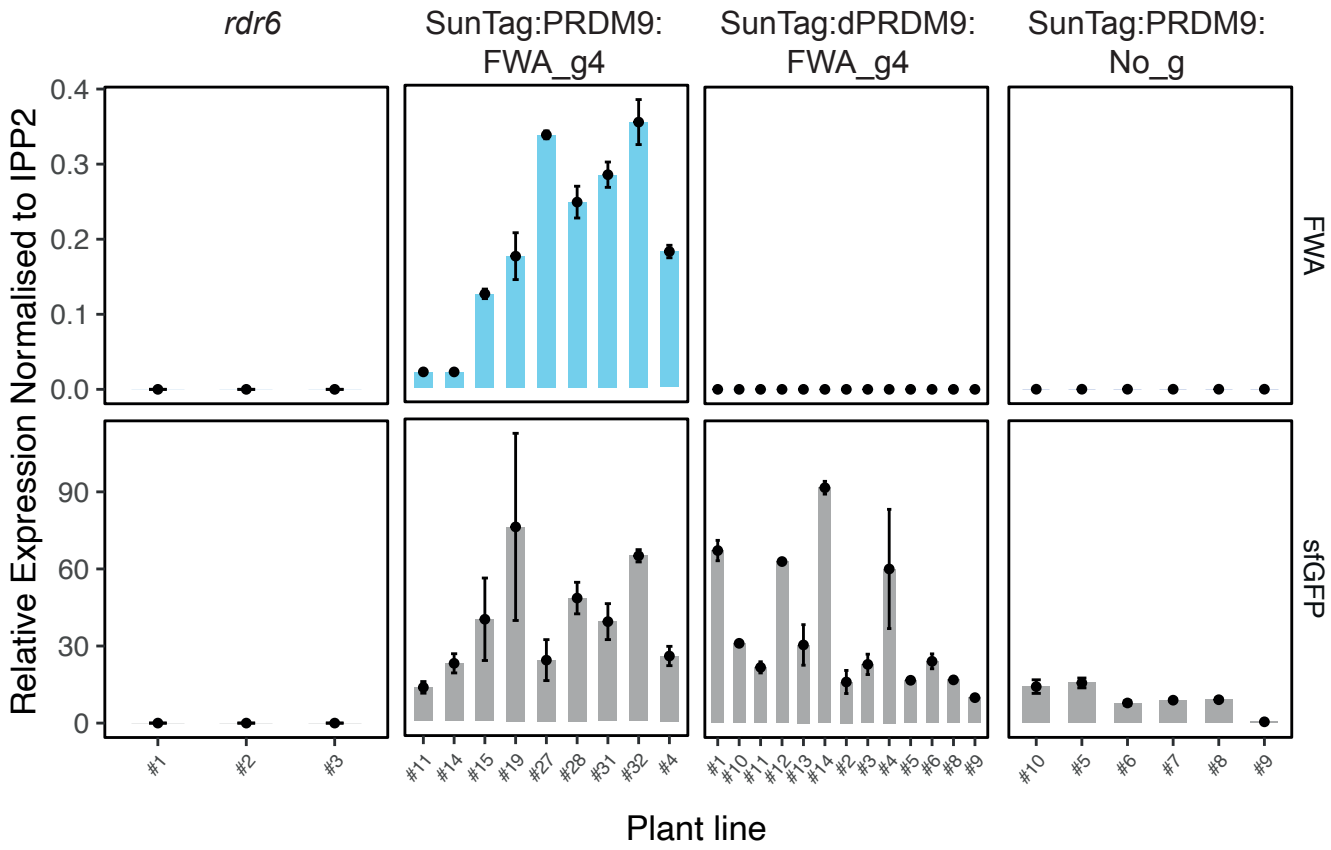
b

Overexpression of SunTag-SDG2**dcat** effector can activate FWA



S11

a



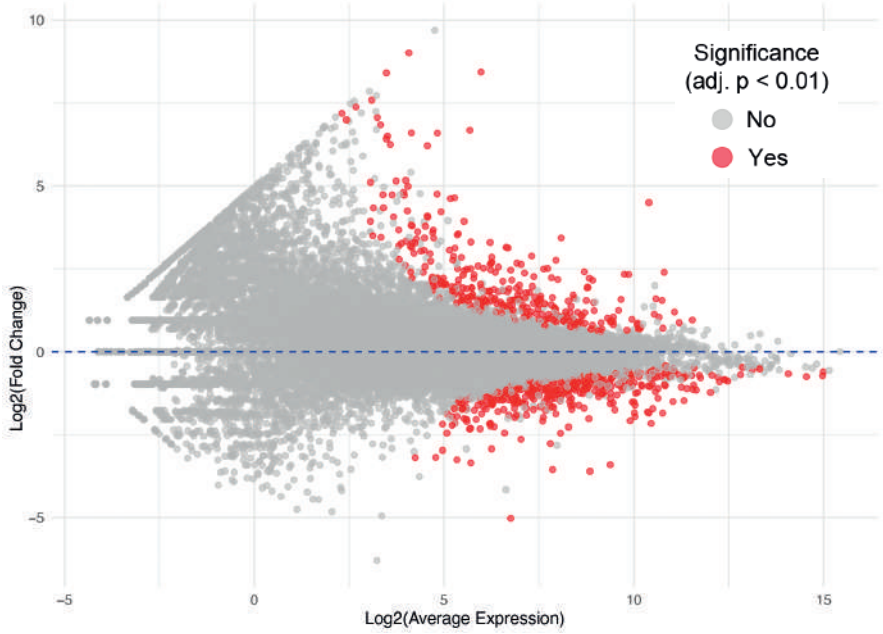
b

Genotype	T1's screened (no.)	FWA expressing (no.)	FWA expressing (%)
SunTag:PRDM9:FWA_g4	32	23	72%
SunTag:dPRDM9:FWA_g4	19	0	0%
SunTag:PRDM9:No_g	16	0	0%

S12

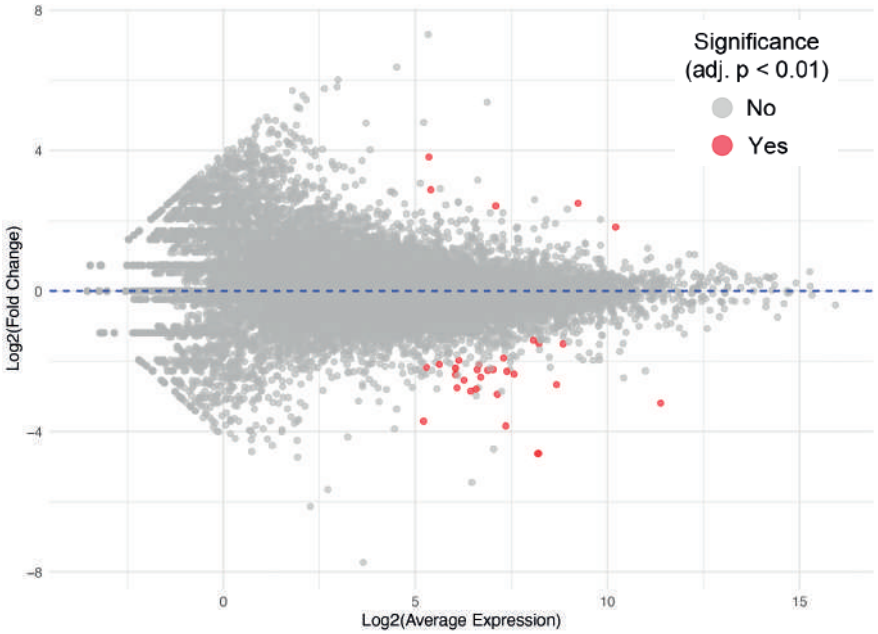
a

SunTag:SDG2:NoGuide vs. *rdr6*



b

SunTag:PRDM9:NoGuide vs. *rdr6*



*Note that FWA is not a differentially expressed gene in either comparisons