Atypical epigenetic and small RNA control of transposons in clonally reproducing *Spirodela polyrhiza.*

Rodolphe Dombey[1,2], Verónica Barragán-Borrero[1], Daniel Buendía-Ávila[1,2], Arturo Ponce-Mañe[1], José Mario Vargas-Guerrero[1,3], Rana Elias[1] and Arturo Marí-Ordóñez[1].*

[1] Gregor Mendel Institute of Molecular Plant Biology (GMI) of the Austrian Academy of Sciences, Vienna, 1030, Austria.

[2] Vienna BioCenter PhD Program, Doctoral School of the University of Vienna and Medical University of Vienna, Austria.

[3] Current address: Biology School, Costa Rica University, San José, 11501-2060, Costa Rica.

* The author responsible for correspondence and distribution of materials integral to the findings presented in this article is: Arturo Marí-Ordóñez (arturo.mari-ordonez@gmi.oeaw.ac.at)

Abstract

Background: A handful of model plants have provided insight into silencing of transposable elements (TEs) through RNA-directed DNA methylation (RdDM). Guided by 24-nt long small-interfering RNAs (siRNAs), this epigenetic regulation installs DNA methylation and histone modifications like H3K9me2, which can be subsequently maintained independently of siRNAs. However, the genome of the clonally propagating duckweed *Spirodela polyrhiza* (*Lemnaceae*) has low levels of DNA methylation, very low expression of RdDM components, and near absence of 24-nt siRNAs. Moreover, some genes encoding RdDM factors, DNA methylation maintenance, and RNA silencing mechanisms are missing from the genome.

Results: We investigated the distribution of TEs and their epigenetic marks in the Spirodela genome. While abundant degenerated TEs have largely lost DNA methylation and H3K9me2 is low, they remain transcriptionally silenced and are marked by H3K9me1. By contrast, we found high levels of DNA methylation and H3K9me2 in the relatively few intact TEs which are source of 24-nt siRNAs like RdDM-controlled TEs in other angiosperms. Some intact TEs are also regulated by post-transcriptional gene silencing (PTGS), producing 22-nt siRNAs despite the absence of a DCL2, required for their biogenesis in other angiosperms.

Conclusions: The data suggest that, potentially as adaptation to vegetative propagation, RdDM extent, components, and targets are different from other angiosperms, preferentially focused on potentially intact TEs. It also provides evidence for heterochromatin maintenance independently of DNA methylation. These discoveries highlight the diversity of silencing mechanisms that exist in plants and the importance of using disparate model species to discover these mechanisms.

Keywords

epigenetics, transposons, DNA methylation, small RNAs, heterochromatin, angiosperms, clonal reproduction, duckweeds.

52 # Introduction

53

54 Transposable elements (TEs) are mobile genetic elements which can change
55 location and generate new copies in their host genomes. As a result, TEs populate
56 eukaryotic genomes and can contribute to a significant fraction of their nuclear DNA
57 [1]. Owing to their mobile nature, TEs have an important role in evolution and genetic
58 innovation, but also pose a threat to the genome integrity [2,3]. To control TEs, various
59 mechanisms have evolved to prevent or reduce their mobility.

60 In eukaryotes, transcriptional gene silencing (TGS) of TEs relies on the
61 deposition of chromatin modifications leading to the formation of compact chromatin
62 states, collectively known as heterochromatin, incompatible with canonical Pol II
63 transcription [4]. Heterochromatin is generally characterized by repressive epigenetic
64 marks, prominently DNA methylation (5'methylcytosine; 5mC), histone modifications
65 such as histone H3 lysine-9 mono- and dimethylation (H3K9me1/H3K9me2,
66 constitutive heterochromatin) or histone H3 lysine-27 triple methylation H3K27me3
67 (facultative or developmentally regulated heterochromatin), and the presence of
68 specific histone variants [5,6]. Upon silencing, transposons accumulate mutations and
69 degenerate into non-autonomous TEs and fragmented remnants or relics [7]. Despite
70 their inability to mobilize, degenerated TEs provide alternative DNA and RNA
71 regulatory sequences (e.g., transcription factor binding sites, splicing information,
72 transcriptional start or termination sites) or induce recombination due to their
73 repetitiveness [8–12]. Therefore, maintaining epigenetic control is important even for
74 degenerated TEs, and in consequence, TE abundance and location largely shapes
75 the epigenetic landscape of genomes [13–17].

76 Often involved and largely conserved in mechanisms to suppress TEs and
77 maintain epigenetic silencing are ~18-30-nt long small RNAs (sRNAs) and members
78 of the ARGONAUTE (AGO) protein family. Loaded into AGOs, TGS-associated
79 sRNAs guide chromatin modifying complexes to transposons through sequence
80 complementarity [18,19]. In flowering plants (angiosperms), where TE silencing is
81 mostly associated with DNA methylation and H3K9me2, this mechanism is known as
82 RNA-directed DNA methylation (RdDM). Best studied in *Arabidopsis thaliana,* during
83 RdDM, H3K9me2 and DNA methylation recruit the plant-specific RNA Polymerases
84 IV and V (Pol IV, V) to respectively provide the source and target for 24-nt siRNAs,

85 loaded into members of the AGO4/6 clade, to deposit DNA methylation in all cytosine
86 contexts (CG, CHG and CHH, where H is any nucleotide but G). Once stablished,
87 DNA methylation can reinforce the deposition of H3K9me2 through the action of
88 SUPRESSOR OF VARIEGATION 3-9 HOMOLOGUE (SUVH) family H3K9
89 methyltransferases, which can read DNA methylation. In turn, CHROMO-METHYL
90 TRANSFERASE 3 and 2 (CMT3, CMT2) can bind to H3K9me2 to mediate CHG and
91 CHH methylation respectively. This results in a self-reinforcing loop between non-CG
92 methylation and H3K9me2 independent of siRNAs, although in Arabidopsis most of
93 the CHH methylation depends on the RdDM pathway. Symmetrical CG methylation is
94 maintained across cell divisions through the action of METHYL-TRANSFERASE 1
95 (MET1) (Reviewed in [20]). Loss of RdDM, 5mC or H3K9me2 results in the
96 transcriptional reactivation of TEs [21–25]. Therefore, DNA methylation, together with
97 H3K9me2 and in cooperation with sRNAs, keep TEs silenced.

98 Loss of TE control in somatic tissue is likely less deleterious as long as it does
99 not affect the germline to the next generation in which stability of the genome is highly
100 relevant. In accordance, small RNA-guided TGS pathways are prominently active in
101 reproductive tissues and silence active TEs encountered during fertilization [26–28].
102 Nevertheless, as opposed to metazoans, RdDM is not constrained to reproductive
103 organs in most flowering plants: 24-nt siRNAs are also present in vegetative, somatic
104 organs such as leaves or roots across all developmental stages [29–31]. With
105 interesting exceptions: besides the aquatic carnivorous plant *Utricularia gibba*, the
106 monocot *Spirodela polyrhiza* (L.) Scheid. contains much less 24-nt siRNAs and
107 reduced genome-wide levels of DNA methylation [32–34]. In contrast to the small *U.*
108 *gibba* genome, where repeats represent less than 3% of the genome [35], the TE
109 content in the genome of Spirodela (~139 Mbp) is similar to Arabidopsis [36] excluding
110 the reduced genomic TE load as explanation for the low levels of DNA methylation
111 and small RNAs. Furthermore, several sRNA silencing components including some
112 involved in RdDM have been shown to be missing, low or not expressed. In addition,
113 CMT2, which maintains CHH methylation, is lost [37–39].

114 The *Spirodela* genus is the most basal of the *Lemnaceae* family (order
115 Alismatales), a monophyletic group of free-floating freshwater monocot plants,
116 commonly known as duckweeds, composed of 5 genera (*Spirodela*, *Landoltia*, *Lemna*,
117 *Wolfiella* and *Wolffia*) [40,41] (**Fig. 1a, Supplemental Fig. S1a, b**). Comprised of a
118 leaf-like structure known as a frond, ranging from ~2 cm to 1 mm, duckweeds display

119 an evolutionary reduction and simplification of their body plan compared to most
120 angiosperms, including vestigiality and loss of organs such as roots and vasculature
121 [42–45] (**Supplemental Fig. S1c**). Although flowering and seed production has been
122 reported, it is rare in most species [42,46,47]. Duckweeds mostly propagate through
123 rapid clonal reproduction (doubling of individuals ~1-5 days), known as vegetative
124 budding: new fronds originate from budding pockets where meristems containing stem
125 cells are located [42,48]. Mother fronds (MF) continuously give rise to daughter fronds
126 (DF), which carry grand-daughter fronds (GDF) in their budding pockets even before
127 detaching from the mother frond [42,43] (**Fig. 1b**). This extremely efficient amplification
128 of somatic tissue, omitting regular sexual propagation, and the free-floating way of
129 living, make Spirodela a unique model to explore if the differences in epigenetic
130 parameters are connected with these differences in lifestyle compared to other
131 angiosperms.
132 Our results expand the notion of simplified silencing pathways in Spirodela.
133 DNA methylation and H3K9me2 profiles indicate that maintenance and feedback loops
134 of RdDM are still functioning at the few long, intact TEs but are weakened at the
135 majority of degenerated TEs, explaining the generally poorly methylated Spirodela
136 genome. Nonetheless, H3K9me1 remains at demethylated TE sequences. Despite
137 the absence of an obvious DCL2 ortholog in the genome, the siRNA fraction contains
138 22-nt siRNAs, produced from TEs as well as other endogenous and exogenous
139 sources of dsRNA. Our data document the variability of TE control and open the
140 possibility to study tissue-specific or developmentally regulated mechanisms, in
141 connection with the lifestyle, and reproduction strategy of Spirodela.
142

143 # Results

144

145 **Genome sequencing of *S. polyrhiza* #9509 from the Landolt collection.**

146 The genomes of several *Spirodela polyrhiza* accessions (clones) have been
147 sequenced [32,36,49], including a high-resolution reference genome assembly of its
148 20 chromosomes from the *S. polyrhiza* 9509 and 7498 clones (Sp9509, Sp7498)
149 [32,50,51]. For our study, we retrieved Sp9509 from the Landolt duckweed collection
150 (Zurich, Switzerland), now managed at the Institute of Agriculture Biology and
151 Biotechnology (IABB, CNR, Milan, Italy). We performed 88X coverage short-read
152 Illumina genome sequencing to account for potential genetic differences due to

153 independent clonal propagation over thousands generations between this accession
154 and that of the Rutgers Duckweed Stock Cooperative (RDSC, Rutgers University, NJ,
155 USA) used to generate the reference genome [50]. The Landolt Sp9509 genome
156 showed ~163k indels and ~512k SNPs (~0.4% of the Spirodela genome) compared to
157 the Sp9509 reference (**Supplemental Fig. S2**). Given our special interest on TEs, we
158 performed a *de novo* TE annotation using a combination of EDTA and
159 RepeatModeler2 to identify repeats that were then further classified using DeepTE
160 [52–54].
161 Silencing pathways and epigenetic mechanisms in general are well
162 characterized by molecular and genetic means in Arabidopsis. Similar genome size
163 and TE content of Spirodela and Arabidopsis (**Fig. 1c**) allow therefore valid
164 comparisons despite the evolutionary distance between the species (**Fig. 1a,**
165 **Supplemental Fig. S1a**).
166
167 **Small RNA and DNA methylation patterns in Spirodela imply low RdDM activity**
168 **in Spirodela.**
169 We compared the AGO-loaded small RNAs, using TraPR [55], of Spirodela
170 confluent cultures (including fronds in all developmental stages) to those of 10 day-old
171 Arabidopsis seedlings as the closest equivalent developmental stage. In contrast to
172 Arabidopsis, where TE-derived 24-nt siRNAs are highly abundant, Spirodela displays
173 very low levels of 24-nt siRNAs and its sRNA profile is dominated by 21-22-nt sRNAs
174 (**Fig. 1d, Supplemental Fig. S3**), predominantly contributed by regulatory microRNAs
175 (miRNA). The lack of 24-nt siRNAs in Spirodela is more pronounced than their loss in
176 Arabidopsis RdDM siRNA biogenesis defective Pol IV mutant (*nrpd1*) (**Supplemental**
177 **Fig. S3**). Furthermore, the reduced amounts of 24-mers is not compensated by TE-
178 derived 21-22-nt siRNAs as in DNA methylation-deficient Arabidopsis mutants [56–
179 58]. Hence, low levels of 24-nt siRNAs in Spirodela are compatible with the hypothesis
180 of reduced or lack of RdDM activity at TEs (**Supplemental Fig. S4**).
181 Methylome analysis by Enzymatic Methyl-seq (EM-seq) [59] showed similar
182 results to those previously published using bisulfite sequencing [32]. Transposon 5mC
183 in all contexts is lower in Spirodela compared to Arabidopsis, with very low mCHG
184 levels and hardly any mCHH (**Fig. 1e**). This is in agreement with the partial
185 dependency of mCHH on 24-nt siRNAs and seconds a weak RdDM in *S. polyrhiza*. In

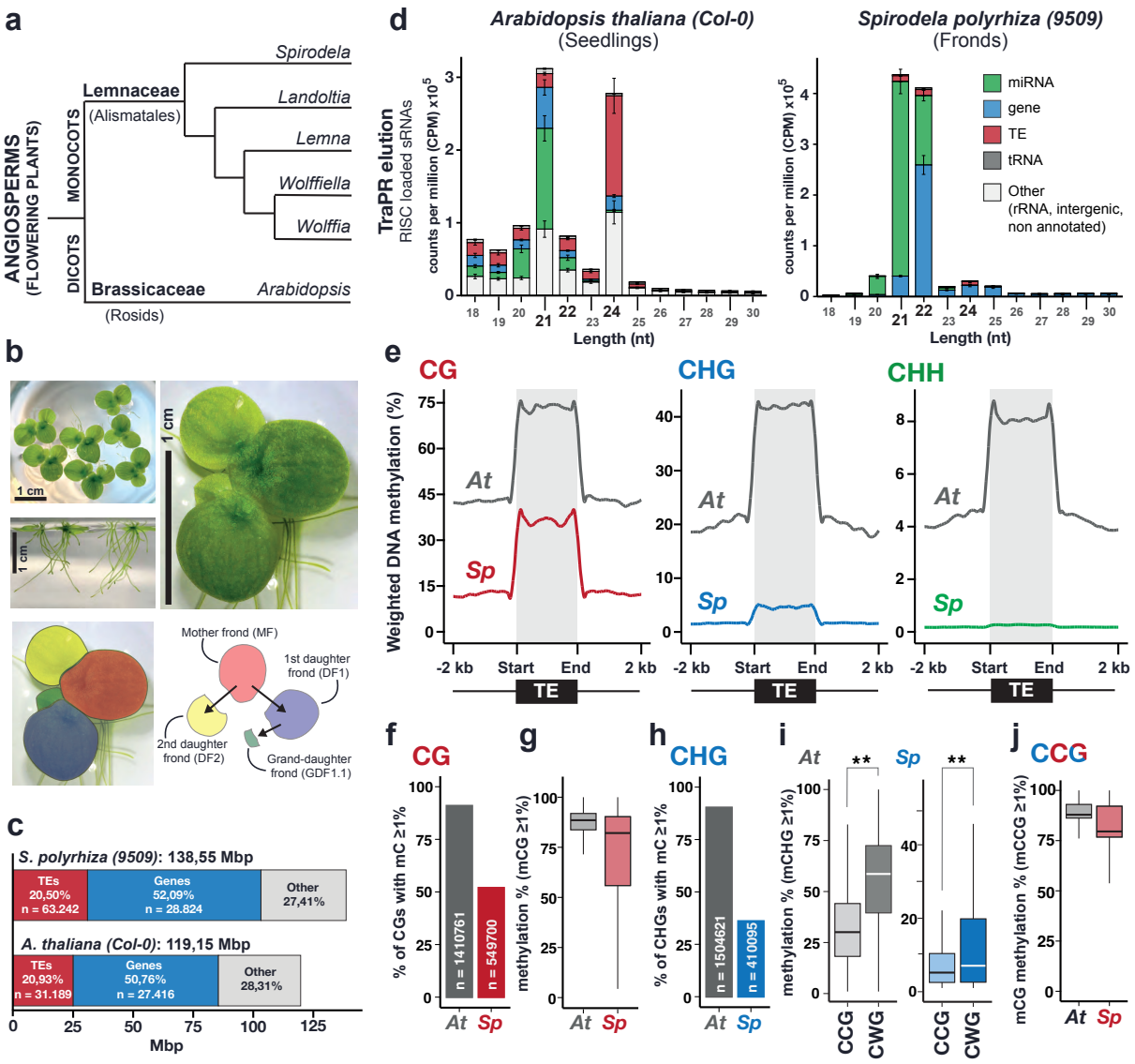

Figure 1: *Spirodela polyrhiza* general developmental, small RNA and DNA methylation patterns. **a** Simplified phylogenetic relationship of Spirodela and other duckweeds relative to Arabidopsis. **b** Representative images of Spirodela fronds in culture and schematic representation of the organization of clonal generations in a group of fronds during vegetative growth. **c** Genome size, number and genomic occupancy of genes and TEs in Spirodela and Arabidopsis. **d** Size distribution and genomic feature annotation of TraPR-purified small RNA from Arabidopsis and Spirodela. **e** Metaplots of averaged weighted DNA methylation in all three contexts (CG, CHG and CHH, where H is any but G) over transposons (≥100 bp) and 2 kb flanking regions in Arabidopsis and Spirodela. **f** number and percentage of transposon CG sites displaying methylation levels equal or above 1% in Arabidopsis and Spirodela. **g** DNA methylation levels at CG sites from (f). **h** number and percentage of transposon CHG sites displaying methylation levels equal or above 1% in Arabidopsis and Spirodela. **i** DNA methylation levels at CCG and CWG in Arabidopsis and Spirodela. **h** Inner 5mC (CmCG) DNA methylation levels at CCG sites with outer 5mC (mCCG) equal or higher than 1% in Arabidopsis and Spirodela. (**) indicates p-value <2.22 x 10^{-16} (Wilcoxon rank sum test). In all boxplots: median is indicated by solid bar, the boxes extend from the first to the third quartile and whiskers stretch to the furthest value within 1.5 times the interquartile range. At: Arabidopsis, Sp: Spirodela.

addition, Spirodela lacks gene body CG methylation (gbM) [32] (**Supplemental Fig. S4a**). Although the role of gbM is unclear, it generally associates with highly expressed genes. In contrast, mCG near the transcription start site (TSS) is associated with gene silencing [60,61]. Like in Arabidopsis, mCG around the TSS is higher in non-

190 expressed genes of Spirodela (**Supplemental Fig. S4b, c**), indicating that mCG in
191 gene regulatory sequences influences transcription also in Spirodela.
192
193 **CG and CHG methylation signatures are compatible with functional**
194 **maintenance of DNA methylation independently of small RNAs.**
195 Once established, DNA methylation can be further maintained independently
196 of small RNAs through several pathways depending on the sequence context. CG
197 methylation is maintained across cell divisions through the action of MET1 which, after
198 DNA replication, copies mCG patterns in daughter strands [62–64] . In Spirodela, the
199 ratio of mCG sites at TEs (51,9%) is nearly half of that in Arabidopsis (90,9%) (**Fig. 1f**).
200 However, at CG sites with 5mC levels equal or higher than 1%, mCG is similarly stable
201 as in Arabidopsis, displaying median methylation levels above 80% (**Fig. 1g**). Hence,
202 the reduced global mCG level at TEs in Spirodela is the result of less individual CG
203 sites being methylated. However, a higher variation in Spirodela than in Arabidopsis
204 (**Fig. 1g**) does not exclude that MET1 might be less efficient, not active in all cells, or
205 not all CG sites being as faithfully maintained.
206 As for mCG, the number of mCHG sites at TEs is much lower in Spirodela
207 (36,1%) than in Arabidopsis (90,3%) (**Fig. 1h**). CHG sites can be further divided into
208 CCG and CWG (CAG/CTG) sub-contexts. In Arabidopsis, mCWG is higher than
209 mCCG (**Fig. 1i**) due to the combined differential affinities of SUVH4/5/6 and CMT3 to
210 read and methylate individual cytosine sequence contexts [63,65–67]. This applies
211 also to Spirodela, despite lower mCHG in both sub-contexts (**Fig. 1i**). mCCG
212 maintenance relies additionally on MET1, as the sites have a C in both CG and CHG
213 contexts, are generally observed as mCmCG due to the outer cytosine methylation
214 (mCCG) being largely dependent on the internal CG methylation (CmCG) mediated
215 by MET1 [65,66,68,69]. Consequently, mCCG sites in Arabidopsis display a high
216 internal CG methylation level, a pattern also observed in Spirodela (**Fig. 1j**). The
217 similar mCHG patterns between the two species indicate that the responsible
218 enzymes, SUVHs and ZMET, the CMT3 homolog in monocots [70] probably share the
219 same sequence preference albeit their activity seems to be lower than in Arabidopsis.
220 Nonetheless, given that CMT3/ZMET has been shown to be involved in gbM
221 [60,71,72], lower mCHG and lack of gbM in Spirodela might be related.
222 Altogether, our analysis indicates that mCG and mCHG maintenance
223 mechanisms in Spirodela might operate in a similar fashion as in Arabidopsis.

224 However, while low mCG is a consequence of reduced number of CG sites being
225 methylated, low mCHG appears to be the combined result of lower number of targeted
226 CHG sites and decreased maintenance activity.
227
228 **Genetic composition and expression of silencing pathways partially account for**
229 **sRNAs and DNA methylation features in Spirodela.**
230 In addition to low signatures of RdDM in Spirodela, coding potential for several
231 of its components have been shown to be missing [37,39,73]. To gain further insights
232 why levels of 24-nt siRNA and DNA methylation are reduced, we set to identify and
233 examine the expression of known core components of angiosperm silencing pathways
234 in *S. polyrhiza* (**Fig. 2, Supplemental Fig. S5**). The identification of orthologues and
235 paralogues in the genome of *S. polyrhiza* 9509 was based on the presence of
236 conserved functional domains. Annotations were further curated based on phylogeny
237 trees built with proteins from several angiosperms and the prediction of known key
238 domains and their organization (**Supplemental Fig. S6-S16**). For gene expression,
239 we performed short-read transcriptome analysis in Spirodela and compared it to that
240 of recently published Arabidopsis seedlings [74] (**Fig. 2**). We further performed long-
241 read RNAseq (PacBio Isoform-Sequencing, Iso-Seq) in Spirodela to validate and
242 eventually improve gene model annotations.
243 As expected from the prevalence of miRNAs in Spirodela TraPR analysis,
244 conserved plant miRNA biogenesis factors [75] are expressed at similar levels as
245 those in Arabidopsis (**Fig. 2a**). Likewise, genes required for the production of
246 regulatory 21-nt siRNA trans-acting (ta)siRNA [76,77]: SUPRESSOR OF GENE
247 SILENCING 3 (SGS3), RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) and
248 DICER-LIKE 4 (DCL4), are present and expressed (**Fig. 2b**). Remarkably, Spirodela
249 lacks many components, effectors and regulators of 21-22-nt siRNA pathways, such
250 as RNASE THREE-LIKE 1 (RTL1), DOUBLE-STRANDED RNA BIDING PROTEIN 7
251 (DRB7) [78–80], and orthologues of the RDRγ clade highly conserved in plants and
252 fungi [81,82] (**Fig. 2b,c, Supplemental Fig. S6, S7**). Gene absences include the
253 antiviral factors DICER-LIKE 2 (DCL2) and ARGONAUTE 2 (AGO2) [83–86] (**Fig. 2b,**
254 **c, Supplemental Fig. S8, S9**). In contrast, two RDR1 and eight tandemly repeated
255 copies of AGO5 are present (**Fig. 2b, c, Supplemental Fig. S7, S9, S10**), which have
256 been implicated in virus exclusion and TE silencing in meristems [87–90].

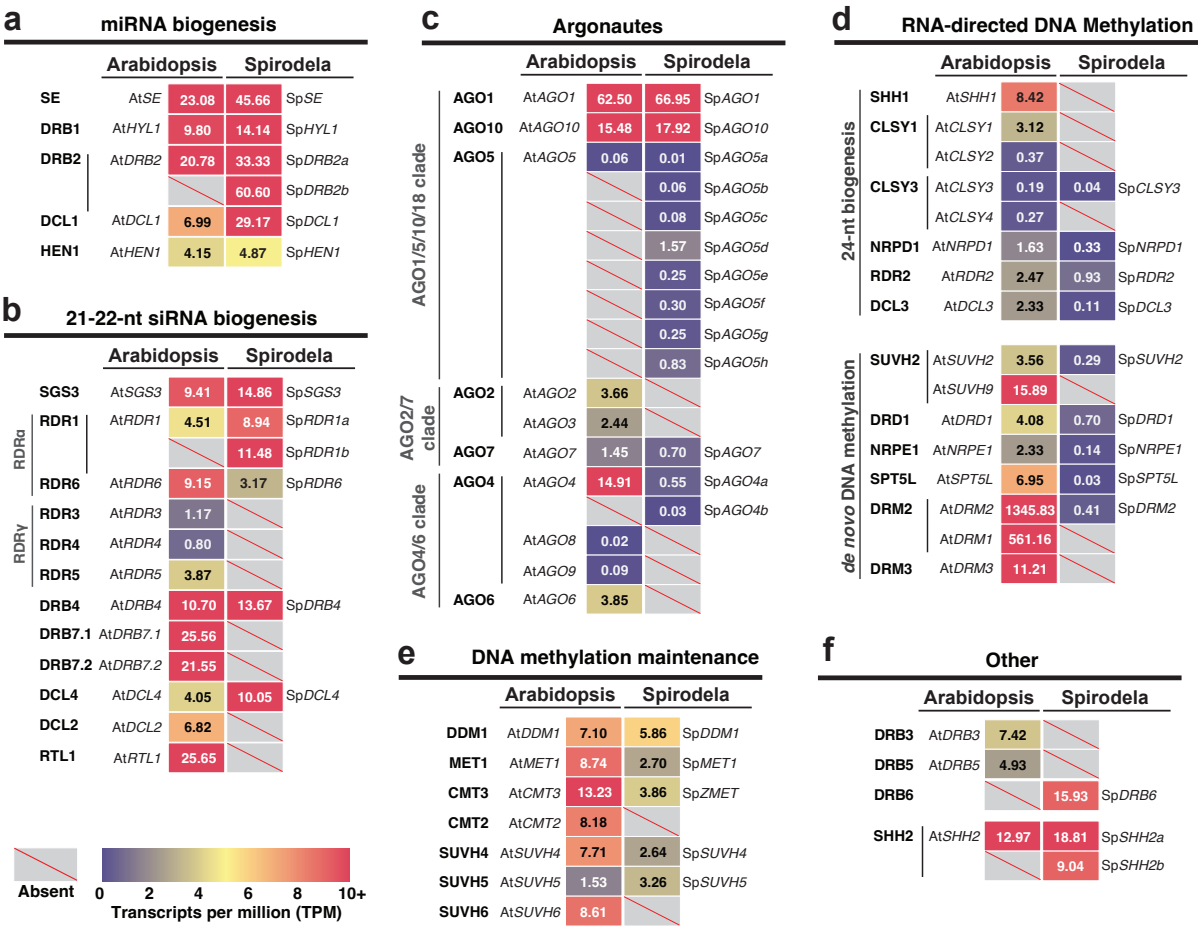

Figure 2: Genetic composition and expression of silencing pathways in Spirodela. a-f Expression in transcripts per million (TPM) of silencing components in Arabidopsis and their identified orthologues and paralogues in the *Spirodela polyrhiza* #9509 (Sp9509) genome. Gene IDs for both Arabidopsis and Spirodela can be found in Supplementary Figure 5.

Regarding the RdDM pathway, Spirodela contains the minimal set of its components, either with no or with a reduced number of paralogues compared to Arabidopsis (**Fig. 2c, d**). In Spirodela, SAWADEE HOMEODOMAIN HOMOLGUE 1 (SHH1), a dual reader for repressive H3K9me2 and unmethylated H3K4(me0), and the chromatin remodeler CLASSY 1/2 (CLSY1/2) are absent (**Fig. 2d, Supplemental Fig. S11, S12**). However, a single orthologue of CLSY3 was found in Spirodela (**Fig. 2d, Supplemental Fig. S11**). In Arabidopsis, the SHH1-CLSY1/2 module mainly recruits Pol IV to short TEs within gene-rich regions, while CLSY3 and its paralogue CLSY4 recruit Pol IV to TE-rich pericentromeric regions in a SHH1-idependent but DNA methylation-dependent manner [31,91–93]. However, Spirodela has two SHH2 orthologues (**Supplemental Fig. S11**). Part of a nucleosome positioning complex during transcription initiation in Arabidopsis [94], in maize, SHH2 binds preferentially

270 to H3K9me1 *in vitro* and interacts *in vivo* with both Pol IV and V in maize, possibly
271 connecting RdDM to H3K9me1 installation [95,96].
272 Similarly, no AGO6 orthologue was found, yet two AGO4, both required non-
273 redundantly to maintain DNA methylation at most RdDM target loci in Arabidopsis [97]
274 (**Fig. 2d, Supplemental Fig. S9**). A single SUVH2 orthologue is present in Spirodela
275 (**Fig. 2d, Supplemental Fig. S13**), which together with its paralogue SUVH9
276 redundantly recruit Pol V to RdDM targets by binding to methylated DNA in
277 Arabidopsis [98–100]. Regarding the RdDM-associated DNA methyltransferases,
278 DOMAIN REARRANGED METHYLTRASFERASE 2 (DRM2) but no DRM3, a
279 conserved catalytically inactive paralogue that acts as cofactor of DRM2 [101–105],
280 orthologues were found (**Fig. 2d, Supplemental Fig. S14**). Additionally, although
281 NRPE1 (Pol V) and its partner SUPPRESSOR OF TY INSERTION-LIKE (SPT5L) (a
282 paralogue of the general transcription elongation factor SPT5) are conserved in
283 Spirodela; we noticed that both contain fewer tryptophan/glycine (GW/WG) repeats
284 than those of other angiosperms (**Fig. 2d, Supplemental Fig. S15, S16**) [106–110].
285 GW/WG repeats, or AGO-hooks, allow AGO4/6 binding to Pol V to scan its nascent
286 transcripts for positive siRNA matches to recruit DRM2. These, together with NRPD1,
287 RDR2, DCL3, and the rest of conserved RdDM components, might suffice for a
288 functional pathway in Spirodela (**Fig. 2, Supplemental Fig. S7, S8, S12, S15**).
289 Nonetheless, their low expression levels (**Fig. 2 c, d**), rationalize the low abundance
290 of 24-nt siRNAs and associated CHH methylation in Spirodela (**Fig. 1d, e**).
291 Lastly, we inspected DNA methylation maintenance pathways. As predicted
292 from the patterns of mCG and mCHG (**Fig. 1f-j**), MET1, CMT3/ZMET, SUVH4, SUVH5
293 and DECREASE IN DNA METHYLATION 1 (DDM1) orthologues were found present
294 and expressed in Spirodela, albeit at a lower degree than in Arabidopsis. On the
295 contrary, CMT2, required to maintain mCHH together with SUVH4 [23], is missing as
296 previously reported [73] (**Fig. 2e, Supplemental Fig. S13, S14**).
297 Thus, reduced 24-nt siRNA and mCHH levels in Spirodela can be explained by
298 the combined reduction in RdDM composition and expression in addition to the
299 absence of CMT2. However, it does not account for the global decrease in mCG and
300 mCHG as their maintenance pathways seem intact and expressed in Spirodela.
301
302

303 **Few transposable elements from different superfamilies display high levels of**
304 **DNA methylation in Spirodela.**
305 To further gain insight into the patterns of TE DNA methylation, we examined
306 its distribution along the Spirodela genome. To exclude any confounding effects of
307 gbM, we calculated mean DNA methylation levels exclusively on TEs and their density
308 in 100-kb windows. At gene-rich regions, DNA methylation is very scarce (**Fig. 3a,**
309 **Supplemental Fig. S17a**). As in Arabidopsis, Spirodela DNA methylation is
310 concentrated at TE-dense regions (presumably centromeric and pericentromeric
311 domains). However, not all TE-dense regions are highly methylated (**Supplemental**
312 **Fig. S17b**). Hence, we investigated if local DNA methylation levels correlate with TE
313 density as in other angiosperms [14,16,111,112]. At equal TE-content, DNA
314 methylation is lower in all contexts in Spirodela compared to Arabidopsis, especially
315 at CHG and CHH (**Fig. 3b, Supplemental Fig. S18a**). Moreover, at the CG context,
316 Spirodela displays a broad variation range for a given TE content: even at TE densities
317 above 50% or higher, average mCG still spans between 0 and 75% (**Fig. 3b**).
318 To address the origin of such variation, we investigated the DNA methylation
319 levels of individual transposons. This revealed that the majority of TEs in Spirodela
320 display low or very low levels of DNA methylation in all contexts, especially CHG and
321 CHH (**Fig. 3c, Supplemental Fig. S18b**). Nonetheless, despite this substantially
322 reduced DNA methylation, increased expression at TEs was not observed (**Fig. 3d**).
323 TEs with high or low DNA methylation levels did not belong to any specific transposon
324 class or superfamily. For all the main plant TE superfamilies [113] (TE classification
325 and codes used here are summarized in **Supplemental Fig. S19**), DNA methylation
326 was nearly absent except for a few elements (**Fig. 3e**). Thus, low genome-wide
327 methylation levels in Spirodela result from only a few methylated TEs, although these
328 can display high mCG levels similar to those found in Arabidopsis (**Fig. 3e**).

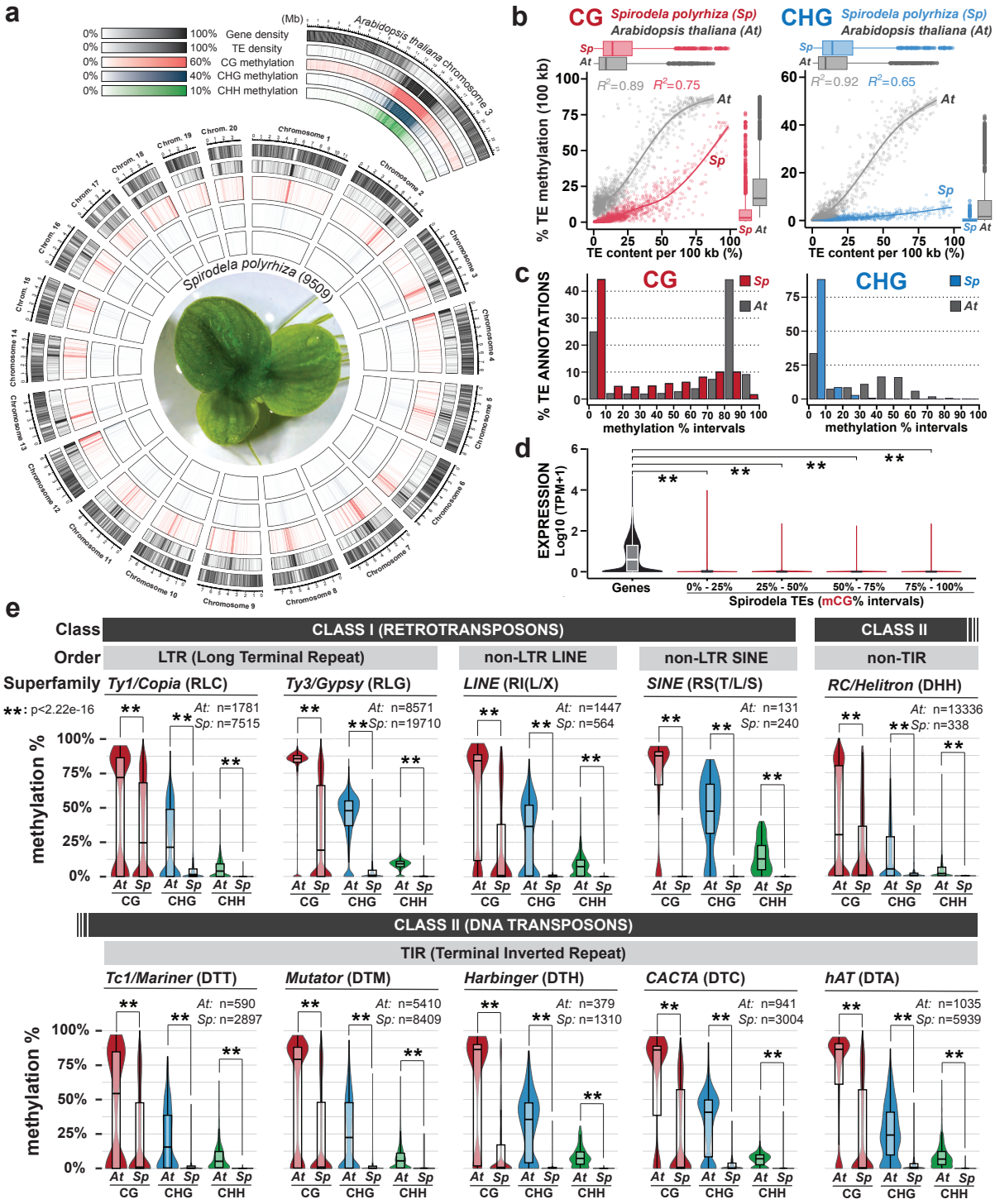

Figure 3: Distribution of DNA methylation in Spirodela chromosomes and transposable elements. a Circos plot of genome-wide distribution of DNA methylation densities by chromosome in Spirodela. From outside to inside: gene density, TE density, CG, CHG and CHH methylation. Mean DNA methylation, gene and TE densities values within the indicated ranges are displayed for 100 kbp intervals in 50 kbp sliding windows. Arabidopsis chromosome 3 is shown for comparison. **b** Scatter plot and best fit line of average TE CG and CHG DNA methylation against TE content per 100 kb genomic bins in Arabidopsis and Spirodela. Colored shadows indicate the standard error of the mean. **c** Bar plot of the distribution (%) of TE annotations (≥100 bp) within 10% incremental intervals of CG and CHG methylation in Arabidopsis and Spirodela. **d** Analysis of Spirodela TE expression within 25% incremental intervals of CG methylation. Global gene expression values are shown for comparison. **e** Distribution of CG, CHG and CHH methylation across the main TE superfamilies in Arabidopsis and Spirodela. The number of TE annotations belonging to each superfamily by species is indicated (n). (**) indicates p-value <2.22 x 10^{-16} (Wilcoxon rank sum test). In boxplots, median is indicated by solid bar, the boxes extend from the first to the third quartile and whiskers stretch to the furthest value within 1.5 times the interquartile range. At: Arabidopsis, Sp: Spirodela.

329

330 **Potentially intact transposons are the conspicuous targets of DNA methylation**
331 **in Spirodela.**

332 Although TE location and density influences their epigenetic regulation [15], in
333 Arabidopsis and tomato DNA methylation of longer, evolutionary younger and
334 therefore potentially mobile TE insertions are differentially regulated than often
335 fragmented TE relics resulting from the accumulation of mutations and truncations
336 over evolutionary times following their silencing [7,15,114]. On one hand, DNA
337 methylation in long TEs is maintained by a combination of RdDM acting on their
338 boundaries (mainly in young active TEs) and CMT activity along their body. On the
339 other hand, short, degenerated TE remnants do not need CMT to maintain their
340 methylation [115–117]. To explore if transposon length influences DNA methylation in
341 Spirodela, we sorted TEs into short TE fragments (100 to 500 bp), intermediate TEs
342 (0.5 to 4 kb) and long TEs (above 4 kb). This revealed that only long TEs are heavily
343 methylated at CG and, to a lesser extent, in CHG. Intermediate TEs display a wide
344 range of mCG levels, whereas TE fragments are minimally or not methylated (**Fig. 4a**).

345 To further explore the connection between TE length and DNA methylation we
346 focused on the two superfamilies of long terminal repeat (LTR) retroelements (RTEs):
347 *Ty1/Copia* (RLC) and *Ty3/Gypsy* (RLG). LTR-RTEs are the most abundant class of
348 TEs in the Spirodela genome, collectively representing more than 50% of the
349 transposon fraction of the genome and largely contributing as well to that of
350 Arabidopsis (**Fig. 4b, Supplemental Fig. S20**). In addition, the genetic structure and
351 domain organization of complete functional elements are well known [118–120]. Thus,
352 LTR-RTEs offer a good case study to investigate any link between DNA methylation
353 and TE length, using the latter as a proxy for their completeness and potential
354 autonomy and mobility. For better resolution, we split TE sizes into further size
355 categories in the intermediary range between 0.5 and 4 kb. Within each size bin,
356 Spirodela DNA methylation is lower than in Arabidopsis. However, in Spirodela,
357 average mCG and mCHG gradually increase with TE size, although mCHG does not

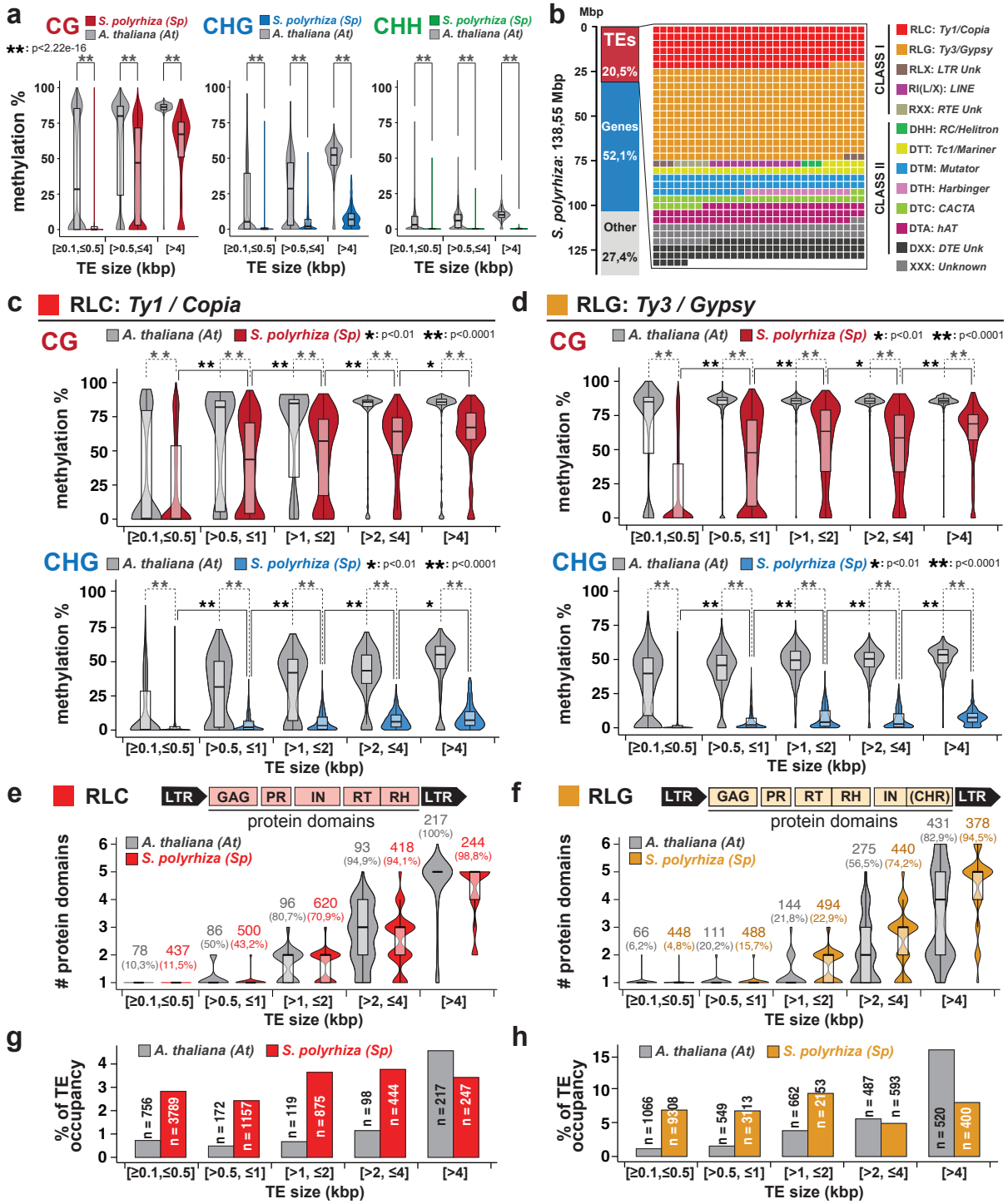

Figure 4: Length-dependent TE DNA methylation in *Spirodela polyrhiza*. a CG, CHG and CHH methylation levels (%) across all TEs (≥100 bp) in Arabidopsis and Spirodela split by size into short (≥100 bp, <500 bp), intermediate (>500 bp, ≤ 4 kbp) and long (> 4kbp). **b** Waffle chart of the relative contribution of different TE superfamilies over the total genomic space occupied by TEs in Spirodela. **c-d** CG and CHG methylation levels (%) between different length range groups for RLC: *Ty1/Copia* (c) and RLG: *Ty3/Gypsy* (d) in Arabidopsis and Spirodela. **e-f** Presence and number of RCL (e) and RLG (f) TE protein domains across TE annotations of varying. Violin and box plots depict domain count distributions for TEs with at least one domain. Numbers above denote TE annotations in each size category with at least one domain, with percentages shown in brackets relative to total TE annotations per size category (g-h). Schematic representation of RLC and RLG protein and domain structure is included on top of each plot. GAG: Nucleocapsid; PR: Protease; IN: Integrase; RT: Reverse transcriptase; RH: RNase H; CHR: Chromodomain, present only in the Chromoviridae genus of RLG. **g-h** Number of TE annotations and relative contribution to total genomic TE load (TE occupancy %) at each size category for RLC (g) and RLG (h). (*) indicates p-value <0.01 (**) indicates p-value <0.0001 (Wilcoxon rank sum test). In all boxplots: median is indicated by solid bar; the boxes extend from the first to the third quartile and whiskers stretch to the furthest value within 1.5 times the interquartile range. At: Arabidopsis, Sp: Spirodela.

358 achieve levels comparable to those of Arabidopsis (**Fig. 4c, d, Supplemental**
359 **Fig. S21a, b**). The proportion of TEs displaying high CG methylation rises across TE-
360 size groups. While very few TE-fragments below 500 bp are highly methylated, the
361 majority of long TEs (>4 kb) have mCG above 60% (**Fig. 4c, d**). At the same time, the
362 fraction of TEs with at least one intact cognate protein domain and the number of
363 domains identified per TE annotation increase with each size category, with the
364 majority of long LTR-RTEs containing all domains associated with each superfamily
365 (**Fig. 4e, f**).
366 Therefore, in Spirodela, most potentially intact TEs display high levels of DNA
367 methylation, which decreased as they degenerated and decayed. Given that long
368 intact LTR-RTEs represent a minority of Spirodela TEs (**Supplemental Fig. S21c, d**),
369 their low abundance helps explain the few TE numbers with high DNA methylation
370 (**Fig. 3c, e**). In addition, the proportion of hypomethylated shorter LTR-RTE fragments
371 relative to methylated intact TEs is higher in Spirodela than in Arabidopsis (**Fig. 4g, h,**
372 **Supplemental Fig. S21c-f**), resulting in the global reduction of TE methylation
373 (**Fig. 1e, Fig. 3a, b**).
374
375 **Levels of H3K9me2, but not H3K9me1, are reduced in Spirodela.**
376 Considering the association between DNA and H3K9 methylation in plants, we
377 explored how far DNA methylation in Spirodela would impact the histone methylation
378 landscape. H3K9me2 and H3K9me1 are commonly linked with heterochromatin
379 formation at silenced TEs [6,23,121,122]. Western blot analysis revealed that the
380 H3K9me2 level in Spirodela is lower than in Arabidopsis, whereas H3K9me1 displayed
381 similar levels (**Fig. 5a**). This was unexpected as we anticipated in Spirodela a
382 decrease of both marks, given their connection with DNA methylation and reduction
383 upon loss of non-CG methylation in Arabidopsis [23,123]. To corroborate the
384 observation, we investigated histone post-translational modifications (PTMs) by mass-
385 spectrometry [124,125], which confirmed an approximately 8-10-times lower level of
386 H3K9me2 but similar H3K9me1 levels in Spirodela relative to Arabidopsis
387 (**Supplemental Fig. S22**).
388
389
390
391

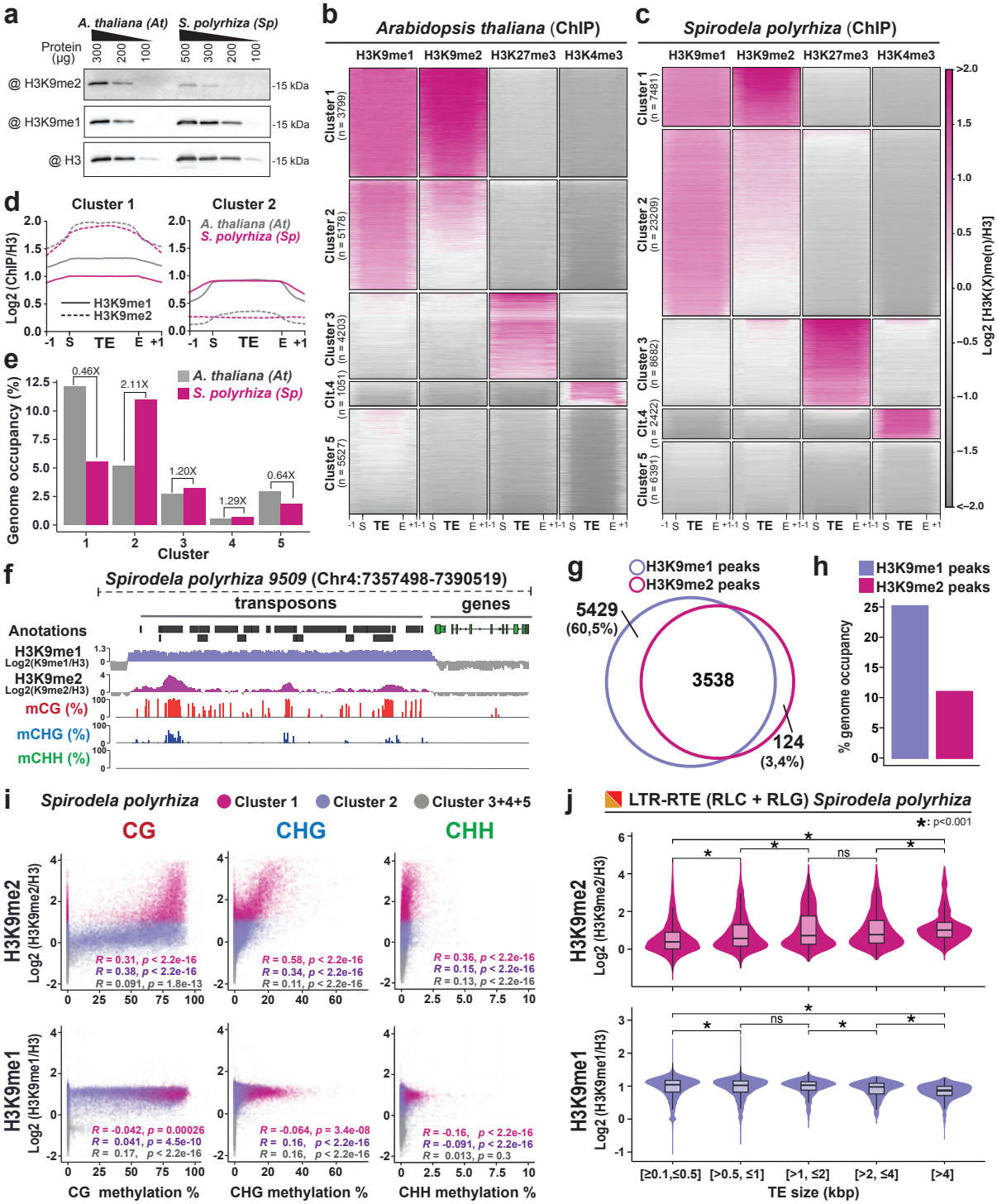

Figure 5: Patterns of H3K9me1/2 distribution in *Spirodela polyrhiza* transposons. **a** Protein blot analysis of H3K9me1 and H3K9me2 abundance in Spirodela compared to Arabidopsis at different protein amounts (µg). Total H3 protein levels are shown as loading control. **b-c** Heatmap of k-means clustering of transposons in Arabidopsis (b) and Spirodela (c) using H3K9me1, H3K9me2, H3K27me3 and H3K4me3 enrichment over H3, presented as log2[ChIP(X)/ChIP(H3)]. Enrichment is represented by color, while depletion is shown in gray. Each row corresponds to a length-normalized TE annotation ± 1 kb. Transposons are grouped into clusters 1 to 5 based on the abundance of examined marks. **d** Metaplots of average weighted H3K9me1 (solid line) and H3K9me2 (dashed line) enrichments over H3 at cluster 1 and 2 transposons ± 1 kb in Arabidopsis (colored) and Spirodela (gray). **e** Relative genome (%) occupancy of each of the TE clusters defined in b-c in Arabidopsis (grey) and Spirodela (colored). **f** Genome browser capture of the distribution of H3K9me1 and H3K9me2 enrichments, together with DNA methylation, along genes and transposons in Spirodela. **g** Venn diagram of the overlap between H3K9me1 and H3K9me2 peaks along Spirodela genome. **h** Genome occupancy (%) of H3K9me1 and H3K9me2 in Spirodela genome. **i** Scatter plots of DNA methylation in all three contexts against H3K9me1 and H3K9me2 enrichments of TEs in Cluster 1, 2 and 3/4/5 combined in Spirodela. *R* and *p* indicate Pearson correlation coefficients and P-values. **j** Analysis of H3K9me1 and H3K9me2 enrichments in different TE length groups for RLC and RLG combined in Spirodela. In all boxplots: mean is indicated by solid bar; the boxes extend from the first to the third quartile and whiskers stretch to the furthest value within 1.5 times the interquartile range. (*) indicates p-value <0.001. n.s.: not significant.

392

393 **A low proportion of TEs show high H3K9me2 enrichment in Spirodela.**

394 To further delve into the distribution of H3K9me1/2 in Spirodela and how the
395 general organization of heterochromatic marks is affected by its reduced DNA
396 methylation levels, we performed chromatin immunoprecipitation (ChIP) in Spirodela
397 fronds and Arabidopsis 10 days-old seedlings (**Supplemental Fig. S23**) and focused
398 on the association of these modifications with the TEs. In addition to H3, H3K9me1
399 and H3K9me2, we generated genomic profiles through ChIP-seq for H3K27me3 and
400 H3K4me3. Although Polycomb-mediated H3K27me3 is generally mediates
401 developmentally regulated gene silencing, its association with TEs was investigated
402 given that it also regulates TEs in the absence or loss of DNA methylation [126–129],
403 and may contribute directly to TE silencing [130]. H3K4me3, was taken as an
404 euchromatic mark control [13,131–134].

405 As observed for DNA methylation (**Fig. 3a**), H3K9me1/2 is concentrated at TE-
406 dense regions in Spirodela (**Supplemental Fig. S24**), likely centromeric and
407 pericentromeric heterochromatic regions [13,73,121]. The smaller size of such
408 regions, compared to Arabidopsis, and their distribution in the Spirodela 20
409 chromosomes, fits the speckled pattern of small dense DNA foci in DAPI staining of
410 interphase nuclei of Spirodela (**Supplemental Fig. S25**) [135].

411 A more detailed K-means clustering analysis of histone mark profiles relative to
412 H3 over TEs, identified 5 clusters with similar features in both Arabidopsis and
413 Spirodela (**Fig. 5b, c**). Clusters 1 and 2 contain TEs enriched for H3K9me1/2 while
414 clusters 3 and 4 consist of those marked by H3K27me3 and H3K4me3 respectively.
415 Lastly, cluster 5 encompasses TEs depleted of any of the profiled marks (**Fig. 5b, c**).

416 Although cluster 1 and 2 TEs are marked by H3K9me1/2, they display
417 homogeneous H3K9me1 but contrasting H3K9me2 levels (**Fig. 5d, Supplemental**
418 **Fig. S26**). TEs in cluster 1 are more enriched in H3K9me2 than in H3K9me1 while in
419 cluster 2 TEs H3K9me2 is below H3K9me1 (**Fig. 5d**). Remarkably, the two clusters
420 comprise very different fractions of the two genomes and TE space. In Spirodela,

421

422 cluster 1 only comprises about 20% of the TE space, representing less than 6% of the
423 genome, while in Arabidopsis is the most abundant one, encompassing ~35% of the
424 TE space and ~12% of the genome. Cluster 2, on the contrary, occupies 45% of the
425 Spirodela TE space (more than 10% of its genome), but only about 20% in Arabidopsis

426 (5% of the genome) (**Fig. 5e, Supplemental Fig. S27a**). Given the similar genome
427 size and TE content in both species, the opposed contribution of cluster 1 and cluster
428 2 to their genomes is in line with the levels of H3K9me2 observed by western blot and
429 mass-spectrometry. Furthermore, similar H3K9me1 levels in both clusters concur with
430 its even abundance in Spirodela and Arabidopsis. Hence, reduced DNA methylation
431 levels in Spirodela correlate with a smaller fraction of the genome displaying a high
432 H3K9me2 enrichment over TEs but it does not impact H3K9me1 distribution, nor it
433 leads to a widespread compensatory regulation of TEs by H3K27me3.
434
435 **H3K9me1 marks TEs independently of their length and methylation status in**
436 **Spirodela.**
437 A closer inspection of the distribution of H3K9me1/2 along Spirodela TEs
438 revealed that H3K9me1 uniformly covers TEs, while H3K9me2 peaks are found only
439 at individual TEs (**Fig. 5f**). Furthermore, while more than 95% of H3K9me2 peaks
440 overlapped with H3K9me1 domains, a significant fraction of H3K9me1 peaks (60,5%)
441 did not with H3K9me2-associated regions (**Fig. 5g**). Moreover, most H3K9me2 peaks
442 were embedded within H3K9me1 (**Supplemental Fig. S27b**). Adding to that,
443 H3K9me1-enriched regions occupy 25% of the Spirodela genome whereas H3K9me2
444 peaks represent less than 12% (**Fig. 5h**). This indicates that, similarly to DNA
445 methylation, high levels of H3K9me2 are only found within few TEs. It prompted us to
446 investigate the association between H3K9me1/2 and DNA methylation in Spirodela
447 and indeed, H3K9me2 levels in Spirodela positively correlate with DNA methylation,
448 including mCHG, like in Arabidopsis (**Fig. 5i, Supplemental Fig. S28**). Parallel to
449 DNA methylation (**Fig. 3**), H3K9me2 increased with TE size (**Fig. 5j**). This further
450 supports that, as previously inferred, the positive feedback loop between mCHG and
451 H3K9me2 operates in Spirodela. However, this did not apply to H3K9me1. While in
452 Arabidopsis TEs enriched in H3K9me1 invariably displayed high levels of DNA
453 methylation (**Supplemental Fig. S28**), they did not so in Spirodela with its broad range
454 of DNA methylation levels at TEs (**Fig. 5i**). Accordingly, H3K9me1 levels vary little
455 across TE sizes (**Fig. 5j**).
456 These results indicate that, in Spirodela, H3K9me1 serves as broad transposon
457 mark but does not rely on DNA methylation as H3K9me2, suggesting that they might
458 be differentially regulated. Although Arabidopsis H3K9 histone methyl-transferases
459 (HMTs) can catalyze both H3K9me1 and me2 in vitro [121,136], it is unclear how their

460 mono- and di- methyl-transferase activity is regulated in vivo or if both marks are
461 deposited by the same HMTs. Loss of SUVH4, the main HMT in Arabidopsis, results
462 in a strong reduction of H3K9me2 and loss of epigenetic silencing, only marginally
463 affecting H3K9me1 and global heterochromatin organization [121,137,138]. This
464 originally led to the proposition that H3K9me2 and H3K9me1 play different roles in
465 silencing and heterochromatin maintenance respectively [121]. Indeed, increased
466 H3K9me2 and DNA methylation at long TEs, in contrast to shorter TE fragments, hints
467 that their self-reinforcement is focused at potentially functional or recently active ones
468 to ensure their silencing.
469
470 **Long, complete TEs show signatures of RdDM activity.**
471 For the reasons explained above, we focused the detailed analysis on the
472 complete LTR-RTEs of the *Copia* (RLC) and *Gypsy* (RLG) superfamilies in which
473 H3K9me2 levels were higher than those of H3K9me1. However, the patterns differed
474 depending on the TE superfamily. While in RLCs had more H3K9me2 towards the
475 center of the TE, the same modification peaked on the edges of RLGs (**Fig. 6a**). Such
476 patterns resembled those found in Arabidopsis intact LTR-RTEs, (**Supplemental**
477 **Fig. S29a**). DNA methylation at CG and CHG followed a similar trend as H3K9me2 in
478 both cases (**Fig. 6b**), although mCHG was comparatively lower (**Fig. 6b,**
479 **Supplemental Fig. S29b**). Furthermore, albeit very low, peaks of mCHH were also
480 observable at RLGs edges (**Fig. 6b**). Hence, the layout of epigenetic silencing marks
481 along complete TEs in Spirodela are very similar to those in Arabidopsis.
482 Given that in Arabidopsis these patterns are partially determined by RdDM
483 [23,68,116], we inspected if small RNA mapped to complete TEs in Spirodela. Indeed,
484 siRNAs derived from regions with DNA methylation and H3K9me2 were present in
485 Spirodela, specially 24-nt siRNAs, despite at lower levels than in Arabidopsis (**Fig. 6c,**
486 **Supplemental Fig. S29c**). Moreover, 24-nt siRNAs mapping to complete TEs
487 displayed an adenine 5'-nucleotide bias (5'A), characteristic of small RNAs loaded into
488 the RdDM effector AGO4 [29,139] (**Fig. 6d**), although not as preferential as in
489 Arabidopsis (**Supplemental Fig. S29d**). Complementarity of siRNAs with DNA
490 methylation and H3K9me2 within the TE body and/or their LTRs were validated for
491 individual complete TEs (**Fig. 6e-g, Supplemental Fig. S30**).
492
493

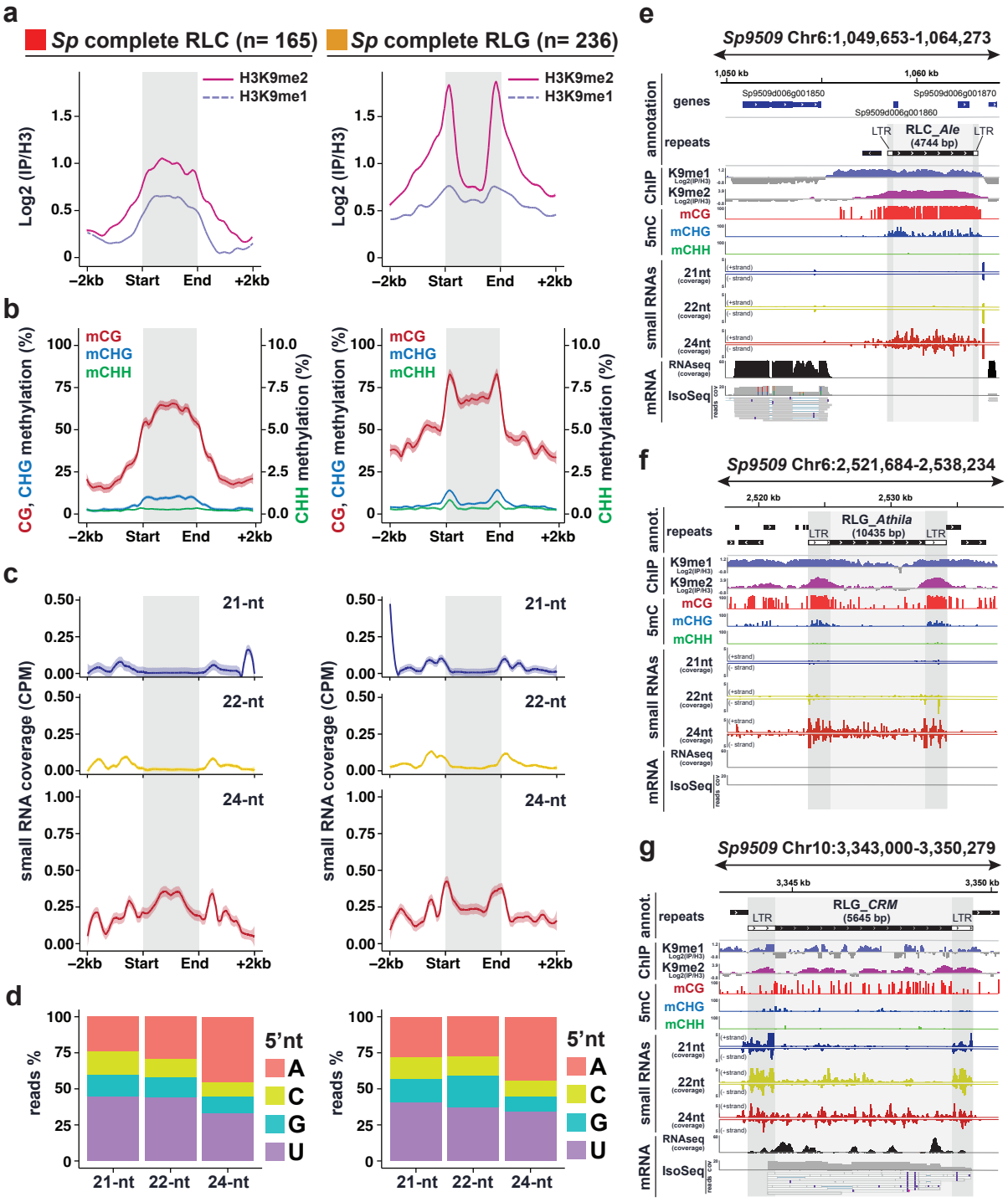

Figure 6: RdDM signatures in *Spirodela polyrhiza* complete TEs. a Metaplots of average weighted H3K9me1 (solid line) and H3K9me2 (dashed line) enrichments over H3 at complete RLC (left) and RLG (right) transposon annotations ± 2 kb in Spirodela. Complete TEs were defined as those for which all RLC and RLG protein domains where identified. n indicates number of complete elements in each superfamily. **b** Metaplots of averaged weighted DNA methylation in all three contexts over RLC and RLG transposons ± 2 kb flanking regions in Spirodela. mCHH values are scaled to 10% max values instead of 100%. **c** Metaplots of averaged weighted 21-, 22- and 24-nt small RNA abundance over intact RLC and RLG ± 2 kb flanking regions in Spirodela. In all metaplots (a-d), solid lines represent the mean and colored shadows the standard error. **d** 5'-nucleotide (5'nt) composition distribution (as % of reads) of 21-, 22- and 24-nt siRNAs mapping to intact RLC and RLG ± 2 kb flanking regions in Spirodela. **e-g** Genome browser captures of selected complete TEs in Spirodela. From top to bottom tracks display: gene and TE annotations, H3K9me1 and H3K9me2 enrichment profiles; mCG, mCHG and mCHH methylation; 21-, 22- and 24-nt siRNA mapping to Watson (+) and Crick (-) DNA strands; RNA expression by short (illumina) and long reads (IsoSeq). Superfamily, clade and length (in bp) of TEs is indicated above their annotation. Light and dark gray shadow delimit TEs and LTRs respectively across all tracks.

494

495 We also detected the presence of very low amounts of 21-22-nt siRNAs at the
496 edges of complete TEs (**Fig. 6c**). In some instances, 24-nt siRNAs overlapped with
497 equally abundant 21-22-nt siRNAs, especially at the flanking LTR sequences (**Fig. 6g,**
498 **Supplemental Fig. S31**). 21-22-nt siRNAs can arise from TE edges in Arabidopsis
499 (**Supplemental Fig. S29c**) as result of DCL2/4 processing of Pol IV transcripts and
500 direct DNA methylation [57,140–142].

501 To assess if TE-derived siRNAs might participate in RdDM also in Spirodela,
502 we tested if its AGO4 (SpAGO4) has the same small RNA loading preference as
503 Arabidopsis AGO4 (AtAGO4). To do so, we cloned genomic *AGO4* loci from both
504 species with a Flag-HA tag for heterologous transient expression in *Nicotiana*
505 *benthamiana* leaves. No Sp*AGO4* protein was detectable due to intron retention
506 events when expressed in *N. benthamiana* (**Supplemental Fig. S332a-c**). This was
507 likely due to the inability of dicots (*N. benthamiana*) to properly process introns of the
508 monocots (Spirodela) [143–145]. To circumvent this limitation, we amplified and cloned
509 full length cDNA from one of the two Spirodela AGO4, *AGO4a*, and expressed it in *N.*
510 *benthamiana* leaves (**Supplemental Fig. S32d**). To assess their small RNA loading
511 preferences, AtAGO4 and SpAGO4a were immunoprecipitated (IP) from *N.*
512 *benthamiana* leaf lysates [146] (**Fig. 7a**) and small RNAs extracted and sequenced.
513 Like AtAGO4, SpAGO4a displayed a strong preference for 5'A 24-nt long siRNA and
514 reduced affinity for 21-nt siRNA, compared to the population of siRNAs present in *N.*
515 *benthamiana* leaves (**Fig. 7b, c**).

516 Therefore, small RNAs might be indeed involved in maintaining DNA
517 methylation and H3K9me2 in Spirodela, at long complete TEs through canonical
518 RdDM. Our analysis of silencing pathways shows that core components required for
519 RdDM are present in the Spirodela genome and, although poorly expressed, probably
520 active during vegetative clonal propagation.

521

522 **Post-transcriptional silencing of TEs likely involves DCL2-independent 22-nt-**
523 **siRNAs in Spirodela.**

524 It is unlikely that the observed TE-derived 21-22-nt siRNAs participate in RdDM,
525 due to the low affinity of SpAGO4 to bind them (**Fig. 7b**), and those mapping to TEs
526 have less 5'A bias but more 5'U (**Fig. 6d, Supplemental Fig. S29c**), a signature of
527 AGO1-loaded small RNAs [29,139]. This hints to a role rather in post-transcriptional

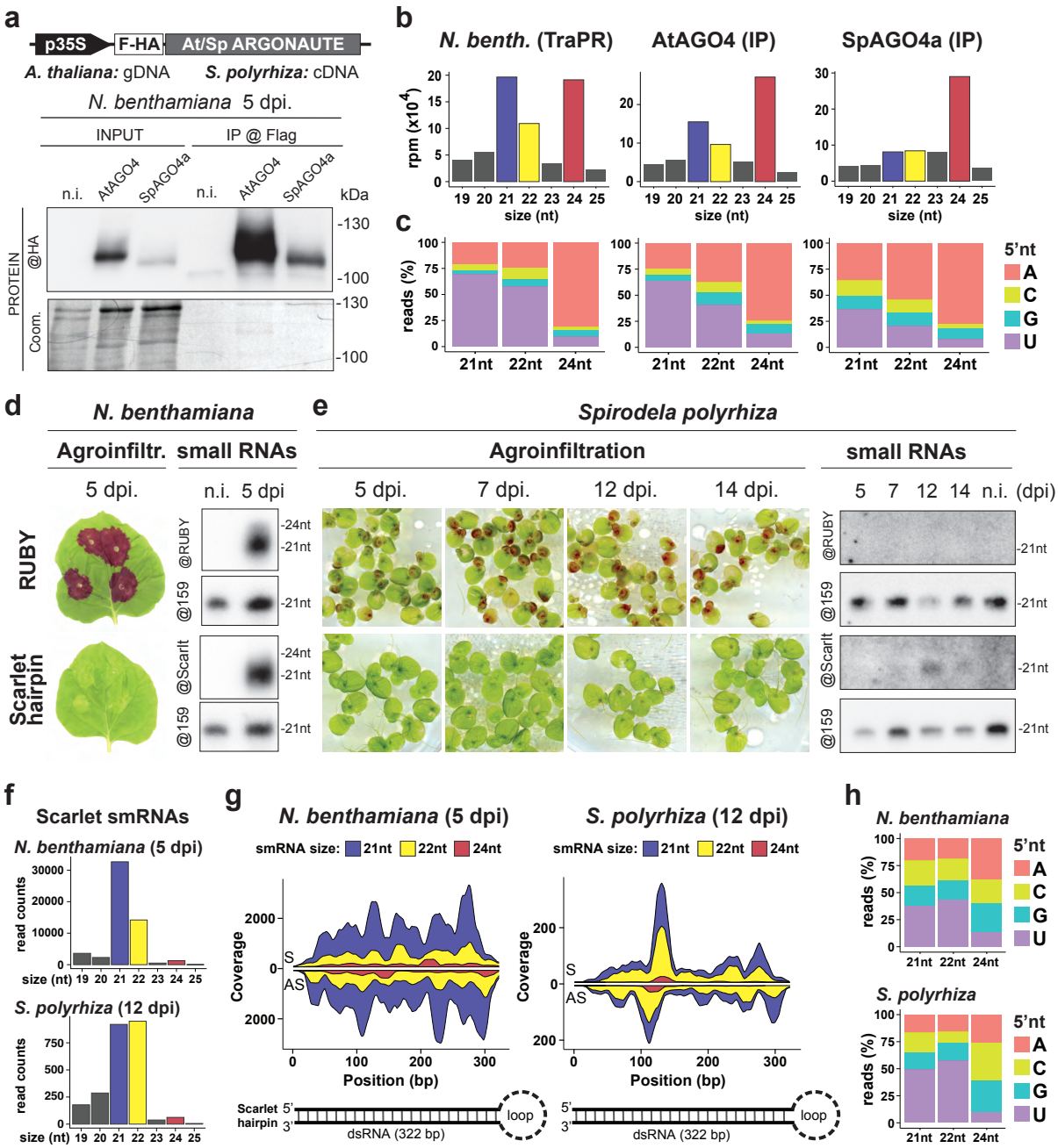

Figure 7: *Spirodela polyrhiza* AGO4a siRNA loading preferences and IR-derived siRNAs in transient expression. a Protein blot analysis of Arabidopsis and Spirodela AGO4 (AtAGO4 and SpAGO4a) input and IP fractions from transient expression in *N. benthamiana* 5 days post infiltration (dpi). n.i.: non-infiltrated. Coomassie (Coom.) staining of blot is shown as control for input and IP fractions. Scheme of constructs infiltrated in *N. benthamiana* leaves for transient expression is shown above. p35S: CaMV 35S promoter; F-HA: in-frame N-term Flag and HA peptide tags. Genomic (gDNA) or cDNA origin of the AGO4 sequences is indicated. **b** Size distribution and abundance in reads per million (rpm) of TraPR purified *N. bethamiana* siRNAs (as control) and siRNAs extracted from AtAGO4 and spAGO4a IPs. **c** 5'-nucleotide (5'nt) composition distribution (as % of reads) of 21-, 22- and 24-nt siRNAs from b. **d** Representative pictures of *N. benthamiana* leaves agroinfitrated for transient expression of RUBY or the scarlet hairpin (*hpScarlet*) IR next to small RNA blot analysis in non-infiltrated and 5 dpi leaves. miRNA159 is shown as loading control. **e** Representative pictures of Spirodela fronds agroinfitrated for transient expression of *RUBY* or *hpScarlet* and small RNA blot analysis in non-infiltrated and infiltrated samples at different timepoints. miRNA159 as loading control. **f** Size distribution and abundance of small RNAs mapping to *hpScarlet* from 5 dpi *N. bethamiana* and 12 dpi Spirodela. **g** Stacked distribution of sense (S) and antisense (AS) small RNA coverage by size along *hpScarlet* hairpin in *N. bethamiana* and Spirodela. Schematic representation of the IR is depicted below. **h** 5'-nucleotide (5'nt) composition distribution of 21-, 22- and 24-nt siRNAs mapping to *hpScarlet* in *N. bethamiana* and Spirodela.

528 silencing (PTGS). During PTGS, 21-22-nt siRNAs loaded into AGORGONAUTE
529 proteins of the AGO1/5/10/18 and AGO2/7 clades target Pol II transcripts to guide
530 mRNA cleavage or translational repression (reviewed in [147]) Coincidentally, we
531 observed expression of TEs producing 21-22-nt siRNAs in our transcriptome datasets,
532 including full length transcripts by Iso-Seq (**Fig. 6g, Supplemental Fig. S31**).
533 Although PTGS is a well-documented control mechanism against transcriptionally
534 active TEs in Arabidopsis [27,56,58,148–150], we were intrigued by the presence of
535 22-nt siRNAs, given that any identifiable ortholog of DCL2, their cognate Dicer in
536 plants [151–154], is missing in Spirodela and *Lemnaceae* genomes in general (**Fig. 2b,**
537 **Supplemental Fig. S8**) [39,73].

538 To rule out artifacts and to investigate if Spirodela can produce 22-nt siRNAs in
539 the absence of DCL2, we attempted to stimulate small RNA biogenesis through
540 expression of suitable exogenous substrate RNAs. As stable transformation in
541 Spirodela is difficult, we chose Agrobacterium-mediated transient expression. This
542 technique has been widely used to trigger and investigate silencing mechanisms,
543 owing to the quick onset of siRNA production in response to expression of double-
544 stranded RNA (dsRNA) following infiltration of plant tissue with Agrobacterium carrying
545 the corresponding transgenes [155–159]. We developed a modified protocol for frond
546 transformation in *Lemna minor* [160] and applied it successfully to induce transient
547 transgene expression in Spirodela fronds (**Supplemental Fig. S33**). We used the
548 visual reporter *RUBY* [161] and an inverted repeat (IR) derived from the coding
549 sequence of the fluorescent reporter *Scarlet* (*Scarlet* hairpin, *hpScarlet*) [89]. Both
550 constructs triggered siRNAs when infiltrated into *N. benthamiana* as expected
551 (**Fig. 7d**). In Spirodela, however, we could not detect *RUBY*-derived siRNAs even after
552 two weeks post-infiltration. On the other hand, *hpScarlet*-derived siRNAs were
553 observed 12- and 14-days post-infiltration (**Fig. 7e**). Sequencing sRNAs from both *N.*
554 *benthamiana* and Spirodela infiltrated with *hpScarlet* showed that the IR is processed
555 into 21-22-nt siRNAs in both plants. However, Spirodela produceed relatively equal
556 amounts of 21- and 22-nt siRNAs, as observed for TEs, whereas the profile in *N.*
557 *benthamiana* is dominated by 21-nt siRNAs (**Fig. 7f**). Both siRNA sizes were
558 distributed along the length of the IR dsRNA (**Fig. 7g**) and displayed the 5'U bias
559 signature of AGO1 loading (**Fig. 7h**). In Arabidopsis, trans-acting sRNAs (tasiRNAs)
560 are processed from transcripts of a specific locus (TAS3) mainly into 21-nt-siRNA by
561 DCL4 [76]. As the TAS3 locus had been previously identified in Spirodela [32], we

562 scored for matching sRNAs. In Spirodela, abundant 22-nt siRNAs account for about
563 half of TAS3 tasiRNAs (**Supplemental Fig. S34**). Given the similar 21-22-nt siRNA
564 patterns observed between TEs, miRNA-triggered secondary tasiRNA and IR-derived
565 siRNAs, PTGS on TEs is likely the result of of: (i) a shared miRNA targeting event to
566 initiate 2ry siRNAs, shown to regulate some TEs in other plants [58,162], or; (ii) direct
567 processing of RNA hairpins commonly present in LTRs [163].
568 Hence, Spirodela is capable of processing 22-nt siRNAs in the absence of
569 DCL2, allowing to conclude that 21-22-nt siRNAs mapped to TEs are *bona-fide*
570 siRNAs and that control of TEs in Spirodela combines transcriptional and post-
571 transcriptional silencing mechanisms.
572

573 # Discussion

574

575 Based on high resolution genomic information of a duckweed species during its
576 clonal reproduction, we have provided a detailed analysis of TE content, of the
577 epigenetic configuration at TEs, and several components expected to contribute to the
578 silencing pathways that control TE activity in other angiosperms. Our results reinforce
579 the idea that, with few exceptions, the Spirodela genome encodes only a reduced set
580 of both PTGS and TGS components compared to Arabidopsis, some of which are
581 absent or display a reduced set of paralogues, partially accounting for its reduced
582 levels of DNA methylation and other epigenetic marks generally associated with TEs.
583 The unique epigenome of Spirodela is the result of the loss of DNA methylation, and
584 reduced levels of H3K9me2 associated to it, in highly abundant degenerated TE
585 remnants scattered across the genome which remain nonetheless associated to the
586 heterochromatic mark H3K9me1. In contrast, potentially intact or recently active TEs
587 display all the epigenetic signatures associated with silencing in angiosperms. These
588 changes in epigenetic regulation and silencing might be a consequence of the
589 simplified morphology or reduced tissue and cell types in duckweeds compared to
590 other angiosperms [36,39,45,164,165]. Alternatively, it could be a consequence of
591 their different lifestyle or propagation mode.
592 Spirodela displays a broad range of adaptations to their free-floating aquatic
593 lifestyle, such as the loss of complete stomatal closure response to prevent water loss
594 [166]. Besides TE control, RdDM plays a role in the regulation of gene expression and
595 physiological adaptation to several abiotic stresses, such as heat, drought, salinity, or

596 low environmental humidity [20,167,168]. Given that water supply is not limiting in
597 aquatic environments, and their higher stability to fluctuations in these parameters
598 compared to terrestrial ones, RdDM activity in adult somatic tissues to respond to
599 stress might have become less relevant in Spirodela.

600 Alternatively, but not mutually exclusive, the epigenetic landscape in the
601 Spirodela genome might be a consequence of their reproduction strategy. Asexual
602 reproduction has been linked to lower TE activity and content. Transposons require
603 sexual reproduction to spread horizontally throughout populations, instead of being
604 locked in vertical clonal lineages where they might become extinct if their proliferation
605 becomes detrimental for the host without increased purifying selection of new TE
606 insertions facilitated by sex [169–172]. Furthermore, RdDM activity over short TE
607 remnants, leading to the production of 24-nt siRNAs, is seen as an immunity memory
608 against mobilization or invasion by related intact TEs [173], akin to the piRNA pathway
609 in animals [174]. Thus, TEs potentially being less active during Spirodela clonal
610 propagation might have led to the loss of DNA methylation, RdDM expression ,and
611 lack of components required for its activity during vegetative development such as
612 SHH1 and CLSY1/2 [31] over TE fragments. Maintenance of heterochromatin at such
613 repetitive TEs remnants through H3K9me1 independently of DNA methylation ensures
614 proper chromatin organization, preventing their interference with transcriptional and
615 RNA processes [9,11,12,175] or causing genome instability through illegitimate
616 recombination [176,177]. Deeper characterization of the chromatin structure (e.g.,
617 chromatin accessibility) and landscape (additional histone modifications, histone
618 variants …), together with functional studies, might shed light onto the regulation of
619 H3K9me1 and its role in shaping heterochromatin and TE regulation in Spirodela and
620 other plants.

621 Furthermore, in contrast to vegetative propagation in other plants where clones
622 originate from differentiated organs such as stems, roots or leaves, with tissue-specific
623 and developmentally regulated TEs, new individuals in duckweeds originate from a
624 dedicated group of stem cells [42]. This might have further favored a relaxation of
625 epigenetic and RdDM-mediated TE control in somatic tissues which do not participate
626 in vegetative reproduction in Spirodela. Nevertheless, potentially intact TEs are under
627 several layers of epigenetic control, including RdDM, suggesting the pathway might
628 be tissue-specific or developmentally regulated.

629 In contrast to stem cells in other angiosperms, contained in a well-organized
630 shoot apical meristem (SAM) that generates new organs [178], meristems in Spirodela
631 and all duckweeds are formed by unstructured groups of few undifferentiated cells,
632 which generate new meristems that develop into daughter fronds. [42,43,179–181]. In
633 order to maintain genome integrity along the clonal lineage, RdDM might be expressed
634 in the few stem cells within the budding pockets or during the developmental window
635 between the formation of a new meristem and its development into a new frond with
636 its own stem cells.
637 Furthermore, loss of expression of RdDM and weak CMT3/ZMET activity in
638 somatic tissues, in combination with the absence of CMT2, during development will
639 suffice to explain the global loss of mCHH and low levels of mCHG respectively, which
640 in some instances, might permit transcription of some TEs that then triggers PTGS.
641 Alternatively, transient relaxation of TE control and TE expression has been shown to
642 take place in Arabidopsis meristematic stem cells during developmental transitions
643 [182]. A similar phenomenon taking place during the development of new meristems
644 and fronds might as well explain the paradox of observed expression from TEs
645 covered in silencing marks and their apparent simultaneous regulation by TGS and
646 PTGS. Further tissue- and cell-specific investigation of epigenetic parameters will be
647 able to address these hypotheses.
648 Nonetheless, although infrequent, Spirodela does flower occasionally [47] and
649 its population genetic structure shows signs of sexual reproduction [183,184]. Hence,
650 RdDM might still be relevant and active during sexual reproduction as suggested by
651 the presence of CLSY3, which recruits RdDM to TEs during flowering in Arabidopsis
652 [31].
653 Adding to the unique epigenetic regulation of TEs in Spirodela here described,
654 one of the intriguing results of our study is the constatation that, despite lack of DCL2,
655 Spirodela can produce 22-nt siRNAs nearly as abundantly as 21-nt siRNAs, not only
656 from TEs but also from endogenous and exogenous sources of dsRNA. As these
657 sources of siRNAs are generally targets of DCL4 in plants [151,185,186], we consider
658 processing dsRNA into 22-nt siRNAs by Spirodela DCL4 as the most obvious
659 explanation. In Arabidopsis, DCL2 and its capacity to process dsRNA into 22-nt
660 siRNAs have been proposed to play a dual role in antiviral immunity. On one hand its
661 processing of viral dsRNA acts as an innate immunity sensor [187,188]. On the other
662 hand, 22-nt siRNAs in plants guide the recruitment of RDRs to boost siRNAs

663 production from their targets and amplify silencing [189–192] . Although future genetic
664 and biochemical characterization of Spirodela DCLs will be required to understand the
665 factors responsible for the biogenesis of 22-nt siRNAs, their presence of independently
666 of DCL2 and the lack of several other RNAi factors associated with antiviral defense
667 in angiosperms raises several intriguing questions about RNA-based defense
668 mechanisms in duckweeds.
669 Spirodela might represent a unique model to investigate not only transposon
670 dynamics and the epigenetic and genetic consequences of long periods of asexual
671 reproduction but also non-canonical PTGS pathways.
672

673 # Conclusions

674

675 By investigating TE silencing landscape during the clonal reproduction phase
676 of the angiosperm *Spirodela polyrhiza*, we have shown that despite the lack of
677 ubiquitous expression of RdDM and loss of DNA maintenance over degenerated TE-
678 relics, certain heterochromatic marks such as H3K9me1 can be maintained
679 independently of DNA methylation to ensure proper genome stability and chromatin
680 organization. Indicating that H3K9me1 and H3K9me2 might be differentially regulated
681 and play independent roles in heterochromatin biology. Potentially functional TEs are,
682 however, targets of several silencing mechanisms, including RdDM and PTGS,
683 probably due to tissue-specific regulation of such pathways during Spirodela clonal
684 reproduction. Interestingly, despite the absence of a DCL2 orthologue, Spirodela
685 produced 22-nt siRNAs from TEs and other sources of dsRNA, which has implications
686 not only for TE silencing but in antiviral immunity as well. Hence, Spirodela represents
687 a unique model system to gain further insights into PTGS and TGS pathways in plants.
688

689 # Methods

690

691 **Spirodela culture and growth conditions:**
692 ***Spirodela polyrhiza*.** Spirodela was grown on half-strength Schenk and Hildebrandt
693 (½ SH, Duchefa Biochemie, #S0225) liquid medium in Magenta™ GA-7 containers
694 (Sigma Aldrich, #V8505), culture dishes (Greiner Bio-One #664160), 6-wells culture
695 plates (Greiner Bio-One #657185), or in 1L glass beakers (covered with a sterile
696 plastic lid), all sealed with Leucopore tape (Duchefa Biochemie #L3301), under long-

day conditions (16h light /8h dark) at 21°C with a light intensity of 85 µM m^{-2} s^{-1}. The *Spirodela polyrhiza* #9509 clone was obtained from the Landolt collection (Zürich, Switzerland), recently transferred to the CNR-IBBA_MIDW collection (Milano, Italy: https://biomemory.cnr.it/collections/CNR-IBBA-MIDW).
Further information regarding Spirodela sterilization for axenic cultures and long-term storage, Arabidopsis and *N. benthamiana* material information and culture can be found within Supplementary Methods in the Supplementary Information files.

Genome sequencing and gene annotation:
Genomic DNA from Spirodela was purified using Dneasy® Plant mini kit (Qiagen #69104) following the manufacturer's instructions. Library preparation was carried using NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NEB #E7103). Sequencing was performed on an Illumina HiSeq4A to produce paired-end reds of 125 bp. The Landolt Sp9505 genome was generated by pseudo-assembly with MaSuRCA [193] (version 3.4.1, default parameters), using as scaffold the Sp9509_Oxforrd_v3 reference assembly[50]. The pseudo-assembly was polished using POLCA [194]. Variation, between the two assemblies, was further evaluated using GAKT[195] (version 4.0.1.2), in particular with HaplotypeCaller [196] (--standard-min-confidence-threshold-for-calling 30). Variants identified were summarize into a table using GAKT VariantsToTable (-F CHROM -F POS -F TYPE -GF AD). Gene annotations were transferred from the latest annotated version of Sp9509 [39] obtained from (https://www.lemna.org). Annotation transfer was performed using LiftOff (v1.6.3) [197] using default parameters.

Transposon annotation:
The annotation of transposable elements (TEs) in Spirodela polyrhiza was performed by combining EDTA [52] and RepeatModeler2 [53] tools. To ensure accuracy, both tools were run in triplicate and the resulting annotation libraries were aligned and merged. The classification of TEs was further refined using DeepTE [54] (-sp P). The presence of the protein domain for LTR-RT was evaluated using TEsorter [198] (version 1.3, -db rexdb-plant -st nucl).

731 **Transcriptome analysis:**

732 To extract RNA from Spirodela, 100 mg of plant material was flash-frozen upon
733 harvesting and ground. The samples were homogenized in 1 mL of TRIzol™
734 (ThermoFisher, #15596018) using Silamat S6 (Ivoclar Vivadent). RNA was extracted
735 according to manufacturer's instructions. RNA library preparation was performed
736 using the NEBNext® Ultra™ II RNA Library Prep Kit for Illumina® (NEB #E7775) or
737 the SMRTbell Prep Kit 3.0 for Illumina and Iso-Seq respectively. Illumina sequencing
738 was performed on HiSeq5A to generate 125 bp paired-end reads. Long read
739 sequencing was performed on PacBio® Sequel® II. For Illumina data, base calls were
740 performed using bcl2fastq (v2.17). RNA-seq reads were trimmed using trimgalore
741 (v0.6, --illumina -phred33 -stringency 1 --fastqc --length 20). Trimmed reads were
742 mapped using hiSat2 (v2.1.0). Gene and transposable element expression levels were
743 quantified separately using Kallisto [199] (v0.46) with default parameters. The resulting
744 expression levels (measured in TPM) were then averaged across the three replicates.
745 For Arabidopsis seedlings, data was obtained from (GSM6892967, GSM6892968,
746 GSM6892969).Iso-Seq long reads were processed using nf-core/isoseq
747 (https://github.com/nf-core/isoseq/tree/1.0.0). High quality transcripts were then
748 mapped to Spirodela genome using minimap2 [200] (v2.17, default parameters).

749

750 **DNA methylation analysis:**

751 For EMseq, DNA extraction was performed as above. Libraries were prepared using
752 the NEBNext® Enzymatic Methyl-seq Kit (NEB, #M7634) following the manufacturer's
753 instructions. Sequencing was performed on an Illumina NovaSeq SP to produce
754 paired-end reads of 150 bp. Briefly, Briefly, basecalls were performed using bcl2fastq
755 (v2.17) and sequenced reads were quality filtered and adaptor trimmed using
756 TrimGalore version 0.6.2 (https://github.com/FelixKrueger/TrimGalore). Enzymatic-
757 converted reads were aligned either to the TAIR10 genome or Spirodela polyrhiza
758 genome using Bismark version 0.22.2 with *bismark* (--non_directional *-q* --score-
759 min *L,0,–0.4*) [201] to generate coverage files per cytosine. Conversion rates were
760 assessed by mapping (samtools v1.9) the reads to the unmethylated lambda phage
761 DNA, with the same parameters used for the genomes BAM files containing clonal
762 deduplicated and uniquely mapped reads were then used to extract weighted
763 methylation rates at each cytosine as described in [202]. Replicates were merged into
764 a single bedFile using the BEDtools [203] merge function and only cytosines with

765 coverage with ≥4 were considered for further analyses. 5mC was calculated using
766 ViewBS [204] (v0.1.11) (https://github.com/xie186/ViewBS) using default parameters.
767
768 **Small RNA extraction:**
769 **Total RNA:** For small RNA extraction from total RNA and TraPR input fraction, 10µg
770 of total RNA (as described above) or RNA extracted as described above from TraPR
771 inputs was separated on a 15% denaturing polyacrylamide-urea gel. For size
772 reference, 20µl of microRNA marker (New England Biolabs, #N2102S) were loaded in
773 the same gel. Following gel staining with peqGreen (Peqlab, #732-3196), and using
774 the marker as reference, small RNAs ranging from 17-to-25 nt were excised and
775 recovered using the ZR small-RNA PAGE Recovery Kit (Zymo Research, #R1070).
776 **TraPR:** Purification of Argonaute-loaded small RNAs was performed according to the
777 protocol described in [55]. In brief, 100mg of flash-frozen plant tissue was ground to
778 fine powder using a Silamat S6 (Ivoclar Vivadent) twice for 3s in 1.5 mL
779 microcentrifuge tubes with ~20 1 mm ø glass beads (Carl Roth, #A554.1). Frozen
780 powder was incubated with TraPR lysis buffer and the homogenate loaded into the
781 TraPR columns. TraPR eluates were collected and small RNAs extracted by adding
782 and mixing 1 volume of acidic Phenol:Chloroform:Isoamyl alcohol (50:49:1, pH 4.5-5)
783 (Carl Roth, #X985.1). After centrifugation (12000*g* for 10 min at 4ºC) RNA was
784 precipitated from the aqueous phase with 1 volume of isopropanol and 10% (v/v) of
785 sodium actetate 3M pH 5.2, overnight at -20°C. RNA was collected by centrifugation
786 (12000*g* for 30 min at 4ºC) and the resulting pellet was washed with 75% ethanol and
787 resuspended in 10-20 µL RNase free water.
788
789 **Small RNA library preparation and analysis:**
790 The libraries were prepared according to [205]. In summary, small RNAs purified using
791 TraPR or total small RNA as described above, were ligated to 3′ barcoded DNA
792 adapters using truncated T4 RNA ligase 2 (New England Biolabs, #M0373). These
793 fragments were separated in an 12% denaturing polyacrylamide-urea gel, excised,
794 and purified using ZR small-RNA PAGE Recovery Kit (Zymo Research, # R1070).
795 The elution was then ligated to 5′ barcoded RNA adapters using T4 RNA ligase 1 (New
796 England Biolabs, #M0204). To reduce ligation biases, the sequence of both adaptors
797 included 4 random nucleotides at the ends. The resulting fragments were reverse
798 transcribed and amplified by PCR using primers compatible with the Illumina Platform.

799 The size distribution of the libraries was assessed using the 5200 Fragment Analyser
800 System (Agilent, #M5310AA). Sequencing was performed on an Illumina NextSeq550
801 to produce single-end reads of 75 bp. Basecalls were performed using bcl2fastq
802 (v2.17). Reads were trimmed, quality checked using Trimmomatic [206] (v0.39).
803 Trimmed reads were mapped to Arabidopsis TAIR10 and Spirodela genomes, and
804 *RUBY* or *hpScarlet* sequences using Bowtie2 [207] (v1.2.2; -e 50 -a -v 0 --best --strata
805 --nomaqround -y --phred33-quals --no-unal --sam). Mapped reads were then
806 normalized using bamCoverage [208] (Deeptools v,--normalizeUsing CPM --binSize 1
807 --smoothLength 10 --maxFragmentLength 30). To visualise strand, mapped reads
808 were split using samtools view [209] (samtools v1.9, using -h -b -F 16 or -h -b -f 16
809 options).
810
811 **Protein extraction and western blot:**
812 **Protein extraction:** For protein extraction from Arabidopsis and Spirodela, plant
813 material was flash-frozen upon harvesting and ground with a Silamat S6 as described
814 in the RNA extraction section. For *N. benthamiana*, leaf tissue was grounded in liquid
815 nitrogen using mortar and pestle. 100mg of ground tissue was mixed and
816 homogenized for 5 min. with 500µl of extraction buffer (0.7 M sucrose, 0.5 M Tris-HCl,
817 pH 8, 5 mM EDTA, pH 8, 0.1 M NaCl, 2% b-mercaptoethanol) and cOmplete™ mini
818 EDTA-free protease inhibitor cocktail (Roche, #04693159001) following
819 manufacturer's instructions. 1 volume of TRIS-buffered phenol (pH 8) was added and
820 incubated with agitation for 5 minutes. The upper phenolic phase was recovered after
821 centrifugation (12000*g* for 10 min at 4ºC) and proteins precipitated with 5 volumes of
822 0.1 M ammonium acetate dissolved in methanol at -20ºC overnight. Following
823 centrifugation (5000*g* for 15 min at 4ºC), the resulting protein pellet obtained was
824 washed twice with 0.1 M ammonium acetate dissolved in methanol and resuspended
825 in 50-100 µL of resuspension buffer (3% SDS, 62.3 mM Tris-HCl, pH 8, 10% glycerol).
826 **Protein blots:** Western blot was carried following standard denaturing SDS-PAGE
827 and blotting procedures [210]. Unless otherwise specified, 150µg of protein were
828 resolved on SDS-polyacrylamide gels, transferred onto Inmobilon-P PVDF membrane
829 (Millipore, # IPVH00010) and incubated with the appropriate antibodies (see below) in
830 PBS with 0.1% Tween-20 and 5% non-fat dried milk. The primary antibody was
831 incubated overnight at 4ºC, and the HRP-conjugated secondary antibody for 1 hour at
832 room temperature. Detection was performed using Clarity Max™ Western ECL

833 Substrate (Biorad, #1705062), and imaged with the ChemiDoc Imaging system (Bio-
834 rad). Following detection, loading was verified by coomassie staining of membranes.
835 **Primary antibodies used for western blot:** monoclonal (mouse) anti-HA tag [12CA5]
836 (Abcam, #ab1424), dilution 1:10000. Polyclonal (rabbit) anti-H3 (Abcam #ab1791),
837 dilution 1:10000. Polyclonal (rabbit) anti-H3K9me1 (Abcam #ab8896), dilution 1:2000.
838 Monoclonal (mouse) anti-H3K9me2 (Abcam #ab1220), dilution 1:2000.
839 **Secondary antibodies used for western blot:** Polyclonal (goat) anti-mouse HRP-
840 conjugated (ThermoFisher, #62-6520), dilution 1:10000. Polyclonal (goat) anti-rabbit
841 HRP-conjugated (ThermoFisher, #65-6120), dilution 1:10000.
842
843 **Chromatin immunoprecipitation (ChIP):**
844 Chromatin immunoprecipitation (ChIP) was conducted on 2 g of Arabidopsis 10-day-
845 old seedlings (WT Col-0) and *Spirodela polyrhiza* #9509 fronds. Following standard
846 procedures (see detailed method in Supplementary Methods in the Supplementary
847 Information files) using the following antibodies: anti-H3 (rabbit polyclonal, Abcam,
848 #ab1791), anti-H3K27me3 (rabbit polyclonal, Millipore, #07–449), anti-H3K4me3
849 (rabbit polyclonal, Abcam, #ab8580), anti-H3K9me1 (rabbit polyclonal, Abcam,
850 3ab8896), or anti-H3K9me2 (mouse monoclonal, Abcam, #ab1220).
851
852 **ChIP-seq library prep and data processing**:
853 For ChIP-seq, libraries were prepared using the NEBNext® Ultra™ II DNA Library
854 Prep Kit for Illumina® (NEB, #E7103) according to the manufacturer's instructions.
855 Sequencing was performed on an Illumina NovaSeq S4 to produce paired-end reads
856 of 150 bp. The raw BAM files were converted to fastq using BEDTools [203] (v2.27.1,
857 default parameters) for quality assessment of the sequenced samples. FastQC
858 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) (v0.11.8, default
859 parameters) was used to generate quality reports for all sequencing data. Basecalls
860 were performed using bcl2fastq (v2.17). Trimming of reads was performed using
861 trim_galore (https://zenodo.org/doi/10.5281/zenodo.5127898) (v0.6.5, --dont_gzip --
862 stringency 1 --fastqc --length 5'). The aligned reads were then mapped to either the
863 TAIR10 Arabidopsis genome or Spirodela polyrhiza genome using Bowtie2 [207]
864 (v2.4.1, default parameters). Duplicate reads were removed using Picard
865 (http://broadinstitute.github.io/picard/) (v2.22.8, default parameters). Mapped reads
866 were then normalized using bamCoverage (Deeptools v3.3.1). Correlations between

867 ChIP samples were evaluated using deepTools [208] (v3.1.2, default parameters). The
868 bamCompare function from deepTools was used to normalize ChIP samples to Input
869 or H3 and produce log2 ratio (ChIP/H3) bigwig files.
870
871
872 **Agrobacterium-mediated transient expression:**
873 ***N. benthamiana:*** Agrobacterium-mediated transient expression in *N. benthamiana*
874 was performed as in [211]. In brief, *A. tumefaciens* strain GV3101 harbouring the
875 vectors described below were cultured at 28°C in LB medium with appropriate
876 antibiotics. After 36 hours, cultures were pelleted at 2000g, 10 minutes at 20°C and
877 resuspended to an OD_{600} of 0.5 in filtered sterilized agroinfiltration buffer: 2mM MES,
878 50mM NaH_2PO_4, 0.5% (w/v) glucose, pH 5.6. Acetosyringone was added prior to use
879 at 100 µM. After 1 hour of incubation at room temperature in darkness, the bacterial
880 solutions were infiltrated into the abaxial side of young *N. benthamiana* true leaves,
881 using a 1ml syringe without needle. At 3 or 5 days after infiltration (dpi), leaf discs were
882 collected, flash-frozen in liquid nitrogen and ground in liquid nitrogen using mortar and
883 pestle. The ground samples were stored at -80C.
884 ***Spirodela polyrhiza*:** For Agrobacterium-mediated transient expression in Spirodela,
885 Agrobacterium strain EHA105 was used. Bacteria cultures were prepared as for *N.*
886 *benthamiana* infiltrations with the following modifications: Bacteria was resuspended
887 to OD_{600} 0.8 in Spirodela agroinfiltration media (10 mM MgCl2, 5% sucrose (w/v), pH
888 5.6). Spirodela fronds were prepared as it follows: Cultures were initiated from axenic
889 stocks by inoculating fronds into 2L glass beakers containing 0.5-1L of ½ SH media,
890 covered with the lid of a 145 x 20 mm polystyrene culture dish (Greiner Bio-One
891 #639102) sealed with 2.5 cm wide Leucopore tape (Duchefa Biochemie #L3301).
892 Cultures were grown to 50-80% confluence (as surface occupancy). Prior to
893 agroinfiltration, fronds were pretreated as in [160] by transferring them to an ice-cold
894 20 mg/L L-Glutamine (PanReac AppliChem #A1420) solution for 20 min. Following
895 pretreatment, fronds were infiltrated with the *Vir*-induced Agrobacterium solution either
896 (i) manually, at the budding pouch area with a 1 mL syringe without needle by gently
897 applying pressure with the plunger on the upper side of fronds or, (ii) trough vacuum
898 infiltration by transferring the fronds to 100 mL agrobacterium solution complemented
899 with 0.02% (v/v) Silwet L-77 (Kurt-Obermieier GmbH #7060-10) in a 200 mL beaker in
900 a vacuum desiccator. -70 kPa vacuum was applied twice for 10 min. Following either

901 manual or vacuum agroinfiltration, excess of bacteria was removed by rinsing the
902 fronds in ddH_2O and plants transferred to 100 x 20 mm culture dishes (Greiner Bio-
903 One #664160) containing 30-50 mL ½ SH media + 100 µM acetosyringone for a 3 day
904 co-culture. 4 days post-infiltration (dpi) transfer fronds to plates containing ½ SH
905 media. At 7 dpi agrobacterium elimination was initiated by or replacing the media with
906 ½ SH + 250 µM ticarcillin (Duchefa #T0190.0010, lab stock at 250 mg/mL in H2O) +
907 300 µM cefotaxime (Duchefa #C0111.0005, lab stock at 300 mg/mL in H2O) for a
908 week. Plants were cultured with regular (1/week) ½ SH media replacements. See
909 Supplementary Fig. S36 for a graphic depiction of the protocol and workflow.
910
911 **Small RNA blot analysis:**
912 AGO-loaded small RNAs were purified from 100mg of starting material using TraPR
913 columns, as described above. The resulting small RNAs were mixed with an equal
914 volume of 2X Novex™ TBE-Urea sample buffer (ThermoFisher, #LLC6876),
915 separated on a 17.5% polyacrylamide-urea gel, transferred onto Hybond-NX
916 membrane (Cytiva, #RPN303T) and chemically crosslinked (12.5M 1-methylimidazol
917 (Merk, #M50834), 31.25mg/mL N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide
918 hydrochloride (Merk, #E7750) as in [212]. Probes produced from PCR templates were
919 radiolabelled with [α-^{32}P]-dCTP (Hartmann Analytic, #SRP-305) using Prime-it II
920 Random Primer Labelling kit (Agilent, #300385) and oligonucleotides used as a probe
921 were radiolabelled with [γ-^{32}P]-ATP (Hartmann analytic #SRP-501), using T4
922 Polynucleotide kinase (Thermofisher, #EK0031) (see Supplementary figures for a list
923 of oligonucleotides used). The probes were incubated in PerfectHyb™ Plus
924 Hybridization Buffer (Sigma Aldrich, #H7033) and washed with 2XSSC, 0.1%SDS.
925 The signal was detected with the Amersham™Typhoon™ (Cytivia). Probes and
926 oligonucleotides used are listed in Supplementary figures.
927
928 **Immunoprecipitation of Flag-tagged AGO4:**
929 100-150mg of *N. benthamiana* ground tissue was homogenized in 1 ml of IP buffer
930 (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 0.1% NP-40) containing 2 µM
931 MG-132 and cOmplete™ mini EDTA-free protease inhibitor cocktail (Roche,
932 #04693159001) for 30 minutes at 4 °C. After lysis, debris were removed by
933 centrifugation (12000*g* for 10 min at 4ºC). An aliquot of lysate was kept as input, while
934 the remainder was incubated with ANTI-FLAG^{R} M2 Magnetic beads (20µl of

935 slurry/sample) (Sigma-Aldrich, #M8823) for 3 hours at 4°C, on a orbital shaker. After
936 incubation, the resin was separated from supernatant in a magnetic rack (DynaMag-
937 5, Thermofisher, #12321D) and washed with 1 ml of IP buffer for 20 minutes at 4°C,
938 for 3 times. After the final wash, 100µl of PBS were added and 10µl put aside, used to
939 check immunoprecipitation quality by western blot by adding Laemli buffer and boiling
940 the sample for 10 min. at 95ºC. The remaining 90µl was used to analyze the small
941 RNA composition upon extraction using TRIzol™ (ThermoFisher, #15596018). The
942 isolated small RNAs were used to generate libraries and sequenced as described
943 above. TraPR purified small RNAs (described above) from not-infiltrated plants were
944 used to assess *N. benthamiana* small RNA profile.
945
946 **Visualization and Statistical analysis:**
947 Data visualization files were generated according to each specific sections. Statistical
948 tests were performed using R-dependent ggpubr (v0.6.0)
949 (https://rpkgs.datanovia.com/ggpubr/), statistical test performed are indicated
950 accordingly on each figure legend figure. Plots under R (v4.2.2) (https://www.r-
951 project.org) were generated using Circlize [213] (v0.4.15), ggplot2 [214] (v3.4.4),
952 profileplyr (v1.14.1)
953 (https://bioconductor.org/packages/release/bioc/html/profileplyr.html), ViewBS [204]
954 (v0.1.11), or pheatmap (v 1.0.12)
955 (https://rdocumentation.org/packages/pheatmap/versions/1.0.12).
956
957 **Supplementary Methods:**
958 Further information regarding plant culture, plant imaging, phylogenetic analysis, mass
959 spectrometry of histone PTMs, chromatin immunoprecipitation, DAPI staining of nuclei
960 and molecular cloning procedures can be found within the Supplementary Methods in
961 the Supplementary Information files.
962
963
964 # Ethics approval and consent to participate
965
966 Not applicable
967
968 # Consent for publication

Not applicable

Availability of data and materials

An inventory of all the raw data used to generate each figure panel (as well as those in Supplementary information) along with all raw image files (including RNA, protein blots and agarose gels), DNA and protein sequences, alignments, and phylogenetic tree files can be found in Zenodo (www.zenodo.org) under the doi: 10.5281/zenodo.10911533. NGS data generated for this study are accessible NCBI SRA under ID number PRJNA1095698, submission ID SUB14339214. Publicly available Arabidopsis seedling transcriptome from [74] is found under the accession number GSM6892968. Ready-to-use IGV genome browser tracks for all the NGS datasets generated here is available in the associated Zenodo dataset. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortiumvia the PRIDE partner repository with the dataset identifier PXD050443. Additional Excel files with the MS processed data can be found in the associated Zenodo dataset. All plasmids generated have been submitted to Addgene and can be retrieved under the following IDs: p*35S*:*FHA*-At*AGO4*_gDNA (#216838); p*35S*:*FHA*-Sp*AGO4a*_gDNA (#216841); p*35S*::*FHA*-Sp*AGO4a*_cDNA (#216842). The following plasmids previously originated in [161] and [215] were obtained from Addgene: p*35S*:*RUBY* (#160908); p*ZmUbq*:*RUBY* (#160909); p*35S*:*GFP-GUS* (#167122). p3At*UBQ*:*hpScarlet* was a gift from Marco Incarbone (Max-Planck-Institute of Plant Physiology, Golm, Germany).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Gregor Mendel Institute of the Austrian Academy of Sciences core funding attributed to A.M-O.

Author contributions

AM-O and RD conceived and designed the study. RD and VB-B performed most experiments. AM-O and RD performed identification and phylogenetic analysis of silencing components in Spirodela. VB-B cloned and carried out transient AGO4 expression experiments and prepared small RNA libraries. DB-A made nuclei DAPI stainings. JMV-G performed western blot analysis of HPTMs. AM-O, DB-A, and AP-M optimized agroinfiltration-mediated transient expression in duckweeds. AP-M carried out transient expression experiments in Spirodela. VB-B and RE cultured and maintained Spirodela stocks and axenic cultures and provided technical support. RD performed computer and statistical analysis. AM-O and RD analyzed the data and wrote the manuscript.

Acknowledgements

We thank current and past members of the Marí-Ordóñez group and colleagues from the Gregor Mendel Institute and the Vienna BioCenter (VBC) campus for fruitful discussions, ideas, and feedback. We are specially thankful to Ortrun Mitettelsten Scheid for critical feedback and comments on the manuscript. We would also like to thank Klaus Appenroth from the Friedrich Schiller University of Jena for useful advice on duckweed cultivation and storage, Thomas Grentzinger for technical assistance with the TraPR procedure, and Marco Incarbone from the Max-Planck-Institute of Plant Physiology in Golm for kindly providing the pAt*UBQ*:*hpScarlet* plasmid. We are grateful to the Vienna BioCenter Core Facilities (VBCF): NGS for DNA, EM-seq, PacBio, RNA-seq and ChIP-seq library preparation and sequencing, including sequencing of small RNA libraries; the IMP/IMBA/GMI Proteomics Facility, in special Ines Steinmacher for sample preparation, mass-spectrometry using the VBCF instrument pool, and Richard Imre and Elisabeth Rothinger for assistance with data

1035 analysis of histone PTMs; the Plant Science Facility for assistance with plant work; the
1036 IMP Molecular Biology Service for providing reagents and Sanger sequencing. Lastly,
1037 we would like to thank the VBC in-house COVID-19 testing facility for enabling a safe
1038 working environment during the pandemic.

References

1. Wells JN, Feschotte C. A Field Guide to Eukaryotic Transposable Elements. Annu Rev Genet. 2020;54:1–23.
2. Biémont C, Vieira C. Genetics: junk DNA as an evolutionary force. 2006 Oct 5;521–4.
3. Feschotte C. Transposable elements and the evolution of regulatory networks. Nature. 2008;9:397–405.
4. Allshire RC, Madhani HD. Ten principles of heterochromatin formation and function. Nature Reviews Molecular Cell Biology. 2018;19:229–44.
5. Henikoff S, Smith MM. Histone Variants and Epigenetics. Cold Spring Harb Perspect Biol. 2015;7:a019364.
6. Jamge B, Lorković ZJ, Axelsson E, Osakabe A, Shukla V, Yelagandula R, et al. Histone variants shape chromatin states in Arabidopsis. eLife. 2023;12:RP87714.
7. Blumenstiel JP. Birth, School, Work, Death, and Resurrection: The Life Stages and Dynamics of Transposable Element Proliferation. Genes. 2019;10:336.
8. Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature. 2011;472:115–9.
9. Kim MY, Zilberman D. DNA methylation as a system of plant genomic immunity. Trends Plant Sci. 2014;19:320–6.
10. Zervudacki J, Yu A, Amesefe D, Wang J, Drouaud J, Navarro L, et al. Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons. The EMBO journal. 2018;37:e98482.
11. Sammarco I, Pieters J, Salony S, Toman I, Zolotarov G, Placette CL. Epigenetic targeting of transposon relics: beating the dead horses of the genome? Epigenetics. 2022;17:1331–44.
12. Ilık İA, Glažar P, Tse K, Brändl B, Meierhofer D, Müller F-J, et al. Autonomous transposons tune their sequences to ensure somatic suppression. Nature. 2024;1–9.
13. Houben A, Demidov D, Gernand D, Meister A, Leach CR, Schubert I. Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. Plant J. 2003;33:967–73.
14. Seymour DK, Koenig D, Hagmann J, Becker C, Weigel D. Evolution of DNA Methylation Patterns in the Brassicaceae is Driven by Differences in Genome Organization. PLoS Genet. 2014;10:e1004785.
15. Sigman MJ, Slotkin RK. The First Rule of Plant Transposable Element Silencing: Location, Location, Location. Plant Cell. 2016;28:304–13.
16. Wyler M, Stritt C, Walser J-C, Baroux C, Roulin AC. Impact of transposable elements on methylation and gene expression across natural accessions of Brachypodium distachyon. Genome Biol Evol. 2020;12:evaa180-.
17. Klein SP, Anderson SN. The evolution and function of transposons in epigenetic regulation in response to the environment. Curr Opin Plant Biol. 2022;69:102277.
18. Moazed D. Small RNAs in transcriptional gene silencing and genome defence. Nature. 2009;457:413–20.
19. Malone CD, Hannon GJ. Small RNAs as Guardians of the Genome. Cell. 2009;136:656–68.
20. Erdmann RM, Picard CL. RNA-directed DNA Methylation. Plos Genet. 2020;16:e1009034.
21. Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, et al. Selective epigenetic control of retrotransposition in Arabidopsis. Nature. 2009;461:427–30.
22. Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. Bursts of retrotransposition reproduced in Arabidopsis. Nature. 2010;461:423–6.
23. Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, et al. Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. Nat Struct Mol Biol. 2014;21:64–72.
24. He L, Zhao C, Zhang Q, Zinta G, Wang D, Lozano-Durán R, et al. Pathway conversion enables a double-lock mechanism to maintain DNA methylation and genome stability. Proc Natl Acad Sci. 2021;118:e2107320118.
25. Osakabe A, Jamge B, Axelsson E, Montgomery SA, Akimcheva S, Kuehn AL, et al. The chromatin remodeler DDM1 prevents transposon mobility through deposition of histone variant H2A.W. Nat Cell Biol. 2021;23:391–400.
26. Bourc'his D, Voinnet O. A Small-RNA Perspective on Gametogenesis, Fertilization, and Early Zygotic Development. Science. 2010;330:617–22.
27. Marí-Ordóñez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. Reconstructing de novo silencing of an active plant retrotransposon. Nature Genetics. 2013;45:1029–39.

1087 28. Parhad SS, Theurkauf WE. Rapid evolution and conserved function of the piRNA pathway. R Soc Open Biol.
1088 2019;9:180181.
1089 29. Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, et al. The Arabidopsis RNA-Directed DNA
1090 Methylation Argonautes Functionally Diverge Based on Their Expression and Interaction with Target Loci. THE PLANT
1091 CELL ONLINE. 2010;22:321–34.
1092 30. Patel P, Mathioni SM, Hammond R, Harkess AE, Kakrana A, Arikit S, et al. Reproductive phasiRNA loci and DICER-
1093 LIKE5, but not microRNA loci, diversified in monocotyledonous plants. Plant Physiol. 2021;185:kiab001.
1094 31. Zhou M, Coruh C, Xu G, Martins LM, Bourbousse C, Lambolez A, et al. The CLASSY family controls tissue-specific
1095 DNA methylation patterns in Arabidopsis. Nat Commun. 2022;13:244.
1096 32. Michael TP, Bryant D, Gutierrez R, Borisjuk N, Chu P, Zhang H, et al. Comprehensive definition of genome features in
1097 Spirodela polyrhiza by high-depth physical mapping and short-read DNA sequencing strategies. Plant J. 2017;89:617–
1098 35.
1099 33. Fourounjian P, Tang J, Tanyolac B, Feng Y, Gelfand B, Kakrana A, et al. Post-transcriptional adaptation of the aquatic
1100 plant Spirodela polyrhiza under stress and hormonal stimuli. Plant J. 2019;98:1120–33.
1101 34. Cervantes-Pérez SA, Yong-Villalobos L, Florez-Zapata NMV, Oropeza-Aburto A, Rico-Reséndiz F, Amasende-Morales
1102 I, et al. Atypical DNA methylation, sRNA-size distribution, and female gametogenesis in Utricularia gibba. Sci Rep.
1103 2021;11:15725.
1104 35. Ibarra-Laclette E, Lyons E, Hernández-Guzmán G, Pérez-Torres CA, Carretero-Paulet L, Chang T-H, et al. Architecture
1105 and evolution of a minute plant genome. Nature. 2013;498:94–8.
1106 36. Wang W, Haberer G, Gundlach H, Gläßer C, Nussbaumer T, Luo MC, et al. The Spirodela polyrhiza genome reveals
1107 insights into its neotenous reduction fast growth and aquatic lifestyle. Nat Commun. 2014;5:3311.
1108 37. An D, Zhou Y, Li C, Xiao Q, Wang T, Zhang Y, et al. Plant evolution and environmental adaptation unveiled by long-
1109 read whole-genome sequencing of Spirodela. Proc Natl Acad Sci. 2019;116:18893–9.
1110 38. Harkess A, Bewick AJ, Lu Z, Fourounjian P, Messing J, Michael TP, et al. Unusual predominance of maintenance DNA
1111 methylation in Spirodela polyrhiza. bioRxiv. 2020;2020.12.03.410332.
1112 39. Ernst E, Abramson B, Acosta K, Hoang PTN, Mateo-Elizalde C, Schubert V, et al. The genomes and epigenomes of
1113 aquatic plants (Lemnaceae) promote triploid hybridization and clonal reproduction. bioRxiv. 2023;2023.08.02.551673.
1114 40. Bog M, Appenroth K-J, Sree KS. Duckweed (Lemnaceae): Its Molecular Taxonomy. Front Sustain Food Syst.
1115 2019;3:117.
1116 41. Tippery NP, Les DH, Appenroth KJ, Sree KS, Crawford DJ, Bog M. Lemnaceae and Orontiaceae Are Phylogenetically
1117 and Morphologically Distinct from Araceae. Plants. 2021;10:2639.
1118 42. Landolt E. Biosystematic investigations in the family of duckweeds (Lemnaceae) (Vol 2). The family of Lemnaceae a
1119 monographic study. ETH V des GI der, editor. Stiftung Ruebel, Switzerland: Veroeffentlichungen des Geobotanischen
1120 Instituts der ETH; 1986.
1121 43. Lemon GD, Posluszny U. Comparative Shoot Development and Evolution in the Lemnaceae. Int J Plant Sci.
1122 2000;161:733–48.
1123 44. Lam E, Michael TP. Wolffia, a minimalist plant and synthetic biology chassis. Trends Plant Sci. 2022;27:430–9.
1124 45. Ware A, Jones DH, Flis P, Chrysanthou E, Smith KE, Kümpers BMC, et al. Loss of ancestral function in duckweed
1125 roots is accompanied by progressive anatomical reduction and a re-distribution of nutrient transporters. Curr Biol.
1126 2023;33:1795-1802.e4.
1127 46. Pieterse AH. Is flowering in Lemnaceae stress-induced? A review. Aquat Bot. 2013;104:1–4.
1128 47. Fourounjian P, Slovin J, Messing J. Flowering and Seed Production across the Lemnaceae. Int J Mol Sci.
1129 2021;22:2733.
1130 48. Ziegler P, Adelmann K, Zimmer S, Schmidt C, Appenroth K -J. Relative in vitro growth rates of duckweeds (Lemnaceae)
1131 – the most rapidly growing higher plants. Plant Biol. 2015;17:33–41.
1132 49. Xu S, Stapley J, Gablenz S, Boyer J, Appenroth KJ, Sree KS, et al. Low genetic variation is associated with low mutation
1133 rate in the giant duckweed. Nat Commun. 2019;10:1243.
1134 50. Hoang PNT, Michael TP, Gilbert S, Chu P, Motley ST, Appenroth KJ, et al. Generating a high-confidence reference
1135 genome map of the Greater Duckweed by integration of cytogenomic, optical mapping, and Oxford Nanopore
1136 technologies. Plant J. 2018;96:670–84.

1137 51. Harkess A, McLoughlin F, Bilkey N, Elliott K, Emenecker R, Mattoon E, et al. Improved Spirodela polyrhiza genome
1138 and proteomic analyses reveal a conserved chromosomal structure with high abundance of chloroplastic proteins
1139 favoring energy production. J Exp Bot. 2021;72:2491–500.
1140 52. Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, et al. Benchmarking transposable element annotation
1141 methods for creation of a streamlined, comprehensive pipeline. Genome Biol. 2019;20:275.
1142 53. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, et al. RepeatModeler2 for automated genomic
1143 discovery of transposable element families. Proc Natl Acad Sci. 2020;117:9451–7.
1144 54. Yan H, Bombarely A, Li S. DeepTE: a computational method for de novo classification of transposons with convolutional
1145 neural network. Bioinformatics. 2020;36:4269–75.
1146 55. Grentzinger T, Oberlin S, Schott G, Handler D, Svozil J, Barragan-Borrero V, et al. A universal method for the rapid
1147 isolation of all known classes of functional silencing small RNAs. Nucleic Acids Res. 2020;48:gkaa472-.
1148 56. Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, Ciaudo C, et al. A Role for RNAi in the Selective Correction
1149 of DNA Methylation Defects. Science. 2009;323:1600–4.
1150 57. Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN, et al. The Initiation of Epigenetic Silencing of
1151 Active Transposable Elements Is Triggered by RDR6 and 21-22 Nucleotide Small Interfering RNAs. PLANT
1152 PHYSIOLOGY. 2013;162:116–31.
1153 58. Creasey KM, Zhai J, Borges F, Ex FV, Regulski M, Meyers BC, et al. miRNAs trigger widespread epigenetically
1154 activated siRNAs from transposons in Arabidopsis. Nature. 2014;508:411–5.
1155 59. Feng S, Zhong Z, Wang M, Jacobsen SE. Efficient and accurate determination of genome-wide DNA methylation
1156 patterns in Arabidopsis thaliana with enzymatic methyl sequencing. Epigenetics Chromatin. 2020;13:42.
1157 60. Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q, et al. Widespread natural variation of DNA methylation
1158 within angiosperms. Genome Biol. 2016;17:194.
1159 61. Bewick AJ, Schmitz RJ. Gene body DNA methylation in plants. Curr Opin Plant Biol. 2017;36:103–10.
1160 62. Mathieu O, Reinders J, Čaikovski M, Smathajitt C, Paszkowski J. Transgenerational Stability of the Arabidopsis
1161 Epigenome Is Coordinated by CG Methylation. Cell. 2007;130:851–62.
1162 63. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the
1163 *Arabidopsis* genome reveals DNA methylation patterning. Nature. 2008;452:215–9.
1164 64. Reinders J, Wulff BBH, Mirouze M, Ordóñez AM, Dapp M, Rozhon W, et al. Compromised stability of DNA methylation
1165 and transposon immobilization in mosaic Arabidopsis epigenomes. Genes & development. 2009;23:939–50.
1166 65. Gouil Q, Baulcombe DC. DNA Methylation Signatures of the Plant Chromomethyltransferases. PLoS Genet.
1167 2016;12:e1006526.
1168 66. Li X, Harris CJ, Zhong Z, Chen W, Liu R, Jia B, et al. Mechanistic insights into plant SUVH family H3K9
1169 methyltransferases and their binding to context-biased non-CG DNA methylation. Proc Natl Acad Sci.
1170 2018;115:E8793–802.
1171 67. Fang J, Jiang J, Leichter SM, Liu J, Biswal M, Khudaverdyan N, et al. Mechanistic basis for maintenance of CHG DNA
1172 methylation in plants. Nat Commun. 2022;13:3877.
1173 68. Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE. Comprehensive Analysis of Silencing Mutants
1174 Reveals Complex Regulation of the Arabidopsis Methylome. Cell. 2013;152:352–64.
1175 69. Yaari R, Noy-Malka C, Wiedemann G, Gershovitz NA, Reski R, Katz A, et al. DNA METHYLTRANSFERASE 1 is
1176 involved in mCG and mCCG DNA methylation and is essential for sporophyte development in Physcomitrella patens.
1177 Plant Mol Biol. 2015;88:387–400.
1178 70. Bewick AJ, Niederhuth CE, Ji L, Rohr NA, Griffin PT, Leebens-Mack J, et al. The evolution of CHROMOMETHYLASES
1179 and gene body DNA methylation in plants. Genome Biol. 2017;18:65.
1180 71. Bewick AJ, Ji L, Niederhuth CE, Willing E-M, Hofmeister BT, Shi X, et al. On the origin and evolutionary consequences
1181 of gene body DNA methylation. Proc Natl Acad Sci. 2016;113:9111–6.
1182 72. Wendte JM, Zhang Y, Ji L, Shi X, Hazarika RR, Shahryary Y, et al. Epimutations are associated with
1183 CHROMOMETHYLASE 3-induced de novo DNA methylation. eLife. 2019;8:e47891.
1184 73. Harkess A, Bewick AJ, Lu Z, Fourounjian P, Michael TP, Schmitz RJ, et al. Unusual predominance of maintenance
1185 DNA methylation in Spirodela polyrhiza. G3: Genes, Genomes, Genet. 2024;jkae004
1186 74. Li Z, Wang M, Zhong Z, Gallego-Bartolomé J, Feng S, Jami-Alahmadi Y, et al. The MOM1 complex recruits the RdDM
1187 machinery via MORC6 to establish de novo DNA methylation. Nat Commun. 2023;14:4135.

75. Zhan J, Meyers BC. Plant Small RNAs: Their Biogenesis, Regulatory Roles, and Functions. Annu Rev Plant Biol. 2023;74:21–51.
76. Yoshikawa M, Peragine A, Park MY, Poethig RS. A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. Genes Development. 2005;19:2164–75.
77. Liu Y, Teng C, Xia R, Meyers BC. PhasiRNAs in Plants: Their Biogenesis, Genic Sources, and Roles in Stress Responses, Development, and Reproduction. THE PLANT CELL ONLINE. 2020;32:3059–80.
78. Shamandi N, Zytnicki M, Charbonnel C, Elvira-Matelot E, Bochnakian A, Comella P, et al. Plants Encode a General siRNA Suppressor That Is Induced and Suppressed by Viruses. PLoS Biol. 2015;13:e1002326.
79. Tschopp M-A, Iki T, Brosnan CA, Jullien PE, Pumplin N. A complex of Arabidopsis DRB proteins can impair dsRNA processing. RNA. 2017;23:782–97.
80. Montavon T, Kwon Y, Zimmermann A, Hammann P, Vincent T, Cognat V, et al. A specific dsRNA-binding protein complex selectively sequesters endogenous inverted-repeat siRNA precursors and inhibits their processing. Nucleic Acids Res. 2017;45:1330–44.
81. Zong J, Yao X, Yin J, Zhang D, Ma H. Evolution of the RNA-dependent RNA polymerase (RdRP) genes: Duplications and possible losses before and after the divergence of major eukaryotic groups. Gene. 2009;447:29–39.
82. Willmann MR, Endres MW, Cook RT, Gregory BD. The Functions of RNA-Dependent RNA Polymerases in Arabidopsis. Arab Book. 2011;2011:e0146.
83. Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science. 2006;313:68–71.
84. Garcia-Ruiz H, Takeda A, Chapman EJ, Sullivan CM, Fahlgren N, Brempelis KJ, et al. Arabidopsis RNA-Dependent RNA Polymerases and Dicer-Like Proteins in Antiviral Defense and Small Interfering RNA Biogenesis during Turnip Mosaic Virus Infection. THE PLANT CELL ONLINE. 2010;22:481–96.
85. Harvey JJW, Lewsey MG, Patel K, Westwood J, Heimstädt S, Carr JP, et al. An Antiviral Defense Role of AGO2 in Plants. Bendahmane M, editor. PLoS ONE. 2011;6:e14639.
86. Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert KB, Montgomery TA, Nguyen T, et al. Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. THE PLANT CELL ONLINE. 2012;24:3613–29.
87. Yu D, Fan B, MacFarlane SA, Chen Z. Analysis of the Involvement of an Inducible Arabidopsis RNA-Dependent RNA Polymerase in Antiviral Defense. Mol Plant-Microbe Interact. 2003;16:206–16.
88. Brosseau C, Moffett P. Functional and Genetic Analysis Identify a Role for Arabidopsis ARGONAUTE5 in Antiviral RNA Silencing. Plant Cell. 2015;27:1742–54.
89. Incarbone M, Bradamante G, Pruckner F, Wegscheider T, Rozhon W, Nguyen V, et al. Salicylic acid and RNA interference mediate antiviral immunity of plant stem cells. Proc Natl Acad Sci. 2023;120:e2302069120.
90. Bradamante G, Nguyen VH, Incarbone M, Meir Z, Bente H, Donà M, et al. Two ARGONAUTE proteins loaded with transposon-derived small RNAs are associated with the reproductive cell lineage in Arabidopsis. Plant Cell. 2023;koad295.
91. Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE. SHH1, a Homeodomain Protein Required for DNA Methylation, As Well As RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA Polymerase IV. PLoS Genet. 2011;7:e1002195.
92. Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AMS, et al. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. Nature. 2013;1–7.
93. Zhou M, Palanca AMS, Law JA. Locus-specific control of the de novo DNA methylation pathway in Arabidopsis by the CLASSY family. Nat Genet. 2018;50:865–73.
94. Diego-Martin B, Pérez-Alemany J, Candela-Ferre J, Corbalán-Acedo A, Pereyra J, Alabadí D, et al. The TRIPLE PHD FINGERS proteins are required for SWI/SNF complex-mediated +1 nucleosome positioning and transcription start site determination in Arabidopsis. Nucleic Acids Res. 2022;50:10399–417.
95. Haag JR, Brower-Toland B, Krieger EK, Sidorenko L, Nicora CD, Norbeck AD, et al. Functional Diversification of Maize RNA Polymerase IV and V Subtypes via Alternative Catalytic Subunits. Cell Rep. 2014;9:378–90.
96. Wang Y, Zhou X, Luo J, Lv S, Liu R, Du X, et al. Recognition of H3K9me1 by maize RNA-directed DNA methylation factor SHH2. J Integr Plant Biol. 2021;63:1091–6.
97. Duan C-G, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, et al. Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. EMBO. 2014;e201489453.

98. Johnson LM, Law JA, Khattar A, Henderson IR, Jacobsen SE. SRA-Domain Proteins Required for DRM2-Mediated De Novo DNA Methylation. PLoS Genet. 2008;4:e1000280.
99. Johnson LM, Du J, Hale CJ, Bischof S, Feng S, Chodavarapu RK, et al. SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. Nature. 2014;507:124–8.
100. Liu Z-W, Shao C-R, Zhang C-J, Zhou J-X, Zhang S-W, Li L, et al. The SET Domain Proteins SUVH2 and SUVH9 Are Required for Pol V Occupancy at RNA-Directed DNA Methylation Loci. PLoS Genetics. 2014;10:e1003948.
101. Chan SW-L, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE. RNA Silencing Genes Control de Novo DNA Methylation. Science. 2004;303:1336–1336.
102. Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. Nature Genetics. 2009;41:630–4.
103. Zhong X, Hale CJ, Nguyen M, Ausin I, Groth M, Hetzel J, et al. DOMAINS REARRANGED METHYLTRANSFERASE3 controls DNA methylation and regulates RNA polymerase V transcript abundance in Arabidopsis. Proc Natl Acad Sci. 2015;112:911–6.
104. Liu W, Duttke SH, Hetzel J, Groth M, Feng S, Gallego-Bartolome J, et al. RNA-directed DNA methylation involves co-transcriptional small-RNA-guided slicing of polymerase V transcripts in Arabidopsis. Nat Plants. 2018;4:181–8.
105. Fang J, Leichter SM, Jiang J, Biswal M, Lu J, Zhang Z-M, et al. Substrate deformation regulates DRM2-mediated DNA methylation in plants. Sci Adv. 2021;7:eabd9224.
106. El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, et al. Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Development. 2007;21:2539–44.
107. Bies-Etheve N, Pontier D, Lahmy S, Picart C, Vega D, Cooke R, et al. RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. EMBO Rep. 2009;10:649–54.
108. Rowley MJ, Avrutsky MI, Sifuentes CJ, Pereira L, Wierzbicki AT. Independent Chromatin Binding of ARGONAUTE4 and SPT5L/KTF1 Mediates Transcriptional Gene Silencing. PLoS Genet. 2011;7:e1002120.
109. Huang Y, Kendall T, Forsythe ES, Dorantes-Acosta A, Li S, Caballero-Pérez J, et al. Ancient Origin and Recent Innovations of RNA Polymerase IV and V. Mol Biol Evol. 2015;32:1788–99.
110. Trujillo JT, Beilstein MA, Mosher RA. The Argonaute-binding platform of NRPE1 evolves through modulation of intrinsically disordered repeats. N Phytol. 2016;212:1094–105.
111. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen H, et al. Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. Cell. 2006;126:1189–201.
112. Regulski M, Lu Z, Kendall J, Donoghue MTA, Reinders J, Llaca V, et al. The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. Genome Res. 2013;23:1651–62.
113. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. Nature Reviews Genetics. 2007;8:973–82.
114. Zattera ML, Bruschi DP. Transposable Elements as a Source of Novel Repetitive DNA in the Eukaryote Genome. Cells. 2022;11:3373.
115. Zhong X, Hale CJ, Law JA, Johnson LM, Feng S, Tu A, et al. DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. Nat Struct Mol Biol. 2012;19:870–5.
116. Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, et al. The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. Cell. 2013;153:193–205.
117. Wang Z, Baulcombe DC. Transposon age and non-CG methylation. Nat Commun. 2020;11:1221.
118. Kumar A, Bennetzen JL. Plant retrotransposons. Annual Review of Genetics. 1999;33:479–532.
119. Gorinšek B, Gubenšek F, Kordiš D. Evolutionary Genomics of Chromoviruses in Eukaryotes. Mol Biol Evol. 2004;21:781–98.
120. Neumann P, Novák P, Hoštáková N, Macas J. Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. Mob DNA. 2019;10:1.
121. Jackson JP, Johnson L, Jasencakova Z, Zhang X, PerezBurgos L, Singh PB, et al. Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. Chromosoma. 2004;112:308–15.

1289 122. Fuchs J, Demidov D, Houben A, Schubert I. Chromosomal histone modification patterns – from conservation to
1290 diversity. Trends Plant Sci. 2006;11:199–208.
1291 123. Choi J, Lyons DB, Zilberman D. Histone H1 prevents non-CG methylation-mediated small RNA biogenesis in
1292 Arabidopsis heterochromatin. eLife. 2021;10:e72676.
1293 124. Johnson L, Mollah S, Garcia BA, Muratore TL, Shabanowitz J, Hunt DF, et al. Mass spectrometry analysis of
1294 Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. Nucleic Acids Res.
1295 2004;32:6511–8.
1296 125. Scheid R, Dowell JA, Sanders D, Jiang J, Denu JM, Zhong X. Histone Acid Extraction and High Throughput Mass
1297 Spectrometry to Profile Histone Modifications in Arabidopsis thaliana. Curr Protoc. 2022;2:e527.
1298 126. Mathieu O, Probst AV, Paszkowski J. Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG
1299 methylation in Arabidopsis. 2005;24:2783–91.
1300 127. Walter M, Teissandier A, Pérez-Palacios R, Bourc'his D. An epigenetic switch ensures transposon repression upon
1301 dynamic loss of DNA methylation in embryonic stem cells. eLife. 2016;5:e11418.
1302 128. Montgomery SA, Tanizawa Y, Galik B, Wang N, Ito T, Mochizuki T, et al. Chromatin Organization in Early Land Plants
1303 Reveals an Ancestral Association between H3K27me3, Transposons, and Constitutive Heterochromatin. Curr Biol.
1304 2020;30:573-588.e7.
1305 129. Rougée M, Quadrana L, Zervudacki J, Hure V, Colot V, Navarro L, et al. Polycomb mutant partially suppresses DNA
1306 hypomethylation–associated phenotypes in Arabidopsis. Life Sci Alliance. 2021;4:e202000848.
1307 130. Hisanaga T, Romani F, Wu S, Kowar T, Wu Y, Lintermann R, et al. The Polycomb repressive complex 2 deposits
1308 H3K27me3 and represses transposable elements in a broad range of eukaryotes. Curr Biol. 2023;33:4367-4380.e9.
1309 131. Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide analysis of mono-, di- and
1310 trimethylation of histone H3 lysine 4 in Arabidopsis thaliana. Genome Biology. 2009;10:R62.
1311 132. Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, et al. Integrative epigenomic mapping defines four
1312 main chromatin states in Arabidopsis. EMBO J. 2011;30:1928–38.
1313 133. Ding Y, Fromm M, Avramova Z. Multiple exposures to drought "train" transcriptional responses in Arabidopsis. Nat
1314 Commun. 2012;3:740.
1315 134. Pratx L, Wendering P, Kappel C, Nikoloski Z, Bäurle I. Histone retention preserves epigenetic marks during heat stress-
1316 induced transcriptional memory in plants. EMBO J. 2023;42:e113595.
1317 135. Cao HX, Vu GTH, Wang W, Messing J, Schubert I. Chromatin organisation in duckweed interphase nuclei in relation
1318 to the nuclear DNA content. Plant Biol. 2015;17:120–4.
1319 136. Ebbs ML, Bender J. Locus-Specific Control of DNA Methylation by the Arabidopsis SUVH5 Histone Methyltransferase.
1320 Plant Cell. 2006;18:1166–76.
1321 137. Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3
1322 methyltransferase. Nature. 2002;416:556–60.
1323 138. Jasencakova Z, Soppe WJJ, Meister A, Gernand D, Turner BM, Schubert I. Histone modifications in Arabidopsis– high
1324 methylation of H3 lysine 9 is dispensable for constitutive heterochromatin. Plant J. 2003;33:471–80.
1325 139. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, et al. Sorting of Small RNAs into Arabidopsis Argonaute Complexes Is
1326 Directed by the 5′ Terminal Nucleotide. Cell. 2008;133:116–27.
1327 140. McCue AD, Panda K, Nuthikattu S, Choudury SG, Thomas EN, Slotkin RK. ARGONAUTE 6 bridges transposable
1328 element mRNA-derived siRNAs to the establishment of DNA methylation. The EMBO journal. 2015;34:20–35.
1329 141. Panda K, McCue AD, Slotkin RK. Arabidopsis RNA Polymerase IV generates 21–22 nucleotide small RNAs that can
1330 participate in RNA-directed DNA methylation and may regulate genes. Philos Trans R Soc B. 2020;375:20190417.
1331 142. Sigman MJ, Panda K, Kirchner R, McLain LL, Payne H, Peasari JR, et al. An siRNA-guided ARGONAUTE protein
1332 directs RNA polymerase V to initiate DNA methylation. Nat Plants. 2021;7:1461–74.
1333 143. Keith B, Chua N-H. Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco.
1334 EMBO J. 1986;5:2419–25.
1335 144. Goodall GJ, Filipowicz W. Different effects of intron nucleotide composition and secondary structure on pre-mRNA
1336 splicing in monocot and dicot plants. EMBO J. 1991;10:2635–44.
1337 145. Lou H, McCullough AJ, Schuler MA. Expression of maize Adh1 intron mutants in tobacco nuclei. Plant J. 1993;3:393–
1338 403.

1339 146. Jullien PE, Grob S, Marchais A, Pumplin N, Chevalier C, Bonnet DMV, et al. Functional characterization of Arabidopsis
1340 ARGONAUTE 3 in reproductive tissues. Plant J. 2020;103:1796–809.
1341 147. Fang X, Qi Y. RNAi in Plants: An Argonaute-Centered View. Plant Cell. 2016;28:272–85.
1342 148. Slotkin RK, Vaughn M, Borges F, TanurdZiC M, Becker JoD, FeijO JA, et al. Epigenetic Reprogramming and Small
1343 RNA Silencing of Transposable Elements in Pollen. Cell. 2009;136:461–72.
1344 149. Lee SC, Ernst E, Berube B, Borges F, Parent J-S, Ledon P, et al. Arabidopsis retrotransposon virus-like particles and
1345 their regulation by epigenetically activated small RNA. Genome Research. 2020;30:576–88.
1346 150. Oberlin S, Rajeswaran R, Trasser M, Barragán-Borrero V, Schon MA, Plotnikova A, et al. Innate, translation-dependent
1347 silencing of an invasive transposon in Arabidopsis. Embo Rep. 2022;23:e53400.
1348 151. Xie Z, Allen E, Wilken A, Carrington JC. DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and
1349 vegetative phase change in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United
1350 States of America. 2005;102:12984–9.
1351 152. Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, et al. Dissecting Arabidopsis thaliana DICER function
1352 in small RNA processing, gene silencing and DNA methylation patterning. Nature Genetics. 2006;38:721–5.
1353 153. Wang T, Deng Z, Zhang X, Wang H, Wang Y, Liu X, et al. Tomato DCL2b is required for the biosynthesis of 22-nt small
1354 RNAs, the resulting secondary siRNAs, and the host defense against ToMV. Hortic Res. 2018;5:62.
1355 154. Jia J, Ji R, Li Z, Yu Y, Nakano M, Long Y, et al. Soybean DICER-LIKE2 Regulates Seed Coat Color via Production of
1356 Primary 22-Nucleotide Small Interfering RNAs from Long Inverted Repeats. Plant Cell. 2020;32:3662–73.
1357 155. Hamilton AJ, Voinnet O, Chappell L, Baulcombe DC. Two classes of short interfering RNA in RNA silencing. Embo
1358 Journal. 2002;21:4671–9.
1359 156. Bertazzon N, Raiola A, Castiglioni C, Gardiman M, Angelini E, Borgo M, et al. Transient silencing of the grapevine gene
1360 VvPGIP1 by agroinfiltration with a construct for RNA interference. Plant Cell Rep. 2012;31:133–43.
1361 157. Andrieu A, Breitler JC, Siré C, Meynard D, Gantet P, Guiderdoni E. An in planta, Agrobacterium-mediated transient
1362 gene expression method for inducing gene silencing in rice (Oryza sativa L.) leaves. Rice. 2012;5:23.
1363 158. Sun S, Kang X-P, Xing X-J, Chen Z-F, Zheng S-W, Xing G-M. Transient expression of siRNA targeted against the
1364 TYLCV AV1, AC1 and AC3 genes for high resistance in tomato. Sci Hortic. 2014;179:321–7.
1365 159. Jay F, Brioudes F, Voinnet O. A contemporary reassessment of the enhanced transient expression system based on
1366 the tombusviral silencing suppressor protein P19. Plant J. 2023;113:186–204.
1367 160. Yang G-L, Fang Y, Xu Y-L, Tan L, Li Q, Liu Y, et al. Frond transformation system mediated by Agrobacterium
1368 tumefaciens for Lemna minor. Plant Mol Biol. 2018;98:319–31.
1369 161. He Y, Zhang T, Sun H, Zhan H, Zhao Y. A reporter for noninvasively monitoring gene expression and plant
1370 transformation. Hortic Res. 2020;7:152.
1371 162. Šurbanovski N, Brilli M, Moser M, Si-Ammour A. A highly specific microRNA-mediated mechanism silences LTR
1372 retrotransposons of strawberry. The Plant Journal. 2016;85:70–82.
1373 163. Benachenhou F, Sperber GO, Bongcam-Rudloff E, Andersson G, Boeke JD, Blomberg J. Conserved structure and
1374 inferred evolutionary history of long terminal repeats (LTRs). Mob DNA. 2013;4:5.
1375 164. Michael TP, Ernst E, Hartwick N, Chu P, Bryant D, Gilbert S, et al. Genome and time-of-day transcriptome of Wolffia
1376 australiana link morphological minimization with gene loss and less growth control. Genome Res.
1377 2020;31:gr.266429.120.
1378 165. Abramson BW, Novotny M, Hartwick NT, Colt K, Aevermann BD, Scheuermann RH, et al. The genome and preliminary
1379 single-nuclei transcriptome of Lemna minuta reveals mechanisms of invasiveness. Plant Physiol. 2021;188:879–97.
1380 166. Fang Y, Tian X, Jin Y, Du A, Ding Y, Liao Z, et al. Duckweed evolution: from land back to water. bioRxiv.
1381 2023;2023.03.22.533731.
1382 167. Liu J, He Z. Small DNA Methylation, Big Player in Plant Abiotic Stress Responses and Memory. Front Plant Sci.
1383 2020;11:595603.
1384 168. Urquiaga MC de O, Thiebaut F, Hemerly AS, Ferreira PCG. From Trash to Luxury: The Potential Role of Plant LncRNA
1385 in DNA Methylation During Abiotic Stress. Front Plant Sci. 2021;11:603246.
1386 169. Arkhipova I, Meselson M. Transposable elements in sexual and ancient asexual taxa. Proc Natl Acad Sci.
1387 2000;97:14473–7.
1388 170. Wright S, Finnegan D. Genome evolution: Sex and the transposable element. Curr Biol. 2001;11:R296–9.

1389 171. Bast J, Jaron KS, Schuseil D, Roze D, Schwander T. Asexual reproduction reduces transposable element load in
1390 experimental yeast populations. eLife. 2019;8:e48548.
1391 172. Dechaud C, Volff J-N, Schartl M, Naville M. Sex and the TEs: transposable elements in sexual development and
1392 function in animals. Mob DNA. 2019;10:42.
1393 173. Nguyen V, Gutzat R. Epigenetic regulation in the shoot apical meristem. Curr Opin Plant Biol. 2022;69:102267.
1394 174. Czech B, Munafò M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E, et al. piRNA-Guided Genome Defense: From
1395 Biogenesis to Silencing. Annu Rev Genet. 2018;52:131–57.
1396 175. Huff JT, Zilberman D, Roy SW. Mechanism for DNA transposons to generate introns on genomic scales. Nature.
1397 2016;538:533–6.
1398 176. Chénais B, Caruso A, Hiard S, Casse N. The impact of transposable elements on eukaryotic genomes: From genome
1399 size increase to genetic adaptation to stressful environments. Gene. 2012;509:7–15.
1400 177. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, et al. Ten things you should know about
1401 transposable elements. Genome Biol. 2018;19:199.
1402 178. Gaillochet C, Lohmann JU. The never-ending story: from pluripotency to plant developmental plasticity. Development.
1403 2015;142:2237–49.
1404 179. Sree KS, Maheshwari SC, Boka K, Khurana JP, Keresztes Á, Appenroth K-J. The duckweed Wolffia microscopica: A
1405 unique aquatic monocot. Flora - Morphol, Distrib, Funct Ecol Plants. 2015;210:31–9.
1406 180. Yang J, Zhao X, Li G, Hu S, Hou H. Frond architecture of the rootless duckweed Wolffia globosa. BMC Plant Biol.
1407 2021;21:387.
1408 181. Li F, Yang J-J, Sun Z-Y, Wang L, Qi L-Y, A S, et al. Plant-on-chip: Core morphogenesis processes in the tiny plant
1409 Wolffia australiana. PNAS Nexus. 2023;2:pgad141.
1410 182. Gutzat R, Rembart K, Nussbaumer T, Hofmann F, Pisupati R, Bradamante G, et al. Arabidopsis shoot stem cells
1411 display dynamic transcription and DNA methylation patterns. EMBO J. 2020;39:e103667.
1412 183. Ho EKH, Bartkowska M, Wright SI, Agrawal AF. Population genomics of the facultatively asexual duckweed Spirodela
1413 polyrhiza. N Phytol. 2019;224:1361–71.
1414 184. Wang Y, Duchen P, Chávez A, Sree SK, Appenroth KJ, Zhao H, et al. Population genomics and epigenomics provide
1415 insights into the evolution of facultative asexuality in plants. bioRxiv. 2023;2023.07.18.549461.
1416 185. Fusaro AF, Matthew L, Smith NA, Curtin SJ, Dedic-Hagan J, Ellacott GA, et al. RNA interference-inducing hairpin RNAs
1417 in plants act through the viral defence pathway. EMBO Rep. 2006;7:1168–75.
1418 186. Parent J, Bouteiller N, Elmayan T, Vaucheret H. Respective contributions of Arabidopsis DCL2 and DCL4 to RNA
1419 silencing. Plant J. 2015;81:223–32.
1420 187. Nielsen CPS, Arribas-Hernández L, Han L, Andersen SU, Pumplin N, Brodersen P. Evidence for an RNAi-independent
1421 role of DICER-LIKE2 in conferring growth inhibition and basal antiviral resistance. The Plant Cell. 2024; koae0676.
1422 188. Nielsen CPS, Han L, Arribas-Hernández L, Karelina D, Petersen M, Brodersen P. Sensing of viral RNA in plants via a
1423 DICER-LIKE Ribonuclease. bioRxiv. 2023;2023.01.10.523395.
1424 189. Chen H-M, Chen L-T, Patel K, Li Y-H, Baulcombe DC, Wu S-H. 22-Nucleotide RNAs trigger secondary siRNA
1425 biogenesis in plants. Proceedings of the National Academy of Sciences. 2010;107:15269–74.
1426 190. Zhang X, Singh J, Li D, Qu F. Temperature-Dependent Survival of Turnip Crinkle Virus-Infected Arabidopsis Plants
1427 Relies on an RNA Silencing-Based Defense That Requires DCL2, AGO2, and HEN1. Journal of Virology.
1428 2012;86:6847–54.
1429 191. Taochy C, Gursanscky NR, Cao J, Fletcher SJ, Dressel U, Mitter N, et al. A Genetic Screen for Impaired Systemic
1430 RNAi Highlights the Crucial Role of DICER-LIKE 2 . Plant Physiol. 2017;175:1424–37.
1431 192. Chen W, Zhang X, Fan Y, Li B, Ryabov E, Shi N, et al. A Genetic Network for Systemic RNA Silencing in Plants . Plant
1432 Physiol. 2018;176:2700–19.
1433 193. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. The MaSuRCA genome assembler. Bioinformatics.
1434 2013;29:2669–77.
1435 194. Zimin AV, Salzberg SL. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies.
1436 PLoS Comput Biol. 2020;16:e1007981.
1437 195. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A
1438 MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297–303.

196. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Auwera GAV der, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv. 2018;201178.
197. Shumate A, Salzberg SL. Liftoff: accurate mapping of gene annotations. Bioinformatics. 2021;37:btaa1016.
198. Zhang R-G, Li G-Y, Wang X-L, Dainat J, Wang Z-X, Ou S, et al. TEsorter: An accurate and fast method to classify LTR-retrotransposons in plant genomes. Hortic Res. 2022;9:uhac017.
199. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34:525–7.
200. Li H. New strategies to improve minimap2 alignment accuracy. Bioinformatics. 2021;37:btab705-.
201. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;27:1571–2.
202. Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. Nature. 2015;523:212–6.
203. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26:841–2.
204. Huang X, Zhang S, Li K, Thimmapuram J, Xie S, Wren J. ViewBS: a powerful toolkit for visualization of high-throughput bisulfite sequencing data. Bioinformatics. 2018;34:708–9.
205. Hayashi R, Schnabl J, Handler D, Mohn F, Ameres SL, Brennecke J. Genetic and mechanistic diversity of piRNA 3′-end formation. Nature. 2016;539:588–92.
206. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
207. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature Methods. 2012;9:357–9.
208. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44:W160–5.
209. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.
210. Sambrook J, Russell DW. Molecular Cloning. CSHL Press; 2001.
211. Felippes FF de, Weigel D. Transient assays for the analysis of miRNA processing and function. Methods in molecular biology (Clifton, NJ). 2010;592:255–64.
212. Pall GS, Hamilton AJ. Improved northern blot method for enhanced detection of small RNA. Nature Protocols. 2008;3:1077–84.
213. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize implements and enhances circular visualization in R. Bioinformatics. 2014;30:2811–2.
214. Wickham H. Ggplot2: Elegant graphics for data analysis. Springer; 2016.
215. Oberlin S, Sarazin A, Chevalier C, Voinnet O, Ordóñez AM. A genome-wide transcriptome and translatome analysis of Arabidopsis transposons identifies a unique and conserved genome expression strategy for Ty1/Copia retroelements. Genome Research. 2017;27:1549–62.