Atypical epigenetic and small RNA control of degenerated transposons and their fragments in clonally reproducing *Spirodela polyrhiza*

Rodolphe Dombey[1], Daniel Buendía-Ávila[1,2*], Verónica Barragán-Borrero[1], Laura Diezma-Navas[1], 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.

* These two authors contributed equally.

** 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

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. Here, 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 marked by the heterochromatin associated H3K9me1 and H3K27me1 modifications. 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. The data suggest that, potentially as adaptation to vegetative propagation, RdDM extent, silencing 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 in flowering plants. These discoveries highlight the diversity of silencing mechanisms that exist in plants and the importance of using disparate model species to discover these mechanisms.

Introduction

Transposable elements (TEs), or transposons, are mobile genetic elements which can change location and generate new copies in their host genomes. As a result, TEs populate eukaryotic genomes and can contribute to a significant fraction of their nuclear DNA (Wells and Feschotte 2020). Owing to their mobile nature, TEs have an important role in evolution and genetic innovation, but also pose a threat to the genome integrity (Biémont and Vieira 2006; Feschotte 2008). To control TEs, various mechanisms have evolved to prevent or reduce their mobility.

In eukaryotes, transcriptional gene silencing (TGS) of TEs relies on the deposition of chromatin modifications leading to the formation of compact chromatin states, collectively known as heterochromatin, incompatible with canonical Pol II transcription (Allshire and Madhani 2018). Heterochromatin is generally characterized by repressive epigenetic marks, prominently DNA methylation (5'methylcytosine; 5mC), histone modifications such as histone H3 lysine-9 mono- and dimethylation (H3K9me1/H3K9me2, constitutive heterochromatin) or histone H3 lysine-27 triple methylation H3K27me3 (facultative or developmentally regulated heterochromatin), and the presence of specific histone variants (Henikoff and Smith 2015; Jamge et al. 2023). Upon silencing, TEs accumulate mutations and degenerate into non-autonomous TEs and fragmented remnants or relics (Blumenstiel 2019). Despite their inability to mobilize, degenerated TEs provide alternative DNA and RNA regulatory sequences (e.g., transcription factor binding sites, splicing information, transcriptional start or termination sites) or induce recombination due to their repetitiveness (Ito et al. 2011; Kim and Zilberman 2014; Zervudacki et al. 2018; Sammarco et al. 2022; Ilık et al. 2024). Therefore, maintaining epigenetic control is important even for degenerated TEs, and in consequence, TE abundance and location largely shapes the epigenetic landscape of genomes (Houben et al. 2003; Seymour et al. 2014; Sigman and Slotkin 2016; Wyler et al. 2020; Klein and Anderson 2022).

68 Another conserved mechanism to suppress TEs and maintain their epigenetic
69 silencing depends on ~18-30-nt long small RNAs (sRNAs) and members of the ARGONAUTE
70 (AGO) protein family. Loaded into AGOs, TGS-associated sRNAs guide chromatin modifying
71 complexes to TEs through sequence complementarity (Moazed 2009; Malone and Hannon
72 2009). In flowering plants (angiosperms) TE silencing is mostly associated with DNA
73 methylation and H3K9me2, and the mechanism by which sRNAs guide chromatin
74 modifications is known as RNA-directed DNA methylation (RdDM). Best studied in
75 *Arabidopsis thaliana,* during RdDM, H3K9me2 and DNA methylation recruit the plant-specific
76 RNA Polymerases IV and V (Pol IV, V). Pol IV transcripts are processed into 24-nt siRNAs
77 which, loaded into members of the AGO4/6 clade, target Pol V transcripts to guide the
78 sequence-specific deposition of DNA methylation in all cytosine contexts (CG, CHG and CHH,
79 where H is any nucleotide but G). Once stablished, DNA methylation can reinforce the
80 deposition of H3K9me2 through the action of SUPRESSOR OF VARIEGATION 3-9
81 HOMOLOGUE (SUVH) family H3K9 methyltransferases, which can read DNA methylation. In
82 turn, CHROMO-METHYL TRANSFERASE 3 and 2 (CMT3, CMT2) can bind to H3K9me2 to
83 mediate CHG and CHH methylation respectively. This results in a self-reinforcing loop
84 between non-CG methylation and H3K9me2 independent of siRNAs. Symmetrical CG
85 methylation is maintained across cell divisions through the action of METHYL-
86 TRANSFERASE 1 (MET1), recruited to hemi-methylated CG sites after DNA replication by
87 VARIANT IN METHYLATION (VIM) proteins (reviewed in (Erdmann and Picard 2020; Tirot
88 et al. 2021)). Loss of RdDM, 5mC or H3K9me2 results in the transcriptional reactivation of
89 TEs (Mirouze et al. 2009; Tsukahara et al. 2010; Stroud et al. 2014; He et al. 2021;
90 Osakabe et al. 2021). Therefore, DNA methylation, together with H3K9me2 and in
91 cooperation with sRNAs, keep TEs silenced.
92 Loss of TE control in somatic tissue is likely less deleterious as long as it does not
93 affect the germline or, in plants, stem cells in the meristems or gamete precursor cells giving
94 rise to the next generation, in which stability of the genome is highly relevant. In accordance,

95 small RNA-guided TGS pathways are prominently active in reproductive tissues and silence
96 active TEs encountered during fertilization (Bourc'his and Voinnet 2010; Marí-Ordóñez et
97 al. 2013; Parhad and Theurkauf 2019). However, unlike in metazoans, where such pathways
98 are predominantly active in the gonads, in most flowering plants RdDM is not constrained to
99 reproductive organs. 24-nt siRNAs are also present in vegetative, somatic organs such as
100 leaves or roots across all developmental stages (Havecker et al. 2010; Patel et al. 2021;
101 Zhou et al. 2022). With interesting exceptions: the aquatic dicot carnivorous plant *Utricularia*
102 *gibba* (L.), and the monocot *Spirodela polyrhiza* (L.) Scheid. Both contain much less 24-nt
103 siRNAs and display reduced genome-wide levels of DNA methylation in vegetative tissues
104 compared to other angiosperms (Michael et al. 2017; Fourounjian et al. 2019; Cervantes-
105 Pérez et al. 2021). While in the *U. gibba* such features can be a consequence of low TE
106 content as repeats represent less than 3% of genome (Ibarra-Laclette et al. 2013), the TE
107 content of the *Spirodela* genome is similar to that of *Arabidopsis* (~20%) (Wang et al. 2014).
108 Therefore, reduced genomic TE load is unlikely to be the cause behind the low levels of DNA
109 methylation and small RNAs in *Spirodela*. Furthermore, in *Spirodela*, several sRNA silencing
110 components including some involved in RdDM have been shown to be missing, low or not
111 expressed. In addition, CMT2, which maintains CHH methylation, is lost in duckweeds (An et
112 al. 2019; Ernst et al. 2023; Harkess et al. 2024).

113 The *Spirodela* genus is the most basal of the *Lemnaceae* family (order Alismatales), a
114 monophyletic group of free-floating freshwater monocot plants, commonly known as
115 duckweeds, composed of 5 genera (*Spirodela*, *Landoltia*, *Lemna*, *Wolfiella* and *Wolffia*) (Bog
116 et al. 2019; Tippery et al. 2021). Comprised of a leaf-like structure known as a frond, ranging
117 from ~2 cm to 1 mm, duckweeds display an evolutionary reduction and simplification of their
118 body plan compared to most angiosperms, including vestigiality and loss of organs such as
119 roots and vasculature (Landolt 1986; Lemon and Posluszny 2000; Lam and Michael 2022;
120 Ware et al. 2023). Although flowering and seed production has been reported, it is rare in

121 most species (Landolt 1986; Pieterse 2013; Fourounjian et al. 2021). Duckweeds mostly
122 propagate through rapid clonal reproduction (doubling of individuals ~1-5 days), known as
123 vegetative budding: new fronds originate from budding pockets where meristems containing
124 stem cells are located (Landolt 1986; Ziegler et al. 2015). Mother fronds (MF) continuously
125 give rise to daughter fronds (DF), which carry grand-daughter fronds (GDF) in their budding
126 pockets even before detaching from the mother frond (Landolt 1986; Lemon and Posluszny
127 2000). This extremely efficient amplification of somatic tissue, omitting regular sexual
128 propagation, and the free-floating way of living, make *Spirodela* a unique model to explore if
129 the differences in epigenetic parameters are connected with these differences in lifestyle
130 compared to other angiosperms. Thus, the aim of this study is to investigate the TE silencing
131 landscape in *Spirodela polyrhiza* by profiling its genetic composition of silencing pathways,
132 small RNAs and epigenetic modifications commonly associated with TE control in
133 angiosperms.
134

135 ## Results

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

137 The genomes of several *Spirodela polyrhiza* accessions (clones) have been
138 sequenced (Wang et al. 2014; Michael et al. 2017; Xu et al. 2019), including a high-
139 resolution reference genome assembly of its 20 chromosomes from the *S. polyrhiza* 9509 and
140 7498 clones (Sp9509, Sp7498) (Michael et al. 2017; Hoang et al. 2018; Harkess et al.
141 2021). For our study, we retrieved Sp9509 from the Landolt duckweed collection (Zurich,
142 Switzerland), now managed at the Institute of Agriculture Biology and Biotechnology (IABB,
143 CNR, Milan, Italy) (Morello et al. 2024) . We performed 88× coverage short-read Illumina

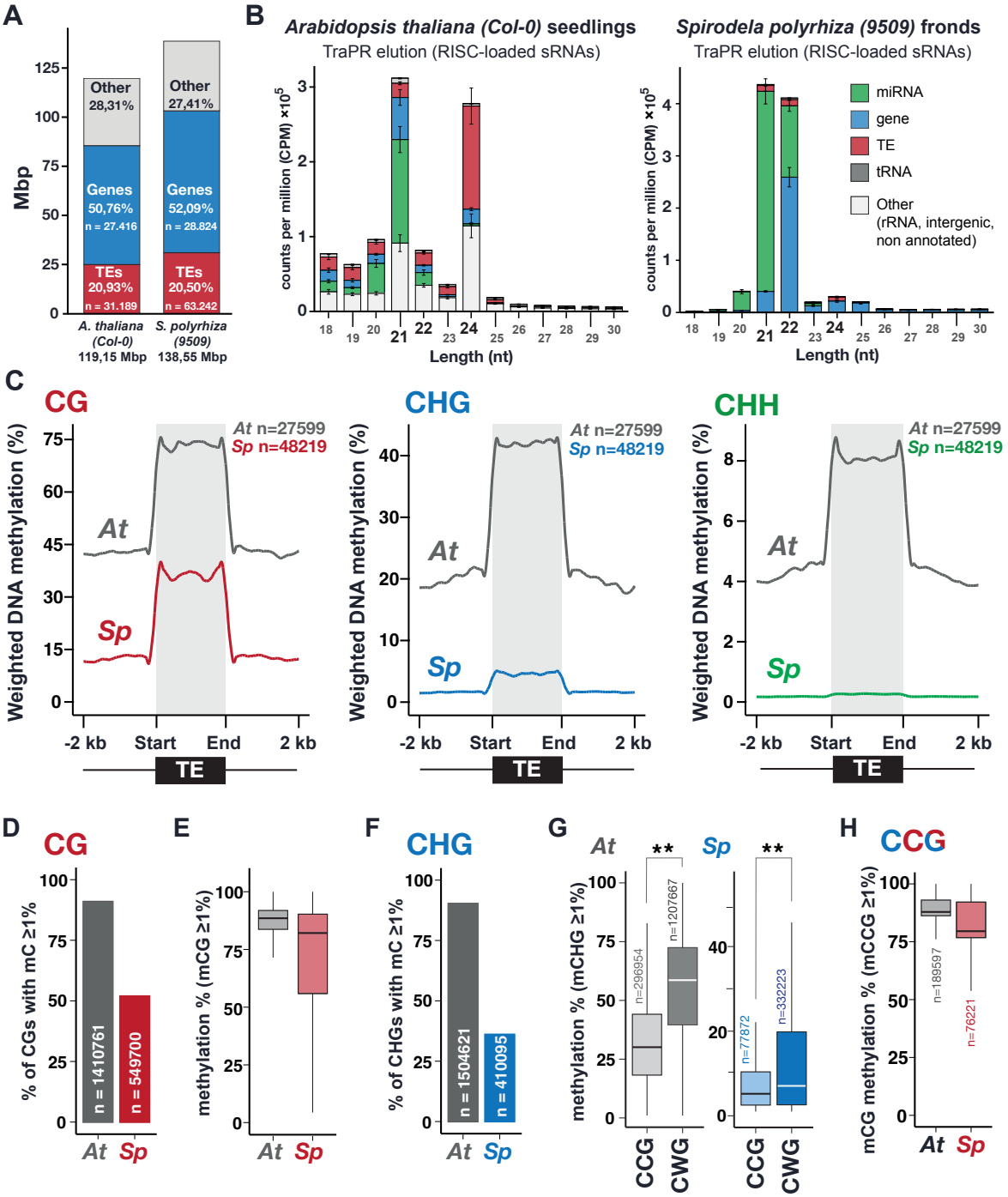

Figure 1: *Spirodela polyrhiza* general developmental, small RNA and DNA methylation patterns. (A) Genome size, number and genomic occupancy of genes and TEs in *Spirodela* and *Arabidopsis*. **(B)** Size distribution and genomic feature annotation of TraPR-purified small RNA from *Arabidopsis* and *Spirodela*. Error bars represent standard deviation of the mean of three biological replicates. **(C)** Metaplots of averaged weighted DNA methylation in all three contexts (CG, CHG and CHH, where H is any but G) over transposons (TEs) (≥100 bp) and 2 kb flanking regions in *Arabidopsis* and *Spirodela*. **(D)** number and percentage of transposon CG sites displaying methylation levels equal or above 1% in *Arabidopsis* and *Spirodela*. **(E)** DNA methylation levels at CG sites from (D). **(F)** number and percentage of transposon CHG sites displaying methylation levels equal or above 1% in *Arabidopsis* and *Spirodela*. **(G)** 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 $<2.22 \times 10^{-16}$ (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*.

144

145 genome sequencing to account for potential genetic differences due to independent clonal
146 propagation over thousands generations between this accession and that of the Rutgers
147 Duckweed Stock Cooperative (RDSC, Rutgers University, NJ, USA) used to generate the
148 reference genome (Hoang et al. 2018). The Landolt Sp9509 genome showed ~163k indels
149 and ~512k SNPs (~0.4% of the *Spirodela* genome) compared to the Sp9509 reference
150 (**Supplemental Fig. S1**). Given our special interest on TEs, we performed a *de novo* TE
151 annotation using a combination of EDTA and RepeatModeler2 to identify repeats that were
152 then further classified using DeepTE (Ou et al. 2019; Flynn et al. 2020; Yan et al. 2020).

153 To explore the TE silencing landscape during *Spirodela* asexual clonal reproduction,
154 we used *Arabidopsis* as a comparative system to investigate *Spirodela* TE-derived sRNAs,
155 TE epigenetic marks and the genetic composition and expression of silencing pathways.
156 Although TE regulatory mechanisms might differ between the two species given their
157 evolutionary distance (**Supplemental Fig. S2**), we choose *Arabidopsis* as its silencing
158 pathways and epigenetic mechanisms are well characterized by molecular and genetic
159 means. Furthermore, the two species display similar genome size and TE content (**Fig. 1A**).
160

161 **AGO-loaded small RNA and DNA methylation patterns in *Spirodela* imply low RdDM**
162 **activity in *Spirodela*.**

163 Given that 24-nt sRNAs are considered one of the hallmarks of RdDM, we first
164 compared the sRNA profiles of *Spirodela* confluent cultures (including fronds in all
165 developmental stages) to those of 10 day-old *Arabidopsis* seedlings as the closest equivalent
166 developmental stage. As sRNAs on the 21-to-24-nt range were very low abundant in profiles
167 generated from total RNA (**Supplemental Fig. S3**), we enriched our small RNA libraries by
168 investigating AGO-loaded small RNAs using TraPR (Grentzinger et al. 2020). In contrast to
169 *Arabidopsis*, where AGO-loaded TE-derived 24-nt siRNAs are highly abundant, *Spirodela*
170 TraPR profile displays very low levels of 24-nt siRNAs and it is dominated by 21-22-nt sRNAs
171 (**Fig. 1B, Supplemental Fig. S3**), predominantly contributed by regulatory microRNAs
172 (miRNA). The lack of 24-nt siRNAs in *Spirodela* TraPR profile is more pronounced than their

173 loss in *Arabidopsis* RdDM siRNA biogenesis defective Pol IV mutant (*nrpd1*) (**Supplemental**
174 **Fig. S3**) (Teixeira et al. 2009; Nuthikattu et al. 2013; Creasey et al. 2014) Hence, low levels
175 of AGO-loaded 24-nt siRNAs in *Spirodela* are compatible with the reduced or lack of RdDM
176 activity at TEs during its asexual clonal growth.

177 Given the connection between sRNAs and DNA methylation, we next investigated
178 *Spirodela* DNA methylation patterns. Methylome analysis by Enzymatic Methyl-seq (EM-seq)
179 (Feng et al. 2020) showed similar results to those previously published using bisulfite
180 sequencing (Michael et al. 2017). TE 5mC in all contexts is lower in *Spirodela* compared to
181 *Arabidopsis*, with very low mCHG levels and hardly any mCHH (**Fig. 1C**). This is in agreement
182 with the partial dependency of mCHH on 24-nt siRNAs. In addition, *Spirodela* lacks gene body
183 CG methylation (gbM) (Michael et al. 2017; Harkess et al. 2024) (**Supplemental Fig. S4A**).
184 Although the role of gbM is unclear, it generally associates with highly expressed genes. In
185 contrast, mCG near the transcription start site (TSS) is associated with gene silencing
186 (Niederhuth et al. 2016; Bewick and Schmitz 2017). Although expressed genes do not
187 display gbM, like in *Arabidopsis*, mCG around the TSS is higher in non-expressed genes of
188 *Spirodela* (**Supplemental Fig. S4B,C**), indicating that mCG in gene regulatory sequences
189 influences transcription also in *Spirodela*.
190

191 **CG and CHG methylation signatures are compatible with functional maintenance of**
192 **DNA methylation independently of small RNAs.**

193 Once established, DNA methylation can be further maintained independently of small
194 RNAs through several pathways depending on the sequence context (Mathieu et al. 2007;
195 Cokus et al. 2008; Reinders et al. 2009) In *Spirodela*, the ratio of mCG sites at TEs (51,9%)
196 is nearly half of that in *Arabidopsis* (90,9%) (**Fig. 1D**). However, at CG sites with 5mC levels
197 equal or higher than 1%, mCG is similarly stable as in *Arabidopsis*, displaying median

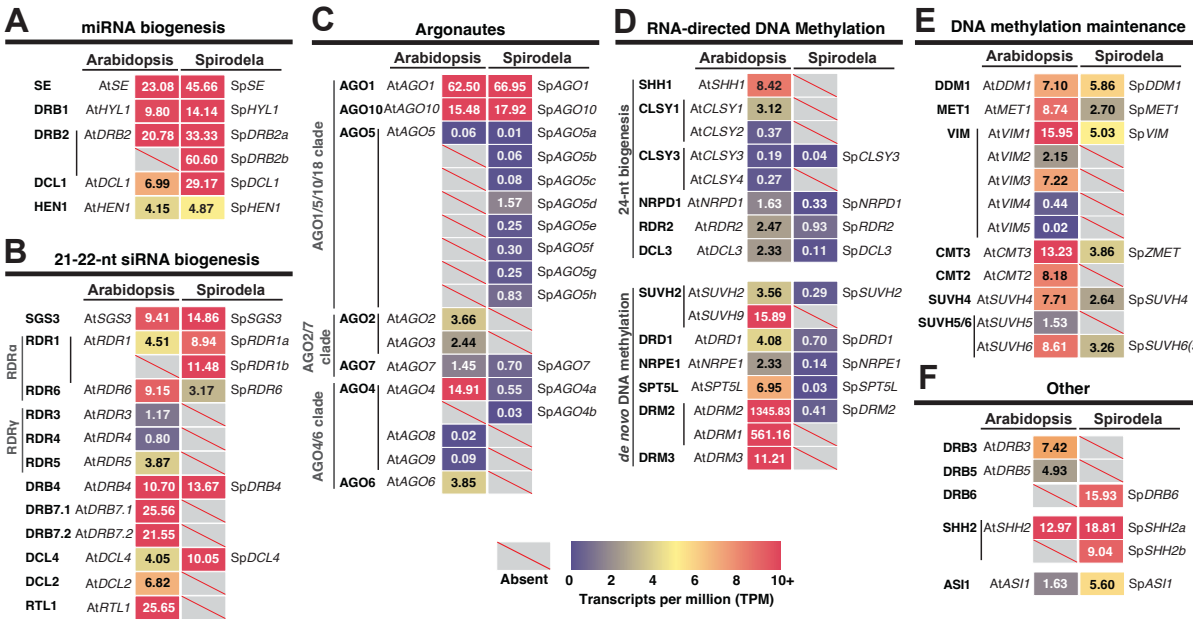

Figure 2: Genetic composition and expression of silencing pathways in *Spirodela*. (A-F) Expression in transcripts per million (TPM) of silencing components in *Arabidopsis* and their identified orthologues and paralogues in the *Spirodela polyrhiza* #9509 (Sp9509) genome. Gene IDs for both *Arabidopsis* and *Spirodela* can be found in Supplemental Figure S5.

methylation levels above 80% (**Fig. 1E**), although a higher variation was observed for *Spirodela* than *Arabidopsis*. Hence, the reduced global mCG level at TEs in *Spirodela* is the result of less individual CG sites being methylated.

As for mCG, the number of mCHG sites at TEs is much lower in *Spirodela* (36,1%) than in *Arabidopsis* (90,3%) (**Fig. 1F**). CHG sites can be further divided into CCG and CWG (CAG/CTG) sub-contexts. In *Arabidopsis*, mCWG is higher than mCCG (**Fig. 1G**) due to the combined differential affinities of SUVH4/5/6 and CMT3 to read and methylate individual cytosine sequence contexts (Cokus et al. 2008; Gouil and Baulcombe 2016; Li et al. 2018; Fang et al. 2022). This applies also to *Spirodela*, although methylation in both sub-contexts is lower than in *Arabidopsis* (**Fig. 1G**). Furthermore, maintenance in the mCCG sub-context relies additionally on MET1. Being both CG and CHG contexts, methylated CCG sites are generally observed as mCmCG due to the outer cytosine methylation (mCCG) being largely dependent on the internal CG methylation (CmCG) mediated by MET1 (Stroud et al. 2013; Yaari et al. 2015; Gouil and Baulcombe 2016; Li et al. 2018). Consequently, mCCG sites in *Arabidopsis* display a high internal CG methylation level, a pattern also observed in *Spirodela*

214 (**Fig. 1H**). The similar mCHG patterns between the two species indicate that the responsible
215 enzymes, SUVHs and ZMET, the CMT3 homolog in monocots (Bewick et al. 2017) probably
216 share the same sequence preference albeit their activity seems to be lower than in
217 *Arabidopsis*. Nonetheless, given that CMT3/ZMET has been shown to be involved in gbM
218 (Niederhuth et al. 2016; Bewick et al. 2016; Wendte et al. 2019), lower mCHG and lack of
219 gbM in *Spirodela* might be related.

220 Altogether, our analysis indicates that mCG and mCHG maintenance mechanisms in
221 *Spirodela* might operate in a similar fashion as in *Arabidopsis*. However, while low mCG is a
222 consequence of reduced number of CG sites being methylated, low mCHG appears to be the
223 combined result of lower number of targeted CHG sites and decreased maintenance activity.
224

225 **Genetic composition and expression of silencing pathways partially account for sRNAs**
226 **and DNA methylation features in *Spirodela*.**

227 In addition to low signatures of RdDM in *Spirodela*, coding potential for several of its
228 components have been shown to be missing (An et al. 2019; Ernst et al. 2023; Harkess et
229 al. 2024). To gain further insights on why levels of 24-nt siRNA and DNA methylation are
230 reduced, we set to identify and examine the expression of known core components of
231 angiosperm silencing pathways in *S. polyrhiza* (**Fig. 2, Supplemental Fig. S5**). The
232 identification of orthologues and paralogues in the genome of *S. polyrhiza* 9509 was based
233 on the presence of conserved functional domains. Annotations were further curated based on
234 phylogeny trees built with proteins from several angiosperms and the prediction of known key
235 domains and their organization (**Supplemental Fig. S6-S18**). For gene expression, we
236 performed short-read transcriptome analysis in *Spirodela* and compared it to that of recently
237 published *Arabidopsis* seedlings (Li et al. 2023) (**Fig. 2**). We further performed long-read
238 cDNA-seq with PacBio full-length transcript Isoform-Sequencing (Iso-Seq) technology in
239 *Spirodela* to validate and eventually improve gene model annotations.

240 As expected from the prevalence of miRNAs in *Spirodela* TraPR analysis, conserved
241 plant miRNA biogenesis factors (Zhan and Meyers 2023) are expressed at similar levels as

242 those in *Arabidopsis* (**Fig. 2A**). Likewise, genes required for the production of regulatory 21-
243 nt siRNA trans-acting (ta)siRNA (Yoshikawa et al. 2005; Liu et al. 2020): SUPRESSOR OF
244 GENE SILENCING 3 (SGS3), RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) and
245 DICER-LIKE 4 (DCL4), are present and expressed (**Fig. 2B**). *Spirodela* lacks many
246 components, effectors and regulators of 21-22-nt siRNA pathways, such as RNASE THREE-
247 LIKE 1 (RTL1), DOUBLE-STRANDED RNA BIDING PROTEIN 7 (DRB7) (Shamandi et al.
248 2015; Tschopp et al. 2017; Montavon et al. 2017), and orthologues of the RDRγ clade highly
249 conserved in plants and fungi (Zong et al. 2009; Willmann et al. 2011) (**Fig. 2B;**
250 **Supplemental Fig. S6, S7**). Gene absences also include the antiviral factors DICER-LIKE 2
251 (DCL2) and ARGONAUTE 2 (AGO2) (Deleris et al. 2006; Garcia-Ruiz et al. 2010; Harvey
252 et al. 2011; Carbonell et al. 2012) (**Fig. 2B,C; Supplemental Fig. S8, S9**). In contrast, two
253 RDR1 and eight tandemly repeated copies of AGO5 are present (**Fig. 2B,C; Supplemental**
254 **Fig. S7, S9, S10**). Both have been implicated in virus exclusion and TE silencing in meristems
255 in *Arabidopsis* (Yu et al. 2003; Brosseau and Moffett 2015; Incarbone et al. 2023;
256 Bradamante et al. 2023) and, in monocots, AGO5 participates in anther development and TE
257 silencing during male gametogenesis (Zhai et al. 2015; Fei et al. 2016; Lee et al. 2021).

258 Regarding the RdDM pathway, *Spirodela* contains the minimal set of its components,
259 either with no or with a reduced number of paralogues compared to *Arabidopsis* (**Fig. 2C, D**).
260 In *Spirodela*, SAWADEE HOMEODOMAIN HOMOLGUE 1 (SHH1), a dual reader for
261 repressive H3K9me2 and unmethylated H3K4(me0), and the chromatin remodeler CLASSY
262 1/2 (CLSY1/2) are absent (**Fig. 2D; Supplemental Fig. S11, S12**). However, a single
263 orthologue of CLSY3 was found in *Spirodela* (**Fig. 2D; Supplemental Fig. S12**). In
264 *Arabidopsis*, the SHH1-CLSY1/2 module mainly recruits Pol IV to short TEs within gene-rich
265 regions, while CLSY3 and its paralogue CLSY4 recruit Pol IV to TE-rich pericentromeric

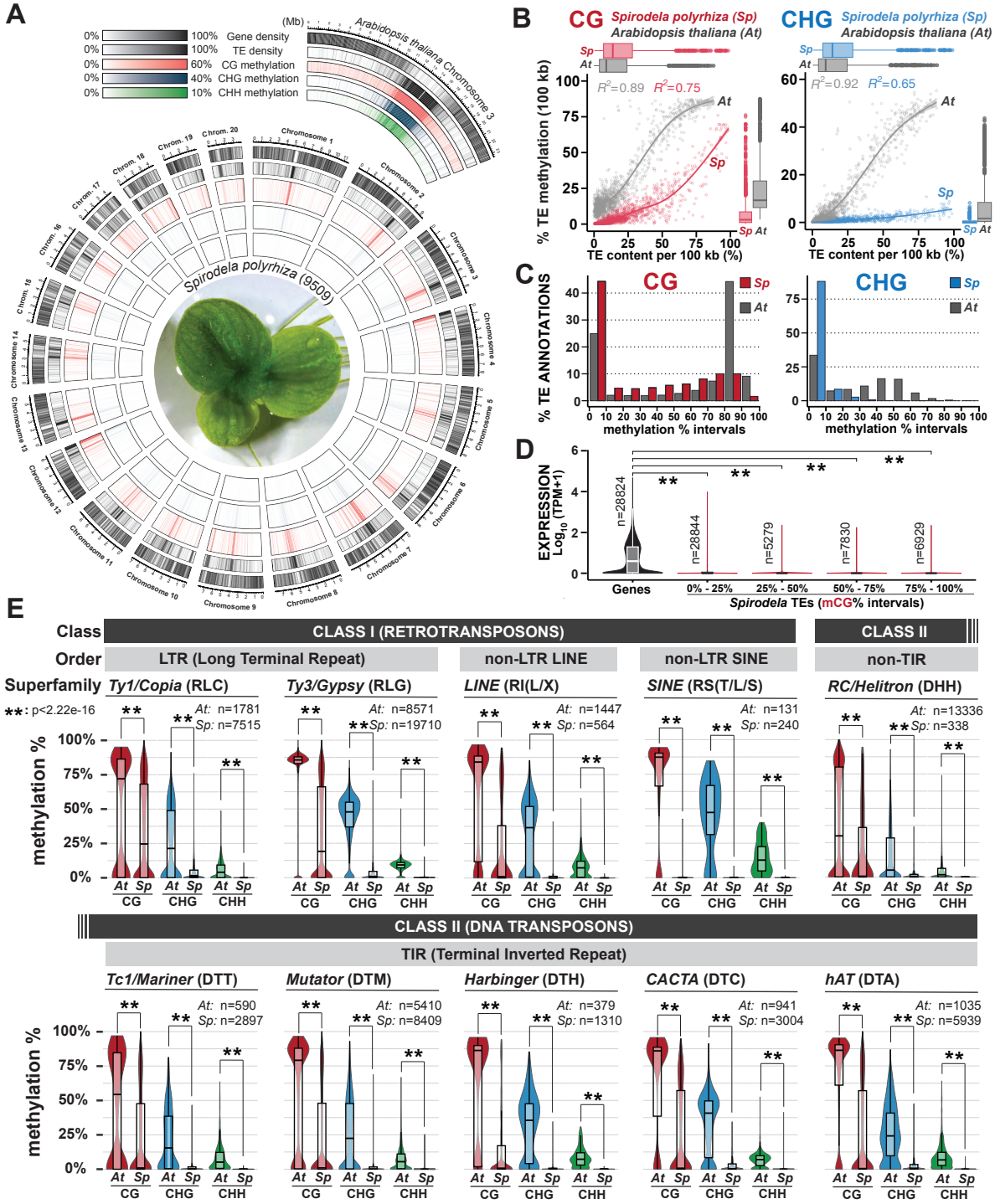

Figure 3: Distribution of DNA methylation in *Spirodela* chromosomes and transposable elements. **(A)** Circos plot of genome-wide distribution of DNA methylation densities by chromosome in *Spirodela*. From outside to inside: gene density, TE density, CG, CHG and CHH methylation. Mean DNA methylation, 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 × 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*.

266 regions in a SHH1-idependent but DNA methylation-dependent manner (Law et al.
267 2011; 2013; Zhou et al. 2018; 2022). However, *Spirodela* has two SHH2 orthologues (**Fig. 2;**
268 **Supplemental Fig. S11**). Part of a nucleosome positioning complex during transcription
269 initiation in *Arabidopsis* (Diego-Martin et al. 2022), in maize, SHH2 binds preferentially to
270 H3K9me1 *in vitro* and interacts *in vivo* with both Pol IV and V in maize, possibly connecting
271 RdDM to H3K9me1 installation (Haag et al. 2014; Wang et al. 2021).

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* (Duan et
274 al. 2014) (**Fig. 2D; Supplemental Fig. S9**). A single SUVH2 orthologue is present in
275 *Spirodela* (**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 *Arabidopsis*
277 (Johnson et al. 2008; 2014; Liu et al. 2014). Regarding the RdDM-associated DNA
278 methyltransferases, DOMAIN REARRANGED METHYLTRASFERASE 2 (DRM2) but no
279 DRM3, a conserved catalytically inactive paralogue that contributes to DRM2 activity (Chan
280 et al. 2004; Wierzbicki et al. 2009; Zhong et al. 2015; Liu et al. 2018; Fang et al. 2021),
281 orthologues were found (**Fig. 2D, Supplemental Fig. S14**). Additionally, although NRPE1
282 (Pol V) and its partner SUPPRESSOR OF TY INSERTION-LIKE (SPT5L) (a paralogue of the
283 general transcription elongation factor SPT5) are conserved in *Spirodela*; we noticed that both
284 contain fewer tryptophan/glycine (GW/WG) repeats than those of other angiosperms (**Fig. 2D;**
285 **Supplemental Fig. S15, S16**) (El-Shami et al. 2007; Bies-Etheve et al. 2009; Rowley et al.
286 2011; Huang et al. 2015; Trujillo et al. 2016). GW/WG repeats, or AGO-hooks, allow AGO4/6
287 binding to Pol V to scan its nascent transcripts for positive siRNA matches to recruit DRM2.
288 These, together with NRPD1, RDR2, DCL3, and the rest of conserved RdDM components,
289 might suffice for a functional pathway in *Spirodela* (**Fig. 2; Supplemental Fig. S7, S8, S12,**
290 **S15**). Nonetheless, their low expression levels (**Fig. 2C,D**) rationalize the low abundance of
291 AGO-bound 24-nt siRNAs and associated CHH methylation during clonal propagation in
292 *Spirodela* (**Fig. 1B,C**).

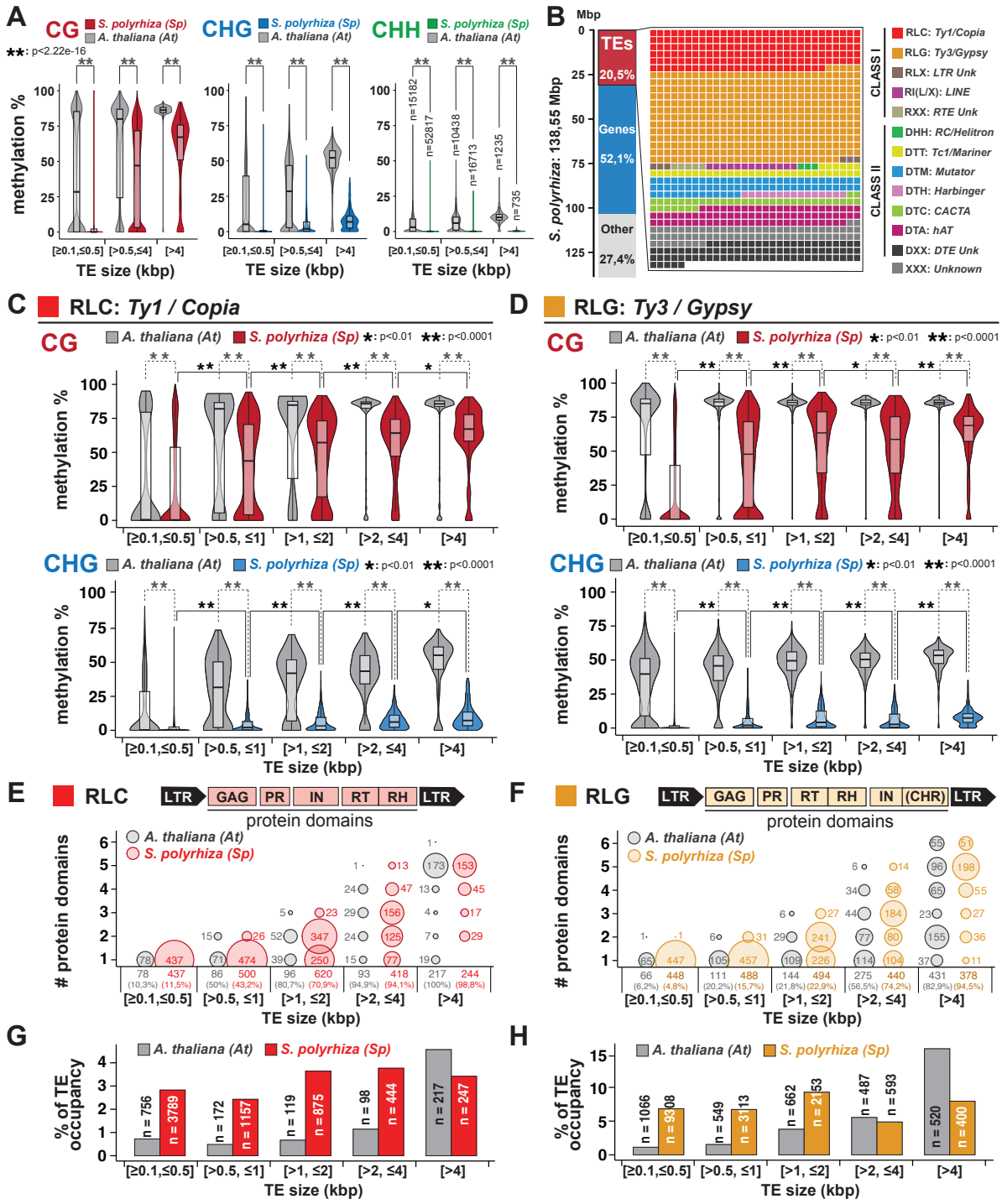

Figure 4: Length-dependent TE DNA methylation in *Spirodela polyrhiza*. (A), CHG and CHH methylation levels (%) across all TEs (≥100 bp) in *Arabidopsis* and *Spirodela* split by size into short (≥100 bp, <500 bp), intermediate (>500 bp, ≤ 4 kbp) and long (> 4kbp). n values for each size category in each genome is indicated within the plot for mCHH **(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*. n values are indicated in panels G-H **(E-F)** Presence and number of RCL (E) and RLG (F) TE protein domains, determined by TEsorter, across TE annotations of varying sizes in *Arabidopsis* and *Spirodela*. Bubble plots depict domain count distributions for TEs with at least one domain. Numbers above TE size category denotes TE annotations with at least one domain, with percentages shown in brackets relative to total TE annotations per size category (G-H). Bubble size is relative to the number of TEs (indicated within or next to the bubble). 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*.

293 Lastly, we inspected DNA methylation maintenance pathways. As predicted from the
294 patterns of mCG and mCHG (**Fig. 1D-H**), MET1, VIM (ORTH), CMT3/ZMET, SUVH4,
295 SUVH5/6 and DECREASE IN DNA METHYLATION 1 (DDM1) orthologues were found
296 present and expressed in *Spirodela*, albeit at a lower degree than in *Arabidopsis* (**Fig. 2E;**
297 **Supplemental Fig. S12, S13, S14, S17**). A single orthologue of the SUVH5/6 orthogroup was
298 found in *Spirodela* (**Fig. 2E; Supplemental Fig. S13**). In *Arabidopsis*, SUVH5 and SUVH6
299 display different 5mC sub-context binding affinities as well as genomic targets (Johnson et al.
300 2007; Rajakumara et al. 2011; Li et al. 2018; Zhang et al. 2023). Absence of either could
301 have implications for the DNA methylation landscape if their functions are conserved outside
302 of *Arabidopsis*. *Spirodela* SUVH5/6 othologue contains a conserved motif present in
303 *Arabidopsis* SUVH6, but absent in SUVH5, required for interaction with ANTI-SILENCING 1
304 (ASI1) to regulate silencing at intronic TEs and expression of their host genes (Saze et al.
305 2013; Wang et al. 2013; Duan et al. 2017; Zhang et al. 2023). As an ASI1 orthologue was
306 also found in *Spirodela* (**Fig. 2F; Supplemental Fig. S18**), we named the *Spirodela* SUVH5/6
307 orthologue as spSUVH6(5) until its sequence binding preference is determined. In addition,
308 with regards to mCHH maintenance, CMT2, required to maintain mCHH together with SUVH4
309 (Stroud et al. 2014), is missing like in maize (Bewick et al. 2017), as previously reported
310 (Harkess et al. 2024) (**Fig. 2E, Supplemental Fig. S13, S14**).
311 Thus, reduced 24-nt siRNA and mCHH levels in *Spirodela* can be explained by the
312 combined reduction in RdDM composition and expression in addition to the absence of CMT2.
313 However, it does not account for the global decrease in mCG and mCHG as their maintenance
314 pathways seem intact and expressed in *Spirodela*.
315
316 **Few transposable elements from different superfamilies display high levels of DNA**
317 **methylation in *Spirodela*.**
318 To further gain insight into the patterns of TE DNA methylation, we examined its
319 distribution along the *Spirodela* genome. To exclude any confounding effects of gbM, we

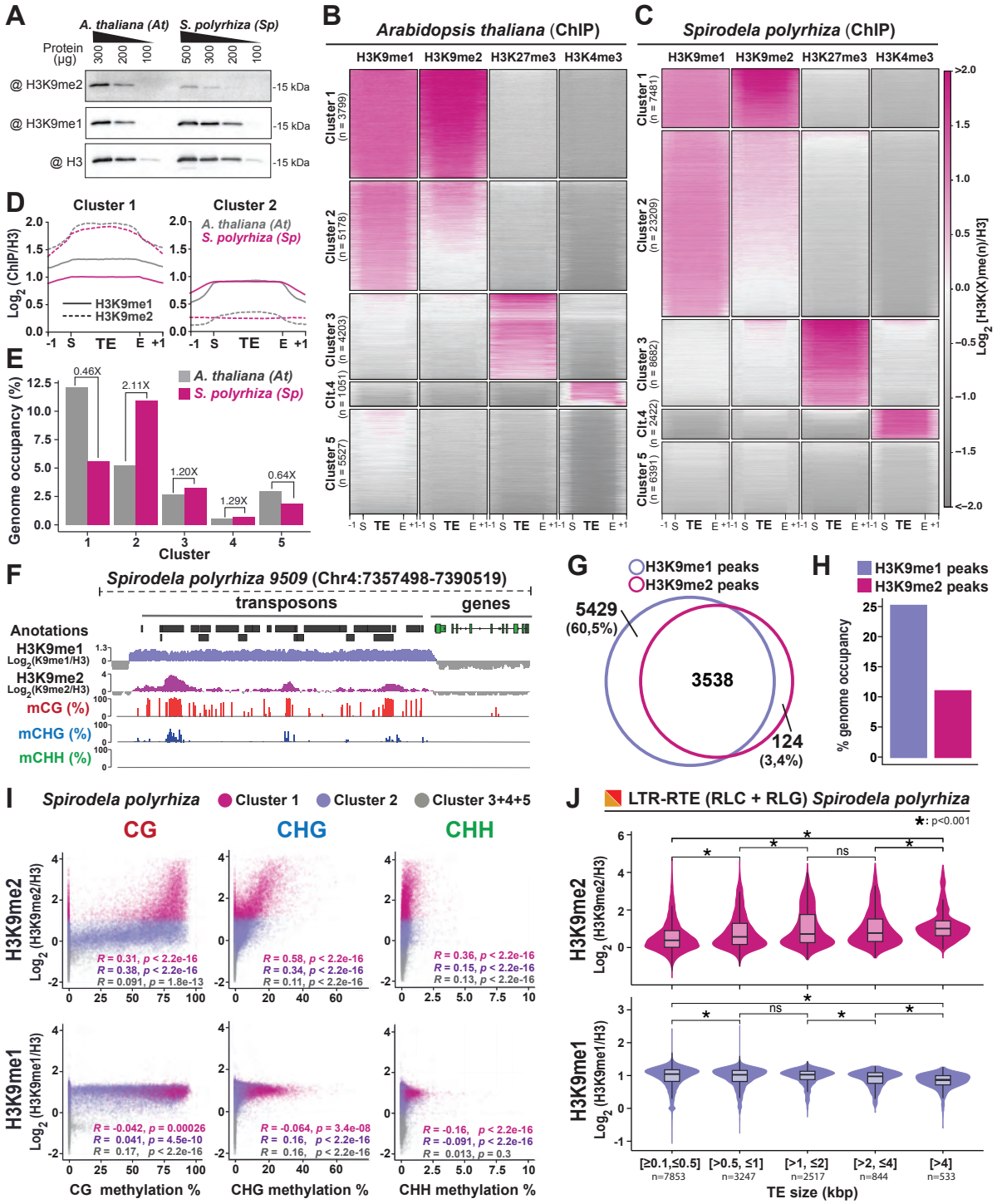

Figure 5: Patterns of H3K9me1/2 distribution in *Spirodela polyrhiza* transposons. (A) Protein blot analysis of H3K9me1 and H3K9me2 abundance in *Spirodela* compared to *Arabidopsis* at different protein amounts (µg). Total H3 protein levels are shown as loading control. **(B-C)** Heatmap of *k*-means clustering of transposons in *Arabidopsis* (B) and *Spirodela* (C) using H3K9me1, H3K9me2, H3K27me3 and H3K4me3 enrichment over H3, presented as $\log_2$[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 (violin and box plots) 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.

320 calculated mean DNA methylation levels exclusively on TEs and their density in 100-kb
321 windows. At gene-rich regions, DNA methylation is very scarce (**Fig. 3A; Supplemental**
322 **Fig. S19A**). As in *Arabidopsis*, *Spirodela* DNA methylation is concentrated at TE-dense
323 regions (presumably centromeric and pericentromeric domains). However, not all TE-dense
324 regions are highly methylated (**Supplemental Fig. S19B**). Hence, we investigated if local
325 DNA methylation levels correlate with TE density as in other angiosperms (Zhang et al. 2006;
326 Regulski et al. 2013; Seymour et al. 2014; Wyler et al. 2020). At equal TE-content, DNA
327 methylation is lower in all contexts in *Spirodela* compared to *Arabidopsis*, especially at CHG
328 and CHH (**Fig. 3B; Supplemental Fig. S20A**). Moreover, at the CG context, *Spirodela*
329 displays a broad variation range for a given TE content: even at TE densities above 50% or
330 higher, average mCG still spans between 0 and 75% (**Fig. 3B**).
331 To address the origin of such variation, we investigated the DNA methylation levels of
332 individual TEs. This revealed that the majority of TEs in *Spirodela* display low or very low
333 levels of DNA methylation in all contexts, especially CHG and CHH (**Fig. 3C; Supplemental**
334 **Fig. S20B,C**). Nonetheless, despite this substantially reduced DNA methylation, increased
335 expression at TEs was not observed (**Fig. 3D; Supplemental Fig. S20D**). TEs with high or
336 low DNA methylation levels did not belong to any specific TE class or superfamily. For all the
337 main plant TE superfamilies (Wicker et al. 2007) (TE classification and codes used here are
338 summarized in **Supplemental Fig. S21**), DNA methylation was nearly absent except for a few
339 elements (**Fig. 3E**). Thus, low genome-wide methylation levels in *Spirodela* result from only a
340 few methylated TEs, although these can display high mCG levels similar to those found in
341 *Arabidopsis* (**Fig. 3C,E**).
342
343 **Potentially intact transposons are the conspicuous targets of DNA methylation in**
344 ***Spirodela*.**
345 Although TE location and density influences their epigenetic regulation (Sigman and
346 Slotkin 2016), in *Arabidopsis* and tomato DNA methylation of longer, evolutionary younger
347 and therefore potentially autonomous TE insertions are differentially regulated than often

348 fragmented TE relics resulting from the accumulation of mutations and truncations over
349 evolutionary times following their silencing (Sigman and Slotkin 2016; Blumenstiel 2019;
350 Zattera and Bruschi 2022). On one hand, DNA methylation in long TEs is maintained by a
351 combination of RdDM acting on their boundaries (mainly in young active TEs) and CMT activity
352 along their body. On the other hand, short, degenerated TE remnants do not need CMT to
353 maintain their methylation (Zhong et al. 2012; Zemach et al. 2013; Wang and Baulcombe
354 2020). To explore if TE length influences DNA methylation in *Spirodela*, we sorted TEs into
355 short TE fragments (100 to 500 bp), intermediate TEs (0.5 to 4 kb) and long TEs (above 4 kb).
356 This revealed that only long TEs are heavily methylated at CG and, to a lesser extent, in CHG.
357 Intermediate TEs display a wide range of mCG levels, whereas TE fragments are minimally
358 or not methylated (**Fig. 4A**).

359 To further explore the connection between TE length and DNA methylation we focused
360 on the two superfamilies of long terminal repeat (LTR) retroelements (RTEs): *Ty1/Copia* (RLC)
361 and *Ty3/Gypsy* (RLG). LTR-RTEs are the most abundant class of TEs in the *Spirodela*
362 genome, collectively representing more than 50% of the TE fraction of the genome and largely
363 contributing as well to that of *Arabidopsis* (**Fig. 4B; Supplemental Fig. S22**). In addition, the
364 genetic structure and domain organization of complete functional elements are well known
365 (Kumar and Bennetzen 1999; Gorinšek et al. 2004; Neumann et al. 2019). Thus, LTR-RTEs
366 offer a good case study to investigate any link between DNA methylation and TE length, using
367 the latter as a proxy for their completeness and potential autonomy for transposition under the
368 assumption that degenerated and fragmented shorter TEs have lost some or all the required
369 domains to mobilize autonomously (non-autonomous elements) (Blumenstiel 2019). For
370 better resolution, we split TE sizes into further size categories in the intermediary range
371 between 0.5 and 4 kb. Within each size bin, *Spirodela* DNA methylation is lower than in
372 *Arabidopsis*. However, in *Spirodela*, average mCG and mCHG gradually increase with TE
373 size, although mCHG does not achieve levels comparable to those of *Arabidopsis* (**Fig. 4C,D;**
374 **Supplemental Fig. S23A,B**). The proportion of TEs displaying high CG methylation rises

375 across TE-size groups. While very few TE-fragments below 500 bp are highly methylated, the
376 majority of long TEs (>4 kb) have mCG above 60% (**Fig. 4C,D**). At the same time, the fraction
377 of TEs with at least one intact cognate protein domain and the number of domains identified
378 per TE annotation increase with each size category, with the majority of long LTR-RTEs
379 containing all domains associated with each superfamily (**Fig. 4E,F**).

380 Therefore, in *Spirodela*, most potentially intact TEs display high levels of DNA
381 methylation, which decreased as they degenerated and decayed. Given that long intact LTR-
382 RTEs represent a minority of *Spirodela* TEs (**Supplemental Fig. S23C,D**), their low
383 abundance helps explain the few TE numbers with high DNA methylation (**Fig. 3C,E**). In
384 addition, the proportion of hypomethylated shorter LTR-RTE fragments relative to methylated
385 intact TEs is higher in *Spirodela* than in *Arabidopsis* (**Fig. 4G,H; Supplemental Fig. S23G,H**),
386 resulting in the global reduction of TE methylation (**Fig. 1E, Fig. 3A,B**).

387

388 **Levels of H3K9me2, but not H3K9me1, are reduced in *Spirodela*.**

389 Considering the association between DNA and H3K9 methylation in plants, we
390 explored how far DNA methylation in *Spirodela* would impact the histone methylation
391 landscape. H3K9me2 and H3K9me1 are commonly linked with heterochromatin formation at
392 silenced TEs (Jackson et al. 2004; Fuchs et al. 2006; Stroud et al. 2014; Jamge et al. 2023).
393 Western blot analysis revealed that the H3K9me2 level in *Spirodela* is lower than in
394 *Arabidopsis*, whereas H3K9me1 displayed similar levels (**Fig. 5A**). This was unexpected as
395 we anticipated in *Spirodela* a decrease of both marks, given their connection with DNA
396 methylation and reduction upon loss of non-CG methylation in *Arabidopsis* (Stroud et al.
397 2014; Choi et al. 2021). To corroborate the observation, we investigated histone post-
398 translational modifications (PTMs) by mass-spectrometry (Johnson et al. 2004; Scheid et al.
399 2022), which confirmed an approximately 8-10-times lower level of H3K9me2 but similar
400 H3K9me1 levels in *Spirodela* relative to *Arabidopsis* (**Supplemental Fig. S24**).

401

402

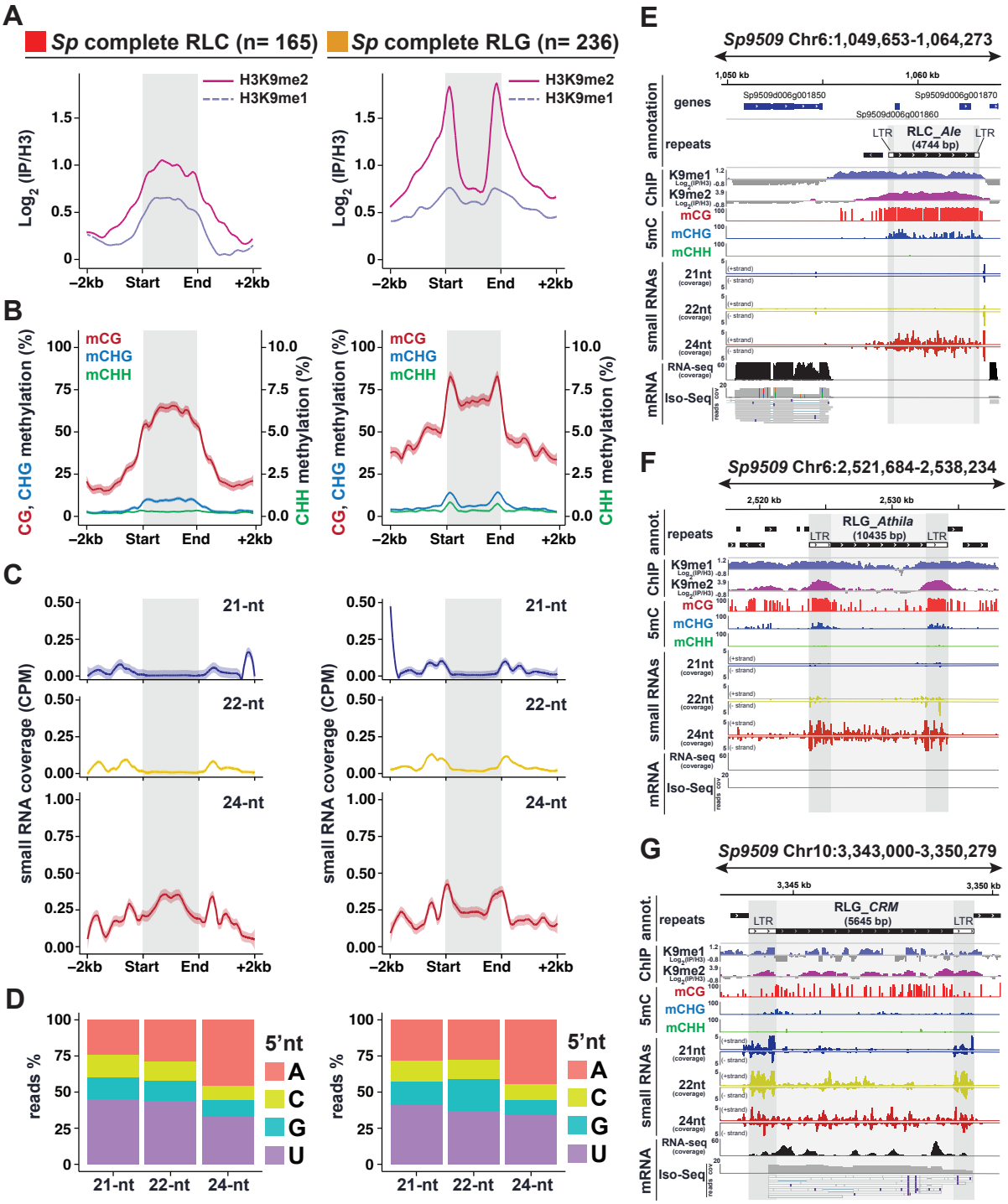

Figure 6: RdDM signatures in *Spirodela polyrhiza* complete TEs. (A) Metaplots of average weighted H3K9me1 (solid line) and H3K9me2 (dashed line) enrichments over H3 at complete RLC (left) and RLG (right) transposon annotations ± 2 kb in *Spirodela*. Complete TEs were defined as those for which all RLC 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 (Iso-Seq). 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.

403

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

405 To further delve into the distribution of H3K9me1/2 in *Spirodela* and how the general
406 organization of heterochromatic marks is affected by its reduced DNA methylation levels, we
407 performed chromatin immunoprecipitation (ChIP) in *Spirodela* fronds and *Arabidopsis* 10
408 days-old seedlings (**Supplemental Fig. S25**) and focused on the association of these
409 modifications with the TEs. In addition to H3, H3K9me1 and H3K9me2, we generated genomic
410 profiles through ChIP-seq for H3K27me3 and H3K4me3. Although Polycomb-mediated
411 H3K27me3 generally mediates developmentally regulated gene silencing, its association with
412 TEs was investigated given that it also regulates TEs in the absence or loss of DNA
413 methylation (Mathieu et al. 2005; Walter et al. 2016; Montgomery et al. 2020; Rougée et
414 al. 2021), and may contribute directly to TE silencing (Hisanaga et al. 2023). H3K4me3, was
415 taken as an euchromatic mark control (Houben et al. 2003; Zhang et al. 2009; Roudier et al.
416 2011; Ding et al. 2012; Pratx et al. 2023).

417 As observed for DNA methylation (**Fig. 3A**), H3K9me1/2 is concentrated at TE-dense
418 regions in *Spirodela* (**Supplemental Fig. S26**), likely centromeric and pericentromeric
419 heterochromatic regions (Houben et al. 2003; Jackson et al. 2004; Harkess et al. 2024). The
420 reduced size of such regions (spanning a few kbps per chromosome, **Supplemental Fig. S26**)
421 in *Spirodela*, compared to *Arabidopsis* larger pericentromeric regions (Mbps long), and their
422 increased number (20 chromosomes in *Spirodela* in contrast to 5 in *Arabidopsis*), fits the
423 speckled pattern of small dense DNA foci in DAPI staining of interphase nuclei of *Spirodela*
424 (**Supplemental Fig. S27**). This is also compatible with previous observations of dispersed
425 H3K9me2 foci in *Spirodela* nuclei immunostainings (Cao et al. 2015).

426 A more detailed *k*-means clustering analysis of histone mark profiles relative to H3
427 over TEs, identified 5 clusters with similar features in both *Arabidopsis* and *Spirodela*
428 (**Fig. 5B,C**). Clusters 1 and 2 contain TEs enriched for H3K9me1/2 while clusters 3 and 4
429 consist of those marked by H3K27me3 and H3K4me3 respectively. Lastly, cluster 5
430 encompasses TEs depleted of any of the profiled marks (**Fig. 5B,C**).

431 Although cluster 1 and 2 TEs are marked by H3K9me1/2, they display homogeneous
432 H3K9me1 but contrasting H3K9me2 levels (**Fig. 5D, Supplemental Fig. S28**). TEs in cluster
433 1 are more enriched in H3K9me2 than in H3K9me1 while in cluster 2 TEs H3K9me2 is below
434 H3K9me1 (**Fig. 5D**). The two clusters comprise very different fractions of the two genomes
435 and TE space. In *Spirodela*, cluster 1 only comprises about 20% of the TE space, representing
436 less than 6% of the genome, while in *Arabidopsis* is the most abundant one, encompassing
437 ~35% of the TE space and ~12% of the genome. Cluster 2, on the contrary, occupies 45% of
438 the *Spirodela* TE space (more than 10% of its genome), but only about 20% in *Arabidopsis*
439 (5% of the genome) (**Fig. 5E, Supplemental Fig. S29A**).
440 Given the similar genome size and TE content in both species, the opposed
441 contribution of cluster 1 and cluster 2 to their genomes is in line with the levels of H3K9me2
442 observed by western blot and mass-spectrometry. Furthermore, similar H3K9me1 levels in
443 both clusters concur with its even abundance in *Spirodela* and *Arabidopsis*. Hence, reduced
444 DNA methylation levels in *Spirodela* correlate with a smaller fraction of the genome displaying
445 a high H3K9me2 enrichment over TEs but it does not impact H3K9me1 distribution, nor it
446 leads to a widespread compensatory regulation of TEs by H3K27me3. We noticed that,
447 despite cluster 3 occupies a comparable fraction of TE sequences in both species (**Fig. 5E**),
448 *Spirodela*, in contrast to *Arabidopsis*, displays TE-rich regions devoid of DNA methylation and
449 H3K9me1/2 but rich in H3K27me3 (e.g. Chromosomes 1, 6, 10, 14, 20) (**Supplemental**
450 **Fig. S26**). No obvious differences in TE-superfamily composition were observed between
451 cluster 3 and clusters 1 and 2 (**Supplemental Fig. S29B**), suggesting no particular TE
452 superfamily is preferentially regulated by Polycomb.
453
454 **H3K9me1 and H3K27me1 mark TEs independently of their length and methylation**
455 **status in *Spirodela*.**
456 A closer inspection of the distribution of H3K9me1/2 along *Spirodela* TEs revealed that
457 H3K9me1 uniformly covers TEs, while H3K9me2 peaks are found only at individual TEs
458 (**Fig. 5F**). Furthermore, while more than 95% of H3K9me2 peaks overlapped with H3K9me1

459 domains, a significant fraction of H3K9me1 peaks (60,5%) did not with H3K9me2-associated
460 regions (**Fig. 5G**). Moreover, most H3K9me2 peaks were embedded within H3K9me1
461 (**Supplemental Fig. S29C**). Adding to that, H3K9me1-enriched regions occupy 25% of the
462 *Spirodela* genome whereas H3K9me2 peaks represent less than 12% (**Fig. 5H**). This
463 indicates that, similarly to DNA methylation, high levels of H3K9me2 are only found within few
464 TEs. Such observations prompted us to investigate the association between H3K9me1/2 and
465 DNA methylation in *Spirodela* and indeed, H3K9me2 levels in *Spirodela* positively correlate
466 with DNA methylation, including mCHG, like in *Arabidopsis* (**Fig. 5I, Supplemental Fig. S30**).
467 Parallel to DNA methylation (**Fig. 3**), H3K9me2 increased with TE size (**Fig. 5J**). This further
468 supports that, as previously inferred, the positive feedback loop between mCHG and
469 H3K9me2 operates in *Spirodela*. However, this did not apply to H3K9me1. While in
470 *Arabidopsis* TEs enriched in H3K9me1 invariably displayed high levels of DNA methylation
471 (**Supplemental Fig. S30**), they did not so in *Spirodela* with its broad range of DNA methylation
472 levels at TEs (**Fig. 5I**). Accordingly, H3K9me1 levels vary little across TE sizes (**Fig. 5J**).

473 These results indicate that, in *Spirodela*, H3K9me1 serves as broad TE mark but does
474 not rely on DNA methylation as H3K9me2, suggesting that they might be differentially
475 regulated. Although *Arabidopsis* H3K9 histone methyl-transferases (HMTs) can catalyze both
476 H3K9me1 and me2 in vitro (Jackson et al. 2004; Ebbs and Bender 2006), it is unclear how
477 their mono- and di- methyl-transferase activity is regulated in vivo or if both marks are
478 deposited by the same HMTs. Loss of SUVH4, the main HMT in *Arabidopsis*, results in a
479 strong reduction of H3K9me2 and loss of epigenetic silencing, only marginally affecting
480 H3K9me1 and global heterochromatin organization (Jackson et al. 2002; Jasencakova et al.
481 2003; Jackson et al. 2004). This originally led to the proposition that H3K9me2 and H3K9me1
482 play different roles in silencing and heterochromatin maintenance respectively (Jackson et al.
483 2004).

484 Given that heterochromatin maintenance is not only relevant for TE silencing by as
485 well for genome stability, we sought to investigate the distribution of other epigenetic marks
486 associated with heterochromatin and genome stability. Specifically, H3K27me1, which in

487 plants is largely associated with maintenance of constitutive heterochromatin, especially
488 during DNA replication, and its distribution is not affected by loss of mCG (Mathieu et al.
489 2005; Shi and Dawe 2006; Jacob et al. 2009; Sequeira-Mendes et al. 2014; Montgomery
490 et al. 2020; Jamge et al. 2023). *Arabidopsis* mutants lacking H3K27me1 do not only show
491 loss of TE silencing but as well defects in chromatin organization and excess of repetitive DNA
492 due to heterochromatin amplification during DNA replication (Jacob et al. 2009; Stroud et al.
493 2012; Dong et al. 2021). To study its distribution in *Spirodela*, we performed H3K27me1,
494 along with H3K9me1/2, ChIP. The resulting profiles and *k*-means clustering analysis showed
495 that H3K27me1 displays an overlapping pattern with H3K9me1, present at TEs independently
496 of DNA methylation and occupying as well a large fraction of the genome (**Supplemental Fig.**
497 **S31**). Hence, in *Spirodela*, despite the reduced levels of DNA methylation and H3K9me2,
498 H3K9me1 and H3K27me1 remain associated with TEs, likely to maintain heterochromatin and
499 genome integrity.
500
501 ***5S rDNA* also display low levels of DNA methylation and H3K9me1/2 in *Spirodela*.**
502 Besides TE silencing, DNA methylation, H3K9me2 and RdDM are also involved in the
503 epigenetic regulation of other repeats in plants. To investigate if the unique patterns seen on
504 TEs also affect other repeats in *Spirodela*, we looked into the epigenetic landscape of the *5S*
505 *rDNA* repeats. Present in hundreds or thousands of copies clustered in several large tandem
506 arrays, most *5S rDNA* repeats are generally silenced in *Arabidopsis* and are well known
507 targets of RdDM (Llave et al. 2002; Xie et al. 2004; Blevins et al. 2009; Simon et al. 2018).
508 Although *Spirodela* and other duckweeds have been shown to have abnormally low numbers
509 of *5s rDNA* repeats (<80-100 copies), two *5S rDNA* clusters in *S. polyrhiza* have been located
510 on Chromosomes 6 and 13 (Hoang et al. 2018; 2020; Chen et al. 2021; 2024). A close
511 inspection to the two *5S rDNA* clusters in *Spirodela* revealed clear differences with the
512 epigenetic patterns found in *Arabidopsis* (**Supplemental Fig. S32A**). In *Spirodela*, *5S rRNA*
513 repeats did not display DNA methylation in any context, nor H3K9me1/2 and H3K27me1/3

514 enrichments (**Supplemental Fig. S32B,C**). Nonetheless, in *Spirodela*, 24-nt siRNAs also
515 mapped to *5S rDNA* repeats (**Supplemental Fig. S32A**). Although, in contrast to *Arabidopsis*,
516 they were exclusively in sense to the *5S rDNA* loci and showed a guanosine 5'-nucleotide
517 (5'G) bias (>75%) (**Supplemental Fig. S32D**) (not known to be preferentially loaded by any
518 of the *Arabidopsis* AGOs), instead of the RdDM effector AGO4 characteristic 5'A signature
519 seen in *Arabidopsis* (Mi et al. 2008; Havecker et al. 2010; Jullien et al. 2020).

520 Hence, lack of DNA methylation and H3K9me2 in *Spirodela* does not only impact
521 degenerated and short TEs, but also other RdDM targets such as *5S rDNA* repeats, and
522 further strengthens the hypothesis that such silencing mechanisms only apply to long,
523 complete, TEs. Indeed, increased H3K9me2 and DNA methylation at long TEs hints that their
524 self-reinforcement is focused at potentially functional or recently active ones to ensure their
525 silencing.

526

527 **Long, complete TEs show signatures of RdDM activity.**

528 To gain further insights into the epigenetic patterns on long TEs, we focused on the
529 complete LTR-RTEs of the *Copia* (RLC) and *Gypsy* (RLG) superfamilies, in which H3K9me2
530 levels were higher than those of H3K9me1. However, the patterns differed depending on the
531 TE superfamily. While in RLCs had more H3K9me2 towards the center of the TE, the same
532 modification peaked on the edges of RLGs (**Fig. 6A**). Such patterns resembled those found
533 in *Arabidopsis* intact LTR-RTEs, (**Supplemental Fig. S33A**). DNA methylation at CG and
534 CHG followed a similar trend as H3K9me2 in both cases (**Fig. 6B**), although mCHG was
535 comparatively lower (**Fig. 6B, Supplemental Fig. S33B**). Furthermore, albeit very low, peaks
536 of mCHH were also observable at RLGs edges (**Fig. 6B**). Hence, the layout of epigenetic
537 silencing marks along complete TEs in *Spirodela* are very similar to those in *Arabidopsis*.

538 Given that in *Arabidopsis* these patterns are partially determined by RdDM (Stroud et
539 al. 2013; Zemach et al. 2013; Stroud et al. 2014), we inspected if small RNA mapped to
540 complete TEs in *Spirodela*. Indeed, siRNAs derived from regions with DNA methylation and
541 H3K9me2 were present in *Spirodela*, specially 24-nt siRNAs, despite at lower levels than in

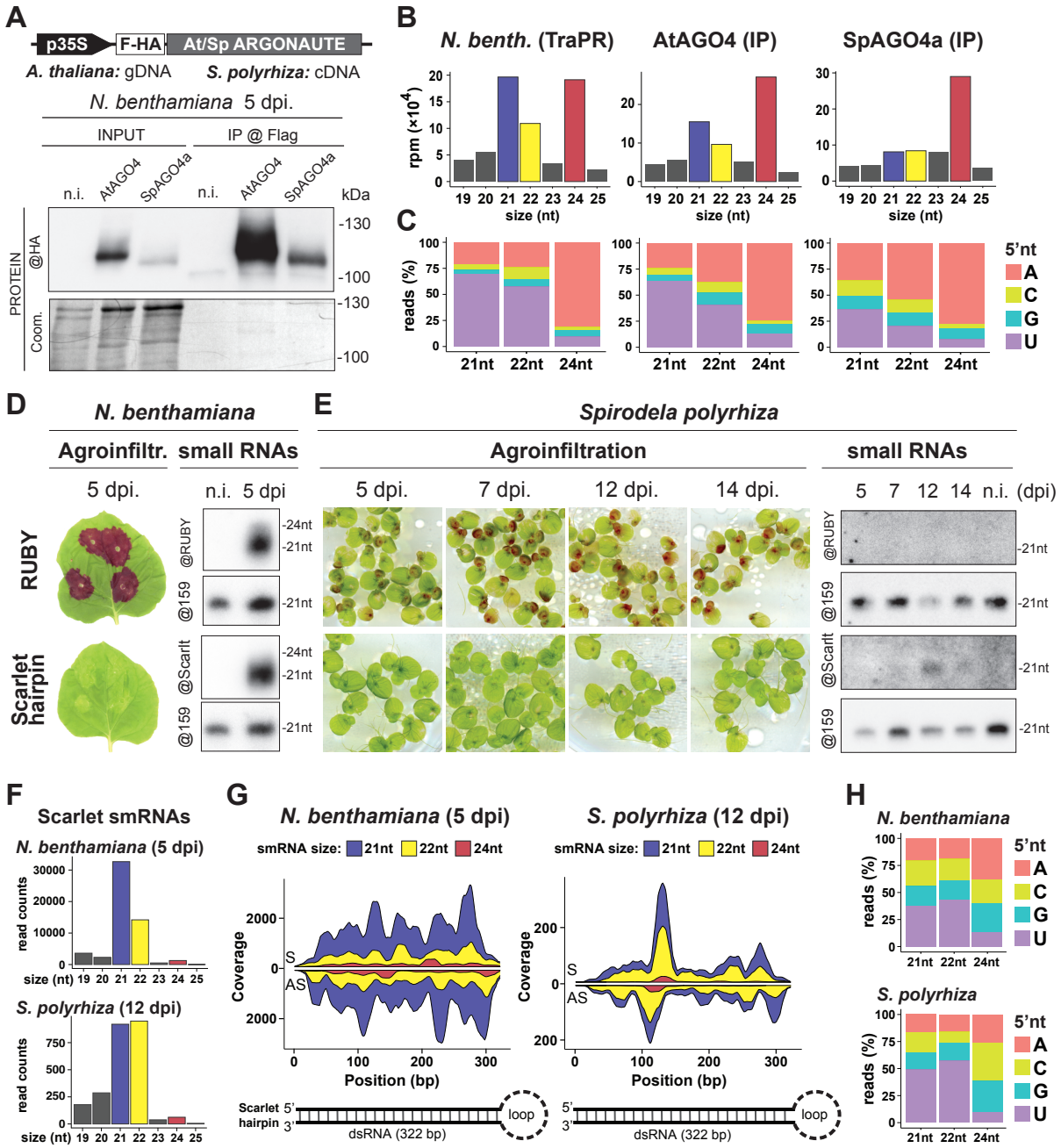

Figure 7: *Spirodela polyrhiza* AGO4a siRNA loading preferences and IR-derived siRNAs in transient expression. (A) Protein blot analysis of *Arabidopsis* and *Spirodela* AGO4 (AtAGO4 and SpAGO4a) input and IP fractions from transient expression in *N. benthamiana* 5 days post infiltration (dpi). n.i.: 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*.

542

543 *Arabidopsis* (**Fig. 6C, Supplemental Fig. S33C**). Moreover, in contrast to *5S rDNA* siRNAs,
544 24-nt siRNAs mapping to complete TEs displayed an increased 5'A bias, characteristic of
545 small RNAs loaded into AGO4 (**Fig. 6D**), although not as preferential as in *Arabidopsis*
546 (**Supplemental Fig. S33D**). Complementarity of siRNAs with DNA methylation and H3K9me2
547 within the TE body and/or their LTRs were validated for individual complete TEs (**Fig. 6E-G,**
548 **Supplemental Fig. S34**).

549 To assess if TE-derived siRNAs might participate in RdDM in *Spirodela*, also given the
550 unconventional 5'G bias of *5S rDNA* 24-nt siRNAs, we tested if its AGO4 (SpAGO4) has the
551 same small RNA loading preference as *Arabidopsis* AGO4 (AtAGO4). To do so, we cloned
552 genomic *AGO4* loci from both species with a Flag-HA tag for heterologous transient
553 expression in *Nicotiana benthamiana* leaves. No Sp*AGO4* protein was detectable due to intron
554 retention events when expressed in *N. benthamiana* (**Supplemental Fig. S35A-C**). This was
555 likely due to the inability of dicots (*N. benthamiana*) to properly process introns of the monocots
556 (*Spirodela*) (Keith and Chua 1986; Goodall and Filipowicz 1991; Lou et al. 1993). To
557 circumvent this limitation, we amplified and cloned full length cDNA from one of the two
558 *Spirodela* AGO4, *AGO4a*, and expressed it in *N. benthamiana* leaves (**Supplemental**
559 **Fig. S35C-D**). To examine their small RNA loading preferences, AtAGO4 and SpAGO4a were
560 immunoprecipitated (IP) from *N. benthamiana* leaf lysates (Jullien et al. 2020) (**Fig. 7A**) and
561 small RNAs extracted and sequenced. Like AtAGO4, SpAGO4a displayed a strong preference
562 for 5'A 24-nt long siRNA and reduced affinity for 21-nt siRNA, compared to the population of
563 siRNAs present in *N. benthamiana* leaves (**Fig. 7B,C**).

564 Therefore, small RNAs might be indeed involved in maintaining DNA methylation and
565 H3K9me2 in *Spirodela* at long complete TEs through canonical RdDM. Our analysis of
566 silencing pathways and epigenetic patterns on TEs shows that core components required for
567 RdDM are present in the *Spirodela* genome and, although poorly expressed, probably active
568 during vegetative clonal propagation to silence long TEs.

569

570

571 **Transcriptionally active TEs are source of 21/22-nt-siRNAs in *Spirodela*.**

572 We also detected the presence of very low amounts of 21-22-nt siRNAs at the edges
573 of complete TEs (**Fig. 6C**). In some instances, 24-nt siRNAs overlapped with equally abundant
574 21-22-nt siRNAs, especially at the flanking LTR sequences (**Fig. 6G, Supplemental**
575 **Fig. S36**). 21-22-nt siRNAs can arise from TE edges in *Arabidopsis* (**Supplemental**
576 **Fig. S33C**), as result of DCL2/4 processing of Pol IV transcripts, and direct DNA methylation
577 (Nuthikattu et al. 2013; McCue et al. 2015; Panda et al. 2020; Sigman et al. 2021). It is
578 unlikely that the observed TE-derived 21-22-nt siRNAs participate in RdDM, due to the low
579 affinity of SpAGO4 to bind such siRNA sizes (**Fig. 7B**), and those mapping to TEs have less
580 5'A bias but more 5'U (**Fig. 6D**), a signature of AGO1-loaded small RNAs (Mi et al. 2008;
581 Havecker et al. 2010). This hints to a potential role rather in post-transcriptional silencing
582 (PTGS). During PTGS, 21-22-nt siRNAs loaded into ARGONAUTE proteins of the
583 AGO1/5/10/18 and AGO2/7 clades target Pol II transcripts to guide mRNA cleavage or
584 translational repression (reviewed in (Fang and Qi 2016)). Coincidentally, we observed
585 expression of TEs producing 21-22-nt siRNAs in our transcriptome datasets, including, in
586 some intances, full-length transcripts in the long-read cDNA sequencing Iso-Seq data
587 (**Fig. 6G, Supplemental Fig. S35**).

588 Although PTGS is a well-documented control mechanism against transcriptionally
589 active TEs in *Arabidopsis* (Teixeira et al. 2009; Slotkin et al. 2009; Marí-Ordóñez et al. 2013;
590 Creasey et al. 2014; Lee et al. 2020; Oberlin et al. 2022), the potential role of 21-22-nt
591 siRNAs in TE silencing in *Spirodela* remains to be experimentally tested.

592

593 **22-nt siRNAs are produced in *Spirodela* despite the lack of a DCL2 orthologue.**

594 We were intrigued by the presence of TE-derived 22-nt siRNAs, given that any
595 identifiable ortholog of DCL2, their cognate Dicer in plants (Xie et al. 2005; Henderson et al.
596 2006; Wang et al. 2018; Jia et al. 2020), is missing in *Spirodela* and duckweed genomes in
597 general (**Fig. 2B, Supplemental Fig. S8**) (Ernst et al. 2023; Harkess et al. 2024).

598 Their co-occurrence with 21-nt siRNAs, (**Fig. 6G, Supplemental Fig. S35**) suggested
599 a common RNA substrate. Given the low amounts of 21-22-siRNAs mapping to individual TEs,
600 to rule out artifacts and inspect if other endogenous siRNA sources are as well processed into
601 22-nt siRNAs, we investigated *Spirodela* trans-acting siRNAs (tasiRNAs). In *Arabidopsis*,
602 trans-tasiRNAs are processed from transcripts of specific loci (TAS), mainly into 21-nt-siRNA
603 by DCL4 to, through their association with AGO1, regulate several target mRNAs by PTGS
604 (Yoshikawa et al. 2005). As the TAS3 locus had been previously identified in *Spirodela*
605 (Michael et al. 2017), we scored for TAS3 matching siRNAs in our TraPR datasets. In contrast
606 to *Arabidopsis*, where TAS3 tasiRNAs are almost exclusively 21-nt long, in *Spirodela*, TAS3
607 tasiRNAs were found as 21- and 22-nt, the later accounting for about one third of TAS3
608 tasiRNAs (**Supplemental Fig. S37A,B**). Their 5'U bias also suggest their preferential loading
609 into AGO1 (**Supplemental Fig. S37C**).
610 Next, to independently test if *Spirodela* can produce 22-nt siRNAs in the absence of
611 DCL2, we attempted to stimulate small RNA biogenesis through expression of suitable
612 exogenous substrate RNAs. As stable transformation in *Spirodela* is difficult, we chose
613 *Agrobacterium*-mediated transient expression. This technique has been widely used to trigger
614 and investigate silencing mechanisms, owing to the quick onset of siRNA production in
615 response to expression of double-stranded RNA (dsRNA) following infiltration of plant tissue
616 with *Agrobacterium* carrying the corresponding transgenes (Hamilton et al. 2002; Bertazzon
617 et al. 2012; Andrieu et al. 2012; Sun et al. 2014; Jay et al. 2023). We developed a modified
618 protocol for frond transformation in *Lemna minor* (Yang et al. 2018) and applied it successfully
619 to induce transient transgene expression in *Spirodela* fronds (**Supplemental Fig. S38**). We
620 used the visual reporter *RUBY* (He et al. 2020) and an inverted repeat (IR) derived from the
621 coding sequence of the fluorescent reporter *Scarlet* (*Scarlet* hairpin, *hpScarlet*) (Incarbone et
622 al. 2023). Both constructs triggered siRNAs when infiltrated into *N. benthamiana* as expected
623 (**Fig. 7D**). In *Spirodela*, however, we could not detect *RUBY*-derived siRNAs by RNA blotting
624 even after two weeks post-infiltration. Given transgene sense PTGS (S-PTGS) is generally

625 associated with high levels of transgene expression, the apparent lack of S-PTGS in *Spirodela*
626 might be due to reduced expression of the reporter compared to *N. benthamiana*. On the other
627 hand, *hpScarlet*-derived siRNAs were observed 12- and 14-days post-infiltration and were lost
628 shortly after (**Fig. 7E**), likely due to loss of transgene expression inherent to the
629 *Agrobacterium*-mediated transient expression method. Sequencing sRNAs from both *N.*
630 *benthamiana* and *Spirodela* infiltrated with *hpScarlet* showed that the IR is processed into 21-
631 22-nt siRNAs in both plants. However, *Spirodela* produced relatively equal amounts of 21- and
632 22-nt siRNAs, as observed for TEs, whereas the profile in *N. benthamiana* is dominated by
633 21-nt siRNAs (**Fig. 7F**). Both siRNA sizes were distributed along the length of the IR dsRNA
634 (**Fig. 7G**) and displayed the 5'U bias signature of AGO1 loading (**Fig. 7H**).

635 Hence, *Spirodela* is capable of producing 22-nt siRNAs in the absence of DCL2,
636 supporting that 21-22-nt siRNAs mapped to TEs are *bona-fide* siRNAs. Given the similar 21-
637 22-nt siRNA patterns observed between TEs, miRNA-triggered secondary tasiRNA and IR-
638 derived siRNAs, 21-22-nt siRNAs mapped to transcriptionally active TEs are likely the result
639 of: (i) a shared miRNA targeting event to initiate secondary siRNAs, shown to regulate some
640 TEs in other plants (Creasey et al. 2014; Šurbanovski et al. 2016), or; (ii) direct processing
641 of RNA hairpins commonly present in LTRs (Benachenhou et al. 2013).

642

643 ## Discussion

644 Based on high resolution genomic information of a duckweed species during its clonal
645 reproduction, we have provided a detailed analysis of TE content, of the epigenetic
646 configuration at TEs, and several components expected to contribute to the silencing
647 pathways that control TE activity in other angiosperms. Our results reinforce the idea that, with
648 few exceptions, the *Spirodela* genome encodes only a reduced set of both PTGS and TGS
649 components compared to *Arabidopsis*, some of which are absent or display a reduced set of
650 paralogues, partially accounting for its reduced levels of DNA methylation and other epigenetic
651 marks generally associated with TEs. The unique epigenome of *Spirodela* is the result of the
652 loss of DNA methylation, and reduced levels of H3K9me2 associated to it, in highly abundant

653 degenerated TE remnants scattered across the genome which remain nonetheless associated
654 to the heterochromatic marks H3K9me1 and. H3K27me1. In contrast, potentially intact or
655 recently active TEs display all the epigenetic signatures associated with silencing in
656 angiosperms. Another interesting observation is the loss of any of the silencing marks
657 inspected here in *5S rDNA* repeats. Although this might reflect the preferential selectivity of
658 RdDM for long TEs, lack of silencing might as well be due to the reduced number of *5S rDNA*
659 repeats in *Spirodela*, compared to other angiosperms, to ensure enough 5S rRNA expression
660 for ribosome biogenesis. These changes in epigenetic regulation and silencing might be a
661 consequence of the simplified morphology or reduced tissue and cell types in duckweeds
662 compared to other angiosperms (Wang et al. 2014; Michael et al. 2020; Abramson et al.
663 2021; Ware et al. 2023; Ernst et al. 2023). However, it could also be a consequence of their
664 different lifestyle or propagation mode.

665 *Spirodela* displays a broad range of adaptations to their free-floating aquatic lifestyle,
666 such as the loss of complete stomatal closure response to prevent water loss (Fang et al.
667 2023). Besides TE control, RdDM plays a role in the regulation of gene expression and
668 physiological adaptation to several abiotic stresses, such as heat, drought, salinity, or low
669 environmental humidity (Erdmann and Picard 2020; Liu and He 2020; Urquiaga et al. 2021).
670 Given that water supply is not limiting in aquatic environments, and their higher stability to
671 fluctuations in these parameters compared to terrestrial ones, RdDM activity in adult somatic
672 tissues to respond to stress might have become less relevant in *Spirodela*.

673 Alternatively, but not mutually exclusive, the epigenetic landscape in the *Spirodela*
674 genome might be a consequence of their reproduction strategy. Asexual reproduction has
675 been linked to lower TE activity and content as TEs require sexual reproduction to spread
676 horizontally within populations. In absence of sexual reproduction, TEs may instead become
677 locked in vertical clonal lineages. Additionally, they might go extinct if their proliferation
678 becomes detrimental for the host without increased purifying selection of new TE insertions
679 facilitated by sex (Arkhipova and Meselson 2000; Wright and Finnegan 2001; Bast et al.

680 2019; Dechaud et al. 2019). Furthermore, RdDM activity over short TE remnants, leading to
681 the production of 24-nt siRNAs, is seen as an immunity memory against mobilization or
682 invasion by related intact TEs (Nguyen and Gutzat 2022), akin to the piRNA pathway in
683 animals (Czech et al. 2018). Thus, TEs potentially being less active during *Spirodela* clonal
684 propagation might have led to the loss of DNA methylation, RdDM expression ,and lack of
685 components required for its activity during vegetative development such as SHH1 and
686 CLSY1/2 (Zhou et al. 2022) over TE fragments. Maintenance of heterochromatin at such
687 repetitive TEs remnants through H3K9me1 and H3K27me1 independently of DNA methylation
688 might ensure proper chromatin organization and stability, preventing TE interference with
689 transcriptional and RNA processes (Kim and Zilberman 2014; Huff et al. 2016; Sammarco
690 et al. 2022; Ilık et al. 2024), illegitimate recombination or re-replication of repeats during cell
691 division (Jacob et al. 2009; Stroud et al. 2012; Chénais et al. 2012; Bourque et al. 2018;
692 Dong et al. 2021). The presence of dense TE clusters marked by H3K27me3 might also
693 prevent genome instability as their alternative regulation by DNA and H3K9 methylation could
694 mimic centromeric and pericentromeric regions, leading to the formation of polycentromeric
695 chromosomes highly unstable during mitosis (Fu et al. 2012). As inactivation of secondary
696 centromeres has been linked to increased H3K27me2/3 in wheat (Zhang et al. 2010), such
697 TE-rich clusters might correspond to inactive centromeres in *Spirodela*. Deeper
698 characterization of the chromatin structure (e.g., chromatin accessibility) and landscape
699 (additional histone modifications, histone variants …), together with functional studies, might
700 shed light onto the regulation of H3K9me1, H3K27me1, H3K27me3 and their role in shaping
701 heterochromatin, TE silencing and genome stability regulation in *Spirodela* and other plants.
702 Furthermore, unlike vegetative propagation in other plants where clones originate from
703 differentiated organs such as stems, roots or leaves, with tissue-specific and developmentally
704 regulated TEs, new individuals in duckweeds originate from a dedicated group of stem cells
705 (Landolt 1986). This might have further favored a relaxation of epigenetic and RdDM-
706 mediated TE control in somatic tissues which do not participate in vegetative reproduction in

707 *Spirodela*. Nevertheless, potentially intact TEs are under several layers of epigenetic control,
708 including RdDM, suggesting the pathway might be tissue-specific or developmentally
709 regulated.

710 In contrast to stem cells in other angiosperms, contained in a well-organized shoot
711 apical meristem (SAM) that generates new organs (Gaillochet and Lohmann 2015),
712 meristems in *Spirodela* and all duckweeds are formed by unstructured groups of few
713 undifferentiated cells, which generate new meristems that develop into daughter fronds.
714 (Landolt 1986; Lemon and Posluszny 2000; Sree et al. 2015; Yang et al. 2021; Li et al.
715 2023). In order to maintain genome integrity along the clonal lineage, RdDM might be
716 expressed in the few stem cells within the budding pockets or during the developmental
717 window between the formation of a new meristem and its development into a new frond with
718 its own stem cells.

719 Furthermore, lack of, or weakened, RdDM activity due to losses such as DRM3,
720 CLSY1, SHH1 and absence of CMT2, in combination with weak CMT3/ZMET activity in
721 somatic tissues during development, will suffice to explain the global loss of mCHH and low
722 levels of mCHG. Additionally, the higher variation in mCG levels in *Spirodela* could be a
723 consequence of mCG maintenance being less efficient, not active in all cells, or not all CG
724 sites being as faithfully maintained. Altogether, this might, in some instances, permit
725 transcription of some TEs that then trigger 21-22-nt siRNAs. Alternatively, transient relaxation
726 of TE control and TE expression has been shown to take place in *Arabidopsis* meristematic
727 stem cells during developmental transitions (Gutzat et al. 2020). A similar phenomenon taking
728 place during the development of new meristems and fronds might as well explain the paradox
729 of observed expression from TEs covered in silencing marks and the apparent simultaneous
730 presence of small RNAs associated with TGS and PTGS. Further tissue- and cell-specific
731 investigation of epigenetic parameters will be able to address these hypotheses.

732 Nonetheless, although infrequent, *Spirodela* does flower occasionally (Fourounjian et
733 al. 2021) and its population genetic structure shows signs of sexual reproduction (Ho et al.

734 2019; Wang et al. 2024). Hence, RdDM might still be relevant and active during sexual
735 reproduction as suggested by the presence of CLSY3, which recruits RdDM to TEs during
736 flowering in *Arabidopsis* (Long et al. 2021; Zhou et al. 2022). Nonetheless, we note that the
737 function of silencing components present and absent in *Spirodela* might play different roles in
738 duckweeds (and other monocots) compared to those described in *Arabidopsis* given their
739 evolutionary distance. Hence, a full understanding of TE silencing in duckweeds will require
740 further molecular genetics and biochemical characterization of their silencing pathways.
741 Adding to the unique epigenetic regulation of degenerated TEs and their relics in *Spirodela*
742 here described, one of the intriguing results of our study is the constatation that, despite lack
743 of DCL2, *Spirodela* can produce 22-nt siRNAs, nearly as abundantly loaded in AGOs as 21-
744 nt siRNAs, not only from TEs but also from endogenous and exogenous sources of dsRNA.
745 As these sources of siRNAs are generally products of DCL4 in plants (Xie et al. 2005; Fusaro
746 et al. 2006; Parent et al. 2015), given the overlap between 22- and 21-nt but not 24-nt (DCL3),
747 we consider processing dsRNA into 22-nt siRNAs by *Spirodela* DCL4 as a plausible
748 explanation. Alternatively, DCL3 in *Arabidopsis* has been shown to be able to process certain
749 dsRNA precursors into 22-nt siRNAs *in vitro* (Loffer et al. 2022) and is responsible for 22-to-
750 24-nt siRNA biogenesis in the moss *Physcomitrium patens* (Cho et al. 2008). Therefore,
751 *Spirodela* DCL3 might as well be responsible for the biogenesis of 22-nt siRNAs in *Spirodela*.
752 Future genetic and biochemical characterization of *Spirodela* DCLs will be required to
753 understand the factors responsible for the biogenesis of 22-nt siRNAs and their role in
754 *Spirodela* and other duckweeds. In *Arabidopsis*, DCL2 and its capacity to process dsRNA into
755 22-nt siRNAs have been proposed to play a dual role in antiviral immunity. On one hand its
756 processing of viral dsRNA acts as an innate immunity sensor (Nielsen et al. 2023; 2024). On
757 the other hand, 22-nt siRNAs in plants guide the recruitment of RDRs to boost siRNAs
758 production from their targets and amplify silencing (Chen et al. 2010; Zhang et al. 2012;
759 Taochy et al. 2017; Chen et al. 2018) . Their presence in *Spirodela* independently of DCL2
760 and the lack of several other RNAi factors associated with antiviral defense in angiosperms,

761 as well as the apparent lack of sense transgene silencing (S-PTGS) triggered by transient
762 expression of the *RUBY* reporter, raises several intriguing questions about RNA-based
763 defense mechanisms in duckweeds that deserve further investigation in the future.
764 Hence, *Spirodela* might represent a unique model to investigate not only transposon
765 dynamics and the epigenetic and genetic consequences of long periods of asexual
766 reproduction but also noncanonical PTGS pathways.
767

768 ## Methods

769 ***Spirodela* culture and growth conditions:**

770 ***Spirodela polyrhiza*.** *Spirodela* was grown on half-strength Schenk and Hildebrandt (½ SH,
771 Duchefa Biochemie, #S0225) liquid medium in Magenta™ GA-7 containers (Sigma Aldrich,
772 #V8505), culture dishes (Greiner Bio-One #664160), 6-wells culture plates (Greiner Bio-One
773 #657185), or in 1L glass beakers (covered with a sterile plastic lid), all sealed with Leucopore
774 tape (Duchefa Biochemie #L3301), under long-day conditions (16h light /8h dark) at 21°C with
775 a light intensity of 85 µM m^{-2} s^{-1}. The *Spirodela polyrhiza* #9509 clone was obtained from the
776 Landolt collection (Zürich, Switzerland), recently transferred to the CNR-IBBA_MIDW
777 collection (Milano, Italy: https://biomemory.cnr.it/collections/CNR-IBBA-MIDW) (Morello et al.
778 2024) .
779 Further information regarding *Spirodela* sterilization for axenic cultures and long-term storage,
780 *Arabidopsis* and *N. benthamiana* material information and culture and plant imaging can be
781 found within Supplemental Methods in Supplemental Material.
782

783 **Genome sequencing and gene annotation:**

784 Genomic DNA from *Spirodela* was purified using Dneasy® Plant mini kit (Qiagen #69104)
785 following the manufacturer's instructions. Library preparation was carried using NEBNext®
786 Ultra™ II DNA Library Prep Kit for Illumina® (NEB #E7103). Sequencing was performed on
787 an Illumina HiSeq4A to produce paired-end reads of 125 bp. The Landolt Sp9505 genome
788 was generated by pseudo-assembly with MaSuRCA (Zimin et al. 2013) (version 3.4.1, default

789 parameters), using as scaffold the Sp9509_Oxforrd_v3 reference assembly(Hoang et al.
790 2018). The pseudo-assembly was polished using POLCA (Zimin and Salzberg 2020).
791 Variation, between the two assemblies, was further evaluated using GAKT(McKenna et al.
792 2010) (version 4.0.1.2), in particular with HaplotypeCaller (Poplin et al. 2018) (--standard-
793 min-confidence-threshold-for-calling 30). Variants identified were summarized into a table
794 using GAKT VariantsToTable (-F CHROM -F POS -F TYPE -GF AD). Gene annotations were
795 transferred from the latest annotated version of Sp9509 (Ernst et al. 2023) obtained from
796 (https://www.lemna.org). Annotation transfer was performed using Liftoff (v1.6.3) (Shumate
797 and Salzberg 2021) using default parameters. Information regarding the phylogenetic
798 analysis and protein domain annotations of *Spirodela* silencing components can be found
799 within Supplemental Methods in Supplemental Material.
800
801 **Transposon annotation:**
802 The annotation of transposable elements (TEs) in *Spirodela* polyrhiza was performed by
803 combining EDTA (Ou et al. 2019) and RepeatModeler2 (Flynn et al. 2020) tools. To ensure
804 accuracy, both tools were run in triplicate and the resulting annotation libraries were aligned
805 and merged. The classification of TEs was further refined using DeepTE (Yan et al. 2020) (-
806 sp P). The presence of the protein domain for LTR-RT was evaluated using TEsorter (Zhang
807 et al. 2022) (version 1.3, -db rexdb-plant -st nucl).
808
809 **Transcriptome analysis:**
810 To extract RNA from *Spirodela*, 100 mg of plant material was flash-frozen upon harvesting
811 and ground. The samples were homogenized in 1 mL of TRIzol™ (ThermoFisher, #15596018)
812 using Silamat S6 (Ivoclar Vivadent). RNA was extracted according to manufacturer's
813 instructions. RNA library preparation was performed using the NEBNext® Ultra™ II RNA
814 Library Prep Kit for Illumina® (NEB #E7775) or the SMRTbell Prep Kit 3.0 for Illumina and Iso-
815 Seq respectively. Illumina sequencing was performed on HiSeq5A to generate 125 bp paired-

816 end reads. Long read sequencing was performed on PacBio® Sequel® II. For Illumina data,
817 base calls were performed using bcl2FASTQ (v2.17). RNA-seq reads were trimmed using
818 Trim Galore! (v0.6, --illumina -phred33 -stringency 1 --FASTQc --length 20). Trimmed reads
819 were mapped using HISAT2 (v2.1.0). Gene and transposable element expression levels were
820 quantified separately using kallisto (Bray et al. 2016) (v0.46) with default parameters. The
821 resulting expression levels (measured in TPM) were then averaged across the three
822 replicates. For *Arabidopsis* seedlings, data was obtained from (GSM6892967, GSM6892968,
823 GSM6892969) (Li et al. 2023). Iso-Seq long reads were processed using nf-core/isoseq
824 (https://github.com/nf-core/isoseq/tree/1.0.0). High quality transcripts were then mapped to
825 *Spirodela* genome using minimap2 (Li 2021) (v2.17, default parameters).
826
827 **DNA methylation analysis:**
828 For EMseq, DNA extraction was performed as above. Libraries were prepared using the
829 NEBNext® Enzymatic Methyl-seq Kit (NEB, #M7634) following the manufacturer's
830 instructions. Sequencing was performed on an Illumina NovaSeq SP to produce paired-end
831 reads of 150 bp. Briefly, basecalls were performed using bcl2FASTQ (v2.17) and sequenced
832 reads were quality filtered and adaptor trimmed using Trim Galore! version 0.6.2
833 (https://github.com/FelixKrueger/TrimGalore). Enzymatic-converted reads were aligned either
834 to the TAIR10 genome or *Spirodela* polyrhiza genome using Bismark version 0.22.2
835 with *Bismark* (--non_directional *-q* --score-min *L,0,–0.4*) (Krueger and Andrews 2011) to
836 generate coverage files per cytosine. Conversion rates were assessed by mapping (SAMtools
837 v1.9) the reads to the unmethylated lambda phage DNA, with the same parameters used for
838 the genomes BAM files containing clonal deduplicated and uniquely mapped reads were then
839 used to extract weighted methylation rates at each cytosine as described in (Schultz et al.
840 2015). Replicates were merged into a single bedFile using the BEDTools (Quinlan and Hall
841 2010) merge function and only cytosines with coverage with ≥4 were considered for further

842 analyses. 5mC was calculated using ViewBS (Huang et al. 2018) (v0.1.11)
843 (https://github.com/xie186/ViewBS) using default parameters.
844
845 **Small RNA library preparation and analysis:**
846 The libraries were prepared according to (Hayashi et al. 2016). In summary, small RNAs
847 purified from plant tissues or immunoprecipitated AGO (Supplemental Methods in
848 Supplemental Material), using TraPR (Grentzinger et al. 2020) or from total small RNA as
849 described in Supplemental Methods (Supplemental Material), were ligated to 3′ barcoded DNA
850 adapters using truncated T4 RNA ligase 2 (New England Biolabs, #M0373). These fragments
851 were separated in an 12% denaturing polyacrylamide-urea gel, excised, and purified using ZR
852 small-RNA PAGE Recovery Kit (Zymo Research, # R1070). The elution was then ligated to
853 5′ barcoded RNA adapters using T4 RNA ligase 1 (New Englland Biolabs, #M0204). To reduce
854 ligation biases, the sequence of both adaptors included 4 random nucleotides at the ends.
855 The resulting fragments were reverse transcribed and amplified by PCR using primers
856 compatible with the Illumina Platform. The size distribution of the libraries was assessed using
857 the 5200 Fragment Analyser System (Agilent, #M5310AA). Sequencing was performed on an
858 Illumina NextSeq550 to produce single-end reads of 75 bp. Basecalls were performed using
859 bcl2FASTQ (v2.17). Reads were trimmed and quality checked using Trimmomatic (Bolger et
860 al. 2014) (v0.39). Trimmed reads were mapped to *Arabidopsis* TAIR10 and *Spirodela*
861 genomes, and *RUBY* or *hpScarlet* sequences using Bowtie 2 (Langmead and Salzberg 2012)
862 (v1.2.2; -e 50 -a -v 0 --best --strata --nomaqround -y --phred33-quals --no-unal --sam). Mapped
863 reads were then normalized using bamCoverage (Ramírez et al. 2016) (DeepTools v,--
864 normalizeUsing CPM --binSize 1 --smoothLength 10 --maxFragmentLength 30). To visualise
865 strand, mapped reads were split using SAMtools view (Li et al. 2009) (SAMtools v1.9, using -
866 h -b -F 16 or -h -b -f 16 options).
867
868 **Chromatin immunoprecipitation (ChIP):**

869 Chromatin immunoprecipitation (ChIP) was conducted on 2 g of *Arabidopsis* 10-day-old
870 seedlings (WT Col-0) and *Spirodela polyrhiza* #9509 fronds. Following standard procedures
871 (see detailed method in Supplemental Methods in Supplemental Material) using the following
872 antibodies: anti-H3 (rabbit polyclonal, Abcam, #ab1791), anti-H3K27me3 (rabbit polyclonal,
873 Merck-Millipore, #07–449), anti-H3K4me3 (rabbit polyclonal, Abcam, #ab8580), anti-
874 H3K9me1 (rabbit polyclonal, Abcam, 3ab8896), anti-H3K9me2 (mouse monoclonal, Abcam,
875 #ab1220) or anti-H3K27me1 (rabbit polyclonal, Merck-Millipore, #17–643). Further detailed
876 information on ChIP sample preparation and histone PTM procedures and analysis by mass
877 spectrometry, can be found within Supplemental Methods in Supplemental Material.
878
879 **ChIP-seq library prep and data processing**:
880 For ChIP-seq, libraries were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for
881 Illumina® (NEB, #E7103) according to the manufacturer's instructions. Sequencing was
882 performed on an Illumina NovaSeq S4 to produce paired-end reads of 150 bp. The raw BAM
883 files were converted to FASTQ using BEDTools (Quinlan and Hall 2010) (v2.27.1, default
884 parameters) for quality assessment of the sequenced samples. FASTQc
885 (https://www.bioinformatics.babraham.ac.uk/projects/FASTQc/) (v0.11.8, default parameters)
886 was used to generate quality reports for all sequencing data. Basecalls were performed using
887 bcl2FASTQ (v2.17). Trimming of reads was performed using trim_galore
888 (https://zenodo.org/doi/10.5281/zenodo.5127898) (v0.6.5, --dont_gzip --stringency 1 --fastqc
889 --length 5'). The aligned reads were then mapped to either the TAIR10 *Arabidopsis* genome
890 or *Spirodela* polyrhiza genome using Bowtie 2 (Langmead and Salzberg 2012) (v2.4.1,
891 default parameters). Duplicate reads were removed using Picard
892 (http://broadinstitute.github.io/picard/) (v2.22.8, default parameters). Mapped reads were then
893 normalized using bamCoverage (deepTools v3.3.1). Correlations between ChIP samples
894 were evaluated using deepTools (Ramírez et al. 2016) (v3.1.2, default parameters). The
895 bamCompare function from deepTools was used to normalize ChIP samples to Input or H3

896 and produce log_2 ratio (ChIP/H3) bigWig files. Peak calling was performed using MACS2
897 (v2.2.5 –nomodel –nolambda –broad -q 0.01 -f BAMPE -g 138550563) (Zhang et al. 2008).
898
899 **Plasmids used in this study**:
900 All plasmids generated have been submitted to Addgene (www.addgene.org) and can be
901 retrieved under the following IDs: p*35S*:*FHA*-At*AGO4*_gDNA (#216838); p*35S*:*FHA*-
902 Sp*AGO4a*_gDNA (#216841); p*35S*::*FHA*-Sp*AGO4a*_cDNA (#216842). The following
903 plasmids previously originated in (He et al. 2020) and (Oberlin et al. 2017) were obtained
904 from Addgene: p*35S*:*RUBY* (#160908); p*ZmUbq*:*RUBY* (#160909); p*35S*:*GFP-GUS*
905 (#167122). p3At*UBQ*:*hpScarlet* was a gift from Marco Incarbone (Max-Planck-Institute of
906 Plant Physiology, Golm, Germany). Cloning and transient expression procedures in both
907 *Spirodela* and Tobacco can be found in Supplemental Methods (Supplemental Material).
908 Primers used for cloning can be found in **Supplemental Fig. S39**.
909
910 **Small RNA and protein blotting**:
911 RNA and protein gels, blotting and detection were performed using standard procedures
912 detailed in Supplemental Methods (Supplemental Material) together with probes
913 (**Supplemental Fig. S39**) and antibodies used in this study.
914
915 **Staining of nuclei**:
916 DAPI-staining of interphase nuclei was performed as described in Supplemental Methods
917 (Supplemental Material).
918
919 **Visualization and Statistical analysis:**
920 Data visualization files were generated according to each specific sections. Statistical tests
921 were performed using R-dependent ggpubr (v0.6.0) (https://rpkgs.datanovia.com/ggpubr/),
922 statistical test performed are indicated accordingly on each figure legend figure. Plots under
923 R (v4.2.2) (https://www.r-project.org) were generated using Circlize (Gu et al. 2014) (v0.4.15),

924 ggplot2 (Wickham 2016) (v3.4.4), profileplyr (v1.14.1)
925 (https://bioconductor.org/packages/release/bioc/html/profileplyr.html), ViewBS (Huang et al.
926 2018) (v0.1.11), or pheatmap (v 1.0.12)
927 (https://rdocumentation.org/packages/pheatmap/versions/1.0.12).
928
929 **Raw data and source files:**
930 An inventory of all the raw data used to generate each figure panel (as well as those in
931 Supplementary information) along with all raw image files (including RNA, protein blots and
932 agarose gels), DNA and protein sequences, alignments, and phylogenetic tree files can be
933 found in Zenodo (www.zenodo.org) under the doi: 10.5281/zenodo.14037413. Additionally,
934 ready-to-use IGV genome browser tracks for all the sequencing datasets generated here are
935 available as well in the associated Zenodo dataset.
936

937 Data access

938 All the sequencing data generated for this study have been submitted to the NCBI
939 BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession number
940 PRJNA1164696. The mass spectrometry proteomics data have been deposited to the
941 ProteomeXchange Consortium via the PRIDE partner repository
942 (https://www.ebi.ac.uk/pride/) with the dataset identifier PXD050443.
943

944 Competing interests

945 The authors declare that they have no competing interests.
946

947 Acknowledgements

948 This work was supported by the Gregor Mendel Institute of Plant Molecular Biology
949 (GMI) of the Austrian Academy of Sciences (ÖAW) core funding attributed to Arturo Marí-
950 Ordóñez. 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 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 analysis of histone PTMs; the Plant Science Facility for assistance with plant work; the IMP Molecular Biology Service for providing reagents and Sanger sequencing. Lastly, we would like to thank the VBC in-house COVID-19 testing facility for enabling a safe working environment during the pandemic.

Author contributions

AM-O conceived and designed the study. RD, VB-B and DB-A performed most experiments. AM-O, RD and D-AB 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 staining. 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 and DB-A performed computer and statistical analysis. AM-O, RD and DB-A analyzed the data and wrote the manuscript. AMO, DB-A and LD-N addressed referee's comments and prepared the revised version of the manuscript.

References

Abramson BW, Novotny M, Hartwick NT, Colt K, Aevermann BD, Scheuermann RH, Michael TP. 2021. The genome and preliminary single-nuclei transcriptome of Lemna minuta reveals mechanisms of invasiveness. *Plant Physiol* **188**: 879–897.

Allshire RC, Madhani HD. 2018. Ten principles of heterochromatin formation and function. *Nature Reviews Molecular Cell Biology* **19**: 229–244.

An D, Zhou Y, Li C, Xiao Q, Wang T, Zhang Y, Wu Y, Li Y, Chao D-Y, Messing J, et al. 2019. Plant evolution and environmental adaptation unveiled by long-read whole-genome sequencing of Spirodela. *Proc Natl Acad Sci* **116**: 18893–18899.

Andrieu A, Breitler JC, Siré C, Meynard D, Gantet P, Guiderdoni E. 2012. An in planta, Agrobacterium-mediated transient gene expression method for inducing gene silencing in rice (Oryza sativa L.) leaves. *Rice* **5**: 23.

Arkhipova I, Meselson M. 2000. Transposable elements in sexual and ancient asexual taxa. *Proc Natl Acad Sci* **97**: 14473–14477.

Bast J, Jaron KS, Schuseil D, Roze D, Schwander T. 2019. Asexual reproduction reduces transposable element load in experimental yeast populations. *eLife* **8**: e48548.

Benachenhou F, Sperber GO, Bongcam-Rudloff E, Andersson G, Boeke JD, Blomberg J. 2013. Conserved structure and inferred evolutionary history of long terminal repeats (LTRs). *Mob DNA* **4**: 5.

Bertazzon N, Raiola A, Castiglioni C, Gardiman M, Angelini E, Borgo M, Ferrari S. 2012. Transient silencing of the grapevine gene VvPGIP1 by agroinfiltration with a construct for RNA interference. *Plant Cell Rep* **31**: 133–143.

Bewick AJ, Ji L, Niederhuth CE, Willing E-M, Hofmeister BT, Shi X, Wang L, Lu Z, Rohr NA, Hartwig B, et al. 2016. On the origin and evolutionary consequences of gene body DNA methylation. *Proc Natl Acad Sci* **113**: 9111–9116.

Bewick AJ, Niederhuth CE, Ji L, Rohr NA, Griffin PT, Leebens-Mack J, Schmitz RJ. 2017. The evolution of CHROMOMETHYLASES and gene body DNA methylation in plants. *Genome Biol* **18**: 65.

Bewick AJ, Schmitz RJ. 2017. Gene body DNA methylation in plants. *Curr Opin Plant Biol* **36**: 103–110.

Biémont C, Vieira C. 2006. Genetics: junk DNA as an evolutionary force. October 5.

Bies-Etheve N, Pontier D, Lahmy S, Picart C, Vega D, Cooke R, Lagrange T. 2009. RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. *EMBO Rep* **10**: 649–654.

Blevins T, Pontes O, Pikaard CS, Jr FM. 2009. Heterochromatic siRNAs and DDM1 Independently Silence Aberrant 5S rDNA Transcripts in Arabidopsis ed. B.A. Sullivan. *PLoS ONE* **4**: e5932.

Blumenstiel JP. 2019. Birth, School, Work, Death, and Resurrection: The Life Stages and Dynamics of Transposable Element Proliferation. *Genes* **10**: 336.

Bog M, Appenroth K-J, Sree KS. 2019. Duckweed (Lemnaceae): Its Molecular Taxonomy. *Front Sustain Food Syst* **3**: 117.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

1017 Bourc'his D, Voinnet O. 2010. A Small-RNA Perspective on Gametogenesis, Fertilization, and Early Zygotic
1018 Development. *Science* **330**: 617–622.

1019 Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, Imbeault M, Izsvák Z, Levin HL,
1020 Macfarlan TS, et al. 2018. Ten things you should know about transposable elements. *Genome Biol* **19**: 199.

1021 Bradamante G, Nguyen VH, Incarbone M, Meir Z, Bente H, Donà M, Lettner N, Scheid OM, Gutzat R. 2023.
1022 Two ARGONAUTE proteins loaded with transposon-derived small RNAs are associated with the
1023 reproductive cell lineage in Arabidopsis. *Plant Cell* koad295.

1024 Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat*
1025 *Biotechnol* **34**: 525–527.

1026 Brosseau C, Moffett P. 2015. Functional and Genetic Analysis Identify a Role for Arabidopsis ARGONAUTE5
1027 in Antiviral RNA Silencing. *Plant Cell* **27**: 1742–1754.

1028 Cao HX, Vu GTH, Wang W, Messing J, Schubert I. 2015. Chromatin organisation in duckweed interphase nuclei
1029 in relation to the nuclear DNA content. *Plant Biol* **17**: 120–124.

1030 Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert KB, Montgomery TA, Nguyen T, Cuperus JT, Carrington JC.
1031 2012. Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. *THE PLANT*
1032 *CELL ONLINE* **24**: 3613–3629.

1033 Cervantes-Pérez SA, Yong-Villalobos L, Florez-Zapata NMV, Oropeza-Aburto A, Rico-Reséndiz F, Amasende-
1034 Morales I, Lan T, Martínez O, Vielle-Calzada JP, Albert VA, et al. 2021. Atypical DNA methylation, sRNA-
1035 size distribution, and female gametogenesis in Utricularia gibba. *Sci Rep* **11**: 15725.

1036 Chan SW-L, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE. 2004. RNA Silencing Genes Control
1037 de Novo DNA Methylation. *Science* **303**: 1336–1336.

1038 Chen G, Stepanenko A, Borisjuk N. 2024. Contrasting patterns of 5S rDNA repeats in European and Asian
1039 ecotypes of greater duckweed, Spirodela polyrhiza (Lemnaceae). *Front Plant Sci* **15**: 1378683.

1040 Chen G, Stepanenko A, Borisjuk N. 2021. Mosaic Arrangement of the 5S rDNA in the Aquatic Plant Landoltia
1041 punctata (Lemnaceae). *Front Plant Sci* **12**: 678689.

1042 Chen H-M, Chen L-T, Patel K, Li Y-H, Baulcombe DC, Wu S-H. 2010. 22-Nucleotide RNAs trigger secondary
1043 siRNA biogenesis in plants. *Proceedings of the National Academy of Sciences* **107**: 15269–15274.

1044 Chen W, Zhang X, Fan Y, Li B, Ryabov E, Shi N, Zhao M, Yu Z, Qin C, Zheng Q, et al. 2018. A Genetic Network
1045 for Systemic RNA Silencing in Plants . *Plant Physiol* **176**: 2700–2719.

1046 Chénais B, Caruso A, Hiard S, Casse N. 2012. The impact of transposable elements on eukaryotic genomes: From
1047 genome size increase to genetic adaptation to stressful environments. *Gene* **509**: 7–15.

1048 Cho SH, Addo-Quaye C, Coruh C, Arif MA, Ma Z, Frank W, Axtell MJ. 2008. Physcomitrella patens DCL3 Is
1049 Required for 22–24 nt siRNA Accumulation, Suppression of Retrotransposon-Derived Transcripts, and
1050 Normal Development. *PLoS Genet* **4**: e1000314.

1051 Choi J, Lyons DB, Zilberman D. 2021. Histone H1 prevents non-CG methylation-mediated small RNA biogenesis
1052 in Arabidopsis heterochromatin. *eLife* **10**: e72676.

1053 Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M,
1054 Jacobsen SE. 2008. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation
1055 patterning. *Nature* **452**: 215–219.

1056 Creasey KM, Zhai J, Borges F, Ex FV, Regulski M, Meyers BC, Martienssen RA. 2014. miRNAs trigger
1057 widespread epigenetically activated siRNAs from transposons in Arabidopsis. *Nature* **508**: 411–415.

1058 Czech B, Munafò M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E, Hannon GJ. 2018. piRNA-Guided Genome
1059 Defense: From Biogenesis to Silencing. *Annu Rev Genet* **52**: 131–157.

1060 Dechaud C, Volff J-N, Schartl M, Naville M. 2019. Sex and the TEs: transposable elements in sexual development
1061 and function in animals. *Mob DNA* **10**: 42.

1062 Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. 2006. Hierarchical action and
1063 inhibition of plant Dicer-like proteins in antiviral defense. *Science* **313**: 68–71.

1064 Diego-Martin B, Pérez-Alemany J, Candela-Ferre J, Corbalán-Acedo A, Pereyra J, Alabadí D, Jami-Alahmadi Y,
1065 Wohlschlegel J, Gallego-Bartolomé J. 2022. The TRIPLE PHD FINGERS proteins are required for SWI/SNF
1066 complex-mediated +1 nucleosome positioning and transcription start site determination in Arabidopsis.
1067 *Nucleic Acids Res* **50**: 10399–10417.

1068 Ding Y, Fromm M, Avramova Z. 2012. Multiple exposures to drought "train" transcriptional responses in
1069 Arabidopsis. *Nat Commun* **3**: 740.

1070 Dong J, LeBlanc C, Poulet A, Mermaz B, Villarino G, Webb KM, Joly V, Mendez J, Voigt P, Jacob Y. 2021.
1071 H3.1K27me1 maintains transcriptional silencing and genome stability by preventing GCN5-mediated histone
1072 acetylation. *Plant Cell* **33**: 961–979.

1073 Duan C-G, Wang X, Zhang L, Xiong X, Zhang Z, Tang K, Pan L, Hsu C-C, Xu H, Tao WA, et al. 2017. A protein
1074 complex regulates RNA processing of intronic heterochromatin-containing genes in Arabidopsis. *Proc Natl*
1075 *Acad Sci* **114**: E7377–E7384.

1076 Duan C-G, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, Wang B, Lang Z, Zhao Y, Wang X, et al. 2014. Specific
1077 but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. *EMBO*
1078 e201489453.

1079 Ebbs ML, Bender J. 2006. Locus-Specific Control of DNA Methylation by the Arabidopsis SUVH5 Histone
1080 Methyltransferase. *Plant Cell* **18**: 1166–1176.

1081 El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, Hakimi MA, Jacobsen SE, Cooke R, Lagrange T.
1082 2007. Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding
1083 platforms in RNAi-related components. *Genes Development* **21**: 2539–2544.

1084 Erdmann RM, Picard CL. 2020. RNA-directed DNA Methylation. *Plos Genet* **16**: e1009034.

1085 Ernst E, Abramson B, Acosta K, Hoang PTN, Mateo-Elizalde C, Schubert V, Pasaribu B, Hartwick N, Colt K,
1086 Aylward A, et al. 2023. The genomes and epigenomes of aquatic plants (Lemnaceae) promote triploid
1087 hybridization and clonal reproduction. *bioRxiv* **14**: 2023.08.02.551673.

1088 Fang J, Jiang J, Leichter SM, Liu J, Biswal M, Khudaverdyan N, Zhong X, Song J. 2022. Mechanistic basis for
1089 maintenance of CHG DNA methylation in plants. *Nat Commun* **13**: 3877.

1090 Fang J, Leichter SM, Jiang J, Biswal M, Lu J, Zhang Z-M, Ren W, Zhai J, Cui Q, Zhong X, et al. 2021. Substrate
1091 deformation regulates DRM2-mediated DNA methylation in plants. *Sci Adv* **7**: eabd9224.

1092 Fang X, Qi Y. 2016. RNAi in Plants: An Argonaute-Centered View. *Plant Cell* **28**: 272–285.

1093 Fang Y, Tian X, Jin Y, Du A, Ding Y, Liao Z, He K, Zhao Y, Guo L, Xiao Y, et al. 2023. Duckweed evolution:
1094 from land back to water. *bioRxiv* 2023.03.22.533731.

1095 Fei Q, Yang L, Liang W, Zhang D, Meyers BC. 2016. Dynamic changes of small RNAs in rice spikelet
1096 development reveal specialized reproductive phasiRNA pathways. *J Exp Bot* **67**: 6037–6049.

1097 Feng S, Zhong Z, Wang M, Jacobsen SE. 2020. Efficient and accurate determination of genome-wide DNA
1098 methylation patterns in Arabidopsis thaliana with enzymatic methyl sequencing. *Epigenetics Chromatin* **13**:
1099 42.

1100 Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet* **9**: 397–405.

1101 Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated
1102 genomic discovery of transposable element families. *Proc Natl Acad Sci* **117**: 9451–9457.

1103 Fourounjian P, Slovin J, Messing J. 2021. Flowering and Seed Production across the Lemnaceae. *Int J Mol Sci*
1104 **22**: 2733.

1105 Fourounjian P, Tang J, Tanyolac B, Feng Y, Gelfand B, Kakrana A, Tu M, Wakim C, Meyers BC, Ma J, et al.
1106 2019. Post-transcriptional adaptation of the aquatic plant Spirodela polyrhiza under stress and hormonal
1107 stimuli. *Plant J* **98**: 1120–1133.

1108 Fu S, Gao Z, Birchler J, Han F. 2012. Dicentric Chromosome Formation and Epigenetics of Centromere
1109 Formation in Plants. *J Genet Genom* **39**: 125–130.

1110 Fuchs J, Demidov D, Houben A, Schubert I. 2006. Chromosomal histone modification patterns – from
1111 conservation to diversity. *Trends Plant Sci* **11**: 199–208.

1112 Fusaro AF, Matthew L, Smith NA, Curtin SJ, Dedic-Hagan J, Ellacott GA, Watson JM, Wang M, Brosnan C,
1113 Carroll BJ, et al. 2006. RNA interference-inducing hairpin RNAs in plants act through the viral defence
1114 pathway. *EMBO Rep* **7**: 1168–1175.

1115 Gaillochet C, Lohmann JU. 2015. The never-ending story: from pluripotency to plant developmental plasticity.
1116 *Development* **142**: 2237–2249.

1117 Garcia-Ruiz H, Takeda A, Chapman EJ, Sullivan CM, Fahlgren N, Brempelis KJ, Carrington JC. 2010.
1118 Arabidopsis RNA-Dependent RNA Polymerases and Dicer-Like Proteins in Antiviral Defense and Small
1119 Interfering RNA Biogenesis during Turnip Mosaic Virus Infection. *THE PLANT CELL ONLINE* **22**: 481–
1120 496.

1121 Goodall GJ, Filipowicz W. 1991. Different effects of intron nucleotide composition and secondary structure on
1122 pre-mRNA splicing in monocot and dicot plants. *EMBO J* **10**: 2635–2644.

1123 Gorinšek B, Gubenšek F, Kordiš D. 2004. Evolutionary Genomics of Chromoviruses in Eukaryotes. *Mol Biol*
1124 *Evol* **21**: 781–798.

1125 Gouil Q, Baulcombe DC. 2016. DNA Methylation Signatures of the Plant Chromomethyltransferases. *PLoS Genet*
1126 **12**: e1006526.

1127 Grentzinger T, Oberlin S, Schott G, Handler D, Svozil J, Barragan-Borrero V, Humbert A, Duharcourt S,
1128 Brennecke J, Voinnet O. 2020. A universal method for the rapid isolation of all known classes of functional
1129 silencing small RNAs. *Nucleic Acids Res* **48**: gkaa472-.

1130 Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. circlize implements and enhances circular visualization in R.
1131 *Bioinformatics* **30**: 2811–2812.

1132 Gutzat R, Rembart K, Nussbaumer T, Hofmann F, Pisupati R, Bradamante G, Daubel N, Gaidora A, Lettner N,
1133 Donà M, et al. 2020. Arabidopsis shoot stem cells display dynamic transcription and DNA methylation
1134 patterns. *EMBO J* **39**: e103667.

1135 Haag JR, Brower-Toland B, Krieger EK, Sidorenko L, Nicora CD, Norbeck AD, Irsigler A, LaRue H, Brzeski J,
1136 McGinnis K, et al. 2014. Functional Diversification of Maize RNA Polymerase IV and V Subtypes via
1137 Alternative Catalytic Subunits. *Cell Rep* **9**: 378–390.

Hamilton AJ, Voinnet O, Chappell L, Baulcombe DC. 2002. Two classes of short interfering RNA in RNA silencing. *Embo Journal* **21**: 4671–4679.

Harkess A, Bewick AJ, Lu Z, Fourounjian P, Michael TP, Schmitz RJ, Meyers BC. 2024. Unusual predominance of maintenance DNA methylation in Spirodela polyrhiza. *G3: Genes, Genomes, Genet* jkae004.

Harkess A, McLoughlin F, Bilkey N, Elliott K, Emenecker R, Mattoon E, Miller K, Czymmek K, Vierstra RD, Meyers BC, et al. 2021. Improved Spirodela polyrhiza genome and proteomic analyses reveal a conserved chromosomal structure with high abundance of chloroplastic proteins favoring energy production. *J Exp Bot* **72**: 2491–2500.

Harvey JJW, Lewsey MG, Patel K, Westwood J, Heimstädt S, Carr JP, Baulcombe DC. 2011. An Antiviral Defense Role of AGO2 in Plants ed. M. Bendahmane. *PLoS ONE* **6**: e14639.

Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC. 2010. The Arabidopsis RNA-Directed DNA Methylation Argonautes Functionally Diverge Based on Their Expression and Interaction with Target Loci. *THE PLANT CELL ONLINE* **22**: 321–334.

Hayashi R, Schnabl J, Handler D, Mohn F, Ameres SL, Brennecke J. 2016. Genetic and mechanistic diversity of piRNA 3′-end formation. *Nature* **539**: 588–592.

He L, Zhao C, Zhang Q, Zinta G, Wang D, Lozano-Durán R, Zhu J-K. 2021. Pathway conversion enables a double-lock mechanism to maintain DNA methylation and genome stability. *Proc Natl Acad Sci* **118**: e2107320118.

He Y, Zhang T, Sun H, Zhan H, Zhao Y. 2020. A reporter for noninvasively monitoring gene expression and plant transformation. *Hortic Res* **7**: 152.

Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE. 2006. Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nature Genetics* **38**: 721–725.

Henikoff S, Smith MM. 2015. Histone Variants and Epigenetics. *Cold Spring Harb Perspect Biol* **7**: a019364.

Hisanaga T, Romani F, Wu S, Kowar T, Wu Y, Lintermann R, Fridrich A, Cho CH, Chaumier T, Jamge B, et al. 2023. The Polycomb repressive complex 2 deposits H3K27me3 and represses transposable elements in a broad range of eukaryotes. *Curr Biol* **33**: 4367-4380.e9.

Ho EKH, Bartkowska M, Wright SI, Agrawal AF. 2019. Population genomics of the facultatively asexual duckweed Spirodela polyrhiza. *N Phytol* **224**: 1361–1371.

Hoang PNT, Michael TP, Gilbert S, Chu P, Motley ST, Appenroth KJ, Schubert I, Lam E. 2018. Generating a high-confidence reference genome map of the Greater Duckweed by integration of cytogenomic, optical mapping, and Oxford Nanopore technologies. *Plant J* **96**: 670–684.

Hoang PTN, Fiebig A, Novák P, Macas J, Cao HX, Stepanenko A, Chen G, Borisjuk N, Scholz U, Schubert I. 2020. Chromosome-scale genome assembly for the duckweed Spirodela intermedia, integrating cytogenetic maps, PacBio and Oxford Nanopore libraries. *Sci Rep-uk* **10**: 19230.

Houben A, Demidov D, Gernand D, Meister A, Leach CR, Schubert I. 2003. Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. *Plant J* **33**: 967–973.

Huang X, Zhang S, Li K, Thimmapuram J, Xie S, Wren J. 2018. ViewBS: a powerful toolkit for visualization of high-throughput bisulfite sequencing data. *Bioinformatics* **34**: 708–709.

Huang Y, Kendall T, Forsythe ES, Dorantes-Acosta A, Li S, Caballero-Pérez J, Chen X, Arteaga-Vázquez M, Beilstein MA, Mosher RA. 2015. Ancient Origin and Recent Innovations of RNA Polymerase IV and V. *Mol Biol Evol* **32**: 1788–1799.

1180 Huff JT, Zilberman D, Roy SW. 2016. Mechanism for DNA transposons to generate introns on genomic scales.
1181 *Nature* **538**: 533–536.

1182 Ibarra-Laclette E, Lyons E, Hernández-Guzmán G, Pérez-Torres CA, Carretero-Paulet L, Chang T-H, Lan T,
1183 Welch AJ, Juárez MJA, Simpson J, et al. 2013. Architecture and evolution of a minute plant genome. *Nature*
1184 **498**: 94–98.

1185 Ilık İA, Glažar P, Tse K, Brändl B, Meierhofer D, Müller F-J, Smith ZD, Aktaş T. 2024. Autonomous transposons
1186 tune their sequences to ensure somatic suppression. *Nature* 1–9.

1187 Incarbone M, Bradamante G, Pruckner F, Wegscheider T, Rozhon W, Nguyen V, Gutzat R, Mérai Z, Lendl T,
1188 MacFarlane S, et al. 2023. Salicylic acid and RNA interference mediate antiviral immunity of plant stem cells.
1189 *Proc Natl Acad Sci* **120**: e2302069120.

1190 Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. 2011. An siRNA pathway prevents
1191 transgenerational retrotransposition in plants subjected to stress. *Nature* **472**: 115–119.

1192 Jackson JP, Johnson L, Jasencakova Z, Zhang X, PerezBurgos L, Singh PB, Cheng X, Schubert I, Jenuwein T,
1193 Jacobsen SE. 2004. Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene
1194 silencing in Arabidopsis thaliana. *Chromosoma* **112**: 308–315.

1195 Jackson JP, Lindroth AM, Cao X, Jacobsen SE. 2002. Control of CpNpG DNA methylation by the KRYPTONITE
1196 histone H3 methyltransferase. *Nature* **416**: 556–560.

1197 Jacob Y, Feng S, LeBlanc CA, Bernatavichute YV, Stroud H, Cokus S, Johnson LM, Pellegrini M, Jacobsen SE,
1198 Michaels SD. 2009. ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure
1199 and gene silencing. *Nat Struct Mol Biol* **16**: 763–768.

1200 Jamge B, Lorković ZJ, Axelsson E, Osakabe A, Shukla V, Yelagandula R, Akimcheva S, Kuehn AL, Berger F.
1201 2023. Histone variants shape chromatin states in Arabidopsis. *eLife* **12**: RP87714.

1202 Jasencakova Z, Soppe WJJ, Meister A, Gernand D, Turner BM, Schubert I. 2003. Histone modifications in
1203 Arabidopsis– high methylation of H3 lysine 9 is dispensable for constitutive heterochromatin. *Plant J* **33**:
1204 471–480.

1205 Jay F, Brioudes F, Voinnet O. 2023. A contemporary reassessment of the enhanced transient expression system
1206 based on the tombusviral silencing suppressor protein P19. *Plant J* **113**: 186–204.

1207 Jia J, Ji R, Li Z, Yu Y, Nakano M, Long Y, Feng L, Qin C, Lu D, Zhan J, et al. 2020. Soybean DICER-LIKE2
1208 Regulates Seed Coat Color via Production of Primary 22-Nucleotide Small Interfering RNAs from Long
1209 Inverted Repeats. *Plant Cell* **32**: 3662–3673.

1210 Johnson L, Mollah S, Garcia BA, Muratore TL, Shabanowitz J, Hunt DF, Jacobsen SE. 2004. Mass spectrometry
1211 analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. *Nucleic*
1212 *Acids Res* **32**: 6511–6518.

1213 Johnson LM, Bostick M, Zhang X, Kraft E, Henderson I, Callis J, Jacobsen SE. 2007. The SRA Methyl-Cytosine-
1214 Binding Domain Links DNA and Histone Methylation. *Curr Biol* **17**: 379–384.

1215 Johnson LM, Du J, Hale CJ, Bischof S, Feng S, Chodavarapu RK, Zhong X, Marson G, Pellegrini M, Segal DJ,
1216 et al. 2014. SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA
1217 methylation. *Nature* **507**: 124–128.

1218 Johnson LM, Law JA, Khattar A, Henderson IR, Jacobsen SE. 2008. SRA-Domain Proteins Required for DRM2-
1219 Mediated De Novo DNA Methylation. *PLoS Genet* **4**: e1000280.

1220 Jullien PE, Grob S, Marchais A, Pumplin N, Chevalier C, Bonnet DMV, Otto C, Schott G, Voinnet O. 2020.
1221 Functional characterization of Arabidopsis ARGONAUTE 3 in reproductive tissues. *Plant J* **103**: 1796–1809.

1222 Keith B, Chua N-H. 1986. Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic
1223 tobacco. *EMBO J* **5**: 2419–2425.

1224 Kim MY, Zilberman D. 2014. DNA methylation as a system of plant genomic immunity. *Trends Plant Sci* **19**:
1225 320–326.

1226 Klein SP, Anderson SN. 2022. The evolution and function of transposons in epigenetic regulation in response to
1227 the environment. *Curr Opin Plant Biol* **69**: 102277.

1228 Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications.
1229 *Bioinformatics* **27**: 1571–1572.

1230 Kumar A, Bennetzen JL. 1999. Plant retrotransposons. *Annual Review of Genetics* **33**: 479–532.

1231 Lam E, Michael TP. 2022. Wolffia, a minimalist plant and synthetic biology chassis. *Trends Plant Sci* **27**: 430–
1232 439.

1233 Landolt E. 1986. *Biosystematic investigations in the family of duckweeds (Lemnaceae) (Vol 2). The family of*
1234 *Lemnaceae—a monographic study*. ed. V. des G.I. Der ETH. Veroeffentlichungen des Geobotanischen
1235 Instituts der ETH, Stiftung Ruebel, Switzerland.

1236 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

1237 Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AMS, Strahl BD, Patel DJ, Jacobsen SE. 2013. Polymerase
1238 IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* 1–7.

1239 Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE. 2011. SHH1, a Homeodomain Protein Required for DNA
1240 Methylation, As Well As RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA
1241 Polymerase IV. *PLoS Genet* **7**: e1002195.

1242 Lee SC, Ernst E, Berube B, Borges F, Parent J-S, Ledon P, Schorn A, Martienssen RA. 2020. Arabidopsis
1243 retrotransposon virus-like particles and their regulation by epigenetically activated small RNA. *Genome*
1244 *Research* **30**: 576–588.

1245 Lee Y-S, Maple R, Dürr J, Dawson A, Tamim S, Genio C del, Papareddy R, Luo A, Lamb JC, Amantia S, et al.
1246 2021. A transposon surveillance mechanism that safeguards plant male fertility during stress. *Nat Plants* **7**:
1247 34–41.

1248 Lemon GD, Posluszny U. 2000. Comparative Shoot Development and Evolution in the Lemnaceae. *Int J Plant*
1249 *Sci* **161**: 733–748.

1250 Li F, Yang J-J, Sun Z-Y, Wang L, Qi L-Y, A S, Liu Y-Q, Zhang H-M, Dang L-F, Wang S-J, et al. 2023. Plant-
1251 on-chip: Core morphogenesis processes in the tiny plant Wolffia australiana. *PNAS Nexus* **2**: pgad141.

1252 Li H. 2021. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* **37**: btab705-.

1253 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup 1000
1254 Genome Project Data Processing. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*
1255 **25**: 2078–2079.

1256 Li X, Harris CJ, Zhong Z, Chen W, Liu R, Jia B, Wang Z, Li S, Jacobsen SE, Du J. 2018. Mechanistic insights
1257 into plant SUVH family H3K9 methyltransferases and their binding to context-biased non-CG DNA
1258 methylation. *Proc Natl Acad Sci* **115**: E8793–E8802.

1259 Li Z, Wang M, Zhong Z, Gallego-Bartolomé J, Feng S, Jami-Alahmadi Y, Wang X, Wohlschlegel J, Bischof S,
1260 Long JA, et al. 2023. The MOM1 complex recruits the RdDM machinery via MORC6 to establish de novo
1261 DNA methylation. *Nat Commun* **14**: 4135.

Liu J, He Z. 2020. Small DNA Methylation, Big Player in Plant Abiotic Stress Responses and Memory. *Front Plant Sci* **11**: 595603.

Liu W, Duttke SH, Hetzel J, Groth M, Feng S, Gallego-Bartolome J, Zhong Z, Kuo HY, Wang Z, Zhai J, et al. 2018. RNA-directed DNA methylation involves co-transcriptional small-RNA-guided slicing of polymerase V transcripts in Arabidopsis. *Nat Plants* **4**: 181–188.

Liu Y, Teng C, Xia R, Meyers BC. 2020. PhasiRNAs in Plants: Their Biogenesis, Genic Sources, and Roles in Stress Responses, Development, and Reproduction. *THE PLANT CELL ONLINE* **32**: 3059–3080.

Liu Z-W, Shao C-R, Zhang C-J, Zhou J-X, Zhang S-W, Li L, Chen S, Huang H-W, Cai T, He X-J. 2014. The SET Domain Proteins SUVH2 and SUVH9 Are Required for Pol V Occupancy at RNA-Directed DNA Methylation Loci. *PLoS Genetics* **10**: e1003948.

Llave C, Kasschau KD, Rector MA, Carrington JC. 2002. Endogenous and Silencing-Associated Small RNAs in Plants[W]. *Plant Cell* **14**: 1605–1619.

Loffer A, Singh J, Fukudome A, Mishra V, Wang F, Pikaard CS. 2022. A DCL3 dicing code within Pol IV-RDR2 transcripts diversifies the siRNA pool guiding RNA-directed DNA methylation. *eLife* **11**: e73260.

Long J, Walker J, She W, Aldridge B, Gao H, Deans S, Vickers M, Feng X. 2021. Nurse cell–derived small RNAs define paternal epigenetic inheritance in Arabidopsis. *Science* **373**.

Lou H, McCullough AJ, Schuler MA. 1993. Expression of maize Adh1 intron mutants in tobacco nuclei. *Plant J* **3**: 393–403.

Malone CD, Hannon GJ. 2009. Small RNAs as Guardians of the Genome. *Cell* **136**: 656–668.

Marí-Ordóñez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. 2013. Reconstructing de novo silencing of an active plant retrotransposon. *Nature Genetics* **45**: 1029–1039.

Mathieu O, Probst AV, Paszkowski J. 2005. Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in Arabidopsis. **24**: 2783–2791.

McCue AD, Panda K, Nuthikattu S, Choudury SG, Thomas EN, Slotkin RK. 2015. ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. *The EMBO journal* **34**: 20–35.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**: 1297–1303.

Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, et al. 2008. Sorting of Small RNAs into Arabidopsis Argonaute Complexes Is Directed by the 5′ Terminal Nucleotide. *Cell* **133**: 116–127.

Michael TP, Bryant D, Gutierrez R, Borisjuk N, Chu P, Zhang H, Xia J, Zhou J, Peng H, Baidouri ME, et al. 2017. Comprehensive definition of genome features in Spirodela polyrhiza by high-depth physical mapping and short-read DNA sequencing strategies. *Plant J* **89**: 617–635.

Michael TP, Ernst E, Hartwick N, Chu P, Bryant D, Gilbert S, Ortleb S, Baggs EL, Sree KS, Appenroth KJ, et al. 2020. Genome and time-of-day transcriptome of Wolffia australiana link morphological minimization with gene loss and less growth control. *Genome Res* **31**: gr.266429.120.

Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, Cao J, Weigel D, Paszkowski J, Mathieu O. 2009. Selective epigenetic control of retrotransposition in Arabidopsis. *Nature* **461**: 427–430.

Moazed D. 2009. Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**: 413–420.

1302 Montavon T, Kwon Y, Zimmermann A, Hammann P, Vincent T, Cognat V, Michel F, Dunoyer P. 2017. A specific
1303 dsRNA-binding protein complex selectively sequesters endogenous inverted-repeat siRNA precursors and
1304 inhibits their processing. *Nucleic Acids Res* **45**: 1330–1344.

1305 Montgomery SA, Tanizawa Y, Galik B, Wang N, Ito T, Mochizuki T, Akimcheva S, Bowman JL, Cognat V,
1306 Maréchal-Drouard L, et al. 2020. Chromatin Organization in Early Land Plants Reveals an Ancestral
1307 Association between H3K27me3, Transposons, and Constitutive Heterochromatin. *Curr Biol* **30**: 573-588.e7.

1308 Morello L, Braglia L, A.Iannelli M, Lauria M, Gavazzi F, Gianì S, Perna C, Bucci G, Ianigro M. 2024. IBBA
1309 Duckweed Dataset (DUCKWEED-01). *Consiglio Nazionale delle Ricerche, Istituto di Bioscienze e*
1310 *BioRisorse (CNR-IBBR) Occurrence dataset*. https://doi.org/10.15468/hstsnr.

1311 Neumann P, Novák P, Hoštáková N, Macas J. 2019. Systematic survey of plant LTR-retrotransposons elucidates
1312 phylogenetic relationships of their polyprotein domains and provides a reference for element classification.
1313 *Mob DNA* **10**: 1.

1314 Nguyen V, Gutzat R. 2022. Epigenetic regulation in the shoot apical meristem. *Curr Opin Plant Biol* **69**: 102267.

1315 Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q, Rohr NA, Rambani A, Burke JM, Udall JA, et al.
1316 2016. Widespread natural variation of DNA methylation within angiosperms. *Genome Biol* **17**: 194.

1317 Nielsen CPS, Arribas-Hernández L, Han L, Reichel M, Woessmann J, Daucke R, Bressendorff S, López-Márquez
1318 D, Andersen SU, Pumplin N, et al. 2024. Evidence for an RNAi-independent role of Arabidopsis DICER-
1319 LIKE2 in growth inhibition and basal antiviral resistance. *Plant Cell* **36**: 2289–2309.

1320 Nielsen CPS, Han L, Arribas-Hernández L, Karelina D, Petersen M, Brodersen P. 2023. Sensing of viral RNA in
1321 plants via a DICER-LIKE Ribonuclease. *bioRxiv* 2023.01.10.523395.

1322 Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN, Slotkin RK. 2013. The Initiation of
1323 Epigenetic Silencing of Active Transposable Elements Is Triggered by RDR6 and 21-22 Nucleotide Small
1324 Interfering RNAs. *PLANT PHYSIOLOGY* **162**: 116–131.

1325 Oberlin S, Rajeswaran R, Trasser M, Barragán-Borrero V, Schon MA, Plotnikova A, Loncsek L, Nodine MD,
1326 Marí-Ordóñez A, Voinnet O. 2022. Innate, translation-dependent silencing of an invasive transposon in
1327 Arabidopsis. *Embo Rep* **23**: e53400.

1328 Oberlin S, Sarazin A, Chevalier C, Voinnet O, Ordóñez AM. 2017. A genome-wide transcriptome and translatome
1329 analysis of Arabidopsis transposons identifies a unique and conserved genome expression strategy for
1330 Ty1/Copia retroelements. *Genome Research* **27**: 1549–1562.

1331 Osakabe A, Jamge B, Axelsson E, Montgomery SA, Akimcheva S, Kuehn AL, Pisupati R, Lorković ZJ,
1332 Yelagandula R, Kakutani T, et al. 2021. The chromatin remodeler DDM1 prevents transposon mobility
1333 through deposition of histone variant H2A.W. *Nat Cell Biol* **23**: 391–400.

1334 Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T, et al.
1335 2019. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive
1336 pipeline. *Genome Biol* **20**: 275.

1337 Panda K, McCue AD, Slotkin RK. 2020. Arabidopsis RNA Polymerase IV generates 21–22 nucleotide small
1338 RNAs that can participate in RNA-directed DNA methylation and may regulate genes. *Philos Trans R Soc B*
1339 **375**: 20190417.

1340 Parent J, Bouteiller N, Elmayan T, Vaucheret H. 2015. Respective contributions of Arabidopsis DCL2 and DCL4
1341 to RNA silencing. *Plant J* **81**: 223–232.

1342 Parhad SS, Theurkauf WE. 2019. Rapid evolution and conserved function of the piRNA pathway. *R Soc Open*
1343 *Biol* **9**: 180181.

Patel P, Mathioni SM, Hammond R, Harkess AE, Kakrana A, Arikit S, Dusia A, Meyers BC. 2021. Reproductive phasiRNA loci and DICER-LIKE5, but not microRNA loci, diversified in monocotyledonous plants. *Plant Physiol* **185**: kiab001.

Pieterse AH. 2013. Is flowering in Lemnaceae stress-induced? A review. *Aquat Bot* **104**: 1–4.

Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Auwera GAV der, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, et al. 2018. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* 201178.

Pratx L, Wendering P, Kappel C, Nikoloski Z, Bäurle I. 2023. Histone retention preserves epigenetic marks during heat stress-induced transcriptional memory in plants. *EMBO J* **42**: e113595.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

Rajakumara E, Law JA, Simanshu DK, Voigt P, Johnson LM, Reinberg D, Patel DJ, Jacobsen SE. 2011. A dual flip-out mechanism for 5mC recognition by the Arabidopsis SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation in vivo. *Genes Dev* **25**: 137–152.

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**: W160–W165.

Regulski M, Lu Z, Kendall J, Donoghue MTA, Reinders J, Llaca V, Deschamps S, Smith A, Levy D, McCombie WR, et al. 2013. The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Res* **23**: 1651–1662.

Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L, et al. 2011. Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J* **30**: 1928–1938.

Rougée M, Quadrana L, Zervudacki J, Hure V, Colot V, Navarro L, Deleris A. 2021. Polycomb mutant partially suppresses DNA hypomethylation–associated phenotypes in Arabidopsis. *Life Sci Alliance* **4**: e202000848.

Rowley MJ, Avrutsky MI, Sifuentes CJ, Pereira L, Wierzbicki AT. 2011. Independent Chromatin Binding of ARGONAUTE4 and SPT5L/KTF1 Mediates Transcriptional Gene Silencing. *PLoS Genet* **7**: e1002120.

Sammarco I, Pieters J, Salony S, Toman I, Zolotarov G, Placette CL. 2022. Epigenetic targeting of transposon relics: beating the dead horses of the genome? *Epigenetics* **17**: 1331–1344.

Saze H, Kitayama J, Takashima K, Miura S, Harukawa Y, Ito T, Kakutani T. 2013. Mechanism for full-length RNA processing of Arabidopsis genes containing intragenic heterochromatin. *Nat Commun* **4**: 2301.

Scheid R, Dowell JA, Sanders D, Jiang J, Denu JM, Zhong X. 2022. Histone Acid Extraction and High Throughput Mass Spectrometry to Profile Histone Modifications in Arabidopsis thaliana. *Curr Protoc* **2**: e527.

Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Urich MA, Chen H, et al. 2015. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523**: 212–216.

Sequeira-Mendes J, Aragüez I, Peiró R, Mendez-Giraldez R, Zhang X, Jacobsen SE, Bastolla U, Gutierrez C. 2014. The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. *Plant Cell* **26**: 2351–2366.

Seymour DK, Koenig D, Hagmann J, Becker C, Weigel D. 2014. Evolution of DNA Methylation Patterns in the Brassicaceae is Driven by Differences in Genome Organization. *PLoS Genet* **10**: e1004785.

1385 Shamandi N, Zytnicki M, Charbonnel C, Elvira-Matelot E, Bochnakian A, Comella P, Mallory AC, Lepère G,
1386 Sáez-Vásquez J, Vaucheret H. 2015. Plants Encode a General siRNA Suppressor That Is Induced and
1387 Suppressed by Viruses. *PLoS Biol* **13**: e1002326.

1388 Shi J, Dawe RK. 2006. Partitioning of the Maize Epigenome by the Number of Methyl Groups on Histone H3
1389 Lysines 9 and 27. *Genetics* **173**: 1571–1583.

1390 Shumate A, Salzberg SL. 2021. Liftoff: accurate mapping of gene annotations. *Bioinformatics* **37**: btaa1016.

1391 Sigman MJ, Panda K, Kirchner R, McLain LL, Payne H, Peasari JR, Husbands AY, Slotkin RK, McCue AD.
1392 2021. An siRNA-guided ARGONAUTE protein directs RNA polymerase V to initiate DNA methylation. *Nat*
1393 *Plants* **7**: 1461–1474.

1394 Sigman MJ, Slotkin RK. 2016. The First Rule of Plant Transposable Element Silencing: Location, Location,
1395 Location. *Plant Cell* **28**: 304–313.

1396 Simon L, Rabanal FA, Dubos T, Oliver C, Lauber D, Poulet A, Vogt A, Mandlbauer A, Le Goff S, Sommer A, et
1397 al. 2018. Genetic and epigenetic variation in 5S ribosomal RNA genes reveals genome dynamics in
1398 Arabidopsis thaliana. *Nucleic Acids Res* **46**: gky163-.

1399 Slotkin RK, Vaughn M, Borges F, TanurdZiC M, Becker JoD, FeijO JA, Martienssen RA. 2009. Epigenetic
1400 Reprogramming and Small RNA Silencing of Transposable Elements in Pollen. *Cell* **136**: 461–472.

1401 Sree KS, Maheshwari SC, Boka K, Khurana JP, Keresztes Á, Appenroth K-J. 2015. The duckweed Wolffia
1402 microscopica: A unique aquatic monocot. *Flora - Morphol, Distrib, Funct Ecol Plants* **210**: 31–39.

1403 Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE. 2014. Non-CG methylation patterns
1404 shape the epigenetic landscape in Arabidopsis. *Nat Struct Mol Biol* **21**: 64–72.

1405 Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE. 2013. Comprehensive Analysis of
1406 Silencing Mutants Reveals Complex Regulation of the Arabidopsis Methylome. *Cell* **152**: 352–364.

1407 Stroud H, Hale CJ, Feng S, Caro E, Jacob Y, Michaels SD, Jacobsen SE. 2012. DNA Methyltransferases Are
1408 Required to Induce Heterochromatic Re-Replication in Arabidopsis. *PLoS Genet* **8**: e1002808.

1409 Sun S, Kang X-P, Xing X-J, Chen Z-F, Zheng S-W, Xing G-M. 2014. Transient expression of siRNA targeted
1410 against the TYLCV AV1, AC1 and AC3 genes for high resistance in tomato. *Sci Hortic* **179**: 321–327.

1411 Šurbanovski N, Brilli M, Moser M, Si-Ammour A. 2016. A highly specific microRNA-mediated mechanism
1412 silences LTR retrotransposons of strawberry. *The Plant Journal* **85**: 70–82.

1413 Taochy C, Gursanscky NR, Cao J, Fletcher SJ, Dressel U, Mitter N, Tucker MR, Koltunow AMG, Bowman JL,
1414 Vaucheret H, et al. 2017. A Genetic Screen for Impaired Systemic RNAi Highlights the Crucial Role of
1415 DICER-LIKE 2 . *Plant Physiol* **175**: 1424–1437.

1416 Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, Ciaudo C, Cruaud C, Poulain J, Berdasco M, Fraga
1417 MF, et al. 2009. A Role for RNAi in the Selective Correction of DNA Methylation Defects. *Science* **323**:
1418 1600–1604.

1419 Tippery NP, Les DH, Appenroth KJ, Sree KS, Crawford DJ, Bog M. 2021. Lemnaceae and Orontiaceae Are
1420 Phylogenetically and Morphologically Distinct from Araceae. *Plants* **10**: 2639.

1421 Tirot L, Jullien PE, Ingouff M. 2021. Evolution of CG Methylation Maintenance Machinery in Plants. *Epigenomes*
1422 **5**: 19.

1423 Trujillo JT, Beilstein MA, Mosher RA. 2016. The Argonaute-binding platform of NRPE1 evolves through
1424 modulation of intrinsically disordered repeats. *N Phytol* **212**: 1094–1105.

1425 Tschopp M-A, Iki T, Brosnan CA, Jullien PE, Pumplin N. 2017. A complex of Arabidopsis DRB proteins can
1426 impair dsRNA processing. *RNA* **23**: 782–797.

1427 Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. 2010. Bursts of retrotransposition
1428 reproduced in Arabidopsis. *Nature* **461**: 423–426.

1429 Urquiaga MC de O, Thiebaut F, Hemerly AS, Ferreira PCG. 2021. From Trash to Luxury: The Potential Role of
1430 Plant LncRNA in DNA Methylation During Abiotic Stress. *Front Plant Sci* **11**: 603246.

1431 Walter M, Teissandier A, Pérez-Palacios R, Bourc'his D. 2016. An epigenetic switch ensures transposon
1432 repression upon dynamic loss of DNA methylation in embryonic stem cells. *eLife* **5**: e11418.

1433 Wang T, Deng Z, Zhang X, Wang H, Wang Y, Liu X, Liu S, Xu F, Li T, Fu D, et al. 2018. Tomato DCL2b is
1434 required for the biosynthesis of 22-nt small RNAs, the resulting secondary siRNAs, and the host defense
1435 against ToMV. *Hortic Res* **5**: 62.

1436 Wang W, Haberer G, Gundlach H, Gläßer C, Nussbaumer T, Luo MC, Lomsadze A, Borodovsky M, Kerstetter
1437 RA, Shanklin J, et al. 2014. The Spirodela polyrhiza genome reveals insights into its neotenous reduction fast
1438 growth and aquatic lifestyle. *Nat Commun* **5**: 3311.

1439 Wang X, Duan C-G, Tang K, Wang B, Zhang H, Lei M, Lu K, Mangrauthia SK, Wang P, Zhu G, et al. 2013.
1440 RNA-binding protein regulates plant DNA methylation by controlling mRNA processing at the intronic
1441 heterochromatin-containing gene IBM1. *Proc Natl Acad Sci* **110**: 15467–15472.

1442 Wang Y, Duchen P, Chávez A, Sree KS, Appenroth KJ, Zhao H, Höfer M, Huber M, Xu S. 2024. Population
1443 genomics and epigenomics of Spirodela polyrhiza provide insights into the evolution of facultative asexuality.
1444 *Commun Biol* **7**: 581.

1445 Wang Y, Zhou X, Luo J, Lv S, Liu R, Du X, Jia B, Yuan F, Zhang H, Du J. 2021. Recognition of H3K9me1 by
1446 maize RNA-directed DNA methylation factor SHH2. *J Integr Plant Biol* **63**: 1091–1096.

1447 Wang Z, Baulcombe DC. 2020. Transposon age and non-CG methylation. *Nat Commun* **11**: 1221.

1448 Ware A, Jones DH, Flis P, Chrysanthou E, Smith KE, Kümpers BMC, Yant L, Atkinson JA, Wells DM, Bhosale
1449 R, et al. 2023. Loss of ancestral function in duckweed roots is accompanied by progressive anatomical
1450 reduction and a re-distribution of nutrient transporters. *Curr Biol* **33**: 1795-1802.e4.

1451 Wells JN, Feschotte C. 2020. A Field Guide to Eukaryotic Transposable Elements. *Annu Rev Genet* **54**: 1–23.

1452 Wendte JM, Zhang Y, Ji L, Shi X, Hazarika RR, Shahryary Y, Johannes F, Schmitz RJ. 2019. Epimutations are
1453 associated with CHROMOMETHYLASE 3-induced de novo DNA methylation. *eLife* **8**: e47891.

1454 Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O,
1455 et al. 2007. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* **8**:
1456 973–982.

1457 Wickham H. 2016. *Ggplot2: Elegant graphics for data analysis*. Springer.

1458 Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. 2009. RNA polymerase V transcription guides ARGONAUTE4
1459 to chromatin. *Nature Genetics* **41**: 630–634.

1460 Willmann MR, Endres MW, Cook RT, Gregory BD. 2011. The Functions of RNA-Dependent RNA Polymerases
1461 in Arabidopsis. *Arab Book* **2011**: e0146.

1462 Wright S, Finnegan D. 2001. Genome evolution: Sex and the transposable element. *Curr Biol* **11**: R296–R299.

1463 Wyler M, Stritt C, Walser J-C, Baroux C, Roulin AC. 2020. Impact of transposable elements on methylation and
1464 gene expression across natural accessions of Brachypodium distachyon. *Genome Biol Evol* **12**: evaa180-.

1465 Xie Z, Allen E, Wilken A, Carrington JC. 2005. DICER-LIKE 4 functions in trans-acting small interfering RNA
1466 biogenesis and vegetative phase change in Arabidopsis thaliana. *Proceedings of the National Academy of*
1467 *Sciences of the United States of America* **102**: 12984–12989.

1468 Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC. 2004.
1469 Genetic and functional diversification of small RNA pathways in plants. *PLoS Biology* **2**: E104.

1470 Xu S, Stapley J, Gablenz S, Boyer J, Appenroth KJ, Sree KS, Gershenzon J, Widmer A, Huber M. 2019. Low
1471 genetic variation is associated with low mutation rate in the giant duckweed. *Nat Commun* **10**: 1243.

1472 Yaari R, Noy-Malka C, Wiedemann G, Gershovitz NA, Reski R, Katz A, Ohad N. 2015. DNA
1473 METHYLTRANSFERASE 1 is involved in mCG and mCCG DNA methylation and is essential for
1474 sporophyte development in Physcomitrella patens. *Plant Mol Biol* **88**: 387–400.

1475 Yan H, Bombarely A, Li S. 2020. DeepTE: a computational method for de novo classification of transposons with
1476 convolutional neural network. *Bioinformatics* **36**: 4269–4275.

1477 Yang G-L, Fang Y, Xu Y-L, Tan L, Li Q, Liu Y, Lai F, Jin Y-L, Du A-P, He K-Z, et al. 2018. Frond transformation
1478 system mediated by Agrobacterium tumefaciens for Lemna minor. *Plant Mol Biol* **98**: 319–331.

1479 Yang J, Zhao X, Li G, Hu S, Hou H. 2021. Frond architecture of the rootless duckweed Wolffia globosa. *BMC*
1480 *Plant Biol* **21**: 387.

1481 Yoshikawa M, Peragine A, Park MY, Poethig RS. 2005. A pathway for the biogenesis of trans-acting siRNAs in
1482 Arabidopsis. *Genes Development* **19**: 2164–2175.

1483 Yu D, Fan B, MacFarlane SA, Chen Z. 2003. Analysis of the Involvement of an Inducible Arabidopsis RNA-
1484 Dependent RNA Polymerase in Antiviral Defense. *Mol Plant-Microbe Interact* **16**: 206–216.

1485 Zattera ML, Bruschi DP. 2022. Transposable Elements as a Source of Novel Repetitive DNA in the Eukaryote
1486 Genome. *Cells* **11**: 3373.

1487 Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D. 2013.
1488 The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing
1489 Heterochromatin. *Cell* **153**: 193–205.

1490 Zervudacki J, Yu A, Amesefe D, Wang J, Drouaud J, Navarro L, Deleris A. 2018. Transcriptional control and
1491 exploitation of an immune-responsive family of plant retrotransposons. *The EMBO journal* **37**: e98482.

1492 Zhai J, Zhang H, Arikit S, Huang K, Nan G-L, Walbot V, Meyers BC. 2015. Spatiotemporally dynamic, cell-
1493 type–dependent premeiotic and meiotic phasiRNAs in maize anthers. *Proc Natl Acad Sci* **112**: 3146–3151.

1494 Zhan J, Meyers BC. 2023. Plant Small RNAs: Their Biogenesis, Regulatory Roles, and Functions. *Annu Rev Plant*
1495 *Biol* **74**: 21–51.

1496 Zhang J, Yuan J, Lin J, Chen L, You L-Y, Chen S, Peng L, Wang C-H, Du J, Duan C-G. 2023. Molecular basis
1497 of locus-specific H3K9 methylation catalyzed by SUVH6 in plants. *Proc Natl Acad Sci* **120**: e2211155120.

1498 Zhang R-G, Li G-Y, Wang X-L, Dainat J, Wang Z-X, Ou S, Ma Y. 2022. TEsorter: An accurate and fast method
1499 to classify LTR-retrotransposons in plant genomes. *Hortic Res* **9**: uhac017.

1500 Zhang W, Friebe B, Gill BS, Jiang J. 2010. Centromere inactivation and epigenetic modifications of a plant
1501 chromosome with three functional centromeres. *Chromosoma* **119**: 553–563.

1502 Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. 2009. Genome-wide analysis of mono-, di-
1503 and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana. *Genome Biology* **10**: R62.

1504 Zhang X, Singh J, Li D, Qu F. 2012. Temperature-Dependent Survival of Turnip Crinkle Virus-Infected
1505 Arabidopsis Plants Relies on an RNA Silencing-Based Defense That Requires DCL2, AGO2, and HEN1.
1506 *Journal of Virology* **86**: 6847–6854.

1507 Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen
1508 SE, et al. 2006. Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in
1509 Arabidopsis. *Cell* **126**: 1189–1201.

1510 Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W,
1511 et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.

1512 Zhong X, Hale CJ, Law JA, Johnson LM, Feng S, Tu A, Jacobsen SE. 2012. DDR complex facilitates global
1513 association of RNA polymerase V to promoters and evolutionarily young transposons. *Nat Struct Mol Biol*
1514 **19**: 870–875.

1515 Zhong X, Hale CJ, Nguyen M, Ausin I, Groth M, Hetzel J, Vashisht AA, Henderson IR, Wohlschlegel JA,
1516 Jacobsen SE. 2015. DOMAINS REARRANGED METHYLTRANSFERASE3 controls DNA methylation
1517 and regulates RNA polymerase V transcript abundance in Arabidopsis. *Proc Natl Acad Sci* **112**: 911–916.

1518 Zhou M, Coruh C, Xu G, Martins LM, Bourbousse C, Lambolez A, Law JA. 2022. The CLASSY family controls
1519 tissue-specific DNA methylation patterns in Arabidopsis. *Nat Commun* **13**: 244.

1520 Zhou M, Palanca AMS, Law JA. 2018. Locus-specific control of the de novo DNA methylation pathway in
1521 Arabidopsis by the CLASSY family. *Nat Genet* **50**: 865–873.

1522 Ziegler P, Adelmann K, Zimmer S, Schmidt C, Appenroth K -J. 2015. Relative in vitro growth rates of duckweeds
1523 (Lemnaceae) – the most rapidly growing higher plants. *Plant Biol* **17**: 33–41.

1524 Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler.
1525 *Bioinformatics* **29**: 2669–2677.

1526 Zimin AV, Salzberg SL. 2020. The genome polishing tool POLCA makes fast and accurate corrections in genome
1527 assemblies. *PLoS Comput Biol* **16**: e1007981.

1528 Zong J, Yao X, Yin J, Zhang D, Ma H. 2009. Evolution of the RNA-dependent RNA polymerase (RdRP) genes:
1529 Duplications and possible losses before and after the divergence of major eukaryotic groups. *Gene* **447**: 29–
1530 39.

1531

1532