DR JOSÉ LUIS REYES (Orcid ID : 0000-0001-5129-9741)
DR ALEJANDRA A COVARRUBIAS (Orcid ID : 0000-0003-0439-3414)

Article type : Original Article

Title

The canonical RdDM pathway mediates the control of seed germination timing under salinity

Author information

V. Miguel Palomar[1,4], Alejandro Garciarrubio[2], Adriana Garay-Arroyo[3], Coral Martínez-Martínez[1], Omar Rosas-Bringas[1], José L. Reyes[1] and Alejandra A. Covarrubias[1]*

[1]Departamento de Biología Molecular de Plantas, [2] Departamento de Ingeniería Celular y Biocatálisis, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apdo. Postal 510-3, Cuernavaca, Mor. C.P. 62250, Mexico. [3]Laboratorio de Genética Molecular, Desarrollo y Evolución de Plantas, Instituto de Ecología, Universidad Nacional Autónoma de México, Circuito Exterior S/N anexo Jardín Botánico exterior, Ciudad Universitaria, Ciudad de México, C.P. 04500, México. [4]Current address: Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, Michigan, United States.

*** Corresponding author**

Alejandra A. Covarrubias. Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apdo. Postal 510-3, Cuernavaca, Mor. C.P. 62250. crobles@ibt.unam.mx

Running title

RdDM pathway, a tuner of seed germination

Keywords

RNA directed DNA methylation, germination, salinity, stress, Arabidopsis

Summary

Plants respond to adverse environmental cues by adjusting a wide variety of processes through highly regulated mechanisms to maintain plant homeostasis for survival. Because of the sessile nature of plants, their response, adjustment and adaptation to the changing environment is intimately coordinated with their developmental programs through the cross talk of regulatory networks. Germination is a critical process in the plant life cycle, such that plants have evolved various strategies to control germination timing according to their local environment. However, the mechanisms involved in these adjustment responses are largely unknown. Here, we report that mutations in core elements of the canonical RNA directed DNA methylation pathway (RdDM) affect germination and post-germination growth of Arabidopsis seeds under salinity. Transcriptomic and whole genome bisulfite sequencing (WGBS) analyses supports the involvement of this pathway in the control of germination timing and post-germination growth under this stress condition by preventing the transcriptional activation of genes implicated in these processes. Subsequent transcriptional effects on genes whose function is also related to these development events support this conclusion.

Key words: RdDM pathway, AGO4 protein, salt stress, germination, Arabidopsis.

Introduction

Different mechanisms are involved in the tuning of an efficient and appropriate gene expression program under stressful conditions. Among the best-characterized control mechanisms are those modulating transcription patterns that involve

pathways controlling the amount and/or the activity of different transcription factors (Zhu, 2002; Yamaguchi-Shinozaki and Shinozaki, 2006; Nakashima et al., 2009). More recently, additional mechanisms involving chromatin modifications, some of them persistent and/or heritable, have been implicated in the modulation of transcriptional activity denoting epigenetic regulation (Cubas et al., 1999; Soppe et al., 2000; Manning et al., 2006; Pecinka and Mittelsten Scheid, 2012). Transcriptional control of gene expression is also exerted through the recruitment of histone variants to different genome regions, histone post-translational modifications and DNA methylation, which leads to the conformation of a wide variety of 'epigenomes' depending on developmental or environmental signals, or those derived from their crosstalk.

Plants have evolved under frequent and/or sudden environmental changes; hence, the control of their developmental programs has co-evolved with their stress responses to be able to develop, grow and reproduce under adverse conditions (Skirycz and Inze, 2010). A frequent observation in plants is their ability to acclimate upon repeated exposure to a particular stress, a response exhibiting a form of short-term stress memory (Chinnusamy and Zhu, 2009; Secco et al., 2015). This capacity to retain a stress memory also occurs for longer periods (Molinier et al., 2006; Chinnusamy and Zhu, 2009; Han and Wagner, 2014). Diverse processes that lead to stable DNA methylation and histone modifications have been identified in this behavior, allowing growth adjustment and reprogramming of developmental decisions (Chinnusamy and Zhu, 2009; Han and Wagner, 2014; Kinoshita and Seki, 2014).

DNA methylation is considered an essential plant contention mechanism against transposable element activation that may occur under hostile conditions such as high temperatures (Ito et al., 2011; Matsunaga et al., 2012). The extent of plant genome methylation is controlled by *de novo* and maintenance DNA methylation, and by demethylation processes. DNA methylation may occur in symmetric (CG and CHG) and asymmetric (CHH) sequence contexts (H = C, A, or T) (Law and Jacobsen, 2010). Four DNA methylation writers have been described to date in Arabidopsis: DNA METHYL TRANSFERASE 1 (MET1) responsible of maintaining CG methylation,

CHROMOMETHYLASE 3 (CMT3) that keeps methylation in CHG contexts, CMT2 does it in CHG and CHH sequences; whereas, RNA-directed DNA methylation (RdDM) mediates the initiation and maintenance of DNA methylation mostly at CHH sites and to a less extent at CG and CHG sequences, with the participation of DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) and other context specific methyltransferases (Law and Jacobsen, 2010).

RdDM consists of complex pathways occurring through the participation of small interfering RNAs (siRNAs) and requires a specific transcriptional assembly. The best characterized RdDM pathway, known as canonical RdDM, depends on two plant-specific RNA polymerases, Pol IV and Pol V, which play key roles (Mahfouz, 2010; Matzke et al., 2015; Wang and Axtell, 2017), together with RNA-DEPENDENT RNA POLYMERASE 2 (RDR2), type III ribonuclease DICER-LIKE 3 (DCL3) (Henderson et al., 2006; Blevins et al., 2015; Zhai et al., 2015), and ARGONAUTE 4 (AGO4) (Zilberman et al., 2003), among other factors. RdDM is not exclusively a nuclear process, as siRNAs produced in the nucleus are exported to the cytoplasm (Ye et al., 2012), where they are bound by AGO4 to then return to the nucleus to form a silencing complex by pairing with a scaffold long non-coding RNA produced by Pol V (Wierzbicki et al., 2009) (reviewed extensively in (Matzke and Mosher, 2014; Matzke et al., 2015)). For the canonical RdDM pathway, there is evidence indicating that the formation of mature AGO4/siRNA complexes requires AGO4 endonucleolytic activity, which triggers the removal of the passenger strand allowing the entry of the mature AGO4/siRNA complex to the nucleus (Ye et al., 2012). AGO4 is closely related to AGO6 and AGO9, proteins belonging to the same AGO family, which has been shown to control DNA methylation through canonical and non-canonical RdDM pathways. AGO9 has a specific function by silencing transposable elements in the female gametophyte by a non-cell autonomous mechanism (Olmedo-Monfil et al., 2010). While AGO6 also participates in controlling DNA methylation via AGO6/RdDM and non-canonical AGO6/RDR6 RdDM pathways; AGO4 and AGO6 have different preferences for siRNAs (Havecker et al., 2010; McCue et al., 2015). Genome-wide methylation analyses of *ago4, ago6* or *ago4 ago6* mutants showed that there is very low redundancy between these proteins,

indicating that they do not have exactly the same function, although they show some common targets (Havecker et al., 2010; Eun et al., 2011; Duan et al., 2015; McCue et al., 2015).

The participation of the RdDM pathway in different plant processes is exposed by phenotypes found in various plant species including maize (Hollick, 2010), rice (Wu et al., 2010; Wei et al., 2014), tomato (Gouil and Baulcombe, 2016; Corem et al., 2018), and two Brassicaceae species *Brassica rapa* (Grover et al., 2018) and *Capsella rubella* (Wang et al., 2020). However, although a dynamic DNA methylation occurs during different developmental stages, few phenotypes have been detected in the absence of RdDM core proteins in *Arabidopsis thaliana* plants growing under optimal conditions (Matzke et al., 2015). Among these are a delay in flowering time (Pontier et al., 2005; Haag and Pikaard, 2011), an anomalous expression of imprinted genes leading to seed arrest (triploid block (Kradolfer et al., 2013; Borges et al., 2018; Martinez et al., 2018)), and defects in endosperm development (Köhler and Lafon-Placette, 2015; Chow et al., 2020). In addition, a role for RdDM during pathogen infection has been evidenced by the phenotypes of RdDM-associated protein mutants during infection by virus (Hamera et al., 2012; Yang et al., 2013; Brosseau et al., 2016), bacteria (Agorio and Vera, 2007; Dowen et al., 2012; Gohlke et al., 2013) or fungi (Lopez et al., 2011; Le et al., 2014). The RdDM pathway is also involved in the plant response to abiotic stresses such as heat (Ito et al., 2011; Popova et al., 2013), low relative humidity (Tricker et al., 2012) and low temperatures (Chan et al., 2016).

Germination is a crucial developmental stage in the plant life cycle, during which plant embryos confront severe environmental changes; as it happens in orthodox seeds when embryos switch from dry to humid conditions, or when embryo revival confronts unfavorable environments. Recently, it was reported that DNA methylation is dynamically modulated during seed development, and germination in *A. thaliana* under optimal conditions, where major changes occur in the methylation of CHH sequence context (Kawakatsu et al., 2017). In this work, we address the question of whether the canonical RdDM pathway is required for the proper germination of *A. thaliana* under high salinity, a common condition present in arid

and semiarid regions, in degraded croplands and in contaminated soils. We demonstrate that the RdDM pathway plays a key role in the plant response to salinity during germination and early seedling growth. Accordingly, RNA-seq comparative analyses from wild-type and *AGO4* null and weak mutants, germinated in media with or without NaCl, show that in response to salinity, RdDM is mostly involved in the control of embryo development and germination genes, and to a lesser degree in the regulation of those genes implicated in stress tolerance. Whole genome bisulfite sequencing analyses support the conclusion that this response is a consequence of the RdDM pathway modulating gene expression mainly by controlling the DNA methylation on its usual targets and preventing the activation of genes in response to salinity. These results agree with changes in the protein levels and localization of RdDM components during Arabidopsis germination and early seedling growth under this environmental adversity.

Results

The canonical RdDM pathway is required for optimal germination under salinity.

Because there is evidence of a dynamic reprogramming of DNA methylation during seed development and germination and considering that germination is a critical process that repeatedly exposes the embryonic plant to extremes, we asked whether the RdDM pathway is involved in the plant response to salinity. To address this, we evaluated the germination rate of different mutant lines lacking the core proteins of the canonical RdDM pathway: DCL3, RDR2, and the largest subunits of Pol IV and Pol V, NRPD1 and NRPE1, respectively (Matzke and Mosher, 2014). Under optimal conditions all, wild type Arabidopsis (Col-0) and mutant lines, germinated at similar rates, reaching 100% germination (Figure S1a). In contrast, when the seeds of these mutants were germinated under high salt concentration (250 mM NaCl), their germination rate significantly decreased (p-value = $1e^{-4}$) compared to wild type (Figure 1a), indicating that the canonical RdDM pathway is required for proper germination under high salinity.

Given these evidences, we analyzed under the same conditions seeds carrying an AGO4 point mutation allele (*ago4-2*) (E641K within the PIWI domain, close to the residues required for AGO4 slicing activity), which behaves as a dominant-negative allele and causes defective DNA methylation by producing a less active protein, according to Agorio and Vera (2007). The results showed that *ago4-2* seeds present lower germination rates under salt stress when compared to wild type seeds (Figure 1b), consistent with the phenotype shown by the other RdDM defective mutants (Figure 1a), while its germination profile under optimal conditions is only slightly changed (Figure S1b). This result confirms that the whole canonical RdDM is involved in germination under this environmental condition.

Because a phenotype similar to that of *ago4-2* would be expected for a mutant lacking AGO4 as part of the RdDM pathway, we analyzed the effect of the absence of AGO4 protein on germination efficiency under salt stress using a null *AGO4* mutant (*ago4-3*). In sharp contrast to the negative effect on germination observed for the rest of the mutants, (Figure 1a, b), under salt the lack of AGO4 protein led to a higher germination rate when compared to wild type seeds (Col-0; Figure 1c). The same effect was detected for other *AGO4* null-mutant alleles (*ago4-6,* Figure 1c), some of them in different genetic backgrounds (*ago4-1,* L*er* ecotype; and *ago4-4,* Ws ecotype. Figure S2a, b, respectively). These AGO4 null alleles also reached similar germination efficiency compared to the wild type under optimal conditions (Figure S1c-e). The absence of AGO4 protein in the different null mutants was confirmed by western blot experiments (Figure S3). The salt concentration dependence of the germination phenotype in the absence of these RdDM components was also tested (germination under 150 and 200 mM NaCl), showing that the germination rate decreased in the null mutants of the RdDM pathway, as well as in the point mutant *ago4-2*; meanwhile, the germination of the three null *ago4* mutants increased compared to the wild type, as expected (Figure S4a – h), confirming that the observed phenotypes are due to the salt exposure.

To validate AGO4 contribution to the observed phenotypes, the wild type gene was introduced into the *ago4-3* null mutant by genetic crosses (Col-0 x *ago4-3*). Analysis of the germination rate under optimal and stress conditions of the resultant F2 lines

showed recovery of the wild type phenotype in those plants where wild type AGO4 protein levels were detected (Figure 1d and Figure S3). Additionally, the *ago4-3* null allele was introgressed into wild type Col-0 plants (*ago4-3* x Col-0) and the resulting homozygous lines (F2), where AGO4 protein was not detected by western blot (Figure S3), showed high germination rate under salt stress, similar to the one shown by the *ago4* null mutants (Figure 1e), confirming that this phenotype resulted from the absence of AGO4. The germination rate of the crossed lines under optimal conditions showed no differences (Figure S1f, g).

The phenotypes of *ago4* mutants extend into seedling stages, with no significant effect when seedlings were grown under optimal conditions (Figure S5c, d). Seedlings carrying AGO4 null alleles showed longer primary roots when exposed to salinity compared to wild type plants (Figure S5a, b), whereas the *ago4-2* mutant showed a detrimental effect on primary root length (Figure S5a). The phenotypic effects of these mutants are analogous to those observed during germination (Figure 1b, c).

AGO6 plays a role during germination under salinity when AGO4 is absent

The contrasting difference between AGO4 null and *ago4-2* mutant phenotypes led us to hypothesize that either AGO4 was playing an additional role, independent of the RdDM pathway, during this plant developmental stage in response to stress; or that in the absence of AGO4, AGO6, expressed in low levels in the seed embryo (Havecker et al., 2010), could play an unusual role under these conditions. To distinguish between these two possibilities, the phenotype of *ago4-6 ago6-2* double mutant was analyzed during germination under the same stress conditions. The results showed that the double mutant germinates with a lower rate than the wild type genotype under salinity, similar to that shown by the mutants lacking other core RdDM elements, while it showed a wild type phenotype under optimal conditions (Figure 1a, f and Figure S1h), and in sharp contrast to the higher germination rate shown by the different *ago4* null mutants (Figure 1c, f and Figure S2). These data indicate that the *ago4* phenotype was not due to an additional AGO4 function but rather to the contribution of AGO6 in this plant response, perhaps by binding siRNAs that in the

absence of AGO4 are preferentially associated to AGO6, producing a synthetic phenotype. In the absence of both AGO proteins, the plant stress response is compromised as in the case of the rest of the RdDM mutants tested causing a deficient germination (Figure 1a, f). This interpretation is supported by the null *ago6* single mutant phenotype, which showed a wild type behavior under the conditions tested (Figure 1f), indicating that AGO6 by itself does not play a role in this process.

AGO4 protein abundance changes during germination and early seedling growth

Given the above evidence indicating that RdDM is needed for an optimal germination under salinity, and the critical role assigned to AGO4 in this pathway, we asked if its abundance is modulated during this process. For this, we tracked AGO4 protein accumulation during germination and throughout early seedling growth under optimal and salt stress conditions. We performed western blot assays using an AGO4-specific antibody against total protein extracts from six developmental stages: dry and stratified seeds, 18 h and 36 h germinating seeds, seedlings where hypocotyl and cotyledons have emerged from the seed coat, and from seedlings with two rosette leaves, all germinated under optimal conditions. The results showed that the accumulated AGO4 protein present in dry seeds decreased after stratification, but then progressively increased until 36 h post-stratification, when its highest levels were achieved. After this time point, AGO4 protein levels gradually declined again until the last stage tested (Figure 2a). Under salinity (150 mM NaCl), we found a markedly different AGO4 accumulation pattern, where the highest protein levels were detected in stratified seeds, with a subsequent gradual reduction, finding the lowest levels in seedlings with two rosette leaves, as observed under non-stress conditions (Figure 2b). We did not detect changes in AGO4 transcript accumulation in response to salinity during germination (Figure S6, see Col green bars), indicating that under these conditions AGO4 protein abundance is not controlled at the transcriptional level. It is worth mentioning that the accumulation patterns of the point mutant protein encoded by

the *ago4-2* allele were similar to those found using wild type seeds, indicating that AGO4-2 protein synthesis and stability are not drastically affected (Figure 2c, d). These data demonstrate not only that AGO4 is present during germination and early seedling growth, but also that its levels are modulated upon stress, in consonance with a functional role for the canonical RdDM pathway during these developmental stages and in response to adverse germination environments.

AGO4 is differentially localized in response to salinity

To explore in which seedling regions RdDM is exerting its function during seed germination, we determined AGO4 tissue distribution during germination and early seedling growth stages. For this end we used transgenic plants containing a *GFP-AGO4* translational fusion under the control of the *AGO4* endogenous promoter (*pAGO4::GFP-AGO4*) expressed in a wild type background (Ye et al., 2012). Protein localization was determined by GFP fluorescence using confocal microscopy of isolated embryos and young seedlings at different times after imbibition. As expected for AGO4 most of the fluorescence signal was localized in cell nuclei across the different tissues (Figure S7). GFP fluorescence showed a high AGO4 abundance and broad distribution in embryos imbibed for 8 h (Figure 3a). Remarkably, just after testa rupture AGO4 distribution changed, maintaining its abundance in cotyledons but with lower levels and more dispersed localization in radicles. This distribution pattern was then maintained throughout radicle protrusion, late germination phase and young seedling growth (60 h after imbibition). At the later stage AGO4 accumulation was even lower in cotyledons, whereas the signal was detected in hypocotyls and root tips (Figure 3a).

When this analysis was performed on embryos and seedlings exposed to salt stress (100 mM NaCl), GFP fluorescence was also found in cotyledons and radicles; however, a strong signal was clearly detected in vascular tissues in both embryo regions (Figure 3b). This pattern was also present after testa rupture and was followed by a significant reduction during subsequent stages, with the fluorescence signal being barely detected in roots or cotyledons of young seedlings.

We noticed that during late germination stages and early seedling growth a sharp signal could be detected in shoot apical regions and in root tips (Figure 3a, b, arrowheads and Figure S8), consistent with previous findings by other groups using transcriptomic data obtained from roots of seedlings growing under optimal conditions (Kawakatsu et al., 2016). This result strongly suggests that the canonical RdDM pathway plays a particular role in response to salinity during seed germination.

RdDM controls CHH methylation during germination under salinity in a locus specific manner

To obtain a mechanistic insight into the role of the canonical RdDM in the plant response to salinity, we explored the DNA methylation differences between RdDM defective (*ago4-3*, *ago4-2,* and *nrpe1-11* a PolV null mutant) and wild type lines germinated under optimal or high salinity conditions by performing whole-genome bisulfite sequencing (WGBS). As we anticipated according to the proposed RdDM mechanism, no strong differences were found in terms of the global proportion of methylated cytosines on any of the three different methylation contexts found in plants (CG, CHG and CHH) in the mutant alleles when compared to the wild type (Figure S9a). The distribution of CHH methylation was slightly decreased in the chromosome arms without noticeable changes in the centromeric regions, while the other methylation contexts (CG and CHG) remained unaffected (Figure S9b). These results are consistent with the activity and distribution previously reported for RdDM in terms of global DNA methylation (Schoft et al., 2009; Wierzbicki et al., 2012; Zhang and Zhu, 2012; Yaari et al., 2019). Moreover, we did not notice any prominent effect as result of the salinity treatment over the global DNA methylation or its distribution along chromosomes in the wild type or in RdDM mutants (Figure S9a, b). Overall, these results suggest that RdDM does not produce a widespread effect at the genome level, but rather that it acts at a locus specific manner in response to salinity during germination, similar to its activity in other developmental stages.

To validate the later hypothesis, we identified Differentially Methylated Regions (DMRs, defined with at least 10% methylation change in each mutant compared to Col-0 with adjusted p-value (FDR) < 0.01 in 100 bp sliding windows, from the two biological replicates, see experimental procedures) in each growth condition. This analysis showed 191, 2597 and 1481 CHH hypomethylated DMRs in *ago4-2*, *ago4-3* and *nrpe1*, respectively. This number increased by nearly twice in the AGO4 mutants and by 20% in *nrpe1,* when seeds were exposed to salinity (Figure 4a and Table S1), strongly supporting an active locus-specific role of the RdDM pathway during germination in response to this stress. Interestingly, we also observed changes in the number of CG and CHG DMRs, in agreement with the participation of RdDM in the maintenance of DNA methylation (Law and Jacobsen, 2010). An overlapping analysis between the identified DMRs and the genomic features (see experimental procedures) showed, as expected, that most of the DMRs are located within Transposable Elements (TE genes and fragments; Figure S10a), and less than 20% were located to protein-coding genes (including up to 2 Kb of their regulatory regions), non-coding genes or pseudo-genes, confirming the main role of RdDM is the control of methylation of TE; although, it can also affect protein coding genes during germination under the conditions tested.

The *ago4-2* allele produces a low-activity AGO4 protein

Because *ago4-2* was considered to be a dominant-negative AGO4 allele encoding a defective protein (Agorio and Vera, 2007), we expected a similar effect on methylation in *ago4-2* as in the null mutant (*ago4-3*). However, DMRs in *ago4-3* showed 10 times more hypomethylated CHH regions compared to *ago4-2* (Figure 4a and Table S1), suggesting that *ago4-2* produces an AGO4 weak allele. To further investigate this possibility, we performed an overlapping analysis of the DMRs identified in both mutants, under optimal and salinity conditions, and classified them into three sets: specific to *ago4-2*, specific to *ago4-3*, or common to both. From control conditions, we found in these three sets: 73, 2479 and 118 DMRs, respectively; whereas, from NaCl treatment we found 179, 3985 and 221 DMRs in the three respective sets. Thus, in both conditions, about 60% of *ago4-2* DMRs are

common with *ago4-3* (Figure S10b). We then looked at the methylation effects in those DMR sets specific to *ago4-2* (Figure S10c, d) and *ago4-3* (Figure S10e, f) in the different genetic backgrounds. The results showed that specific *ago4-2* DMRs present lower methylation than wild type in *ago4-3* and *nrpe1* genotypes from both conditions (Figure S10c, d); a reduction that was not detected before because was lower than the cut-off considered to define a DMR (10%, see experimental procedures). When we looked at the methylation level of the specific *ago4-3* DMRs in *ago4-2,* we found that it was similar to that in wild type, whereas in *nrpe1* it was low as in *ago4-3* (Figure S10e, f). Therefore, these data indicate that the *ago4-2* allele encodes an AGO4 protein, whose activity to methylation levels lower than wild type but not as reduced as in AGO4 or PolV null mutants, consistent with an AGO4 weak allele. Despite the residual activity of *ago4-2* allele in terms of DNA methylation, its phenotype is clearly similar to *nrpe1* and to other RdDM core element null mutants (Figure 1a, b), indicating that, even though the DNA methylation effect in *ago4-2* is not as penetrant as in *nrpe1,* it is sufficient to alter the RdDM impact on germination under salinity.

The RdDM pathway controls locus specific methylation in response to salinity during germination

To identify AGO4-dependent methylated regions, we looked for those wild-type (Col-0) DMRs that gain CHH methylation under salinity. This analysis identified 170 DMRs with higher methylation comparing wild type seeds germinated in salt media with those germinated under non-stress conditions. To determine whether AGO4 was participating in the methylation of such regions in response to salinity, we compared their methylation profiles in *ago4-2* and *ago4-3* mutants from optimal and salinity conditions, using the *nrpe1* methylation profile as RdDM reference. The comparison of the methylation gain between non-stress and stress conditions showed differences for the wild type line, as expected, but not for *nrpe1*, *ago4-2* and *ago4-3.* As anticipated for an AGO4 weak allele, *ago4-2* showed a small methylation decrease compared to *ago4-3* (Figure 4b). These results indicate that most of the

locus-specific changes in CHH methylation in response to this stress depend on AGO4 and the RdDM activity.

Analysis of the genomic features of those DMRs exhibiting the strongest reduction difference in DNA methylation (>10%), when Col-0 was compared to *ago4-2* (20 regions), *nrpe1* (64 regions) or *ago4-3* (127 regions), showed that most of them correspond to TE genes and TE fragments, typical targets of the canonical RdDM (Figure 4c).

RdDM participates in the control of the germination timing

To obtain information on the impact of AGO4 on gene expression during germination under salinity, we performed RNA-seq experiments using Col-0, *ago4-3* and *ago4-2* seedlings germinated under optimal and salinity conditions. We first focused on the hyper-methylated regions previously identified in Col-0 from salinity conditions that also showed lower DNA methylation levels in *ago4-2* (170 DMRs, see Figure 4b). We searched for protein-coding genes containing these regions, as well as for those falling within 5 kb on each side of the recognized TEs (Figure 4c). Out of 78 potential gene targets, we identified a subset of 13 genes, whose expression is lower in Col-0 in response to salt but is higher in *ago4-2,* consistent with the higher CHH methylation occurring nearby in wild type genotypes (Figure 4d and Table 1). Hence, this result suggests that the RdDM-dependent CHH methylation deposition is responsible for the lower expression of these genes during germination under salinity and that this repression is released in the absence of a functional AGO4/RdDM. In accordance with this outcome, these genes are located significantly closer to previously identified PolV target sites, when compared to a set of randomly selected regions (Bohmdorfer et al., 2016) (Figure 4e). In consequence, this set of RdDM regulated genes might be at least partially responsible of the *ago4-2* germination phenotype under this stressful condition (Figure 1b).

Most of the identified genes (8/13) from this analysis are related to signal transduction or other regulatory functions, some of which have been involved in the control of dormancy and/or germination or in functions related to these processes (Table 2). Among them stand out EIN3 BINDING F-BOX1 (*EBF1*), encoding a

component of Ub-protein ligase (E3) that together with EBF2 is essential for proper ethylene signaling (Binder et al., 2007), underscoring the control of seed dormancy through the ethylene-DOG1 pathway under stress conditions (Nishimura et al., 2018; Li et al., 2019). It is also worth noting the *NAC4* gene transcription factor, whose expression is regulated by AFB3 (Vidal et al., 2014; Lee et al., 2017), an auxin receptor involved in various aspects of plant development including the auxin mediated control of seed dormancy (Liu et al., 2013).

When we performed the same analysis with the *ago4-3* null mutant we found a subset of 18 out of 83 candidate genes, whose expression under salt treatment in Col-0 is de-repressed in this mutant, consistent with the loss of CHH methylation on the nearby DMRs (Figure S11a) and their relative proximity to a PolV transcribed region (Bohmdorfer et al., 2016) (Figure S11b). Twelve of these targets are shared with *ago4-2* (Table S2), indicating that they are *bona fide* AGO4/RdDM direct gene targets. Additional targets in this mutant could be a consequence of the complete loss of AGO4 activity and/or the AGO6 interference as demonstrated in the previous sections.

Beyond the direct targets, the phenotype of RdDM mutants should involve the deregulation of many downstream genes. To characterize these broader effects of AGO4/RdDM during germination and in response to salinity, we analyzed the differential accumulation of transcripts in *ago4-2*, *ago4-3* and wild type lines. We found that transcript accumulation profiles overall are similar between wild type and *ago4-3* lines, and slightly different to *ago4-2* (Figure S12a – c). Interestingly, differences were also found between wild type and the mutants even under optimal growth conditions (Figure S12a). This shows that, while RdDM mutants have no detectable phenotype during germination in the absence of salt stress, they still have some transcriptional effects (Figure S1).

To identify the differently accumulated transcripts under stress conditions between wild type, *ago4-2* and *ago4-3* lines, we performed a correlation analysis (Figure S12b, c) and selected outliers as genes strongly upregulated or downregulated by salt in the mutants. As we anticipated many genes are shared by the *ago4-2* and the *ago4-3* subsets (Figure S12d, e). The greater differences in abundance were found

among transcripts associated with seed maturation, dormancy and germination (Figure 5a, and Tables S3 - S6), most of them related to metabolic functions. The comparison analysis between *ago4-3* and wild type expression patterns, during germination under salt, highlights the up-regulation in *ago4-3* of transcripts involved in dormancy, germination and growth (Figure 5a), consistent with its higher germination rate. This is the case of those encoding two 2C type protein phosphatases (AHG1 and AHG3), proteins that are known to interact with DOG1, an ABA independent key regulator of seed germination (Nishimura et al., 2007; Nee et al., 2017; Li et al., 2019); as well as that encoding ILITHYA (ILA), a protein that mediates the phosphorylation of eIF2α by activation of GCN2, and hence involved in the modulation of translation under stress recovery (Faus et al., 2018); also, transcripts involved in growth and metabolism such as ribosomal proteins, sucrose synthase, phosphofructokinase, stay green protein, an SnRK1 subunit, implicated in the regulation of carbon metabolism, among others (Table S3). The same analysis using *ago4-2* showed among the up-regulated transcripts again the one encoding AHG1, indicating that AGO4/RdDM pathway indirectly modulates its transcription; in addition, other transcripts standing out encode for various late embryogenesis abundant proteins, and that for ANAC032 transcription factor, involved in photosynthesis inhibition and reprogramming of carbon and nitrogen metabolism (Sun et al., 2019), processes that might be related to delayed germination (Table S5).

Discussion

Although most of the information on the regulatory roles of the RdDM pathway has been analyzed during reproductive processes (Matzke and Mosher, 2014; Matzke et al., 2015), there are a few examples where the contribution of this pathway in the plant response to abiotic stress has been investigated (Ito et al., 2011; Matsunaga et al., 2012; Tricker et al., 2012; Popova et al., 2013; Chan et al., 2016). In this last case, the control of transposable elements has been described as one of the mechanisms regulating the transcription of a diversity of genes (Secco et al., 2015; Au et al., 2017) but only a small number of target genes have been identified (Kurihara et al.,

2008; Au et al., 2017; Iwasaki et al., 2019). Recently, it has been reported that non-canonical RdDM participates in the maternal control of Arabidopsis seed dormancy, as well as in the enhanced dormancy produced by cold during seed development (Iwasaki et al., 2019). In this work, we addressed the involvement of the canonical RdDM pathway in the regulation of the plant response to abiotic stress imposed by salinity during germination, a critical developmental stage during the plant life cycle.

Our results showed that mutants affected in different canonical RdDM core elements exhibit delayed germination under salinity, demonstrating that this pathway participates in the modulation of the germination process in high salt environments. Although RdDM has been implicated in the plant response to environmental stresses under other circumstances, our results show that it also plays an important role in the decision to germinate, a requirement for a plant optimal adjustment under stress.

Unexpectedly, the analysis of seeds lacking AGO4 (*ago4-3*) showed a contrasting behavior, displaying a higher germination rate than wild type seeds under stress conditions. This observation was further investigated showing that this phenotype is not the result of an additional RdDM-independent AGO4 function but instead is due to the action of the AGO6 protein. These data lead us to propose that, under these conditions, when AGO4 is not available AGO6 is able to bind otherwise AGO4-specific siRNAs. Because AGO6 or AGO9 transcript level does not change in the AGO4 mutants used (Figure S6), and given the AGO6 distribution reported in Arabidopsis seeds (Havecker et al., 2010; Eun et al., 2011), it is possible to predict that AGO6 would not be able to fully compensate for AGO4 absence therefore leading to a synthetic phenotype. However, in the *ago4-2* mutant the action of AGO6 was not detected. A plausible explanation for this outcome is that the defective AGO4 protein produced by this this weak allele is still efficiently binding siRNAs, and thus not being available for AGO6 apocryphal binding.

Consonant with the involvement of RdDM in the germination process in a non-stress environment (this work, Kawakatsu et al., 2017), high AGO4 protein levels were detected in dry seeds, which change throughout germination and post-germination.

The AGO4 tissue localization also changes throughout these periods, and in some stages the protein is also localized in root meristematic regions. The AGO4 post-germination distribution, where AGO4 amount decreases in radicles and do not change in cotyledons, are consistent with its participation during those stages, when seedlings have to confront their transition to a different environment.

Germination in the presence of high salt leads to a drastic change in AGO4 distribution; in particular, it becomes highly abundant in the radicle vascular cylinder, and with less intensity but clearly patent in cotyledon vasculature. Under both germination conditions, it can be appreciated that AGO4 protein levels decrease in roots, whereas they do not change much in cotyledons through post-germination stages, keeping its presence in meristematic regions in the latest periods tested. Although at this point, we do not have enough data to explain this remarkable AGO4 allocation shift towards vascular tissues, it is likely that this is a response to the severe stress condition occurring around vascular cells, where salt is transported once taken from the growth medium. The finding of AGO4 and other RdDM components transcripts in root stele (Kawakatsu et al., 2016) supports the distribution we observed. These results point to a possible RdDM regulatory role in the vascular system in response to stress and/or in a long-distance alarm communication strategy as suggested by the siRNA transport to neighboring cells or to a potential translocation through the phloem (Melnyk et al., 2011; Sarkies and Miska, 2014).

As expected, seed germination under high salt concentrations produced different changes in transcript abundance profiles between all lines, when they are compared with those obtained from non-stress conditions. However, not all of them share transcripts involved in protection to salt stress, indicating that the difference in their germination rates is not necessarily due only to salt defense responses. Noteworthy, differences were found when CHH DNA methylation patterns from salinity were compared between the wild type and the *ago4-2* and *ago4-3* lines. Of special interest are those genes showing lower methylation levels in *ago4-2* and *ago4-3* than in wild type presenting genomic features characteristic of RdDM targets (neighboring to a TE) and their closeness to PolV transcripts (Bohmdorfer et al.,

2016). Most of these genes are shared between the two AGO4 mutants; hence, corroborating their methylation dependence on the RdDM pathway, which prevents their transcriptional activation under salinity. Although not many protein coding gene targets were found in this analysis, this outcome is consistent with previous reports showing a small contribution of RdDM-dependent repressive marks on gene expression (Zheng et al., 2013; Gallego-Bartolome et al., 2019).

The majority of the identified genes with these characteristics encode proteins that have been associated with different regulatory roles during germination. *EBF1* (At2g25490) is of particular interest because it encodes a nuclear F-Box protein, component of an Ub-protein ligase (E3) that targets EIN3 for degradation, a key transcription factor playing a positive role in ethylene perception. Deficiency of EBF1 leads to an increase in EIN3 levels inducing hypersensitivity to ethylene (Binder et al., 2007; Merchante et al., 2013), a plant hormone that plays a pivotal role in the control of dormancy and germination (Corbineau et al., 2014). In Arabidopsis and numerous other species, ethylene promotes the germination of non-dormant seeds under diverse adverse environments, including salinity (Wang et al., 2007; Lin et al., 2013), and interestingly enough, disruption of the ethylene-signaling pathway leads to salinity and osmotic stress sensitivity during germination (Leubner-Metzger et al., 1998; Pirrello et al., 2006). These observations are in consonance with our findings showing that the defective CHH DNA methylation in *ago4-2* mutant during germination under salinity results in *EBF1* up-regulation (compared to wild type), which would have a negative effect on ethylene perception, and in consequence would lead to a lower rate and/or capacity of germination. Despite a similar effect on *EBF1* expression and methylation for *ago4-3*, the effect on germination was contrasting, suggesting that the activity of AGO6 in the AGO4 null mutant counteracts EBF1 downstream effects. Further experimentation is needed to establish whether some of the genes whose DNA methylation is affected in *ago4-2* also participate in the signal transduction pathways related to the ethylene function, and/or in those involved with other hormones with recognized roles in dormancy and germination, such as abscisic acid

(ABA) and gibberellic acid, whose signaling pathways have a close communication with that for ethylene (Corbineau et al., 2014).

Because the high AGO4 abundance in root and cotyledon vascular tissues in response to salinity suggests systemic translocation, it is worth to highlight the *PDLP7* gene (At5g37660), whose product is a plasmodesmata-located protein (PDL) involved in cell-to-cell communication (Thomas et al., 2008), and *SKS6* (At1g41830) up-regulated by ABA and encoding a multi-copper oxidase-like protein that participates in cotyledon vascular patterning in Arabidopsis (Jacobs and Roe, 2005).

The participation of the AGO4/RdDM pathway in germination under salinity demonstrated by the RdDM mutant phenotypes is strengthened by the transcriptomic profiles obtained from the RNA-seq analysis, as the comparison with wild type expression patterns exhibited differential accumulation levels for transcripts mostly implicated in dormancy, germination and post-germination processes. For *ago4-2,* showing low germination rate and capacity as the rest of the RdDM mutants, some mRNAs typically accumulated in mature embryos (LEA and storage proteins) stand out among the up-regulated transcripts, as well as an mRNA encoding a stress responsive transcription factor (ANAC032) induced by carbon starvation and involved in the inhibition of photosynthesis, and regulation of carbon metabolism, nitrogen assimilation and amino acid catabolism (Sun et al., 2019). Also, consistent with the *ago4-2* phenotype, the AHG1 transcript showed higher levels than wild type upon salinity treatment. AHG1 is a type 2C protein phosphatase, usually highly abundant during seed maturation, that together with DOG1 (DELAY OF GERMINATION 1) control seed dormancy and germination independent of ABA (Nee et al., 2017; Nishimura et al., 2018). Recently, it has been shown that ethylene partially controls seed dormancy/germination through DOG1 pathway (Li et al., 2019), in coincidence with our finding that RdDM-dependent methylation might control *EBF1* expression, a participant of ethylene signal transduction.

Among the *ago4-3* up-regulated transcripts, we found two encoding type 2C protein (PP2C) phosphatases involved in the regulation of seed dormancy and germination, AHG1 and AHG3. These PP2C phosphatases seem to have overlapping and

distinctive functions, showing similar activities but different spatial and temporal expression patterns (Nishimura et al., 2007; Nishimura et al., 2018). However, ABA receptor transcripts were not detected among the differentially expressed mRNAs, suggesting that, at least a pool of AHG1 and AHG3 phosphatases could be in their active forms, as expected for germinating seeds (Cutler et al., 2010). Furthermore, in agreement with its higher germination rate under salinity, we identified a higher abundance of transcripts involved in active growth and metabolism than in wild type seeds.

As a whole, the methylome and transcriptome data in this work exhibit the role of the AGO4/RdDM pathway in germination and dormancy processes, the latter involved in the modulation of the germination timing; a function more evident under stressful environments (Figure 5b). Also, the resultant information exposes the complexity of the gene expression changes affecting numerous gene networks leading to the characterized phenotypes. It should be noticed that we did not detect any effect on germination rate and capacity when the mutant seeds affected in different core RdDM elements were exposed to optimal growth conditions, showing that any effect of RdDM on *A. thaliana* seed development (Wang and Köhler, 2017; Kirkbride et al., 2019; Chow et al., 2020) does not impact the germination process.

Even though further experimentation is needed to understand the precise role of the RdDM pathway activity in Arabidopsis seed germination and dormancy in response to salt stress, the findings in this work unveil RdDM gene targets, some of which have been previously involved in signaling pathways controlling germination timing, and others that open original questions regarding the control mechanisms involved in the plant responses to salinity, providing a state-of-the-art potential source for the search of targets for improving future farming on salt-affected soils.

Experimental Procedures

Plant material and growth conditions. In this work we used seeds of *Arabidopsis thaliana* (L.). For propagation or germination experiments, seeds were surface sterilized with absolute ethanol for 2 min, washed with 40% commercial bleach

containing 0.02% Triton X-100 (Sigma) for 8 min, and were rinsed five times with sterile milliQ^{R} water. After sowing, seeds were stratified in MS plates for 2 (for Columbia) or 4 (for L*er* and Ws) days in the dark at 4 °C and then transferred to optimal growth conditions in a chamber at 21 °C, 40 – 60 % relative humidity and 16/8 h photoperiod under white light (80 to 100 µmol m^{-2} s^{-1}). Seeds were germinated in round dishes containing 1X MS salts (4.3 g L^{-1}, with macro- and micro-nutrients, without vitamins, Caisson Labs), 1 % sucrose (Research Organics), 0.5 g L^{-1} MES (Calbiochem) and 0.65 % agar (Sigma). For propagation, two-week-old seedlings were transferred (or directly sown) into soil (Metromix 200, Hummert Intl.) and kept under the described growth conditions and with optimal watering until the siliques were dry.

Mutant lines *nrpd1-1*, *nrpe1-11*, *rdr2-1*, *dcl3-1,* and *ago4-6*, in Col-0 background, and *ago4-1* in L*er* background (Zilberman et al., 2003) were obtained from ABRC or NASC, whereas *ago4-2*, *ago4-4* and *ago4-3*, in Col-0 background, were kindly donated by Pablo Vera (Agorio and Vera, 2007), Michael Axtell (Wang and Axtell, 2017) and Blake Meyers, respectively. The seeds of the double mutant *ago4* x *ago6* were kindly donated by J-K Zhu (Duan et al., 2015). For genetic crosses, plants were grown in the conditions described above until the first floral buds were visible. Crossed flowers were marked and plants were grown under optimal conditions as described until seeds were dry. Offspring from introgressed lines and crosses were identified by marker selection and by PCR.

Germination under optimal and stress conditions. Seeds used for germination experiments were collected from plants growing under optimal conditions until senescence. Harvested seeds were stored under low humidity at 4 °C for at least two weeks before use. Germination under salt conditions was achieved in MS media supplemented with 100 to 250 mM NaCl (J. T. Baker), as indicated. Germination was quantified by radicle emergence from five technical replicates of 50 seeds each. All germination experiments were performed with seeds of the same age and were replicated at least five times.

Root length determination. For root length experiments, stratified seeds were sown in square 10x10 cm vertical dishes containing 0.2X MS salts, 1 % sucrose, and 1 % agar or in medium containing 100 mM NaCl. After sowing, plants were grown under the conditions described above. After radicle protrusion, root length was measured by marking the root tip position of every plant every day at the same hour until the 8th day. The measurements were made using the ImageJ® program.

Statistics applied to phenotypic analyses. The germination experiments included in this work were replicated five times using 50 seeds per replicate and five biological replicates. Data were accumulated over time and fit to sigmoidal dose-response curves with variable slope [Y = bottom + {(top 2 bottom)/1 + 10(LogEC50-X)*Hillslope}] also called four-parameter logistic equation. Bottom is the Y value at the bottom plateau (constrained to zero); top is the Y value at the top plateau; LogEC50 is the X value when the response is halfway between bottom and top; Hillslope describes the steepness of the curve. The null hypothesis was that two curve-fit parameters (Hillslope and LogEC50) from each data set were the same (Olvera-Carrillo et al., 2010).

For root length experiments, length was measured every day (n = 15 for optimal conditions and n = 60 for salinity) with three biological replicates; data was accumulated over time and fit to sigmoidal dose response curves with variable slope, as described above.

Protein extraction and western blot assays. Total protein extracts from seedlings, dry seeds or germinating seeds were obtained according to (Olvera-Carrillo et al., 2010). Protein contents were quantified by the Lowry protocol (Lowry et al., 1951) and separated in 10 % SDS-PAGE. Western blots were performed following standard protocols using 30 to 40 µg of protein extracts. Anti-AGO4 antibody (Agrisera) was used in 1:2000 to 1:4000 dilutions, whereas secondary antibody (anti-rabbit horseradish peroxidase; Zymed) was used in 1:20000 or 1:30000 dilutions. Signals were developed with peroxidase substrate from Supersignal West

Pico (Thermo Scientific) and exposed to blue X-ray films (Kodak). Films were photographed using ImageQuant 300 (GE Healthcare) imager.

Fluorescence microscopy. For these experiments, we used seeds from *pAGO4::GFP-AGO4* transgenic Arabidopsis, which were kindly donated by Dr. Qi (Ye et al., 2012). This construction allows the production of a chimeric AGO4 wild type protein fused at its N-terminal with GFP under the control of the native *AGO4* promoter. Seeds were germinated in plates containing 1X MS medium or 1X MS added with 100 mM NaCl. Whole embryos were isolated under the microscope and immediately stained with 20 μg μL^{-1} propidium iodide (Sigma) for 20 min and observed using an Olympus FV1000 upright confocal microscope (Olympus) with a 10X objective. Embryo images were obtained from stacks of individual digital slices with a thickness of 6.23 μm. In the case of the intracellular analysis, a 60X objective was used, and images were obtained from stacks of 500 x 456 pixels individual digital slices with a thickness of 0.8 μm. GFP and propidium iodide were excited at 488 nm and 568 nm, respectively. Fluorescence brightness in the GFP channel was increased with ImageJ without affecting signal distribution.

RNA isolation and high throughput sequencing. Total RNA from germinated seeds was extracted using the hot phenol method (Mylne et al., 2010). For high throughput sequencing analysis, RNA was obtained from wild type, *ago4-3* and *ago4-2* lines germinated in 1X MS or 1X MS added with 100 mM NaCl. To achieve the most homogeneous representation of the different developmental stages for each line, samples were harvested when each seed population reached 50 % germination in every condition examined. RNA-seq libraries were prepared from high quality RNA as verified using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA). Construction of polyA RNA libraries was carried out using True-Seq stranded mRNA Library Prep kit (Illumina, San Diego, CA). High throughput sequencing was performed on duplicate biological samples using the multiplex paired-end 2x75 configuration. Libraries and sequencing were performed by the *Unidad Universitaria de Secuenciación Masiva de DNA (IBt-UNAM)* in a HiSeq 2500 platform (Illumina).

RNA-seq analysis. For RNA-seq analysis, all libraries contained between 20.5 and 49.2 million high quality reads per library with non-over-represented sequences. The obtained reads were aligned to the TAIR10 reference genome (www.arabidopsis.org) using the SMALT software version 0.7.6 (Genome Research Ltd.), and at least 99.74% of reads mapped to the genome. Hit counting was done using samtools (Li et al., 2009). Differential expression analysis of RdDM direct targets were performed using EdgeR (Robinson and Smyth, 2008).

To select and sort the genes whose repression in salt could depend on RdDM methylation (RdDM direct targets) we used a "deregulation" score, with the following logic: genes normally "repressed" by RdDM in salt conditions should have a negative log2("FC Col-0 NaCl vs Col-0 MS"). If they become "de-repressed" by the lack of RdDM, they should have a positive log2("FC ago4 NaCl vs Col-0 NaCl"). The "deregulation" value captures the effect of both conditions by subtracting the "repression" term from the "de-repression" term as follows: log2("FC ago4 NaCl vs Col-0 NaCl") / log2("FC Col-0 NaCl vs Col-0 MS"). This applies only to genes that meet the "repression" condition; genes with the highest "deregulation" score are reported on Table 1 (*ago4-2*) and Table S2 (*ago4-3*). For the rest of RNA-seq analyses, after evaluating differential expression using EdgeR, genes were selected with an absolute log2FC greater than 1.5 and an FDR smaller than 1e -4. Heat maps were generated using the 'heatmap.2' package in R; logarithms of CPM were standardized across genes (rows), thus converted to S.D. from the mean (z-score). The standardized values were also used to cluster genes by the similarity of their profiles over samples. To find genes whose induction or repression by salt is stronger in a mutant than in the wild type, we used a "the differential response to NaCl" score ("diff") which subtracts the effect of salt in the wild type from the effect of salt in the mutant as follows: diff=log2("FC ago4 NaCl vs ago4 MS") - log2("FC Col-0 NaCl vs Col-0 MS").

To select the most affected genes in the *ago4-2* and *ago4-3* mutants, diff was standardized across all genes for each mutant, and genes whose diff is 3.0 or more S.D. above the mean were considered hyper-induced (up-regulated) and those

whose diff is 3.0 or more S.D. below the mean were considered hyper-repressed (down-regulated). Diff values for the selected genes are reported in Tables S3 - S6. The selected genes are also shown in Figures S12 b – e.

Whole Genome Bisulfite Sequencing (WGBS). DNA from the lines and conditions described for the RNA-seq, plus *nrpe1* was isolated using the DNeasy Plant mini kit (QIAGEN) following the manufacturer protocol with minor changes to handle low input samples. DNA quality was assessed by 0.8 % agarose gel electrophoresis as a single band above 20 Kb. One hundred nanograms of the DNA were used to generate sequencing libraries using the Methyl-seq Pico Library Prep kit (Zymo) following the manufacturer instructions. Non-directional DNA libraries were single barcoded and sequenced with its biological duplicate in two lanes of the HiSeq X-Ten platform (Illumina) with a 2x150 Paired-End (PE) configuration. Sequencing was performed in HudsonAlpha Institute for Biotechnology (Huntsville, AL) company. The WBGS libraries for the *nrpe1* mutant were sequenced in a NovaSeq S4 flow-cell at the Advanced Genomics Core of the University of Michigan with a 2x150 PE configuration.

WGBS analysis. Following sequencing, all libraries contained between 40 and 82 million high quality reads. Reads were adapter-trimmed using Trim Galore (Chen et al., 2014), and subsequently mapped using Bismark (Krueger and Andrews, 2011) with bowtie2 PE alignment option. Genome coverage higher than 20X was obtained with a bisulfite conversion higher than 99.1%, assessed by analyzing the cytosine methylation in the plastid genome. Analysis of DNA methylation and differentially methylated regions (DMR) calling were done using the methyl-Kit R package (Akalin et al., 2012). DMRs were defined as a 10% statistically significant (FDR = 0.01) methylation difference between samples using 100 bp sliding windows (50 bp step size). Regions with at least 10 reads were considered for the downstream analysis. Overlapping and cross-analysis of DMRs and genomic features, and the RNA-seq and WGBS, respectively, were performed using the bedtools software. Genomic features for overlapping analyses were taken from TAIR database (www.arabidopsis.org)

and they are based on gene type, including protein-coding genes, pseudogenes, transposable element genes, transposon fragments, ncRNAs, snoRNAs, miRNA, ribosomal RNAs, etc.

All boxplots and heat-maps were generated using the R package. TAIR 10 version of the genome was used as a reference for mapping and as a reference for the genomic models used in the comparisons performed in this work.

Data availability statement

The RNA-seq data generated in this work can be freely accessed through the Gene Expression Omnibus (GEO, ncbi.nlm.nih.gov/geo/) project GSE112051. The WGBS data can be found in GSE156093 project. RIP-seq dataset used in this work (Bohmdorfer et al., 2016) to analyze the Pol V transcribed region proximity can be accessed in the GEO project GSE70290.

Acknowledgements

This work was partially supported by a grant from Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica de la Universidad Nacional Autónoma de México (PAPIIT-UNAM) (IN211816).

We are grateful to Andrzej Wierzbicki for his support during the development of this work and for his critical reading and comments. We also thank Caspar Chater, Tzvetanka D. Dinkova and M. Hafiz Rothi for the critical reading of the manuscript and for valuable comments, to P. Vera, Y. Qi, M. J. Axtel, J.K. Zhu and B. Meyers for the generous gift of some of the seeds used in this work, as indicated in the text; as well as to the Unidad Universitaria de Secuenciación Masiva y Bioinformática, Unidad de Síntesis y Secuenciación de DNA, to the Laboratorio Nacional de Microscopía Avanzada, to R. Grande, V. Jiménez-Jacinto, K. Estrada-Guerra, C. L. Ibarra-Sánchez, C. A. González-Chávez, R. M. Solórzano and M. B. Pérez-Morales for their technical support. We are also thankful to A. Cruz for the preliminary phenotypical analysis in

Arabidopsis roots. This work was partially supported by a grant from Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica de la Universidad Nacional Autónoma de México (PAPIIT-UNAM) (IN211816). V.M.P. and C.M.-M received PhD and MSc fellowships, respectively, from Consejo Nacional de Ciencia y Tecnología (CONACyT)-México, and O.R.B. was receptor of a predoctoral fellowship from Instituto de Biotecnologia. V. Miguel Palomar is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México and has received a CONACyT fellowship #239759.

Author Contributions

V.M.P and A.A.C. conceived the idea and designed the research. V.M.P., A.A.C. and J.L.R. interpreted the data and wrote the article. V.M.P. performed germination and immunodetection experiments with the help of O.R.B. A.G.A. performed the root length experiments. V.M.P., A.G. and A.A.C. analyzed the RNA-seq data. C.M.-M. performed the fluorescence imaging with the aid of V.M.P. V.M.P. performed, analyzed and interpreted the WGBS data with the aid of A.A.C. All authors reviewed the manuscript.

Conflicts of Interest statement

The authors declare no competing financial interests.

Supporting material information

Figure S1. Germination of the Arabidopsis lines displayed in Figure 1, germinated under optimal conditions.

Figure S2. Germination of additional AGO4 null mutant lines on different genetic backgrounds.

Figure S3. Western blot immunodetection of AGO4 protein in the different Arabidopsis lines used in this study.

Figure S4. Dose-dependency effect of salt treatments on the RdDM core protein mutants during germination.

Figure S5. Effect of AGO4 mutants on root growth.

Figure S6. Comparison of transcript accumulation among AGO4/6/9 clade.

Figure S7. AGO4 intracellular localization in germinating seeds.

Figure S8. AGO4 localization in the root apical region.

Figure S9. Genome-wide DNA methylation analyses comparing RdDM mutants to wild type (Col-0).

Figure S10. Locus-specific DNA methylation differences in RdDM mutants.

Figure S11. Identification of RdDM direct targets in *ago4-3*.

Figure S12. RNAseq transcriptome analysis of *ago4-2* and *ago4-3* mutants in response to salinity.

Table S1: List of DMR numbers identified for AGO4 and NRPE1 mutants.

Table S2. List of candidate genes directly modulated by RdDM in *ago4-3.*

Table S3: List of specific *ago4*- 3 salt up-regulated genes.

Table S4: List of specific *ago4-3* salt down-regulated genes .

Table S5: List of specific *ago4-2* salt up-regulated genes .

Table S6: List of specific *ago4*-2 salt down-regulated genes.

References

Agorio A, Vera P (2007) ARGONAUTE4 is required for resistance to Pseudomonas syringae in Arabidopsis. Plant Cell **19:** 3778-3790

Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE (2012) methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol **13:** R87

Au PCK, Dennis ES, Wang MB (2017) Analysis of Argonaute 4-Associated Long Non-Coding RNA in Arabidopsis thaliana Sheds Novel Insights into Gene Regulation through RNA-Directed DNA Methylation. Genes (Basel) **8**

Binder BM, Walker JM, Gagne JM, Emborg TJ, Hemmann G, Bleecker AB, Vierstra RD (2007) The Arabidopsis EIN3 binding F-Box proteins EBF1 and

EBF2 have distinct but overlapping roles in ethylene signaling. Plant Cell **19:** 509-523

Blevins T, Podicheti R, Mishra V, Marasco M, Wang J, Rusch D, Tang H, Pikaard CS (2015) Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in Arabidopsis. Elife **4:** e09591

Bohmdorfer G, Sethuraman S, Rowley MJ, Krzyszton M, Rothi MH, Bouzit L, Wierzbicki AT (2016) Long non-coding RNA produced by RNA polymerase V determines boundaries of heterochromatin. Elife **5**

Borges F, Parent J-S, van Ex F, Wolff P, Martínez G, Köhler C, Martienssen RA (2018) Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in Arabidopsis. Nature Genetics **50:** 186-192

Brosseau C, El Oirdi M, Adurogbangba A, Ma X, Moffett P (2016) Antiviral Defense Involves AGO4 in an Arabidopsis-Potexvirus Interaction. Mol Plant Microbe Interact **29:** 878-888

Chan Z, Wang Y, Cao M, Gong Y, Mu Z, Wang H, Hu Y, Deng X, He XJ, Zhu JK (2016) RDM4 modulates cold stress resistance in Arabidopsis partially through the CBF-mediated pathway. New Phytol **209:** 1527-1539

Chen C, Khaleel SS, Huang H, Wu CH (2014) Software for pre-processing Illumina next-generation sequencing short read sequences. Source Code Biol Med **9:** 8

Chinnusamy V, Zhu JK (2009) Epigenetic regulation of stress responses in plants. Curr Opin Plant Biol **12:** 133-139

Chow HT, Chakraborty T, Mosher RA (2020) RNA-directed DNA Methylation and sexual reproduction: expanding beyond the seed. Curr Opin Plant Biol **54:** 11-17

Corbineau F, Xia Q, Bailly C, El-Maarouf-Bouteau H (2014) Ethylene, a key factor in the regulation of seed dormancy. Front Plant Sci **5:** 539

Corem S, Doron-Faigenboim A, Jouffroy O, Maumus F, Arazi T, Bouche N (2018) Redistribution of CHH Methylation and Small Interfering RNAs across the Genome of Tomato ddm1 Mutants. Plant Cell **30:** 1628-1644

Cubas P, Vincent C, Coen E (1999) An epigenetic mutation responsible for natural variation in floral symmetry. Nature **401:** 157-161

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR (2010) Abscisic acid: emergence of a core signaling network. Annu Rev Plant Biol **61:** 651-679

Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR (2012) Widespread dynamic DNA methylation in response to biotic stress. Proc Natl Acad Sci U S A **109:** E2183-2191

Duan CG, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, Wang B, Lang Z, Zhao Y, Wang X, Wang P, Zhou J, Liang G, Liu N, Wang C, Zhu JK (2015) Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. EMBO J **34:** 581-592

Eun C, Lorkovic ZJ, Naumann U, Long Q, Havecker ER, Simon SA, Meyers BC, Matzke AJ, Matzke M (2011) AGO6 functions in RNA-mediated transcriptional gene silencing in shoot and root meristems in Arabidopsis thaliana. PLoS One **6:** e25730

Faus I, Ninoles R, Kesari V, Llabata P, Tam E, Nebauer SG, Santiago J, Hauser MT, Gadea J (2018) Arabidopsis ILITHYIA protein is necessary for proper chloroplast biogenesis and root development independent of eIF2alpha phosphorylation. J Plant Physiol **224-225:** 173-182

Gallego-Bartolome J, Liu W, Kuo PH, Feng S, Ghoshal B, Gardiner J, Zhao JM, Park SY, Chory J, Jacobsen SE (2019) Co-targeting RNA Polymerases IV and V Promotes Efficient De Novo DNA Methylation in Arabidopsis. Cell **176:** 1068-1082 e1019

Gohlke J, Scholz CJ, Kneitz S, Weber D, Fuchs J, Hedrich R, Deeken R (2013) DNA methylation mediated control of gene expression is critical for development of crown gall tumors. PLoS Genet **9:** e1003267

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

Grover JW, Kendall T, Baten A, Burgess D, Freeling M, King GJ, Mosher RA (2018) Maternal components of RNA-directed DNA methylation are required for seed development in Brassica rapa. Plant J **94:** 575-582

Haag JR, Pikaard CS (2011) Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol **12:** 483-492

Hamera S, Song X, Su L, Chen X, Fang R (2012) Cucumber mosaic virus suppressor 2b binds to AGO4-related small RNAs and impairs AGO4 activities. Plant J **69:** 104-115

Han SK, Wagner D (2014) Role of chromatin in water stress responses in plants. J Exp Bot **65:** 2785-2799

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. Plant Cell **22:** 321-334

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. Nat Genet **38:** 721-725

Hollick JB (2010) Paramutation and development. Annu Rev Cell Dev Biol **26:** 557-579

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

Iwasaki M, Hyvarinen L, Piskurewicz U, Lopez-Molina L (2019) Non-canonical RNA-directed DNA methylation participates in maternal and environmental control of seed dormancy. Elife **8**

Iwasaki M, Hyvärinen L, Piskurewicz U, Lopez-Molina L (2019) Non-canonical RNA-directed DNA methylation participates in maternal and environmental control of seed dormancy. Elife **8**

Jacobs J, Roe JL (2005) SKS6, a multicopper oxidase-like gene, participates in cotyledon vascular patterning during Arabidopsis thaliana development. Planta **222:** 652-666

Kawakatsu T, Nery JR, Castanon R, Ecker JR (2017) Dynamic DNA methylation reconfiguration during seed development and germination. Genome Biol **18:** 171

Kawakatsu T, Stuart T, Valdes M, Breakfield N, Schmitz RJ, Nery JR, Urich MA, Han X, Lister R, Benfey PN, Ecker JR (2016) Unique cell-type-specific patterns of DNA methylation in the root meristem. Nat Plants **2:** 16058

Kinoshita T, Seki M (2014) Epigenetic memory for stress response and adaptation in plants. Plant Cell Physiol **55:** 1859-1863

Kirkbride RC, Lu J, Zhang C, Mosher RA, Baulcombe DC, Chen ZJ (2019) Maternal small RNAs mediate spatial-temporal regulation of gene expression, imprinting, and seed development in Arabidopsis. Proc Natl Acad Sci U S A **116:** 2761-2766

Köhler C, Lafon-Placette C (2015) Evolution and function of epigenetic processes in the endosperm. Frontiers in Plant Science **6**

Kradolfer D, Wolff P, Jiang H, Siretskiy A, Kohler C (2013) An imprinted gene underlies postzygotic reproductive isolation in Arabidopsis thaliana. Dev Cell **26:** 525-535

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

Kurihara Y, Matsui A, Kawashima M, Kaminuma E, Ishida J, Morosawa T, Mochizuki Y, Kobayashi N, Toyoda T, Shinozaki K, Seki M (2008) Identification of the candidate genes regulated by RNA-directed DNA methylation in Arabidopsis. Biochem Biophys Res Commun **376:** 553-557

Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet **11:** 204-220

Le TN, Schumann U, Smith NA, Tiwari S, Au PC, Zhu QH, Taylor JM, Kazan K, Llewellyn DJ, Zhang R, Dennis ES, Wang MB (2014) DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in Arabidopsis. Genome Biol **15:** 458

Lee MH, Jeon HS, Kim HG, Park OK (2017) An Arabidopsis NAC transcription factor NAC4 promotes pathogen-induced cell death under negative regulation by microRNA164. New Phytol **214:** 343-360

Leubner-Metzger G, Petruzzelli L, Waldvogel R, Vogeli-Lange R, Meins F, Jr. (1998) Ethylene-responsive element binding protein (EREBP) expression

and the transcriptional regulation of class I beta-1,3-glucanase during tobacco seed germination. Plant Mol Biol **38:** 785-795

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:** 2078-2079

Li X, Chen T, Li Y, Wang Z, Cao H, Chen F, Li Y, Soppe WJJ, Li W, Liu Y (2019) ETR1/RDO3 Regulates Seed Dormancy by Relieving the Inhibitory Effect of the ERF12-TPL Complex on DELAY OF GERMINATION1 Expression. Plant Cell **31:** 832-847

Lin Y, Yang L, Paul M, Zu Y, Tang Z (2013) Ethylene promotes germination of Arabidopsis seed under salinity by decreasing reactive oxygen species: evidence for the involvement of nitric oxide simulated by sodium nitroprusside. Plant Physiol Biochem **73:** 211-218

Liu X, Zhang H, Zhao Y, Feng Z, Li Q, Yang HQ, Luan S, Li J, He ZH (2013) Auxin controls seed dormancy through stimulation of abscisic acid signaling by inducing ARF-mediated ABI3 activation in Arabidopsis. Proc Natl Acad Sci U S A **110:** 15485-15490

Lopez A, Ramirez V, Garcia-Andrade J, Flors V, Vera P (2011) The RNA silencing enzyme RNA polymerase v is required for plant immunity. PLoS Genet **7:** e1002434

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. J Biol Chem **193:** 265-275

Mahfouz MM (2010) RNA-directed DNA methylation: mechanisms and functions. Plant Signal Behav **5:** 806-816

Manning K, Tor M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet **38:** 948-952

Martinez G, Wolff P, Wang Z, Moreno-Romero J, Santos-Gonzalez J, Conze LL, DeFraia C, Slotkin RK, Kohler C (2018) Paternal easiRNAs regulate parental genome dosage in Arabidopsis. Nat Genet **50:** 193-198

Matsunaga W, Kobayashi A, Kato A, Ito H (2012) The effects of heat induction and the siRNA biogenesis pathway on the transgenerational transposition of ONSEN, a copia-like retrotransposon in Arabidopsis thaliana. Plant Cell Physiol **53:** 824-833

Matzke MA, Kanno T, Matzke AJ (2015) RNA-Directed DNA Methylation: The Evolution of a Complex Epigenetic Pathway in Flowering Plants. Annu Rev Plant Biol **66:** 243-267

Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet **15:** 394-408

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. EMBO J **34:** 20-35

Melnyk CW, Molnar A, Baulcombe DC (2011) Intercellular and systemic movement of RNA silencing signals. EMBO J **30:** 3553-3563

Merchante C, Alonso JM, Stepanova AN (2013) Ethylene signaling: simple ligand, complex regulation. Curr Opin Plant Biol **16:** 554-560

Molinier J, Ries G, Zipfel C, Hohn B (2006) Transgeneration memory of stress in plants. Nature **442:** 1046-1049

Mylne JS, Wang CK, van der Weerden NL, Craik DJ (2010) Cyclotides are a component of the innate defense of Oldenlandia affinis. Biopolymers **94:** 635-646

Nakashima K, Ito Y, Yamaguchi-Shinozaki K (2009) Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. Plant Physiol **149:** 88-95

Nee G, Kramer K, Nakabayashi K, Yuan B, Xiang Y, Miatton E, Finkemeier I, Soppe WJJ (2017) DELAY OF GERMINATION1 requires PP2C phosphatases of the ABA signalling pathway to control seed dormancy. Nat Commun **8:** 72

Nishimura N, Tsuchiya W, Moresco JJ, Hayashi Y, Satoh K, Kaiwa N, Irisa T, Kinoshita T, Schroeder JI, Yates JR, 3rd, Hirayama T, Yamazaki T (2018) Control of seed dormancy and germination by DOG1-AHG1 PP2C phosphatase complex via binding to heme. Nat Commun **9:** 2132

Nishimura N, Yoshida T, Kitahata N, Asami T, Shinozaki K, Hirayama T (2007) ABA-Hypersensitive Germination1 encodes a protein phosphatase 2C, an essential component of abscisic acid signaling in Arabidopsis seed. Plant J **50:** 935-949

Olmedo-Monfil V, Duran-Figueroa N, Arteaga-Vazquez M, Demesa-Arevalo E, Autran D, Grimanelli D, Slotkin RK, Martienssen RA, Vielle-Calzada JP (2010) Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature **464:** 628-632

Olvera-Carrillo Y, Campos F, Reyes JL, Garciarrubio A, Covarrubias AA (2010) Functional analysis of the group 4 late embryogenesis abundant proteins reveals their relevance in the adaptive response during water deficit in Arabidopsis. Plant Physiol **154:** 373-390

Pecinka A, Mittelsten Scheid O (2012) Stress-induced chromatin changes: a critical view on their heritability. Plant Cell Physiol **53:** 801-808

Pirrello J, Jaimes-Miranda F, Sanchez-Ballesta MT, Tournier B, Khalil-Ahmad Q, Regad F, Latche A, Pech JC, Bouzayen M (2006) Sl-ERF2, a tomato ethylene response factor involved in ethylene response and seed germination. Plant Cell Physiol **47:** 1195-1205

Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi MA, Lerbs-Mache S, Colot V, Lagrange T (2005) Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. Genes Dev **19:** 2030-2040

Popova OV, Dinh HQ, Aufsatz W, Jonak C (2013) The RdDM pathway is required for basal heat tolerance in Arabidopsis. Mol Plant **6:** 396-410

Robinson MD, Smyth GK (2008) Small-sample estimation of negative binomial dispersion, with applications to SAGE data. Biostatistics **9:** 321-332

Sarkies P, Miska EA (2014) Small RNAs break out: the molecular cell biology of mobile small RNAs. Nat Rev Mol Cell Biol **15:** 525-535

Schoft VK, Chumak N, Mosiolek M, Slusarz L, Komnenovic V, Brownfield L, Twell D, Kakutani T, Tamaru H (2009) Induction of RNA-directed DNA

methylation upon decondensation of constitutive heterochromatin. EMBO Rep **10:** 1015-1021

Secco D, Wang C, Shou H, Schultz MD, Chiarenza S, Nussaume L, Ecker JR, Whelan J, Lister R (2015) Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. Elife **4**

Skirycz A, Inze D (2010) More from less: plant growth under limited water. Curr Opin Biotechnol **21:** 197-203

Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, Peeters AJ (2000) The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. Mol Cell **6:** 791-802

Sun L, Zhang P, Wang R, Wan J, Ju Q, Rothstein SJ, Xu J (2019) The SNAC-A Transcription Factor ANAC032 Reprograms Metabolism in Arabidopsis. Plant Cell Physiol **60:** 999-1010

Thomas CL, Bayer EM, Ritzenthaler C, Fernandez-Calvino L, Maule AJ (2008) Specific targeting of a plasmodesmal protein affecting cell-to-cell communication. PLoS Biol **6:** e7

Tricker PJ, Gibbings JG, Rodriguez Lopez CM, Hadley P, Wilkinson MJ (2012) Low relative humidity triggers RNA-directed de novo DNA methylation and suppression of genes controlling stomatal development. J Exp Bot **63:** 3799-3813

Vidal EA, Alvarez JM, Gutierrez RA (2014) Nitrate regulation of AFB3 and NAC4 gene expression in Arabidopsis roots depends on NRT1.1 nitrate transport function. Plant Signal Behav **9:** e28501

Wang F, Axtell MJ (2017) AGO4 is specifically required for heterochromatic siRNA accumulation at Pol V-dependent loci in Arabidopsis thaliana. Plant J **90:** 37-47

Wang G, Köhler C (2017) Epigenetic processes in flowering plant reproduction. J Exp Bot **68:** 797-807

Wang Y, Liu C, Li K, Sun F, Hu H, Li X, Zhao Y, Han C, Zhang W, Duan Y, Liu M, Li X (2007) Arabidopsis EIN2 modulates stress response through abscisic acid response pathway. Plant Mol Biol **64:** 633-644

Wang Z, Butel N, Santos-Gonzalez J, Borges F, Yi J, Martienssen RA, Martinez G, Kohler C (2020) Polymerase IV Plays a Crucial Role in Pollen Development in Capsella. Plant Cell **32:** 950-966

Wei L, Gu L, Song X, Cui X, Lu Z, Zhou M, Wang L, Hu F, Zhai J, Meyers BC, Cao X (2014) Dicer-like 3 produces transposable element-associated 24-nt siRNAs that control agricultural traits in rice. Proc Natl Acad Sci U S A **111:** 3877-3882

Wierzbicki AT, Cocklin R, Mayampurath A, Lister R, Rowley MJ, Gregory BD, Ecker JR, Tang H, Pikaard CS (2012) Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the Arabidopsis epigenome. Genes Dev **26:** 1825-1836

Wierzbicki AT, Ream TS, Haag JR, Pikaard CS (2009) RNA polymerase V transcription guides ARGONAUTE4 to chromatin. Nat Genet **41:** 630-634

Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y (2010) DNA methylation mediated by a microRNA pathway. Mol Cell **38:** 465-475

Yaari R, Katz A, Domb K, Harris KD, Zemach A, Ohad N (2019) RdDM-independent de novo and heterochromatin DNA methylation by plant CMT and DNMT3 orthologs. Nat Commun **10:** 1613

Yamaguchi-Shinozaki K, Shinozaki K (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu Rev Plant Biol **57:** 781-803

Yang LP, Fang YY, An CP, Dong L, Zhang ZH, Chen H, Xie Q, Guo HS (2013) C2-mediated decrease in DNA methylation, accumulation of siRNAs, and increase in expression for genes involved in defense pathways in plants infected with beet severe curly top virus. Plant J **73:** 910-917

Ye R, Wang W, Iki T, Liu C, Wu Y, Ishikawa M, Zhou X, Qi Y (2012) Cytoplasmic assembly and selective nuclear import of Arabidopsis Argonaute4/siRNA complexes. Mol Cell **46:** 859-870

Zhai J, Bischof S, Wang H, Feng S, Lee TF, Teng C, Chen X, Park SY, Liu L, Gallego-Bartolome J, Liu W, Henderson IR, Meyers BC, Ausin I, Jacobsen SE (2015) A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis. Cell **163:** 445-455

Zhang H, Zhu JK (2012) Seeing the forest for the trees: a wide perspective on RNA-directed DNA methylation. Genes Dev **26:** 1769-1773

Zheng Q, Rowley MJ, Bohmdorfer G, Sandhu D, Gregory BD, Wierzbicki AT (2013) RNA polymerase V targets transcriptional silencing components to promoters of protein-coding genes. Plant J **73:** 179-189

Zhu JK (2002) Salt and drought stress signal transduction in plants. Annu Rev Plant Biol **53:** 247-273

Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science **299:** 716-719

Figure and Table legends

Figure 1. RdDM and AGO4 are required for the germination process under salt stress. (a – f) Arabidopsis seed germination curves in media containing 250 mM NaCl. (a) Comparison between RdDM core component mutant (*nrpd1a-1, rdr2-1, dcl3-1* and *nrpe1-11*) and wild type (Col-0) lines (p-value = $1e^{-4}$), under Col-0 background. Differences between *nrpe1-11* and *nrpd1a-1*, and *nrpe1-11* and *rdr2-1* were also significant (p-value = $1e^{-4}$ and p-value = $3e^{-4}$, respectively). (b) Germination of *ago4-2* (point mutant allele) compared to wild type (Col-0; p-value = $1e^{-4}$). (c) Comparison between two *AGO4* null mutant alleles (*ago4-3* and *ago4-6*) and wild type (Col-0; p-value = $1e^{-4}$), under Col-0 background. (d) Germination rate and capacity of *ago4-3* complementation lines compared with mutant and wild type (Col-0) lines (p-value = $1e^{-4}$). The wild type allele was introduced by crossing into *ago4-3* homozygous plants and the resultant F2 plants (homozygous for the wild type allele) were used for the experiment. (e) Germination rate and capacity of *ago4-3* introgression lines compared with *ago4-3* mutant and wild type (Col-0; p-value = $1e^{-4}$). The *ago4-3* mutant allele was introduced into wild type plants by genetic crosses and the resultant F2 plants (homozygous for the mutant allele *ago4-3*) were used for the experiment (f) Comparison between wild type (Col-0), *ago4-6*, and *ago6-2* single mutants, and *ago4-6 ago6-2* double mutant (homozygous for both mutant alleles; p-value = $1e^{-4}$). The germination analyses under optimal conditions are shown in Figure S1.

Figure 2. AGO4 protein levels are modulated during germination and in response to stress. (a and c) Western-blot immunodetection of AGO4 protein during germination and early seedling development from wild type (Col-0), *ago4-3* and *ago4-2* seeds, germinated under optimal conditions ds (dry seeds), ss (stratified seeds), 18 h and 36 h (after stratification), hcs (seedlings with emerged hypocotyls and cotyledons), rls (seedlings with two rosette leaves)(Top); Ponceau staining of the membranes used above, as loading reference (Bottom); (b and d) As before except that seeds were germinated in media containing 150 mM NaCl.

Figure 3. AGO4 shows differential tissue distribution during germination and early seedling development under optimal and high salt conditions. (a) GFP-AGO4 fluorescence detection in isolated embryos of 8, 12, 24, 48 and 60 h after imbibition (GFP, Top); propidium iodide fluorescence of the same samples (PI, Middle), both signals merged (Bottom). (b) Samples as in (a), except that seeds were germinated in 100 mM NaCl containing media. See experimental procedures for details. These images are representative samples of at least 8 individuals analyzed per experiment, and these observations were reproduced in four independent experiments. Arrowheads and arrows indicate apical and root meristematic regions, and vascular tissues showing high AGO4 abundance, respectively.

Figure 4. RdDM is required for the repression of a small subset of genes in response to salinity. (a) Comparison between the amount of identified hypo-methylated DMRs in all the genotypes and conditions tested to Col-0. Y-axes show the number of hypo-DMRs in each methylation context. The precise values for each case are shown in the Table S1. (b) RdDM dependence of the amount of hyper-methylated DMRs in wild type seeds subjected to salinity treatments, Methylation is represented in percentage for each genotype/treatment (n = 170). (c) Genomic features of the RdDM-dependent Col-0 DMRs in NaCl affected in each of the RdDM mutants (< 10% of CHH methylation compared to Col-0 NaCl). Bars represent percentage of the total number of features mentioned in the text. (d) Correlation analysis of the salinity-dependent hyper-methylated genes found in *ago4-2* and its transcriptional release from the repressive status imposed by NaCl in RdDM deficient plants (blue dots). Grey dots represent a random subset (n=200) of genes whose methylation status on its promoter region and transcriptional status is shown as control. (e) Distance to the nearest PolV transcript with the selected subset of genes in (d) compared to the same number of random regions with the same length (n = 13).

Figure 5. RdDM-dependent AGO4 activity participates in the control of the germination process. (a) Proportion of salt responsive transcripts in *ago4-2* and

ago4-3 lines (Tables S3 - S6). The classification into early and late embryogenesis (seed maturation) and early and late germination (initial growth processes and post-germination growth, respectively) categories was made considering the highest expression peak for every gene in the analysis. (b) Proposed model for RdDM participation in the control of the germination process in response to adverse environments: in wild type seeds, under stress conditions, canonical RdDM pathway controls the germination process by favoring or repressing the expression of genes needed for its initiation and/or progression. This control is exerted through the DNA methylation (red dots) on its canonical target regions (transposable elements, blue boxes) that neighbors the controlled genes; those genes may be key modulators of the germination process through the control of a broad transcriptional network (indirect targets). When the RdDM is lost this regulation is eliminated, affecting the appropriate response to stress during germination.

Table 1. RdDM target candidate list with a corresponding change in DNA methylation found in *ago4-2*

Table 2. Proposed AGO4 / RdDM targets and their functions

Figure S1. Germination of the Arabidopsis lines displayed in Figure 1, germinated under optimal conditions. (a) RdDM core component mutant (*nrpd1a-1, rdr2-1, dcl3-1* and *nrpe1-11*) and wild type (Col-0) lines, under Col-0 background. (b) *ago4-2* (point mutant) compared to wild type (Col-0). (c) *AGO4* null mutant alleles (*ago4-3* and *ago4-6*) and wild type (Col-0), under Col-0 background. (d) wild type (L*er*), *ago4-1* and *gl1-1* (selection marker for *ago4-1* seeds) null mutant alleles, under L*er* background. (e) wild type (Ws) and *ago4-4* null mutant allele lines, under Ws background. (f) *ago4-3* complementation F2 lines compared with mutant and wild type (Col-0) lines. (g) *ago4-3* introgression F2 lines compared with *ago4-3* mutant and wild type (Col-0). (h) Comparison between AGO4 and AGO6 single and double mutants**.**

Figure S2. Germination of additional AGO4 null mutant lines on different genetic backgrounds. (a) Comparison between wild type (L*er*), *ago4-1* (p-value = $1e^{-4}$) and *gl1-1* (selection marker for *ago4-1* seeds, which does not show any difference compared to the wild type (Ler)) null mutant alleles, under L*er* background. Non-significant differences were found between L*er* and *gl1-1*. (b) Comparison between wild type (Ws) and *ago4-4* null mutant allele lines (p-value = $1e^{-4}$), under Ws background.

Figure S3. Western blot immunodetection of AGO4 protein in the different Arabidopsis lines used in this study. AGO4 protein accumulation levels in wild type (L*er* and Col-0), mutant, F2 complementing (Col-0 X *ago4-3)* and F2 introgression (*ago4-3* X Col-0) lines as described in the main text (Top). Ponceau staining of the same membrane to show protein loading (Bottom).

Figure S4. Dose-dependency effect of salt treatments on the RdDM core protein mutants during germination. (a)(b) RdDM core component mutant (*nrpd1a-1, rdr2-1, dcl3-1, ago4-3* and *nrpe1-11*) and wild type (Col-0) lines subjected to 150 and 200 mM NaCl during germination, respectively. Germination rate showed significant differences between all mutant lines and wild type (p-value = $1e^{-4}$). Germination rate differences between mutant lines were also significant (p-value = $1e^{-4}$), except for *dcl3-1* vs *nrpe1-11,* when germinated in 150 mM NaCl (p -value= 0.124); (c)(d) wild type (L*er*), *ago4-1* and *gl1-1* null mutant alleles, under L*er* background. Germination rate between L*er* and *ago4-1* showed significant differences (p-value = $1e^{-4}$); (e)(f) wild type (Ws) and *ago4-4* null mutant allele lines, under Ws background. In all cases, germination rate differences were significant (p-value = $1e^{-4}$); (g)(h) wild type (Col-0), *ago4-2* and *ago4-3* null mutant allele lines, under Col-0 background. All germination rates showed significant differences (p-value = $1e^{-4}$).

Figure S5. Effect of AGO4 mutants on root growth. (a) Primary root length of wild type (Col-0), *ago4-2* and *ago4-3* Arabidopsis seedlings grown in MS + NaCl (100 mM). Differences in root growth rate between lines were significant (p-value = $1e^{-4}$). (b) Primary root length of wild type (L*er*), *ago4-1* and *gl1-1* (as *ago4-1* selection marker) Arabidopsis seedlings grown as in A. Differences in root growth rate between lines were also significant (p-value = $1e^{-4}$). (c) and (d) Primary root length of the genotype used in (a) and (b), respectively, of plants grown in media without stress.

Figure S6. Comparison of transcript accumulation among AGO4/6/9 clade. Transcript accumulation of the mRNAs corresponding to AGO4, AGO6 and AGO9 proteins are shown for all comparisons used in the RNA-seq. Reads between samples were RPM normalized and error bars show standard deviation

Figure S7. AGO4 intracellular localization in germinating seeds. (a) GFP-AGO4 fusion protein fluorescence in cotyledon epidermal tissue obtained from 8 h germinating seeds. (b) GFP-AGO4 protein fusion fluorescence in radicle epidermal tissue obtained from 8 h germinating seeds. The green punctate pattern indicates the high accumulation of AGO4 protein in the epidermal cells' nuclei of both organs.

Figure S8. AGO4 localization in the root apical region. GFP-AGO4 fluorescence detection in the root tip of isolated embryos after 60 h of imbibition in MS solid media (a) or after 48 h (b) and 60 h (c) in the same media containing 100 mM NaCl. Images at the left correspond to a stack of three images, separated by 0.265 μm. Images to the right show a virtual midpoint longitudinal section of the root tip. These images correspond to representative samples of at least 8 individuals analyzed per experiment. These observations were reproduced in four independent experiments

Figure S9. Genome-wide DNA methylation analyses comparing RdDM mutants to wild type (Col-0). (a) Total DNA methylation In CG (top), CHG (middle) and CHH (bottom) contexts in the different samples used in this work. Proportion is shown in terms of total percentage of methylated cytosines compared to unmethylated for each context, the two biological replicates per genotype/condition are shown. (b) DNA methylation distribution along the five *Arabidopsis thaliana* chromosomes, color key is displayed at the top of each context: CG (left), CHG (center) and CHH (right), sample order is shown at the bottom of the plots.

Figure S10. Locus-specific DNA methylation differences in RdDM mutants. (a) Genomic feature analysis of called DMRs from all conditions tested. Enrichment for each genomic feature is represented as a percentage of the total number of called DMRs. (b) Intersection between *ago4-2* and *ago4-3* mutant DMRs identified under optimal and salinity conditions, respectively. For visual purposes, the circles are not scaled to the DMR number. (c)(d) DNA methylation profile of *ago4-2* called DMRs in the same regions in Col-0, *ago4-2*, *ago4-3* and *nrpe1* mutants obtained from control and salt treatments (73 and 179 DMRs respectively; the figure shows the two replicates of each genotype). (e)(f) Similar to (c) and (d) but using the DMRs identified in *ago4-3* (2479 and 3985 DMRs, respectively).

Figure *S*11. Identification of RdDM direct targets in *ago4-3*. (a) Correlation analysis of the salinity-dependent hyper-methylated genes found in *ago4-3* and its transcriptional release from the repressive status imposed by NaCl in RdDM deficient plants (blue dots). Grey dots represent a random subset (n=200) of genes whose methylation status on their promoter region and transcriptional status is shown as a control. (b) Distance to the nearest PolV transcript of the selected subset of genes in (a) compared to the same number of random regions with the same length (n = 18).

Figure S12. RNAseq transcriptome analysis of *ago4-2* and *ago4-3* mutants in response to salinity. (a) Heat map of RNA-Seq transcriptome analysis, comparing

NaCl vs MS conditions, and including the 1000 genes with the lowest FDR (FDR $<3.0e^{-4}$) from wild type (Col-0). Adjacent lanes (by pairs) represent sample duplicates. The original values (logarithms of CPM) were standardized across genes (rows) and are expressed as S.D. from the mean (z-score). The dendrogram on the left shows the genes clustered by the similarity of their profiles over samples. (b) Shows the response to salt (logFC of NaCl vs MS) for 17382 genes in the *ago4-2* mutant compared to wild type. The coefficient of determination (R^2) for the linear regression between both genotypes is shown. Blue dots are significantly above or below the regression model (their "diff" value is 3 or more S.D. from the mean. For details, see experimental procedures. They represent genes that are hyper-induced (above) or hyper-repressed (below) in the mutant. (c) As panel (b) for *ago-4-3* mutant. (d) Venn diagram showing the overlap of hyper-induced genes in *ago4-2* with those in *ago4-3*. (e) As panel (d) for hyper-repressed genes. The genes included in panels (e) and (d) correspond to upper and lower blue dots in panels (b) and (c), respectively.

Tables

Table 1.

Targets found in ago4-2

AGI	Deregulation in ago4-2 †	DNA methylation difference ‡	Name
AT2G05786 §	3.23818	12.4681	hypothetical protein
AT5G35830 §	2.28452	11.8333	Ankyrin repeat family protein
AT3G25490 §	1.23218	11.0896	Protein kinase family protein
AT3G15300 §	0.734898	18.8961	VQ motif-containing protein
AT3G28940 §	0.635058	15.7812	AIG2-like (avirulence induced gene) family protein
AT5G37670 §	0.604745	19.6897	HSP20-like chaperones superfamily protein
AT3G26470 §	0.576367	27.0098	Powdery mildew resistance protein, RPW8 domain
AT3G26480 §	0.536978	27.0098	Transducin family protein / WD-40 repeat family protein
AT5G52070 §	0.474015	18.2488	Agenet domain-containing protein
AT5G37660 §	0.331682	19.6897	plasmodesmata-located protein
AT1G41830 §	0.225374	19.6587	SKU5-similar 6
AT5G07680	0.114474	12.8979	NAC domain containing protein
AT2G19830 §	0.0149694	10.619	SNF7 family protein

† Deregulation is represented as follows: log2((FC *ago4-2* NaCl vs Col-0 NaCl) / (FC Col-0 NaCl vs Col-0 MS)).

Only targets with repression in Col-0 NaCl vs Col-0 MS were selected

‡ DNA methylation difference is the percentage result from: (*ago4-2* NaCl - (Col-0 NaCl - Col-0 MS)) CHH DNA methylation

§ Target is also found in ago4-3

Table 2.

AGI	Name	Description	Function	References
At2g25490	EBF1	Ub-protein ligase E3	Essential for proper ethylene perception together with EBF2	Binder et al., 2007
At2g19830	VPS32 or SNF7.2	SFN7-domain protein	Involved in vacuolar protein sorting, and component of the ESCORT complex	Winter and Houser, 2006; Gao et al., 2017
At5g35830	At5g35830	Ankirin-repeat family protein	Implicated in targeting cytosol proteins to organellar outer membranes	Becerra et al., 2004; Gissot et al., 2006
At5g37660	PDLP7	Plasmodesmata-located protein 7	Protein containing two DUF26 motifs and implicated in bacterial immunity	Thomas et al., 2008; Aung et al., 2019
At5g52070	At5g52070	Agenet domain-containing protein	Associated to chromatin remodeling with significant impact on gene expression. It has been proposed as an RdDM target	Maurer-Stroh et al., 2003; Kurihara et al., 2008
At3g15300	MVQP	VQ motif-containing protein	This protein is phosphorilated by MAPKs and interacts with WRKY transcription factors	Cheng et al., 2012; Jing and Lin, 2015

At5g07680	NAC4	NAC domain-containing protein	Its transcript is targeted by miR164 and its expression is regulated by AFB3, an auxin receptor	Vidal et al., 2013; Lee et al., 2017

Table 1. RdDM target candidate list with a corresponding change in DNA methylation found in *ago4-2*

Targets found in ago4-2

AGI	Deregulation in ago4-2 †	DNA methylation difference ‡	Name
AT2G05786 §	3.23818	12.4681	hypothetical protein
AT5G35830 §	2.28452	11.8333	Ankyrin repeat family protein
AT3G25490 §	1.23218	11.0896	Protein kinase family protein
AT3G15300 §	0.734898	18.8961	VQ motif-containing protein
AT3G28940 §	0.635058	15.7812	AIG2-like (avirulence induced gene) family protein
AT5G37670 §	0.604745	19.6897	HSP20-like chaperones superfamily protein
AT3G26470 §	0.576367	27.0098	Powdery mildew resistance protein, RPW8 domain
AT3G26480 §	0.536978	27.0098	Transducin family protein / WD-40 repeat family protein
AT5G52070 §	0.474015	18.2488	Agenet domain-containing protein
AT5G37660 §	0.331682	19.6897	plasmodesmata-located protein
AT1G41830 §	0.225374	19.6587	SKU5-similar 6
AT5G07680	0.114474	12.8979	NAC domain containing protein
AT2G19830 §	0.0149694	10.619	SNF7 family protein

† Deregulation is represented as follows: log2((FC *ago4-2* NaCl vs Col-0 NaCl) / (FC Col-0 NaCl vs Col-0 MS)).
Only targets with repression in Col-0 NaCl vs Col-0 MS were selected

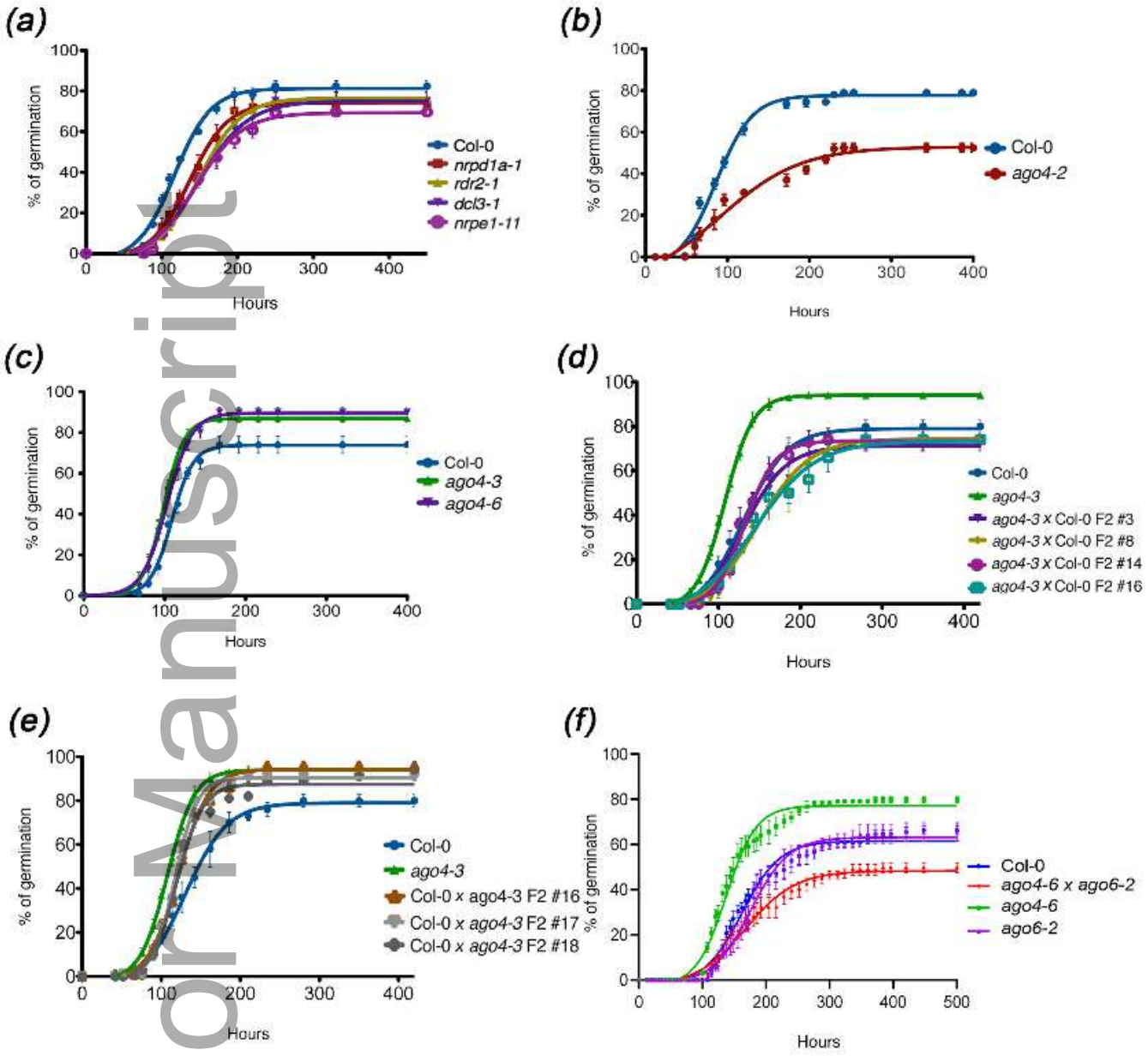
‡ DNA methylation difference is the percentage result from: (*ago4-2* NaCl - (Col-0 NaCl - Col-0 MS)) CHH DNA methylation

§ Target is also found in ago4-3

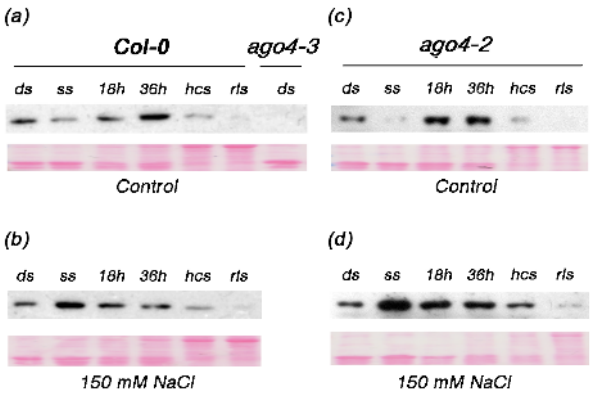
Table 2. Proposed AGO4 / RdDM targets and their functions

AGI	Name	Description	Function	References
At2g25490	EBF1	Ub-protein ligase E3	Essential for proper ethylene perception together with EBF2	Binder et al., 2007
At2g19830	VPS32 or SNF7.2	SFN7-domain protein	Involved in vacuolar protein sorting, and component of the ESCORT complex	Winter and Houser, 2006; Gao et al., 2017
At5g35830	At5g35830	Ankirin-repeat family protein	Implicated in targeting cytosol proteins to organellar outer membranes	Becerra et al., 2004; Gissot et al., 2006
At5g37660	PDLP7	Plasmodesmata-located protein 7	Protein containing two DUF26 motifs and implicated in bacterial immunity	Thomas et al., 2008; Aung et al., 2019
At5g52070	At5g52070	Agenet domain-containing protein	Associated to chromatin remodeling with significant impact on gene expression. It has been proposed as an RdDM target	Maurer-Stroh et al., 2003; Kurihara et al., 2008
At3g15300	MVQP	VQ motif-containing protein	This protein is phosphorilated by MAPKs and interacts with WRKY transcription factors	Cheng et al., 2012; Jing and Lin, 2015

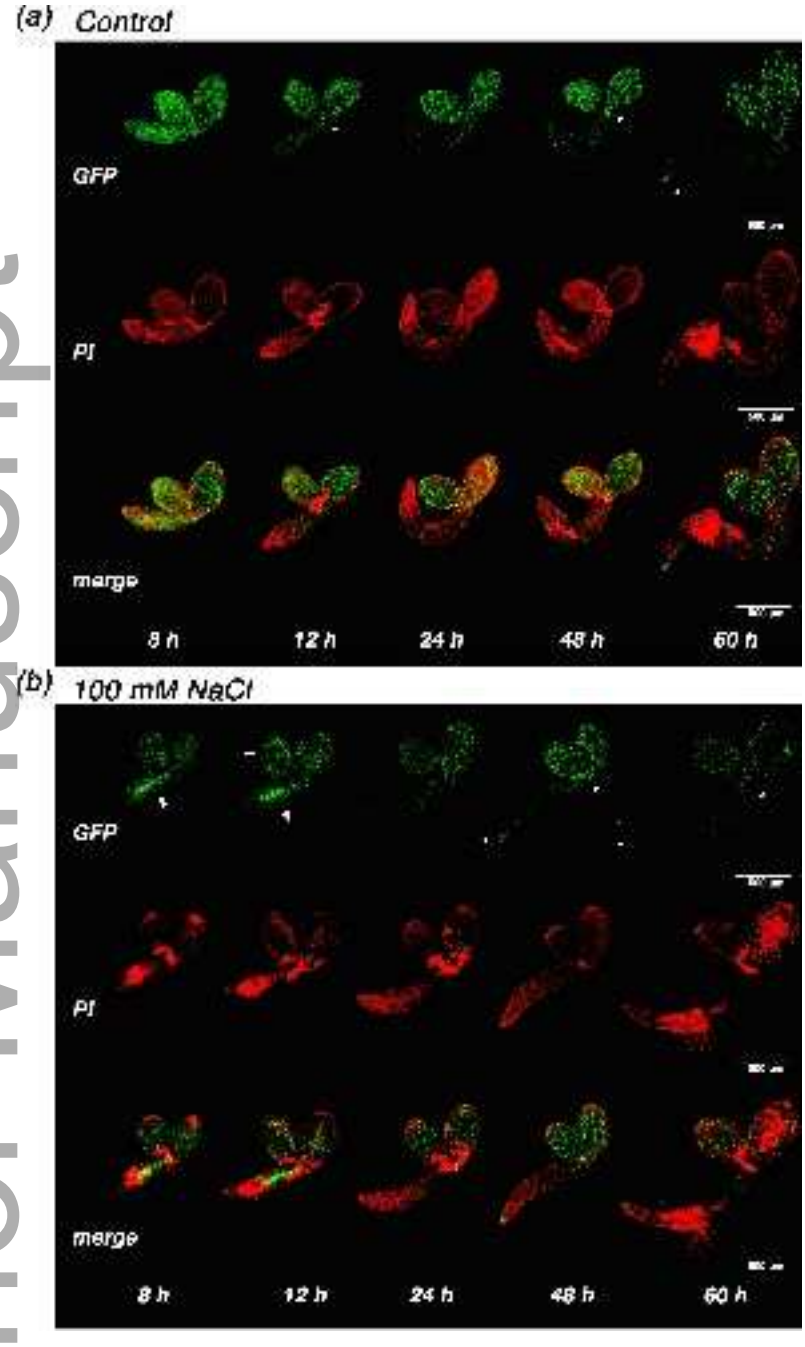
At5g07680	NAC4	NAC domain-containing protein	Its transcript is targeted by miR164 and its expression is regulated by AFB3, an auxin receptor	Vidal et al., 2013; Lee et al., 2017



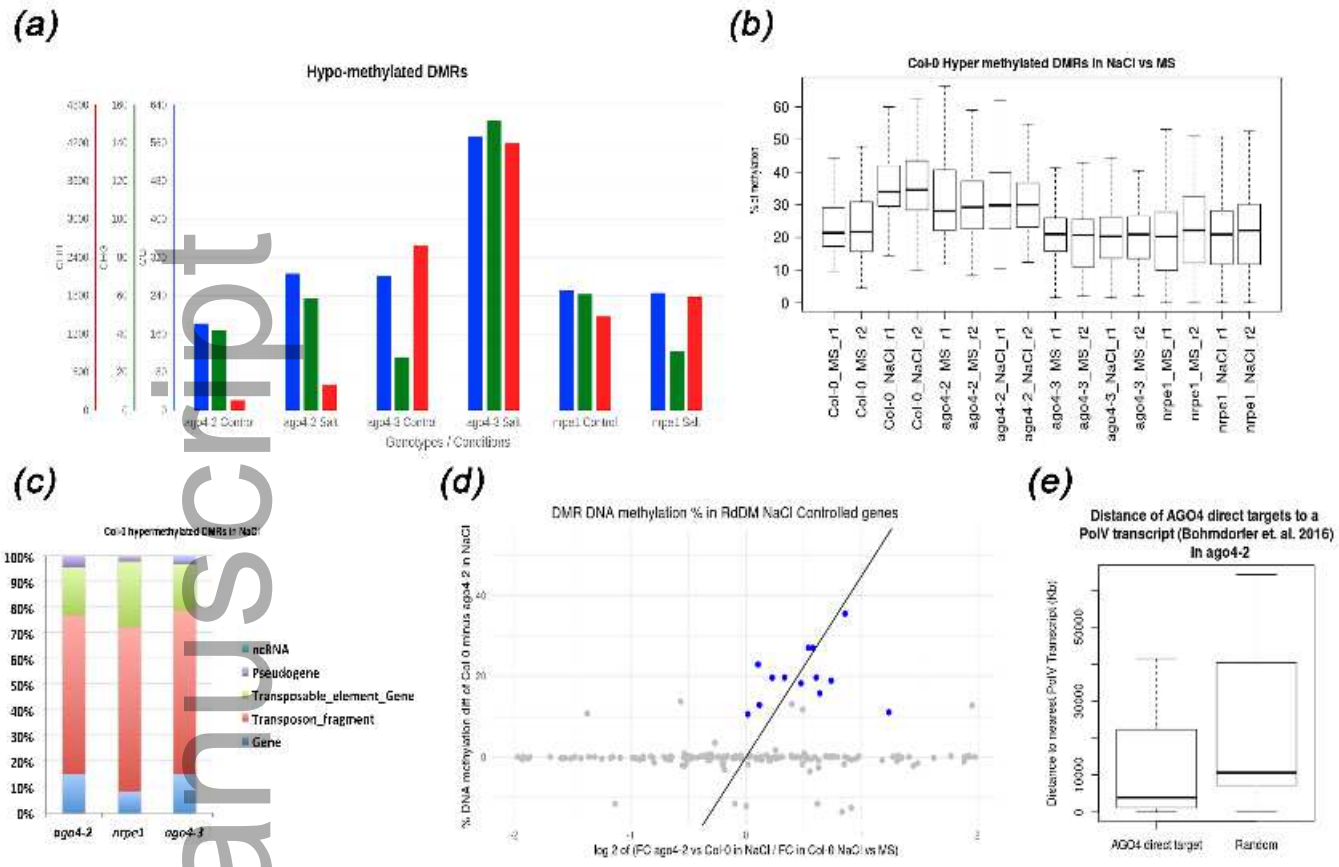
tpj_15064_f1.tif



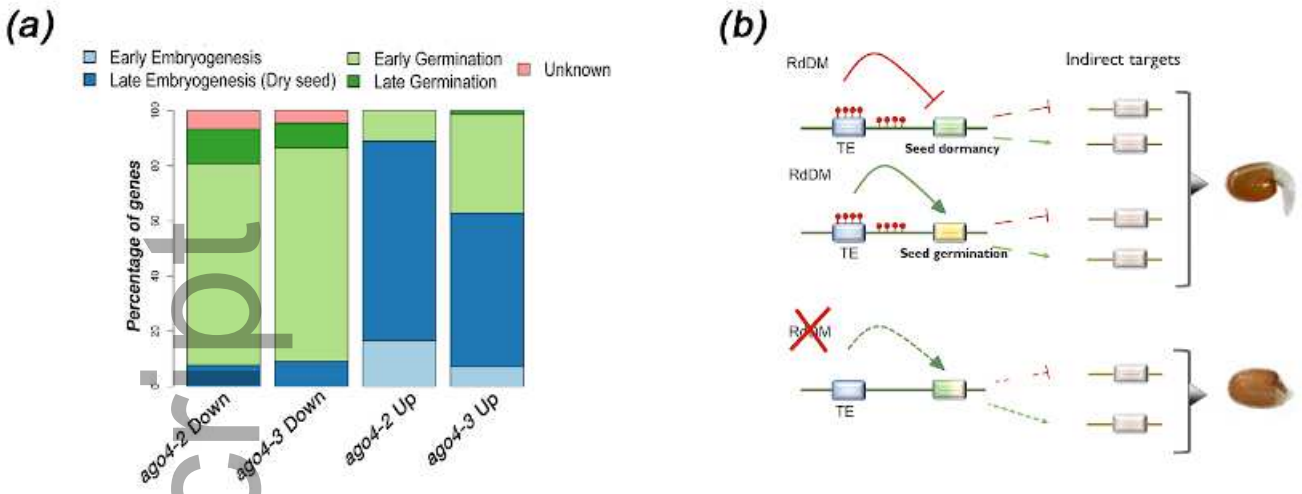
tpj_15064_f2.tif



tpj_15064_f3.tif



tpj_15064_f4.tif



tpj_15064_f5.tif