Titles

Chromatin retained *MUSHER* lncRNA integrates ABA and DOG1 signalling pathways to enhance Arabidopsis seeds dormancy.

Authors

Sebastian Przemyslaw Sacharowski[1], Michal Krzyszton[1]*, Lien Brzezniak[1]*, Karol Jerzy Rogowski[2], Miguel Montez[3], Katarzyna Rosol[1], Magdalena Wrona[1], Ruslan Yatusevich[1], Veena Halale Manjunath[1], Lukasz Szewc[4], Artur Jarmolowski[4], Szymon Swiezewski[1]

1. Laboratory of Seed Molecular Biology, Institute of Biochemistry and Biophysics PAS, Warsaw 02-106, Poland
2. ElementZero Biolabs, Invalidenstraße 65 C/O Edge, Berlin, Germany
3. Department of Cell and Developmental Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK
4. Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan 61-614, Poland

*Equal contribution

Correspondence

Szymon Świeżewski, sswiez@ibb.waw.pl,

ORCID:

Sebastian Sacharowski	0000-0001-8220-9649
Michal Krzyszton	0000-0002-9795-5588
Lien Brzezniak	0000-0002-6932-413X
Karol Jerzy Rogowski	0000-0003-4560-0715
Miguel Montez	0000-0002-1383-0885
Katarzyna Rosol	0000-0001-6663-9884
Magda Wrona	0000-0001-9619-9664
Ruslan Yatusevich	0000-0001-5382-4079
Veena Halale Manjunath	0009-0006-1514-9379
Lukasz Szewc	0000-0002-2219-6513
Artur Jarmolowski	0000-0002-1073-146X
Szymon Swiezewski	0000-0002-2007-2779

Abstract

Many plant lncRNAs regulate gene expression by binding to chromatin, but how they are retained at the target *loci* is unclear. We identify a new, chromatin-localized lncRNA - *MUSHER*, which activates two parallel regulatory pathways to increase Arabidopsis seed dormancy. *MUSHER* is upregulated in response to high temperatures, contributing to the induction of secondary dormancy. It promotes *DOG1* expression by recruitment of the CPSF complex to enhance the proximal cleavage and polyadenylation at the *DOG1* gene. It also increases ABA sensitivity in seeds by activating *PIR1* gene transcription. These genes, located on different chromosomes, are both bound by *MUSHER,* despite lacking sequence homology. The chromatin association of *MUSHER* enables the integration of the DOG1- and ABA pathways to adjust seed germination timing. Additionally, *MUSHER* and other lncRNAs interact with U1 snRNP, which is required for their chromatin localisation, revealing a novel function of U1 snRNP in plants.

Introduction

Multiple long non-coding RNA (lncRNA) are implicated in gene expression regulation in the nucleus by binding to chromatin[1]. Well-known examples of chromatin-bound RNA include *XIST* spreading on the whole X-chromosome and involved in its inactivation in mammals[2,3], and *MALAT1*, tightly bound to chromatin and involved in gene expression and nuclear architecture regulation in humans[4]. There are also many reports of plant lncRNA functioning at the chromatin. This includes *APOLO*, involved in remodelling chromatin structure[5], *SVALKA* proposed to act through induction of RNA polymerase II (Pol II) transcriptional collisions[6] and *PUPPIES*, a lncRNA with a role in delaying Arabidopsis seed germination by controlling Pol II pausing at *DOG1*[7]. *APOLO* was shown to be retained at *trans*-target *loci* via the formation of R-loops[5]. Probably the best-described plant lncRNA *COOLAIR*, involved in flowering time regulation[8,9], has been shown to form a cloud around its locus[10]. Many lncRNAs functioning in the nucleus are involved in environmental sensing in plants. *COOLAIR* regulates vernalisation[9], SVALKA is implicated in cold acclimation in seedlings[6], *ARTA* enhances ABA response[11], *1GOD* affects drought resistance in mature plants[12], *PUPPIES* regulates seed response to salt[7] and *HID1* promotes germination[13]. While an RNA structural element strongly influences *COOLAIR* chromatin association[14], a broader mechanism for lncRNA retention at chromatin remains elusive.

Recently, a U1-based mechanism of human and mouse lncRNA retention at the chromatin was reported[15]. This is based on the physical binding of U1 to a large fraction of lncRNAs[15]. U1 is a highly conserved, small nuclear RNA (snRNA) that, together with conserved proteins including U1C and U1K70, forms the U1 snRNP[16]. U1 snRNP is primarily involved in the first step of splicing by recognising 5' splice sites[17]. However, U1 is also engaged in several other chromatin-related activities: enhancement of Pol II promoters directionality[18]; boosting the speed of RNA polymerase II (Pol II)[19]; and, conserved in plants, suppression of premature cleavage and polyadenylation in introns, known as telescripting[20,21]. Whether U1 snRNP plays a role in chromatin retention in plants is unknown.

Here, we leverage seed biology regulation to explore the functions of lncRNAs. Two key players in seed dormancy are plant hormone abscisic acid (ABA) and a plant-specific protein – DELAY OF GERMINATION 1 (DOG1). Both DOG1 and ABA inhibit seed germination by promoting dormancy, an ability to postpone germination despite favourable external conditions[22]. This allows seeds to ignore short spells of permissive conditions and align germination with seasonal clues. Seed dormancy is initially established when seeds develop on the mother plant[22]. Once this dormancy, known as primary, is released, exposure of imbibed seeds to unfavourable conditions can result in the establishment of secondary seed dormancy[23]. In many plant species, including Arabidopsis, temperature is a key environmental factor required for the efficient establishment of secondary dormancy[24].

DOG1 protein has been suggested to act as part of a thermal-sensing mechanism inducing secondary dormancy[25], but the precise mechanism of how temperature is integrated into the *DOG1* gene regulation during secondary dormancy establishment is not clear[22]. *DOG1* expression is positively correlated with dormancy strength among Arabidopsis accessions, and it has been identified as a major QTL and GWAS for seed dormancy[26,27]. *DOG1* transcription is extensively regulated, including by two independent lncRNA. *DOG1* expression is suppressed

by an antisense transcript *1GOD*, and activated in response to salt by *PUPPIES*, a lncRNA expressed in the *DOG1* promoter[7,28]. In addition, the selection of the *DOG1* proximal polyA site results in the production of *shDOG1* (*short DOG1*), the predominant and functional form of the *DOG1* transcript, while the selection of the distal polyA site results in a presumably non-functional and much less abundant *lgDOG1* (*long DOG1*) transcript[29–31]. Multiple regulators of *DOG1* proximal/distal polyA site selection have been described[30,32], but *DOG1* alternative polyadenylation has so far not been linked to environmental responses.

DOG1 expression is, to a large extent, unaffected by ABA levels, and *dog1* mutations only marginally affect ABA levels in seeds[33]. Interestingly, seed sensitivity to ABA can be modified by both environmental factors and mutations in genes involved in the ABA signalling pathway, without affecting basal ABA levels[34]. ABA is sensed in plants by ABA receptors PYR/PYLs, which once bound by ABA, bind to and inactivate several PP2C phosphatases, including PP2CA[35]. One way plants can modify ABA sensitivity is by interfering with the ABA signal transduction pathway. For example, PP2CA interacting RING finger protein 1 (PIR1) and PIR2 have been shown to mediate PP2CA degradation to enhance abscisic acid response[36]. Although *dog1* mutants do not exhibit altered ABA sensitivity[27], DOG1 protein has been shown to interact with and inhibit the enzymatic activity of PP2CA *in vitro*[37,38]. In summary, DOG1 and ABA pathways are closely intertwined at the protein level, but genetic and transcriptomic analyses indicate that DOG1 function independently from ABA[22,27,33].

Here, we describe *MUSHER*, a novel lncRNA that integrates ABA and DOG1 pathways on the transcriptional level by enhancing *DOG1* and *PIR1* expression. Through genetic analyses, we demonstrate that *MUSHER* acts as a positive regulator of dormancy. Our findings imply that *MUSHER* responds to increased temperature and is essential for high-temperature-induced secondary dormancy in Arabidopsis. At the molecular level, *MUSHER* promotes *DOG1* mRNA proximal polyadenylation by recruiting the CPSF complex. Additionally, through RNA Pol II recruitment, *MUSHER* facilitates the expression of *PIR1 in trans*, thereby enforcing ABA sensitivity. Finally, we show that *MUSHER*, as well as *COOLAIR* and *PUPPIES*, interact with chromatin through U1 snRNP, suggesting the existence of a conserved mechanism of lncRNA-chromatin retention across eukaryotes.

Results

Chromatin-localised lncRNA, *MUSHER*, regulates *DOG1* and seed dormancy.

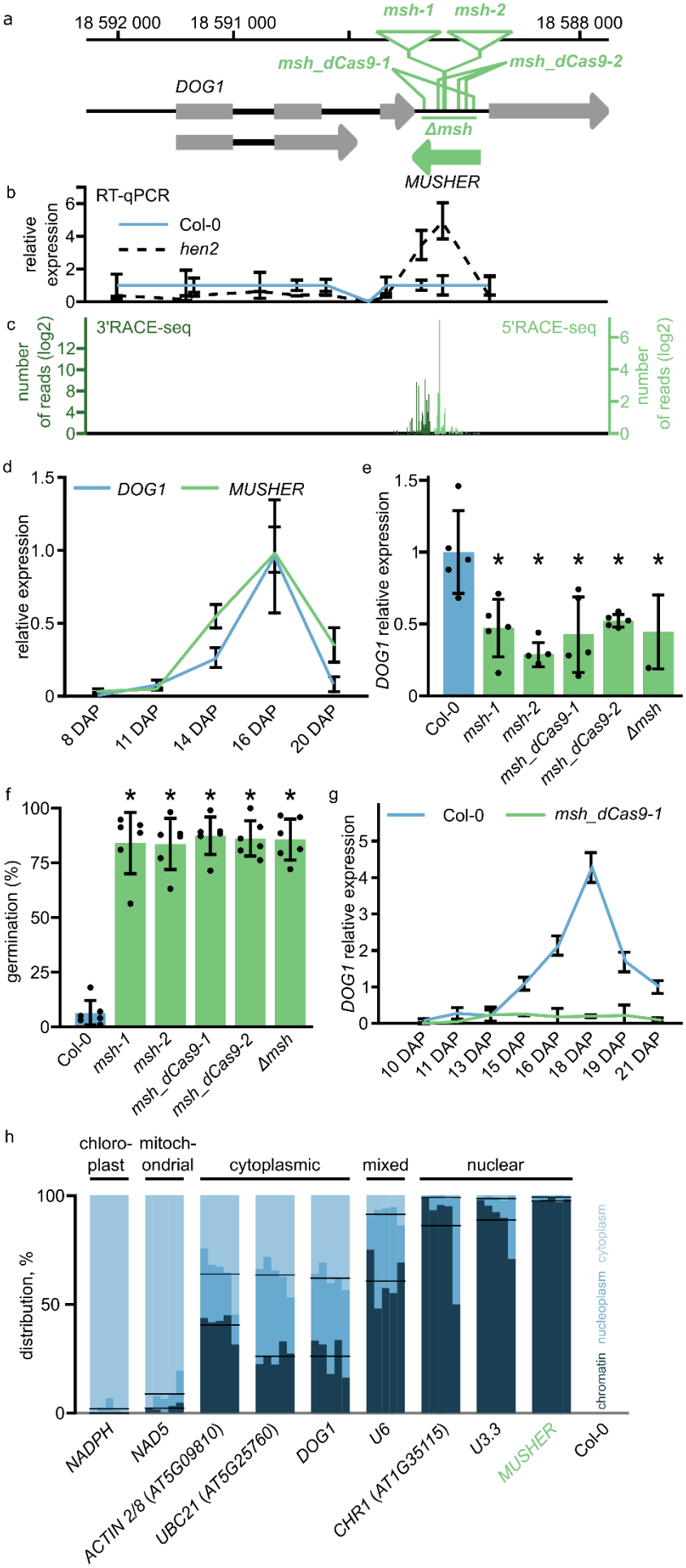
To identify new lncRNA controlling *DOG1*, we performed RT-qPCR analysis with primers spanning the *DOG1* locus in WT and *hen2* mutant seeds (Figs. 1a,b). HEN2 is an RNA helicase that interacts with the exosome complex responsible for the nuclear 3' to 5' degradation of various ncRNAs[39], and its mutation provides a good background to identify unstable lncRNA species[40,41]. We found a *hen2*-dependent signal near the *DOG1* gene 3'end, pointing to transcription of a previously uncharacterised transcript targeted by the exosome in seeds (Fig. 1b). We did not detect increased signals in regions corresponding to *PUPPIES* and *1GOD* transcripts[7,28], indicating that these *DOG1* locus-originating lncRNAs are not targeted by the exosome in seeds. 5' and 3'RACE-seq analyses (Fig. 1c) suggested the existence of an intergenic transcriptional unit between the *SnRK3.6* and *DOG1* genes. Due to its location downstream of *DOG1*, we named this putative lncRNA *MUSHER*.

Measuring this ncRNA level throughout seed development (Fig. 1g) revealed that *MUSHER* transcripts accumulate during seed maturation concomitant to *DOG1* induction. To disrupt the *MUSHER* transcriptional unit, we used two T-DNA insertion lines, *msh-1* and *msh-2* (Fig. 1a), a Cas9 deletion mutant *Δmsh,* and two additional lines: *msh_dCas9-1* and *msh_dCas9-2* with catalytically inactive Cas9 (dCas9). dCas9 acts as a transcriptional repressor by binding to target DNA without inducing cleavage[42]. T-DNA insertion and dCas9 lines, but not the dCas9 controls targeting two intergenomic regions localised 100 kb from *DOG1 locus* showed reduced *MUSHER* levels in seeds (Supplementary Fig. 1a,b). Unexpectedly, *Δmsh,* despite deletion of mapped *MUSHER* start site, showed ectopic expression of an antisense transcript downstream of *DOG1* (Supplementary Fig. 1c and d). Despite this, all tested *musher* mutants demonstrated reduced *DOG1* expression in seeds and increased seed germination (Figs. 1e,f and Supplementary Fig. 1e). The impact of *msh_dCas9-1* on *DOG1* expression is observed throughout seed maturation (Fig. 1g). Because of the observed *MUSHER*-like transcripts in *Δmsh* we decided to exclude this line from downstream analysis.

Since *MUSHER* and *DOG1* are co-expressed during seed maturation (Fig. 1d), we questioned whether *DOG1* expression can have a feedback effect on *MUSHER* transcription. We observed reduced *MUSHER* expression in *dog1* insertional mutants, *dog1-3* and *dog1-4,* but no effect in a *DOG1* large deletion mutant generated using CRISPR/Cas9, *dog1-Δ19* (Supplementary Fig. 1a). This suggested that while *MUSHER* regulates *DOG1* expression, *MUSHER* is not dependent on *DOG1* expression. We speculated that the downregulation of *MUSHER* expression in *dog1* insertional mutants could result from local chromatin state changes caused by T-DNA insertion[43]. Furthermore, *MUSHER* expression was unaffected by mutations in the downstream gene *SnRK3.6* (Supplementary Fig. 1a), confirming that *MUSHER* is an independent transcriptional unit.

Fig. 1| *MUSHER* lncRNA acts as a positive regulator of seed dormancy and *DOG1* expression.

a Schematic representation of the *DOG1-MUSHER* genomic region of chromosome 5 in *A. thaliana*. The direction of transcription is indicated with arrowheads. Grey rectangles represent coding exons and introns with thick black lines. The green arrow represents the newly annotated *MUSHER* transcriptional unit. Numbers along the x-axis on the top indicate genome coordinates (TAIR version 10). Triangles represent T-DNA insertions of two T-DNA lines (*msh-1* and *msh-2*). The green vertical lines show the positions of multiple guides for *msh_dCas9-1* and *msh_dCas9-2* dCas9 lines. The green horizontal line corresponds to the region deleted using Cas9 *(Δmsh)*. **b** Primer walking RT-qPCR analysis across the *DOG1* locus. The position of amplicons corresponds to the *DOG1* locus representation above. Expression levels in the *hen2* mutant were normalised to WT and *UBC21*. Error bars represent the standard deviation of four biological replicates. **c** Positions of *MUSHER* transcript ends found in WT seeds using 3' and 5' RACE-seq, coloured in dark and light green, respectively. **d** Expression profiles of *DOG1* and *MUSHER* transcripts during seed maturation in the WT. Expression levels were normalised to *UBC21* mRNA and 16 DAP. Error bars represent the standard deviation of three biological replicates. **e** *DOG1* relative expression levels, *musher* mutants relative to WT in freshly harvested seeds,



normalised to *UBC21*. *p < 0.05 from two-tailed Student's *t*-test. **f** Freshly harvested seeds of WT and different *musher* mutants were scored for germination. Germination was scored as radical protrusion and counted 4d after sowing. A two-tailed *t-test p*-value was calculated compared to WT (*p < 0.05). **g** *DOG1* expression profile during silique development in WT and *msh_dCas9-1* plants. Expression levels were normalised to *UBC21* and shown at the indicated number of days after pollination (DAP); error bars represent the standard deviation of three biological replicates. **h** Subcellular localisation of selected transcripts in maturing WT Arabidopsis seeds. Each transcript was tested in 5 biological replicates represented as separate bars. Black horizontal lines represent the mean of chromatin and nucleoplasmic fractions contribution of the total amount of indicated transcript.

Next, we tested if *MUSHER* accumulates in the cytoplasm or chromatin, which could provide insight into its mode of action. To test *MUSHER* localisation, we used developing seeds, as this is where *MUSHER* and *DOG1* are co-expressed at high levels, and *DOG1* exerts its function in seed dormancy establishment. We isolated cytoplasmic, nucleoplasmic and chromatin fractions and performed RT-qPCR analysis. Mitochondrial and chloroplast-localised RNA, used as negative controls, are nearly exclusively found in the cytoplasmic fraction. Several tested nuclear-encoded mRNAs showed similar levels in cytoplasm, nucleoplasm and chromatin. U6, a spliceosome component, was detected mainly in chromatin and nucleoplasmic fractions. *MUSHER*, *CHR1*- a transposon-derived RNA and U3.3, involved in co-transcriptional pre-rRNA modification, were predominantly detected in chromatin fraction (Fig. 1h).
In summary, our data showed that *MUSHER* is a lncRNA transcribed downstream of the *DOG1* gene that increases *DOG1* mRNA levels to delay the germination of freshly harvested Arabidopsis seeds. *MUSHER's* chromatin association suggests that it may control *DOG1* expression by binding to its chromatin.

MUSHER lncRNA is required for *DOG1* regulation by CPL1, a Pol II phosphatase

We showed previously that RNA Polymerase II Carboxyl Terminal Domain (CTD) Phosphatase-Like1 (CPL1)-mediated suppression of *DOG1* expression (Supplementary Fig. 2a) requires the 3' end of the *DOG1* locus but not the antisense *1GOD* lncRNA expression from this region[32]. Consequently, the molecular mechanism of CPL1 control over *DOG1* expression remains incompletely understood. Given the effect of *MUSHER* on *DOG1* expression and dormancy, we speculated that CPL1 could act through *MUSHER* regulation. Indeed, multiple *cpl1* alleles, showed elevated *MUSHER* transcript level (Supplementary Fig. 2b), suggesting that *cpl1* is a *MUSHER* lncRNA over-expressor.
To test the requirement of *MUSHER* for CPL1's effect on *DOG1*, we constructed a full-length *DOG1*-luciferase (LUC) fusion with or without *MUSHER* (*pDOG1::LUC-DOG1*[30] and *pDOG1::LUC-DOG1ΔMUSHER*). Transient expression in Arabidopsis seedlings showed that *MUSHER* deletion suppressed *DOG1* upregulation in the *cpl1* plants (Supplementary Fig. 2c).
To cross-validate this, we checked the effect of CPL1 on *MUSHER* locus activity. We generated transgenic plants expressing an internal ribosomal entry site-luciferase gene (IRES-LUC) reporter cassette fused with *MUSHER* and crossed them to *cpl1-7* mutant. The resulting *pMUSHER::MUSHER-IRES-LUC* showed stronger activity in *cpl1-7* when compared to WT, indicating that *MUSHER* is a direct target of CPL1 (Supplementary Fig. 2d). Our data showed that *MUSHER* lncRNA is required for CPL1 to control *DOG1* expression.
Interestingly, inactivation of CPL1 led to preferential expression of proximally over distally polyadenylated *DOG1* mRNA isoform, known respectively as *shDOG1* (*short DOG1*) and *lgDOG1* (*long DOG1*) (Supplementary Fig. 2e). Conversely, *MUSHER* inactivation resulted in reduced

shDOG1 isoform (Supplementary Fig. 2e). Thus, *MUSHER* and CPL1 showed antagonistic effect on the production of *shDOG1*, encoding functional *DOG1* protein and complementing the *dog1* mutant[30]. Importantly, mutating the *cpl1* gene in the *msh* background did not reverse *lgDOG1* polyadenylation preference (Supplementary Fig. 2e), suggesting *DOG1* polyA site selection is directly affected by *MUSHER* and not by CPL1 protein. Taken together, our data showed that CPL1 indirectly affects *DOG1* polyadenylation site selection, by suppressing *MUSHER* expression and activity.

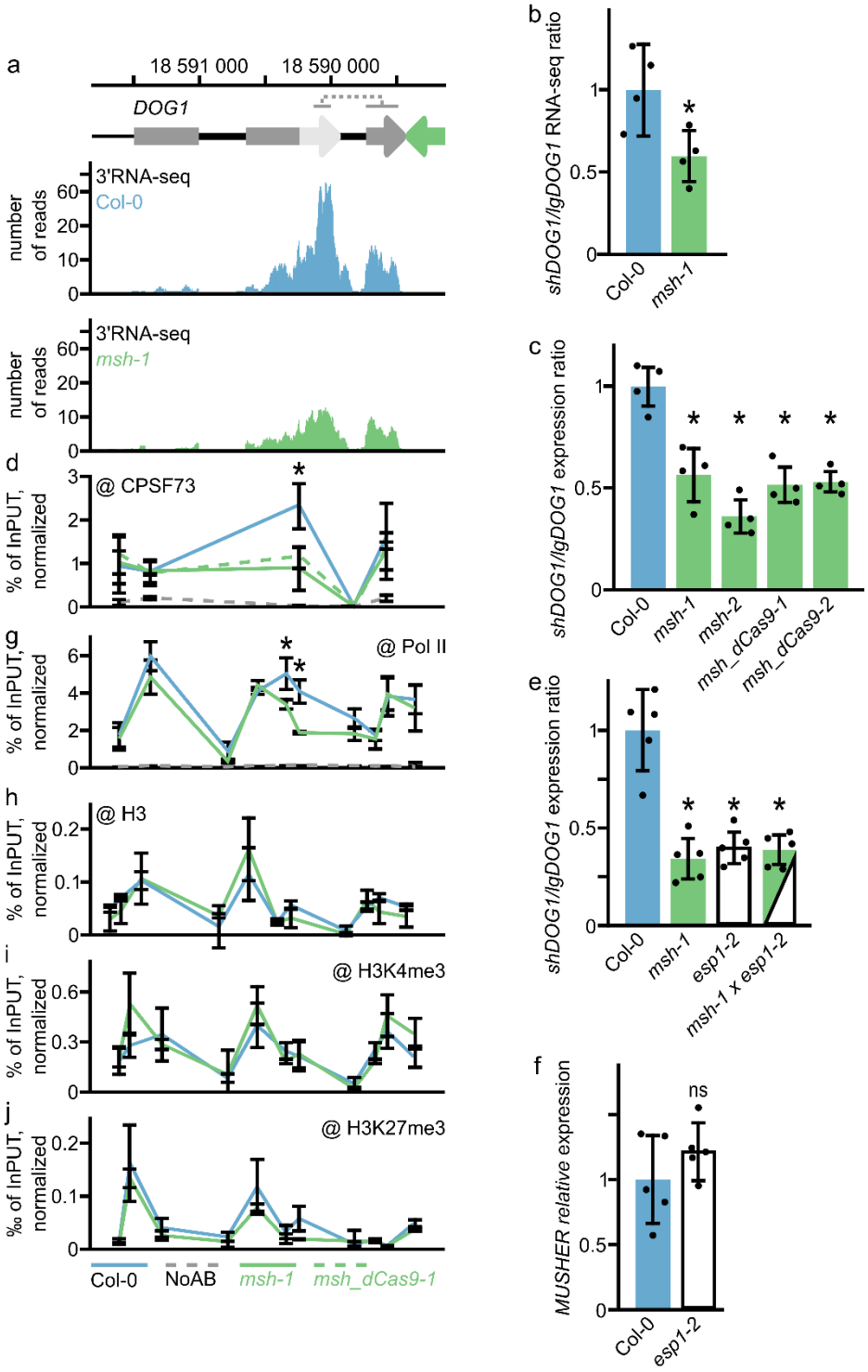


Fig. 2| *MUSHER* enhances *DOG1* proximal polyA site selection. **a** Snapshot of 3'RNA-seq reads for *DOG1* locus in WT and *msh-1* mutant (n=4). Schematic diagram showing *DOG1* gene: exons (grey rectangles), *shDOG1* 3'UTR region (light grey arrow), *lgDOG1* 3'UTR region (grey arrow), and *MUSHER* (green). The regions used to estimate the expression of mRNA isoforms are shown below chromosome coordinates. **b** Ratio of reads corresponding to short and long *DOG1* mRNA isoforms calculated for regions marked on the schematic. Data are the mean of four biological replicates; error bars show standard deviations. **c** The ratio of *shDOG1* and *lgDOG1* isoforms in freshly harvested seeds of WT and *musher* mutants, normalised to *UBC21* and relative to WT (n=4, *p-value < 0.05). **d** CPSF73 occupancy was measured by ChIP-qPCR in WT, *msh-1* and *msh_dCas9-1* mutants. (n=3). **e** The

ratio of *shDOG1* and *lgDOG1* isoforms in seeds of WT and *msh-1, esp1-2* and double *msh-1 esp1-2* mutants, normalised to *UBC21* and relative to WT (n=5, **p*-value < 0.05). The ratio of *shDOG1* and *lgDOG1* isoforms in seeds of WT and *msh-1, esp1-2* and double *msh-1 esp1-2* mutants, normalised to *UBC21* and relative to WT (n=5, **p*-value < 0.05). **f** *MUSHER* relative expression level in seeds of WT and *esp1-2*, normalised to *UBC21* and relative to WT (n=5, ns – not significant). **g** Pol II occupancy was measured by ChIP-qPCR in WT and *msh-1* mutant. (n=3). **h** H3, **i** H3K4me3 and **j** H3K27me3 levels were measured by ChIP-qPCR in WT and *msh-1* mutant. H3K4me3 and H3K27me3 values were normalised to H3 and to *ACTIN7* (*AT5G09810*) gene body. For all ChIP-qPCR experiments presented in this figure: *p*-value from the two-tailed t-test, **p*-value < 0.05, ns- not significant. Error bars represent the standard deviation.

MUSHER promotes *DOG1* proximal polyadenylation by recruiting the CPSF complex

We used 3'RNA-seq to confirm that *shDOG1* is decreased in freshly harvested *msh-1* seeds, while *lgDOG1* remains unchanged when compared to WT (Fig. 2a). Calculation of 3'RNA-seq read density assigned to *DOG1* isoforms revealed a significant reduction of *shDOG1*/*lgDOG1* ratio in *msh-1* (Fig. 2b). Subsequent RT-qPCR analysis confirmed that *shDOG1/lgDOG1* ratio is significantly reduced in all tested *musher* mutants (Fig. 2c).

We previously reported that mutations in pre-mRNA cleavage and polyadenylation factors result in reduced *shDOG1* but not *lgDOG1* expression[30], similar to *MUSHER* knockouts shown here. To confirm changes in *DOG1* mRNA 3'end formation, we performed the ChIP-qPCR analysis of the CPSF73, a subunit of the Cleavage Stimulatory Factor complex[44]. We observed a localised decrease in CPSF73 occupancy at the *DOG1* proximal but not distal termination site in *msh-1* and *msh_dCas9-1* mutants (Fig. 2d). Additionally, we performed a genetic analysis of *MUSHER's* interaction with 3'end formation machinery, by crossing *MUSHER* mutants to *esp-1*. ESP encodes an auxiliary component of Cleavage Stimulatory Factor complex (CSTF64L). We used *esp-1* and not *cpsf73* mutants, because the latter do not produce seeds[45]. Analysis of the *msh-1 esp1-2* double mutant showed no additivity in *DOG1* termination site selection defects suggesting that *MUSHER* controls *DOG1* polyA site selection through CPSF complex (Fig. 2e). This was further supported by no change of *MUSHER* expression in the *esp1* mutant (Fig. 2f). Our results implied that *MUSHER* increases *DOG1* expression by enhancing CPSF complex recruitment to the proximal termination site. This is in agreement with the findings in *cpl1-7* mutant, where *MUSHER* overexpression was accompanied by enhanced *DOG1* proximal polyA site selection (Supplementary Fig. 2). Transcription termination can be coupled with Pol II occupancy changes[46]. We, therefore, profiled Pol II occupancy along the *DOG1* locus in the *musher* mutant. This revealed a significant decrease of Pol II level at the *DOG1* proximal termination region but not at the *DOG1* transcription start site (TSS) or the distal polyA site (Fig. 2g). This suggested that *MUSHER*-promoted proximal polyA site selection is not accompanied by changes in *DOG1* transcription initiation but rather represents a result of *MUSHER* direct activity at *shDOG1* gene 3'end.

The localised decrease of Pol II and CPSF complexes occupancy at the *DOG1* proximal termination site indicates that *MUSHER* promotes the proximal termination site selection rather than controls transcription rates. To test this hypothesis, we took an orthogonal approach and analysed the intensity of *DOG1* transcription using single-molecule RNA FISH (smFISH) in developing embryos[7]. We measured the brightness of nuclear foci identified as transcription sites, as this is generally considered to correlate with the number of RNA molecules present at transcription sites. We also counted the proportion of cells containing transcription sites, which reflects the frequency of promoter bursting[7]. We collected embryos from WT and *msh-1*

346 siliques at ten-time points, corresponding to the experiments showing *DOG1* and *MUSHER*
347 expression during seed maturation (Fig. 1d). *DOG1* smFISH showed a consistent and significant
348 reduction in the number of both nuclear and cytoplasmic *foci* during seed development in *msh-*
349 *1* seeds when compared to WT (Supplementary Figs.3a,b). We observed an increase in *DOG1*
350 transcript number in WT siliques between the 16th and 21st day, which was diminished in *msh-*
351 *1* (Supplementary Figs. 3a,b). Notably, the *msh* mutant showed no significant change in the
352 intensity of nuclear dots identified as transcription sites (Supplementary Fig. 3c)[7] and only
353 minor changes in the frequency of cells with active *DOG1* transcription (Supplementary Fig. 3d).
354 We therefore conclude that, in agreement with Pol II ChIP (Fig. 2g), smFISH indicated no major
355 changes in *DOG1* transcription initiation in *msh-1* seeds.
356 Our data suggest that *MUSHER* promotes *DOG1* proximal termination by locally affecting
357 cleavage and polyadenylation complexes and Pol II occupancy at the proximal transcription
358 termination site without affecting *DOG1* transcription levels. All modes of RNA Pol II activity,
359 including transcription termination, are accompanied and modulated by histone
360 posttranslational modifications. Some lncRNAs may recruit specific histone modelling
361 complexes, including PRC2 involved in H3K27me3 deposition[47,48]. To address the effects of
362 *MUSHER* on the *DOG1* chromatin landscape, we analysed selected active (H3K4me3) or
363 repressive (H3K27me3) chromatin histone marks along the *DOG1* locus. Using ChIP-qPCR, we
364 detected no significant difference in both tested histone modification levels in the *msh-1*
365 mutant throughout the *DOG1* locus (Figs. 2h-j) compared to WT. This result showed that defects
366 at the proximal transcription termination site (TTS) are unrelated to chromatin modifications
367 and agree with no changes in Pol II recruitment for most of the *DOG1* locus. In summary, we
368 propose a model in which *MUSHER* lncRNA specifically enhances *DOG1* proximal termination
369 rather than Pol II recruitment.
370

371 ### *MUSHER* enhances secondary dormancy in response to high temperature

372 *DOG1* is a known regulator of both primary and secondary dormancy[49,50]. Based on the seed's
373 requirement for *MUSHER* to express *DOG1* at high levels and fully develop primary dormancy
374 (Fig. 1), we hypothesised that *MUSHER* may also be involved in secondary dormancy regulation.
375 We used a well-established secondary dormancy induction protocol that consists of prolonged
376 incubation in the darkness at high temperatures followed by germination analysis in permissive
377 conditions[49]. In agreement with defects in primary dormancy, *msh-1* and *msh_dCas9-1* showed
378 weaker secondary seed dormancy, and *cpl1-7*, a mutant with high *MUSHER* expression, showed
379 stronger secondary dormancy when compared to WT (Figs. 3a,b, Supplementary Fig. 4a,b).
380 Consistent with these phenotypes, we observed the down- and up-regulation of the *shDOG1*
381 transcript in *musher* and *cpl1-7* mutants, respectively (Figs. 3c,d).
382

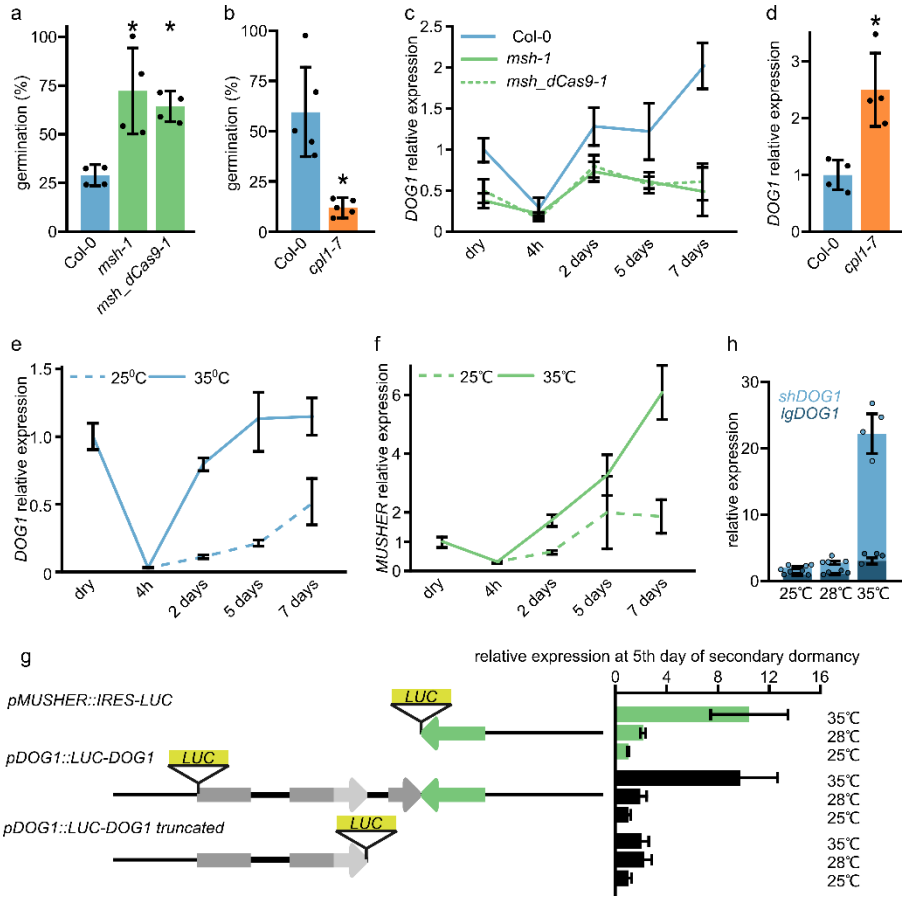


Fig. 3| *MUSHER* activates *DOG1* expression during secondary dormancy induction in response to high temperature. **a** Seed germination of WT and different *musher* mutants after 7 days of secondary dormancy (SD) induction at 35°C. Germination frequency was scored after 5 days. Error bars show the standard deviation of four biological replicates. *p < 0.05. **b** Seed germination of WT and *cpl1-7* mutant seeds after 7 days of SD induction. Germination was scored after 5 days. Error bars show the standard deviation of five biological replicates. *p < 0.05. **c** The expression profile of *DOG1* was analysed using RT-qPCR in seeds of WT, *msh-1* and *msh_dCas9-1* subjected to secondary dormancy induction. Error bars show the standard deviation of four biological replicates. *UBC21* mRNA was used as a reference gene. **d** The *DOG1* relative expression level in seeds of WT and *cpl1-7* mutant, subjected to SD induction. Data normalised to *UBC21* and related to WT. p-value from the two-tailed *t*-test, **p*-value < 0.05. Error bars show standard deviations. **e** *DOG1* and **f** *MUSHER* expression profile in after-ripened seeds subjected to secondary dormancy induction at 25°C and 35°C. Expression was normalised to *UBC21* mRNA and dry seed levels. Bars represent the standard deviations from at least four biological replicates. **g** Schematic representation of constructs introduced to transgenic plants. RT-qPCR analysis of luciferase mRNA levels in seeds subjected to secondary dormancy induction under temperatures indicated on the right. *LUC* expression level measured after 5 days of SD induction, normalised to *UBC21* and related to 25°C for each construct. **h** Long and short *DOG1* expression analysis on the 5th day of secondary dormancy induction in WT at temperatures indicated. Error bars show the standard deviation of four biological replicates. *UBC21* mRNA was used as a reference gene. p-value from the two-tailed t-test, **p*-value < 0.05.

Elevated temperature is a key environmental factor responsible for secondary dormancy establishment[51]. Consistent with previous studies[7,49,52,53], we showed that *DOG1* expression increases in response to high temperatures during secondary dormancy establishment (Fig. 3e). At 35°C, *DOG1* expression dropped initially during imbibition but rose to levels observed in dry seeds on the 2nd day of secondary dormancy induction and remained constant for the rest of the experiment (Fig. 3e). Similarly, high temperatures augmented *MUSHER* induction in imbibed seeds (Fig. 3f). *MUSHER* levels were similarly reduced by imbibition and recovered at 2nd day,

but, in contrast to *DOG1*, continued to rise during the entire experiment length without reaching the plateau. This suggests that *MUSHER* is more responsive to secondary dormancy-inducing conditions than *DOG1* (Fig. 3f). Based on this observation, we speculated that *MUSHER* could be responsible for temperature sensing and the incorporation of this signal into *DOG1* regulation.

Next, to analyse the temperature responsiveness of *MUSHER* during secondary dormancy induction, we took advantage of luciferase reporter lines (Supplementary Fig. 2c,d). We induced secondary dormancy at different temperatures (25, 28, and 35°C) and measured the LUC activity for *DOG1-LUC*, *MUSHER-LUC* and the *DOG1trunc-LUC*, which is devoid of *MUSHER* region. We noted that *MUSHER* and *DOG1* reporters showed similar gradual increases in expression in response to increasing temperatures (Fig. 3g). Interestingly, *DOG1* reporter devoid of *MUSHER* region (*pDOG1::DOG1-LUC truncated*) showed no induction at 35°C (Fig. 3g), indicating *MUSHER* is required for the high temperature-driven *DOG1* upregulation. The responsiveness of *MUSHER* to high temperatures, along with *MUSHER* requirement for high temperature-driven *DOG1* upregulation, indicated that *MUSHER* conveys temperature signals during the induction of secondary dormancy. In agreement with these findings, we show the endogenous *shDOG1* expression in wild-type seeds during secondary dormancy establishment is highly induced at 35°C compared to 28°C (Fig. 3h), with *lgDOG1* transcript being only slightly affected. This is consistent with the concept of *MUSHER* acting through the polyA-site selection (Fig. 2).

In summary, our data are consistent with a model in which *MUSHER* lncRNA is induced during primary and secondary dormancy and activates DOG1 expression by promoting its proximal polyA site selection both during primary and secondary dormancy. In addition, we showed that during secondary dormancy induction, *MUSHER* lncRNA is required for *DOG1* upregulation in response to high temperature.

MUSHER enhances the *PIR1* gene expression independently of its role in *DOG1* regulation.

3'RNA-seq analysis revealed 166 differentially expressed genes, including *DOG1,* between WT and *msh-1* freshly harvested seeds (Fig. 4a). Among the downregulated genes, *PIR1* (*PP2CA INTERACTING RING FINGER PROTEIN 1*) showed the strongest decrease (Fig. 4a). *PIR1* encodes an E3 ligase that enhances the abscisic acid (ABA) response and is highly expressed in maturing siliques[36]. Inspection of 3'RNA-seq coverage plots showed a strong reduction at both *PIR1* and an upstream localised *AT2G35320* gene (Fig. 4b). RT-qPCR confirmed the strong downregulation of *PIR1* expression in *musher* mutant, when tested in dry seeds (Fig. 4c), as well as in secondary dormancy-induced seeds (Fig. 4d), suggesting MUSHER activates *PIR1* expression. Interestingly, *PIR1* expression remained unchanged in the *dog1* mutant. This indicates that *PIR1* is activated by *MUSHER* independently from *DOG1* (Figs. 4c,d, Supplementary Fig. 5a,b). Consistently, increased expression of *MUSHER* in *cpl1-7* led to higher *PIR1* levels (Fig. 4d). ChIP-qPCR analysis revealed a significant decrease in Pol II occupancy throughout the entire *PIR1* locus in *msh*, suggesting that in contrast to *DOG1*, *MUSHER* regulates *PIR1* transcription levels rather than termination of *PIR1* transcripts (Fig. 4e). ChIP-qPCR analysis along *PIR1* locus showed a TSS-localised decrease in H3K4me3, slight decrease in H3 and no change in H3K27me3 levels in *msh-1* (Figs. 4f-h), confirming the lower transcriptional activity of *PIR1* promoter.

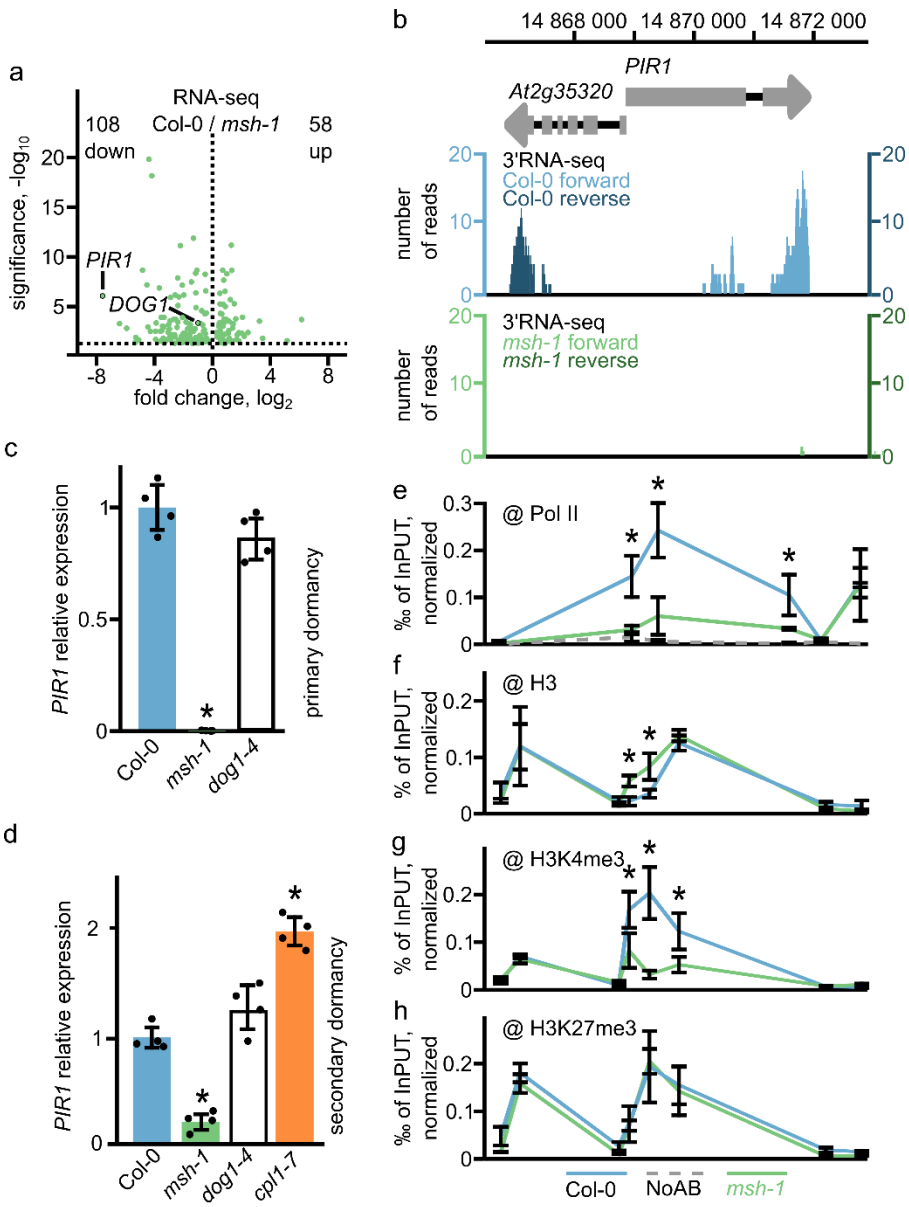


Fig. 4| *MUSHER* enhances *PIR1* expression independently of its role in *DOG1* regulation. **a** Volcano plot showing differentially expressed genes for WT and *msh-1* selected using DESeq2; absolute log2FC >log2 (1.5), FDR<0.05. *DOG1* and *PIR1* genes are highlighted. The number of downregulated (down) and upregulated (up) genes is provided on the plot. **b** Diagram of 3′RNA-seq reads for *PIR1* locus in WT and *msh-1* mutant. Schematic diagram showing *PIR1* locus: exons (grey rectangles), 3'UTR region (grey arrow) **c** The *PIR1* relative expression level in freshly harvested seeds (primary dormancy) of WT, *msh-1* and *dog1-4* mutants, normalised to *UBC21* and relative to WT. Error bars show the standard deviation of four biological replicates. **p*<0.05. **d** The *PIR1* relative expression level in WT *msh-1*, *dog1-4* and *cpl1-7* mutant seeds, after 5 days of SD induction, normalised to *UBC21* and relative to WT. Error bars show the standard deviation of four biological replicates. **p*<0.05. **e** Pol II **f** H3 **g** H3K4me3 and **h** H3K27me3 occupancy was measured by ChIP-qPCR in WT, *msh-1*. (n=3) H3K4me3 and H3K27me3 values were normalised to H3 and to *ACTIN7* (*AT5G09810*); Pol II and H3 were normalised to *ACTIN7* (*AT5G09810*). For all ChIP experiments **P*< 0.05, ns- not significant (two-tailed t-test). Error bars represent the standard deviation.

Our results showed that *MUSHER* activates *PIR1* expression independently of its role in *DOG1* gene regulation. In addition, our data suggest that *MUSHER* regulates transcriptional activity of the *PIR1* promoter, whereas at *DOG1*, it specifically contributes to proximal termination site selection. Importantly, both *DOG1* and *PIR1* regulation by *MUSHER* is compatible with its chromatin association (Fig. 1).

MUSHER lncRNA activates *PIR1* expression through *PIR1-DOG1* chromatin interaction

We showed that *MUSHER* is chromatin-bound (Fig. 1h). To identify *MUSHER* binding sites, we conducted a Chromatin Isolation by *MUSHER*-RNA Purification followed by DNA-seq (ChIRP-seq) in seeds[55]. This revealed 23 DNA regions bound by *MUSHER* lncRNA (Supplementary Table 1). The two most significantly enriched are *PIR1* and *DOG1 loci*. Quantifying the ChIRP signal at *DOG1* and *PIR1 loci* shows that target enrichment is greatly diminished in the *msh-1* mutant (Figs. 5a,b). In addition, ChIRP using *LacZ* probes, a commonly used negative control, also failed to recover significant signals at both targets (Figs. 5a,b). In agreement with our conclusion that *MUSHER* regulates *PIR1* transcription initiation, *MUSHER* peaks centred around *PIR1* TSS. On the *DOG1* locus, *MUSHER* ChIRP-seq detected two strong peaks, one at a site where it is produced and one colocalising with the *DOG1* exon2-intron2 region, just upstream of the proximal polyA site, the selection of which is promoted by *MUSHER* (Fig. 5a).

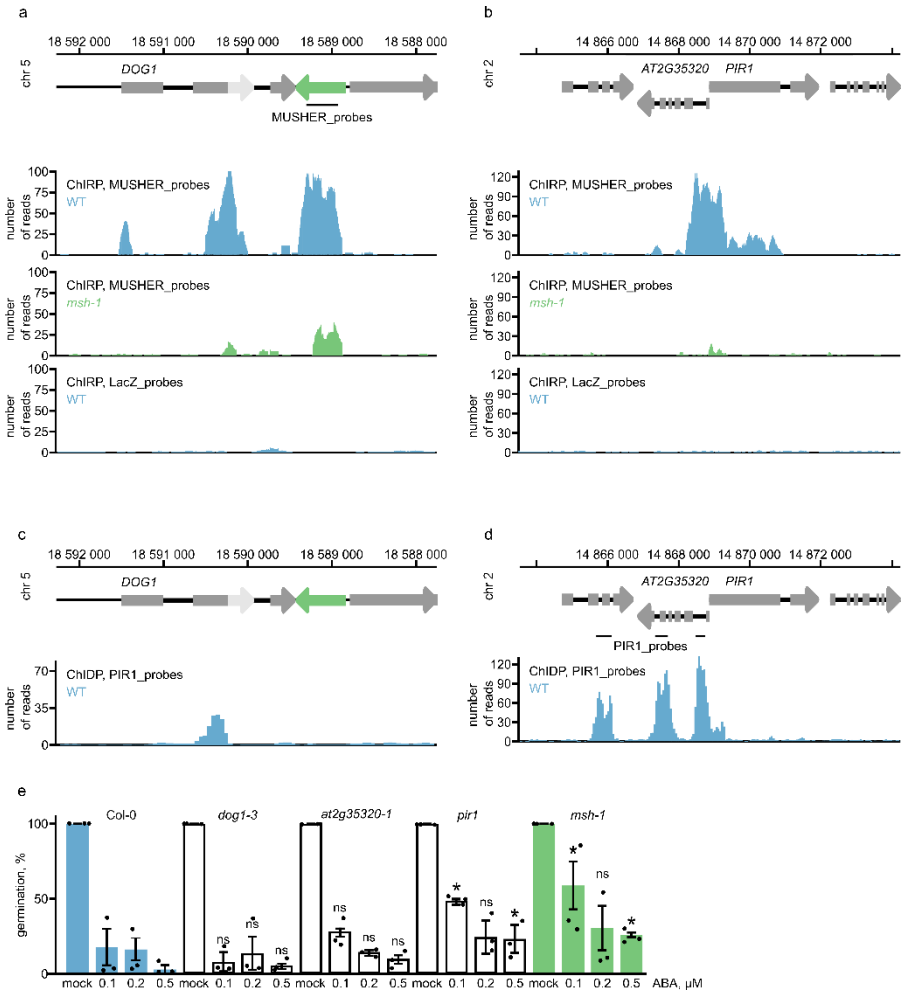


Fig. 5| *MUSHER* lncRNA integrates ABA and DOG1 pathways by interacting with *PIR1* TSS and *DOG1* proximal termination site. Snapshot of ChIRP-seq in WT and *msh-1* mutant using *MUSHER*-specific and *LacZ*-specific probes on **a** *DOG1* and **b** *PIR1 loci*. Schematic diagram showing *DOG1* and *PIR* locus: exons (grey rectangles), *shDOG1* 3'UTR region (light grey arrow), *lgDOG1* 3'UTR region (grey arrow), and *MUSHER* (green). The lines below the scheme mark regions used to design *MUSHER* probes. **c** Snapshot of ChIDP-seq in WT using *PIR1*-specific probes on *DOG1* and **d** *PIR1 loci*. Data are the mean of two biological replicates. The lines below the scheme mark regions used to design *PIR1* probes. **e** The germination rate of WT, *dog1-3*, *at2g35320-1*, *pir1* and *msh-1* mutant seeds on MS media supplemented with various ABA concentrations. *P-value<0.05 from two-tailed Student t-test. Error bars show standard deviations from three biological replicates.

Our data showed that the *DOG1* proximal polyA site and *PIR1* TSS are bound by *MUSHER* lncRNA. Together with the lack of *PIR1* expression changes in the *dog1* mutant (Fig. 4c), this suggested that both *DOG1* and *PIR1* are direct and independent *MUSHER* targets. Of note, we could not find a common DNA motif or R-loop enrichment among *MUSHER* targets, indicating that *MUSHER* may not rely on a sequence-based mechanism for target specification. Whereas

571 the *DOG1 locus* is localised next to *MUSHER* on chromosome 5, the *PIR1 locus* is localised on
572 chromosome 2. This encouraged us to test the physical proximity of *PIR1* and *DOG1 locus*
573 chromatin.

574 To this end, we used *PIR1* probes to fish out crosslinked DNA sequences followed by NGS. This
575 assay is based on a similar principle as PiCh (Proteomics of isolated Chromatin segments)[56] but
576 analyses DNA rather than proteins attached to a specific *locus*. This method was called
577 chromatin-DNA precipitation assay (ChIDP). The use of glutaraldehyde combined with the
578 covalent link between probes and beads allowed us to use more stringent conditions, resulting
579 in a high signal-to-noise ratio (Figs. 5c,d, Supplementary Table 2). Surprisingly, we discovered
580 that *PIR1* DNA interacts with the proximal polyA site of *DOG1* but not with *MUSHER*. This
581 suggests that *MUSHER* activity on *PIR1* is mediated by long-range interaction established by the
582 proximal polyA site of *DOG1*, which is occupied by *MUSHER*.

583 *DOG1* and ABA are the two main regulators of dormancy in Arabidopsis[57]. We confirmed the
584 published results that *pir1* mutants show decreased sensitivity to ABA-mediated germination
585 inhibition (Fig. 5e)[36]. Notably, a mutant in a gene directly upstream of *PIR1* (*at2g35320-1*)
586 showed no significant changes in ABA sensitivity compared to WT. In agreement with published
587 results, *dog1* mutant seed germination was suppressed by ABA in the same way as WT[27].
588 Importantly, multiple *musher* alleles showed much weaker sensitivity to ABA-mediated
589 inhibition of germination across all tested ABA concentrations, similar to the *pir1* mutant (Fig.
590 5e, Supplementary Fig. 6a,b). Therefore, in contrast to the DOG1 protein, *MUSHER* lncRNA does
591 modulate ABA responsiveness. Our data are consistent with a model where *MUSHER* lncRNA
592 enhances the ABA sensitivity by activating *PIR1* gene expression. *MUSHER's* ability to activate
593 *DOG1* and *PIR1* points to its role in integrating the DOG1 and ABA-dependent repression of seed
594 germination.

595 U1 snRNP mediates *MUSHER* association to chromatin

596 *MUSHER* physically binds to and regulates *DOG1* and *PIR1* gene expression (Figs. 1,2,5). Since
597 most *MUSHER* transcripts are localised at the chromatin, it suggests that an active mechanism
598 exists that sequester *MUSHER* at the chromatin (Fig. 1h).
599 In addition to its well-known splicing function, U1 snRNP has been shown in mammals to bind
600 various lncRNAs, facilitating their chromatin tethering[15]. Interestingly, we identified four
601 putative U1 binding sites within the *MUSHER* sequence, even though we failed to detect
602 *MUSHER* splicing (Supplementary Table 3 and Fig. 1c). Therefore, we analysed whether
603 *MUSHER* interacts with U1 snRNP. Using probes specific for *MUSHER*, but not *LacZ* probes, we
604 selectively recovered U1, but not U4 or U6 RNA using formaldehyde-assisted RNA
605 ImmunoPrecipitation (RIP) (Fig. 6a). Supporting the specificity of the approach, U1 recovery was
606 strongly diminished when we repeated the experiment using *MUSHER* probes in *msh-1* lines
607 (Fig. 6b). Next, we validated the *MUSHER*-U1 association by RNA Immunoprecipitation (RIP)
608 using the U1 snRNP component U1-70K. We used a recently reported *u1-70k* complementing
609 line expressing GFP tagged U1-70K driven by endogenous U1-70K promoter (*pU1-70K::U1-70K-*
610 *GFPu1-70k*)[21]. RT-qPCR showed that U1-70K immunoprecipitation recovered spliced control
611 RNA, failed to recover unspliced control mRNA, and, importantly, retrieved a significant amount
612 of *MUSHER* lncRNA (Fig. 6c).
613

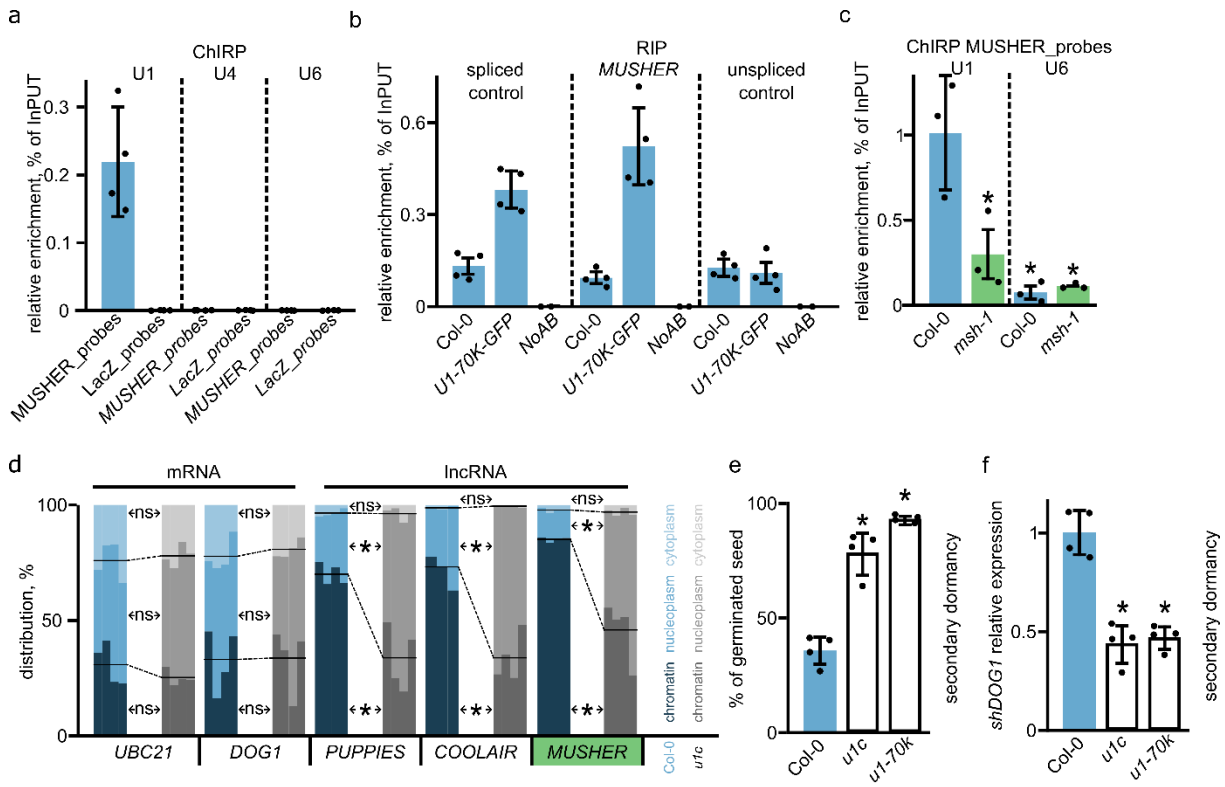


Fig. 6| U1 is required for chromatin tethering of *MUSHER* and other lncRNA.
a *MUSHER* or *LacZ* probes were used in formaldehyde-assisted RIP to enrich associated RNA. U1, U4 and U6 recovery was analysed using RT-qPCR. Error bars show standard deviations. **b** *MUSHER* probes were used to enrich *MUSHER* from WT and as a negative control in *msh-1*. U1 and U6 recovery was analysed using RT-qPCR. Error bars show standard deviations. **c** U1A protein was precipitated side by side with a no-antibodies (NoAB) control from plants expressing *pU1-70K::U1-70K-GFP u1-70k,* and co-precipitated RNA was analysed using RT-qPCR for spliced (AT5G45830) and unspliced (AT5G45820) RNA controls and *MUSHER*. Error bars show standard deviations. **d** Subcellular localisation of mRNAs and lncRNAs in maturing Arabidopsis seeds of WT and *u1c* amiRNA line. Lines represent the mean of chromatin and nucleoplasmic fractions contribution of the total amount of indicated transcript. The blue-shaded colour refers to Col-0, while the grey-shaded colour corresponds to *u1c*. Each transcript was measured in 4 replicates represented as a separate bar. **p*-value<0.01 from two-tailed Student *t*-test. **e** After ripened seeds of WT (WT), *u1c* amiRNA and *u1k-70k* mutants were induced into secondary dormancy, moved to permissive conditions and scored for germination ability. Germination (%) was scored as radical protrusion. A two-tailed t-test p-value was calculated compared to WT (**p*-value < 0.05). Error bars show standard deviations. **f** *shDOG1* expression was analysed in WT, *u1c* amiRNA and *u1-70k* mutants seeds induced into secondary dormancy. A two-tailed t-test p-value was calculated compared to WT (**p*-value < 0.05). Normalised to *UBC21* and relative to WT. Error bars show standard deviations.

Having shown that *MUSHER* is bound by U1 but not U6 or U4 snRNP, we asked if this contributes to *MUSHER* chromatin association. To do this, we used a recently described artificial microRNA (amiRNA) line that shows significantly downregulated *u1c* levels[21]. Subcellular localisation showed that *UBC21* and *DOG1* mRNA are nearly equally distributed between cytoplasmic, nucleoplasm and chromatin-bound fractions, as shown also in Fig. 1 (Fig. 6d). In addition, neither *UBC21* nor *DOG1* mRNA distribution among the cellular fractions was altered in the *u1c* mutant (Fig. 6d). Importantly, the *MUSHER's* association to chromatin was strongly reduced in *u1c* mutant, suggesting U1 tethers *MUSHER* lncRNA at the chromatin. Next, we asked whether the U1 requirement for RNA retention at chromatin is a more general phenomenon in plants.

We started by confirming that *COOLAIR*, shown to be expressed in seeds[58], is predominantly chromatin-bound, like *MUSHER* (Fig. 6d). In addition, unspliced *PUPPIES* isoform, a lncRNA involved in Pol II pausing regulation at *DOG1* in seeds[7], was also strongly enriched in chromatin fraction. Both *PUPPIES* and *COOLAIR* have multiple U1 binding sites (Supplementary Table 3) and we found that in contrast to the control mRNA, *MUSHER* and all other analysed lncRNAs showed reduced chromatin association and a corresponding increase in nucleoplasmic abundancy in the *u1c* mutant (Fig. 6d). We conclude that *MUSHER* binds U1 snRNP and that U1 facilitates *MUSHER* and other tested lncRNA-chromatin association and propose that the U1 role in chromatin retention of lncRNA is conserved between animals and plants.

Finally, we asked if *MUSHER's* chromatin association is required for its function. Both *u1c* and another amiRNA mutant in the U1 snRNP component U1-70K, show pleiotropic phenotypes, including significant dwarfism and delayed development[21]. Thus, it would be difficult to distinguish between primary dormancy phenotype and developmental defects resulting in delayed germination. To circumvent this, we took advantage of the fact *MUSHER* also shows defects during secondary dormancy induction. To synchronise the physiological state of seed, they were matured on mother plants and stored until they lost primary dormancy and then tested for the induced secondary dormancy. *u1c* and *u1-70k* mutant seeds showed low secondary dormancy phenotype with nearly all seeds germinating compared to less than 50% of germinated seeds in the WT (Fig. 6e, Supplementary Fig. 7). The reduction in *u1c* and *u1-70k* secondary dormancy phenotype was accompanied by decreased *shDOG1* expression (Fig. 6f). This agrees with a model where U1 is required for *MUSHER's* association to chromatin and its function in seed dormancy induction. In summary, U1 is required to tether multiple lncRNAs at the chromatin, and in case of *MUSHER,* we showed this chromatin association is required for the lncRNA function.

Discussion

Our results indicate that *MUSHER* modulates seed dormancy through the transcriptional regulation of DOG1 and ABA sensitivity pathways (Fig. 7a). We show that *MUSHER* lncRNA is chromatin-bound and requires U1 snRNP for its retention and activity at chromatin (Fig. 7b).

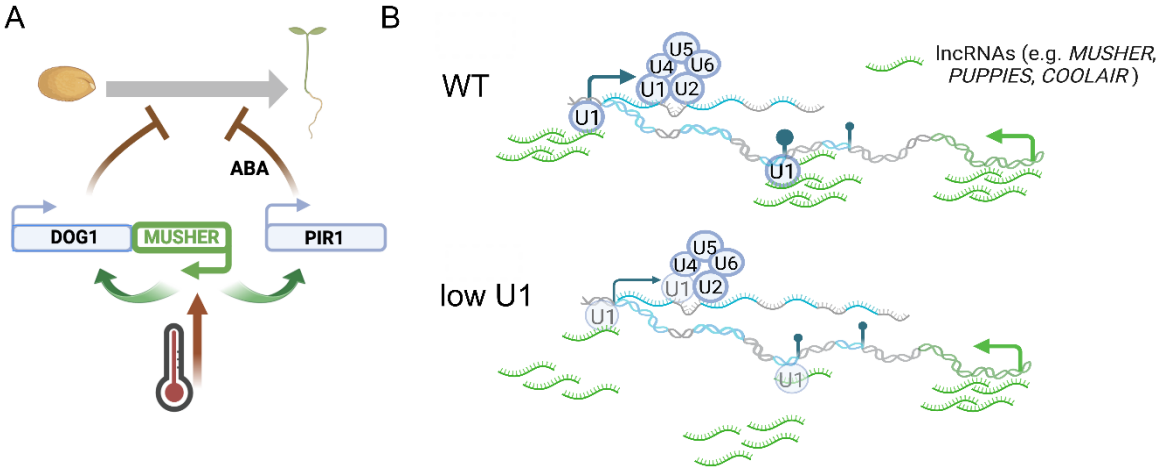


Fig. 7| a *MUSHER* modulates seed dormancy through the transcriptional regulation of DOG1 and ABA pathways. **b** U1 is required for chromatin retention of lncRNA, including *MUSHER* in Arabidopsis.

MUSHER-Mediated Regulation of *DOG1* and *PIR1* Expression

676 Our analysis showed that *MUSHER* employs two different mechanisms to activate *DOG1* and
677 *PIR1*. *MUSHER* promotes *DOG1* expression by enhancing CPSF complex recruitment to the
678 proximal cleavage and polyadenylation site, resulting in enhanced production of a functional
679 short *DOG1* alternative polyadenylation isoform (Fig. 2). In agreement, the *MUSHER* effect on
680 *DOG1* is not accompanied by changes in Pol II recruitment or H3K4me3 levels at the *DOG1* gene
681 (Fig. 2 and Supplementary Fig. 3). In contrast, *MUSHER* activates *PIR1* expression by enhancing
682 Pol II recruitment and consequent H3K4me3 deposition at the *PIR1* promoter (Fig. 4). Similar to
683 *HID1* lncRNA[13], and differently then *COOLAIR*[59] and *APOLO*[5], *MUSHER* promotes H3K4me3
684 deposition without affecting H3K27me3 repressive marks. Notably, *HID1* represses
685 transcription by blocking the binding of ARABIDOPSIS TRITHORAX-RELATED7 (ATXR7), an
686 H3K4me3 methyltransferase, to nucleosomes, leading to increased H3K4me3 levels when HID1
687 is inactivated[13]. In contrast, *MUSHER* inactivation reduces H3K4me3 deposition at the *PIR1*
688 promoter, suggesting that *MUSHER* and *HID1* regulate H3K4me3 through distinct molecular
689 mechanisms. Of note, we did not detect *PIR1* proximal polyadenylation either in our seed-
690 specific 3'RNA-seq data (Figure 4b) or in the published long read-based transcript mapping
691 dataset[39]. It is currently not clear why *MUSHER*-dependent *DOG1* and *PIR1* regulation are based
692 on different modes of activity. It is possible that *MUSHER* acts in the same way on *DOG1* and
693 *PIR1*, but the local chromatin environment of the target or *MUSHER* site of binding with respect
694 to the target transcription start site results in seemingly different mechanisms. For example, the
695 recruitment of CPSF complex to the gene promoter may enhance transcription initiation or
696 transition to efficient elongation. We, however, favour a model where *MUSHER* employs two
697 different mechanisms to enhance *DOG1* and *PIR1* expression. This view is supported by the fact
698 that *DOG1* expression is reduced roughly by 50%, while *PIR1* expression is undetectable in
699 multiple *msh* mutants tested.

700 Chromatin precipitation using *MUSHER* RNA (ChIRP) followed by sequencing revealed only a
701 very limited number of *MUSHER* targets, with *DOG1* and *PIR1* being the most significantly
702 enriched. The high selectivity of *MUSHER* is somehow surprising, as many of the published plant
703 lncRNAs show binding to multiple sites spread across the genome[5,60]. *PIR1* DNA chromatin-DNA
704 precipitation assay (ChIDP) followed by sequencing showed that *PIR1* DNA also has a very
705 limited number of DNA targets, with, excluding *PIR1* bait, *DOG1* being the strongest. This
706 suggests that high *MUSHER* selectivity could be guided by a spatial arrangement of DNA in
707 seeds.

708 ### *MUSHER* as a Link Between ABA Sensitivity and *DOG1*

709 Interestingly, both *MUSHER* targets encode proteins, PIR1 and DOG1, that interact with PP2CA
710 protein, a regulator of ABA signalling[35]. PIR1 binding to PP2CA promotes its degradation, while
711 DOG1 inhibits PP2CA phosphatase activity[35,37,61]. Both *dog1* and ABA synthesis deficient
712 *nced2/5/6/9* mutants fail to enter dormancy[62]. Simultaneously, *dog1* does not show major
713 changes in ABA levels or sensitivity[27,63] and *nced2/5/6/9* shows unaffected *DOG1* induction
714 kinetics during secondary dormancy establishment[53]. This suggests that DOG1 and ABA
715 synthesis pathways function independently in dormancy regulation. We demonstrate that
716 *MUSHER* inactivation leads to decreased transcription of *DOG1* and *PIR1*, which results in

717 decreased ABA sensitivity. We propose that *MUSHER* lncRNA facilitates the concurrent
718 activation of DOG1 and ABA pathways, integrating them into a shared regulatory network.

719 High temperatures are known to enhance *DOG1* expression and ABA biosynthesis in seeds[51,67].
720 We observed altered ABA sensitivity and reduced secondary dormancy induction in response
721 to high temperature in *msh* mutants. Elevated temperatures contribute to two well-described
722 responses: 1) thermoinhibition, a high-temperature-mediated inhibition of germination in
723 light[68] and 2) secondary dormancy induction in non-dormant imbibed seeds exposed to high
724 temperatures in the darkness[31,49,51]. We hypothesise that *MUSHER*-mediated integration of ABA
725 and DOG1 pathways contributes not only to dormancy but also to other ABA/DOG1-related
726 phenotypes (e.g., drought tolerance or thermoinhibition). We conclude that, similar to other
727 long non-coding RNAs involved in temperature sensing, such as *COOLAIR*[8], *COLDWRAP*[69] and
728 *SVALKA* [6] in plants, and *Hsrω*[70] and *Heat*[71] in animals, *MUSHER* also plays a role in temperature
729 sensing in plants.

730 U1 role in chromatin retention of lncRNA in Arabidopsis.

731 In contrast to mature mRNAs, lncRNAs are often enriched in the nucleus[1,15]. Consistently, many
732 lncRNA have been shown to regulate gene expression by binding to chromatin[1,5,6,39]. The
733 molecular mechanism of human lncRNA retention in the nucleus has only recently started to
734 be explored[15,72,73], and remains largely unknown in plants. Mechanisms ensuring human lncRNA
735 nuclear retention encompass the presence of specific sequences[74], including *Alu* transposable
736 elements binding the HNRNPK complex[75]. These protein-RNA interactions may contribute to
737 liquid-liquid phase separation (LLPS) that can also be responsible for the nuclear retention of
738 lncRNA. For example, LLPS by *NEAT1* lncRNA results in the formation of nuclear paraspeckles[76]
739 and *XIST* lncRNA was reported to undergo LLPS with different partners localised at inactive X
740 chromosome[77]. However, nuclear retention may not be enough to guide lncRNA directly to
741 chromatin. This has been proposed to be a result of direct RNA-DNA interaction, either in the
742 form of R-loops or RNA:DNA triplex[15]. In addition, U1 snRNP, involved in 5' splice site
743 recognition, has recently been shown to be responsible for the chromatin binding of a large
744 fraction of lncRNAs[15]. In contrast to *Alu* and other sequence elements identified in humans, U1
745 binding sites are well conserved in plant genomes.

746 Plant lncRNA retention at chromatin has not been explored, but RNA-seq analysis indicates that
747 lncRNAs are enriched in the nucleus[78]. Also, like their human counterparts, plant lncRNAs often
748 regulate gene expression by chromatin binding[6], indicating that their chromatin retention may
749 be actively maintained. Multiple plant lncRNAs have been described, but few of them have been
750 reported to be expressed in seeds and to control gene expression by chromatin-related
751 mechanisms. This includes *COOLAIR* regulating the *FLC* gene chromatin state[9,79], *PUPPIES*
752 regulating Pol II pausing at the *DOG1* gene[7] and, described here, *MUSHER* promoting the *DOG1*
753 gene proximal cleavage and polyadenylation site selection. In seedlings, *COOLAIR* is retained at
754 chromatin, presumably by its ability to form R-loops[80] and a cloud around its transcription site[10].
755 Similarly, *APOLO* lncRNA has been shown to form R-loops at its targets located in *trans* to its
756 transcription site[5]. Here, we propose that U1-mediated chromatin binding of lncRNAs is a
757 largely sequence-independent mechanism used by plants to retain lncRNA at chromatin. We
758 showed that U1 snRNP is required for *MUSHER*, *COOLAIR* and *PUPPIES* to bind the chromatin

759 efficiently. As we did not test the interplay between U1 binding and LLPS, we cannot exclude
760 the possibility that U1 binding is required or otherwise interlinked with the phase separation of
761 lncRNA at chromatin. R-loops have not been reported in seeds, and the one shown in seedlings
762 at *PIR1* locus[81] is unlikely to be formed by *MUSHER* since *MUSHER* is very lowly expressed in
763 seedlings. Therefore, while we cannot exclude this possibility, *MUSHER* is unlikely to form R-
764 loops. As concluded for humans, U1 lncRNA targets[15] do not form direct R-loops with the target
765 DNA. This suggests that in plants, R-loops probably do not contribute to U1-mediated lncRNA
766 retention. The fact that the lncRNA tested here have different modes of action, implies that the
767 U1 role in chromatin retention is not limited to a specific lncRNA activity. Altogether, our data
768 strongly supports the idea that the U1 role in chromatin retention of lncRNA is evolutionarily
769 conserved between animals and plants.

770 Methods

771 Plant material and growth conditions

772 WT used in this work is Columbia (Col-0) *Arabidopsis thaliana* accession. Plants were grown on
773 soil in a greenhouse under long-day conditions (16 h light/8 h dark) at 22°C/18°C. All mutants
774 and transgenic lines were created in Col-0 background and were homozygous for the indicated
775 genotype. *MUSHER* lines *msh-1*(SM_29297) and *msh-2*(SM_29310) were obtained from the
776 NASC centre. Other mutants used in this study have been described previously*: cpl1-7*[82], *cpl1-8*
777 (GK-165H09)[32], *cpl1-9* (GK-849H10)[32], *dog1-3* (SALK_000867)[27], *dog1-4* (SM_3_20886)[30], *fy-2*
778 (SAIL_ 657D4)[83], *esp1-2* (SALK_078793)[84], *pir1* (SALK_078237)[36], *amiR-u1-70k* and *amiR-u1-c*[21].

779 Germination tests

780 Germination assays were performed as described previously to estimate primary dormancy
781 levels[7,30]. Briefly, 100-200 freshly harvested seeds were sown on Petri plates with blue
782 germination paper (HoffmanMfg) and a layer of water-soaked fabric underneath to keep
783 moisture. The plates were sealed and placed in the Percival chamber under long-day conditions
784 (16 h light/8 h dark) at 22°C/18°C. Photographic documentation was performed once a day, and
785 seeds with root protrusion were counted as germinated.

786 To evaluate secondary dormancy induction levels, seeds were sown on Petri plates with blue
787 germination paper (HoffmanMfg) and a layer of water-soaked fabric underneath to keep
788 moisture. The plates were sealed and incubated in the darkness at varying temperatures (25°C,
789 28°C, 35°C) for 7 days. Subsequently, plates were transferred to the Percival chamber under
790 long-day conditions (16 h light/8 h dark) at 22°C/18°C.

791 An abscisic acid (ABA) sensitivity test (Fig. S5) was conducted using the genotypes Col-0, *dCas9-*
792 *control-1*, *msh-1*, *msh-2*, *msh_dCas9-1*, *msh_dCas9-2, pir1, dog1-4 and msh-1xpir1*. Seeds were
793 sown on blue paper soaked with 45ml water supplemented with 0.1 µM ABA and stratified for
794 24h before transferring to 22°C (long day conditions). Differences in germination rates across
795 genotypes were measured after 4 days and assessed using a Kruskal-Wallis test within each
796 treatment, followed by Dunn's post-hoc test for pairwise comparisons.
797

798 Cloning, Plant Transformation and CRISPR-dCas9 mutant generation

799 The generation of the dCas9 transgenic plants was performed according to Čermák et al.[42]. The
800 specific sgRNAs to Arabidopsis TAIR10 genome were designed using the CRISPR-P 2.0
801 (http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR) to select the canonical NGG PAM motif.
802 sgRNAs were cloned to the dCas9 vector by Golden Gate assembly (*dCas9_control-1*,
803 *dCas9_control-2*, *msh_dCas9-1* and *msh_dCas9-2*) according to the protocols 3B and 5[42].
804 Primers used for cloning are listed in Table S4. All constructs were verified by Sanger sequencing.
805 Each construct was transformed into Agrobacterium strain GV3101 and then transformed into
806 Col-0 or an appropriate mutant background using the floral dip method. T1 transformant plants
807 were identified by selecting the resistance to glufosinate. The seeds from the selected T1 plants
808 were used for phenotypic analysis. At least two independent transgenic lines were tested and
809 propagated. As a control for potential indirect effects from dCas9 expression, we also created
810 two dCas9 lines with guides recognising different intergenic regions localised 100 kb from *DOG1*
811 *locus*: *dCas9_control_1* and *dCas9_control_2*, which showed no influence on *MUSHER*
812 transcription.

813 Luciferase measurement

814 100 seeds placed on a microscope glass slide were squashed to release the embryo from the
815 seed coat. After adding D-Luciferin (Beetle luciferin potassium salt, Promega) and 20 minutes
816 of incubation in darkness, the light signal was measured using a NightSHADE camera (Berthold
817 Technologies GmbH & Co.KG) with a 10-minute exposure.

818 RNA extraction

819 RNA extraction protocol was adapted from Kowalczyk et al.[32] After grinding in liquid nitrogen,
820 samples were resuspended in 500 µl of homogenisation buffer (100 mM Tris pH 8.5, 5 mM
821 EDTA, 100 mM NaCl, 0.5% SDS, 1% β-mercaptoethanol) and centrifuged for 5 min at room
822 temperature at 8,000 g. The supernatant was transferred to the new tubes containing 250 µl of
823 chloroform. After 15 min of shaking, 250 µl of phenol was added, and shaking was repeated,
824 followed by centrifugation for 10 min at 10,000 g. The upper phase was transferred to new
825 tubes, and an equal volume of phenol:chloroform:isoamyl alcohol was added. Shaking and
826 centrifugation steps were repeated. The upper phase was mixed with 10% of the volume of 3M
827 sodium acetate (pH 5.2) and an equal volume of isopropanol. After centrifugation for 30 min at
828 20,000 g at 4°C, the pellet was washed with 70% EtOH and resuspended in water.

829 RT–qPCR

830 Trace DNA in the RNA samples was removed using a TURBO DNA-free kit (Thermo Fisher
831 Scientific) according to the manufacturer's protocol. The DNA removal was confirmed by PCR
832 using primers to the *PP2A* gene (AT1G13320). The RNA sample quality and quantity were
833 checked by agarose gel electrophoresis and using a Nanodrop 2000 spectrophotometer.
834 Reverse transcription using 2.5 µg of RNA was performed with a Superscript III cDNA Synthesis
835 kit (Thermo Fisher Scientific) according to the manufacturer's protocol. An equal mixture of
836 oligo(dT) and hexamer primers was used for cDNA synthesis according to the manufacturer's
837 protocol. 5x diluted cDNA was used as a template in the qPCR reaction (LightCycler 480; Roche)
838 with SYBR Green mix (Roche). The relative transcript level was determined with the 2−ΔΔCt

method [28] and normalised to mRNA of the *UBC21* (AT5G25760) gene. The primers used are listed in Table S4.

Chromatin ImmunoPrecipitation (ChIP)

ChIP was performed as described[32] with the following modifications. 100 mg of fresh siliques (13-20 days after pollination) were ground in liquid nitrogen and resuspended in 20 ml of crosslink buffer (10 mM HEPES-KOH pH 7.4, 50 mM NaCl, 100 mM sucrose, 1% formaldehyde). After 10 min incubation on a rotator at 4°C, the crosslinking agent was quenched by adding 1.25 ml of 2 M glycine. After another 5 min incubation on a rotator, samples were centrifuged at 4°C and 2,000 g for 15 min. The pellets were resuspended in 20 ml of cold Honda Buffer (20 mM HEPES-KOH pH 7.4, 0.44 M sucrose, 1.25% Ficoll, 2.5% Dextran T40, 10 mM $MgCl_2$, 5 mM DTT, 0.5% Triton X-100, 10 mM β-mercaptoethanol, 1 mM PMSF, 1x Complete protease inhibitors (Roche)) and filtered through Miracloth. Samples were centrifuged for 15 min at 2,000 g at 4°C. The pellet was resuspended in ChIP lysis buffer (50 mM HEPES-KOH pH 7.4, 10 mM EDTA, 1% SDS, 1x Complete protease inhibitors (Roche)) and sonicated using Bioruptor (Diagenode) to fragments of 300-500 bp. Samples were next diluted 10 times with ChIP Dilution Buffer (1.1% Triton X-100, 1 mM EDTA, 16.7 mM HEPES-KOH pH 7.4, 167 mM NaCl, 1 mM DTT, 1 mM PMSF, 1x Complete protease inhibitors (Roche)). Then antibodies against RNA polymerase II subunit B1 (Agrisera, AS11 1804), H3 (Abcam, ab1791), H3K4me3 (Millipore, 07-473), H3K27me3 (Abcam, ab6002) or CPSF73 (PhytoAB, PHY1043A) were added followed by addition of the mixture of Dynabeads Protein A and Protein G beads (Thermo Fisher Scientific). Next, samples were incubated overnight on a rotator at 4°C. Two washing steps with low salt buffer (150 mM NaCl, 1 mM EDTA, 10 mM HEPES-KOH pH 7.4, 0.1% Triton X-100) and two washing steps using high salt buffer (500m M NaCl, 1 mM EDTA, 10 mM HEPES-KOH pH 7.4, 0.1% Triton X-100) were performed. DNA was released from beads and reverse-crosslinked by incubation at 95°C with 10% Chelex 100 suspension for 15 min. After Proteinase K treatment (55°C) for 90 min and repeated incubation at 95°C for 10 min, samples were centrifuged for 5 min at RT at 10,000 g. Recovered DNA was used in qPCR reaction with primers listed in Table S4.

RNA ImmunoPrecipitation (RIP)

For U1-70K-GFP RIP, we used 100 mg of fresh siliques (13-20 days after pollination), which were ground in liquid nitrogen and resuspended in 20 ml of crosslink buffer (10 mM HEPES-KOH pH 7.4, 50 mM NaCl, 100 mM sucrose, 1% formaldehyde). After 10 min incubation on a rotator at 4°C, the crosslinking agent was quenched by adding 1.25 ml 2 M glycine. After another 5 min incubation on a rotator, samples were centrifuged at 4°C and 2,000 g for 15 min. The pellets were resuspended in 20 ml of cold Honda Buffer (20 mM HEPES-KOH pH 7.4, 0.44 M sucrose, 1.25% Ficoll, 2.5% Dextran T40, 10 mM $MgCl_2$, 5 mM DTT, 0.5% Triton X-100, 10 mM β-mercaptoethanol, 1mM PMSF, 1x Complete protease inhibitors (Roche), and 5 U/ml murine RNase inhibitor (NEB)) and filtered through Miracloth. Samples were centrifuged for 15 min at 2,000 g at 4°C. The pellet was resuspended in Nuclear Lysis Buffer (50 mM HEPES-KOH pH 7.4, 10 mM $MgCl_2$, 1% Triton X-100, 1x Complete protease inhibitors (Roche), 20 U/ml murine RNase inhibitor (NEB)), followed by 20 min incubation at 37°C with TURBO DNAse (Thermo

881 Fisher Scientific). After adding SDS up to 1%, samples were sonicated using Bioruptor
882 (Diagenode). Next, samples were diluted 10 times with RIP Dilution Buffer (0.5 mM EDTA, 10
883 mM HEPES-KOH pH 7.4, 150 mM NaCl, 20 U/ml murine RNase inhibitor (NEB)) and incubated
884 with GFP Agarose beads (Chromotek) for 1h on a rotator at 4°C. After three washing steps with
885 RIP Wash Buffer (150 mM NaCl, 1 mM EDTA, 10 mM HEPES-KOH pH 7.4, 0.05% NP-40, 20 U/ml
886 murine RNase inhibitor (NEB)), RNA was eluted (150 mM NaCl, 1 mM EDTA, 10 mM HEPES-KOH
887 pH 7.4, 1% SDS, 20 U/ml murine RNase inhibitor (NEB)) and reverse-crosslinked by incubation
888 at 95°C for 15 min. After 1h Proteinase K treatment (55°C), RNA was isolated with Trisure
889 (Bioline) according to producer protocol. After cDNA synthesis with SuperScript III (Thermo
890 Fisher Scientific), samples were used in qPCR reaction with appropriate primers listed in Table
891 S4.

892 Nuclei purification for ChIRP, RIRP and ChIDP

893 To uncover RNA-RNA (RNA Isolation by RNA Purification), RNA-DNA (Chromatin Isolation by
894 RNA Purification) and DNA-DNA (Chromatin Isolation by DNA Purification Assay) interactions,
895 we used 400 mg of fresh siliques (13-20 days after pollination), which were ground in liquid
896 nitrogen and resuspended in 20 ml of crosslink buffer (10 mM HEPES-KOH pH 7.4, 50 mM NaCl,
897 100 mM sucrose, 1% formaldehyde (for RNA-RNA and RNA-DNA) or 0.5% glutaraldehyde (for
898 ChIDP)). After 10 min incubation on a rotator at 4°C, the crosslinking reaction was quenched by
899 adding 1.25 ml 2 M glycine. After an additional 5 min incubation on a rotator, samples were
900 centrifuged at 4°C and 2,000 g for 15 min. The pellets were resuspended in 20 ml of cold Honda
901 Buffer (20 mM HEPES-KOH pH 7.4, 0.44M sucrose, 1.25% Ficoll, 2.5% Dextran T40, 10 mM
902 $MgCl_2$, 5 mM DTT, 0.5% Triton X-100, 10 mM β-mercaptoethanol, 1 mM PMSF, 1x Complete
903 protease inhibitors (Roche) and 5 U/ml murine RNase inhibitor (NEB)) and filtered through
904 Miracloth. Subsequently, samples were centrifuged for 15 min at 2,000 g at 4°C. According to
905 downstream assays, the nuclei pellets were resuspended in appropriate buffers (Element Zero
906 Biolabs kit).

907 Chromatin Isolation by RNA Purification Assay (ChIRP)

908 To enrich specific RNA molecules or chromatin fragments, we used MagIC beads with custom
909 probes targeting those sequences (Element Zero Biolabs). Element Zero Biololabs designed
910 probes are delivered with kits and remain confidential as protected intellectual property.

911 For *MUSHER* probes RIRP and ChIRP with *MUSHER* and *lacZ* probes, nuclei pellets were
912 resuspended in 1.1 ml of MagIC Lysis Buffer FA (Element Zero Biolabs kit), supplemented with
913 1x Complete protease inhibitors (Roche) and 5 mM DTT. Next, samples were sonicated using a
914 Bioruptor (Diagenode) and diluted two times with MagIC Lysate Dilution Buffer FA, with 1% of
915 the volume saved as an input sample. MagIC beads were pre-cleared according to the
916 producers' protocol and added to the samples that were next incubated at 40°C for 30 minutes
917 with occasional shaking. This was followed by three washes using MagIC Wash Buffer I FA and
918 nucleic acids elution by incubation at 92°C for 5 min with 10 mM Tris-HCl pH 7.5. After
919 Proteinase K treatment (55°C, 90 min) and phenol:chloroform:isoamyl alcohol extraction, DNA
920 and RNA were precipitated by adding 1 µl glycogen, 10% of the volume of 3 M LiCl, and an equal
921 volume of isopropanol. Samples were centrifuged for 30 min at 4°C at 20,000 g. Pellets were
922 washed with cold 80% ethanol, resuspended in water, and samples were split in half. The

923 presence of specific DNA fragments interacting directly or indirectly with *MUSHER* probes was
924 validated by qPCR analysis. To identify RNAs enriched with *MUSHER* probes, half of the sample
925 was incubated with TURBO DNAse (Thermo Fisher Scientific), extracted with
926 phenol:chloroform: isoamyl alcohol as above, and used for cDNA synthesis and qPCR analysis.
927 For ChIRP library preparation, DNA fragments were eluted from MagIC beads with RNase If
928 (NEB). Libraries were prepared with the NEBNext® Ultra™ DNA Library Prep Kit for Illumina using
929 NEBNext Multiplex Dual Index Primers. Paired-end sequencing was performed on the Illumina
930 NovaSeq 6000.

931 Chromatin Isolation by DNA Purification Assay (ChIDP)

932 For ChIDP, the nuclei pellet was resuspended in 1.1 ml of MagIC Lysis Buffer FA (Element Zero
933 kit) supplemented with 1x Complete protease inhibitors (Roche) and 5 mM DTT. Samples were
934 sonicated using Bioruptor (Diagenode), diluted with MagIC Lysate Dilution Buffer FA, and split
935 into halves (treated with RNaseA at 37°C for 20 min and non-treated). MagIC beads recognising
936 the *PIR1 locus* were mixed with the lysates that were next incubated at 81°C for 5 min with
937 occasional shaking to denature DNA partially. Subsequently, samples were transferred to 40°C
938 for 20 min to enable binding of the probes. This was followed by six washing steps with the
939 MagIC Lysate Wash buffer I and II, and DNA was eluted with 10 mM Tris-HCl pH 7.5 at 92°C for
940 5 min. After adding an equal volume of MagIC Lysate Wash buffer I to eluted samples,
941 proteinase K treatment (55°C, 120 min) was performed. To extract DNA fragments, an equal
942 volume of phenol:chloroform:isoamyl alcohol was added. The upper phase was mixed with 1 µl
943 of glycogen, 10% of the volume of 3 M LiCl and an equal amount of isopropanol. After
944 centrifugation for 30 min at 20000g at 4°C, the pellet was washed with 80% EtOH and
945 resuspended in water. Libraries were prepared with the NEBNext® Ultra™ DNA Library Prep Kit
946 for Illumina using NEBNext Multiplex Dual Index Primers. Paired-end sequencing was performed
947 on the Illumina NovaSeq 6000.

948 ChIRP-seq and ChIDP-seq analysis

949 For ChIRP-seq and ChIDP-seq data analysis, Bowtie2 (v2.5.3)[85] was used to align the reads to
950 the reference genome (TAIR10, with default parameters). To identify significantly enriched
951 regions, MACS2 was used with the default parameters: bandwidth:300; model fold of 5:50; q-
952 value cutoff: 0.05 (v3.0.0a6)[86]. bamCoverage (v3.5.4)[87] with CPM normalisation was used to
953 create bigwig files.

954 5'RACE-seq

955 Analysis was performed as described previously[7]. Shortly, 500 ng of total RNA, after DNAse
956 treatment, was used as the template for cDNA generation using SuperScript II (Thermo Fisher
957 Scientific). The resulting cDNA was purified using AMPure XP magnetic beads (Beckman
958 Coulter) and amplified in a series of three PCR reactions with Phusion polymerase and specific
959 primers (1st PCR: only TSO_n1, 2nd PCR: TSO_n2 and MUSHER_5RACE, 3rd PCR: Illumina indexing
960 primers). The final PCR product was sequenced using Illumina MiSeq. For data analysis, STAR
961 (v2.7.8a)[88] was used to align the reads to the reference genome, followed by UMI-based PCR
962 duplicates removal using UMI-tools (v1.1.0)[89]. The position of the first 5′ nucleotide was
963 extracted using bedtools (v2.30.0)[90].

964 3'RACE-seq

965 Analysis was performed as described previously[7] and based on the earlier procedure[91]. The
966 ligation of the pre-adenylated adaptor to the 3'end of the RNA was performed using truncated
967 T4 RNA Ligase 2 (NEB). RNA ligated with RA3_15N adaptor (containing UMI) was cleaned with
968 AMPure XP magnetic beads and subjected to reverse transcription reaction with SuperScript III
969 (Thermo Fisher Scientific). After three rounds of PCR with specific primers (1st PCR:
970 MUSHER_3RACE and RTPXT, 2nd PCR: mXTf and mXTr,3rd PCR: Illumina indexing primers), each
971 followed with AMPure beads purification, prepared libraries were sequenced using Illumina
972 MiSeq. For data analysis, STAR (v2.7.8a)[88] was used to align the reads to the reference genome,
973 followed by UMI-based PCR duplicates removal using UMI-tools (v1.1.0)[89]. The position of the
974 last 3' nucleotide was extracted using bedtools (v2.30.0).

975 3'RNA-seq

976 Analysis was performed as described previously[49]. Briefly, 0.5 µg of total RNA was reverse
977 transcribed with UMI-containing oligo(dT) primers. Differently barcoded cDNAs were pooled
978 together, and second-strand cDNA was synthesised using the nick translation method. Next,
979 libraries were tagmented with Tn5 transposase, and the 3'cDNA fragments were amplified with
980 Illumina-compatible primers[49]. Paired-end sequencing was performed on the Illumina NextSeq
981 500. Alignment and read demultiplexing were performed using STAR (v2.7.8a)[88] and
982 BRBseqTools (v 1.6)[92], respectively. Differentially expressed genes were identified using DESeq2
983 with a cutoff of FDR < 0.05 and FC > 1.

984 Fractionation

985 Chromatin RNA extraction was performed using the described method[7]. 100 mg of siliques
986 were ground in liquid nitrogen and resuspended in 20 ml of cold Honda Buffer (20 mM HEPES-
987 KOH pH 7.4, 0.44 M sucrose, 1.25% Ficoll, 2.5% Dextran T40, 10 mM $MgCl_2$, 5 mM DTT, 0.5%
988 Triton X-100, 10 mM β-mercaptoethanol, 1 mM PMSF, 1x Complete protease inhibitors
989 (Roche), and 5 U/ml murine RNase inhibitor (NEB)). After rotation at 4°C for 10 min and
990 filtration through Miracloth, samples were centrifuged at 2,000 g for 15 min at 4°C. The nuclei
991 pellet was washed twice with 5 ml of Honda buffer. The pellet was resuspended in 600 µl of
992 Honda buffer and purified on a 40%-75% PercoII density gradient by centrifugation at 10,000 g
993 for 30 min at 4°C. Purified nuclei were collected from the Percoll interface, washed with Honda
994 buffer and resuspended in 500 µl of cold glycerol buffer (20 mM Tris-HCl pH 8.0, 75 mM NaCl,
995 0.5 mM EDTA, 0.85 mM DTT, 50% glycerol, 1% Empigen, 10 mM β-mercaptoethanol, 0.125 mM
996 PMSF, 1x Complete protease inhibitors (Roche), and 5 U/ml murine RNase inhibitor (NEB)).
997 Samples were overlaid on urea lysis buffer (10 mM HEPES-KOH pH 7.4, 7.5 mM $MgCl_2$, 0.2 mM
998 EDTA,300 mM NaCl, 1 M urea, 1% NP-40, 10 mM β-mercaptoethanol, 0.5 mM PMSF, 1x
999 Complete protease inhibitors (Roche), and 5U/ml murine RNase inhibitor (NEB), vortexed for 2
1000 s, and incubated on ice for 30 min, followed by centrifugation at 20,000g for 2 min at 4°C. The
1001 chromatin pellet was washed twice with 600 µl of urea lysis buffer for 30 min on a rotator at
1002 4°C. Chromatin pellet was used for RNA extraction and next for DNAse treatment as described
1003 above. The chromatin-attached RNA quality was tested by agarose gel electrophoresis and
1004 quantified with a Nanodrop 2000 spectrophotometer. cDNA synthesis was performed using a
1005 SuperScript III kit (Invitrogen).

smFISH

smFISH assay was performed as described previously[7] using a manually dissected radicle from dry embryos. The Stellaris probes used for smFISH target the full *shDOG1* sequence, including intron 1 and are labelled with Quasar670 fluorophore (Biosearch Technologies). The embryos were fixed, permeabilised and hybridised with the probes as described[7]. The smFISH signals were gathered using a widefield fluorescence microscope Olympus IX81 (Olympus) and a Hamamatsu Orca-R2 (C10600) CCD camera. The xCellence software (Olympus) was used for image acquisition. Cells were manually segmented using Napari[93] and the PartSeq plugin, which allows foci identification and classification based on subcellular location[94]. Dots number and intensity were analysed using R scripts described earlier[7].

Author contributions

S.P.S., and S.S. designed the experiments. S.P.S. carried out most of the experiments, M.M. and V.H.M. performed smFISH, and L.B. performed ChIDP-seq. M.K. conducted RNA-seq, K.R. performed 3'RACE-seq and obtained *msh-1 x esp1* line, M.W. and R.Y. obtained CRISPR *DOG1* mutant, Ł.S. and A.J. developed *pU1-70K::U1-70K-GFP u1-70k* lines. S.P.S and S.S. wrote the manuscript.

Declaration of interests

K.R. and S.S. have equity interests in ElementZero Biolabs UG. S.S. also serve on the Scientific Advisory Board for ElementZero Biolabs UG. All other authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process

While preparing this work, the authors used Grammarly to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

Data availability

The 3'RNA-seq data generated for this study have been deposited at the Gene Expression Omnibus (GEO) under the accession code GSE274757.

Declaration of interests

K.R. and S.S. have equity interests in ElementZero Biolabs UG. S.S. also serve on the Scientific Advisory Board for ElementZero Biolabs UG. All other authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process

While preparing this work, the authors used Grammarly to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

1. Statello, L., Guo, C.-J., Chen, L.-L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol* **22**, 96–118 (2021).

1043 2. Dror, I. *et al.* XIST directly regulates X-linked and autosomal genes in naive human pluripotent
1044 cells. *Cell* **187**, 110-129.e31 (2024).
1045 3. Li, J. *et al.* Long noncoding RNA XIST: Mechanisms for X chromosome inactivation, roles in sex-
1046 biased diseases, and therapeutic opportunities. *Genes Dis* **9**, 1478–1492 (2022).
1047 4. Arun, G., Aggarwal, D. & Spector, D. L. MALAT1 Long Non-Coding RNA: Functional Implications.
1048 *Noncoding RNA* **6**, 22 (2020).
1049 5. Ariel, F. *et al.* R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Molecular Cell*
1050 **77**, 1055-1065.e4 (2020).
1051 6. Kindgren, P., Ard, R., Ivanov, M. & Marquardt, S. Transcriptional read-through of the long non-
1052 coding RNA SVALKA governs plant cold acclimation. *Nat Commun* **9**, 4561 (2018).
1053 7. Montez, M. *et al.* Promoter-pervasive transcription causes RNA polymerase II pausing to boost
1054 DOG1 expression in response to salt. *EMBO J* **42**, e112443 (2023).
1055 8. Zhao, Y. *et al.* Natural temperature fluctuations promote COOLAIR regulation of FLC. *Genes Dev*
1056 **35**, 888–898 (2021).
1057 9. Swiezewski, S., Liu, F., Magusin, A. & Dean, C. Cold-induced silencing by long antisense transcripts
1058 of an Arabidopsis Polycomb target. *Nature* **462**, 799–802 (2009).
1059 10. Rosa, S., Duncan, S. & Dean, C. Mutually exclusive sense–antisense transcription at FLC facilitates
1060 environmentally induced gene repression. *Nat Commun* **7**, 13031 (2016).
1061 11. Yang, J. *et al.* Long noncoding RNA ARTA controls ABA response through MYB7 nuclear trafficking
1062 in Arabidopsis. *Developmental Cell* **58**, 1206-1217.e4 (2023).
1063 12. Yatusevich, R. *et al.* Antisense transcription represses Arabidopsis seed dormancy QTL DOG1 to
1064 regulate drought tolerance. *EMBO Rep.* **18**, 2186–2196 (2017).
1065 13. Wang, Y. *et al.* The noncoding RNA *HIDDEN TREASURE 1* promotes phytochrome B-dependent
1066 seed germination by repressing abscisic acid biosynthesis. *The Plant Cell* **35**, 700–716 (2023).
1067 14. Yang, M. *et al.* In vivo single-molecule analysis reveals COOLAIR RNA structural diversity. *Nature*
1068 **609**, 394–399 (2022).

1069 15. Yin, Y. *et al.* U1 snRNP regulates chromatin retention of noncoding RNAs. *Nature* **580**, 147–150
1070 (2020).
1071 16. Kondo, Y., Oubridge, C., van Roon, A.-M. M. & Nagai, K. Crystal structure of human U1 snRNP, a
1072 small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. *Elife*
1073 **4**, e04986 (2015).
1074 17. Goguel, V., Liao, X. L., Rymond, B. C. & Rosbash, M. U1 snRNP can influence 3'-splice site
1075 selection as well as 5'-splice site selection. *Genes Dev* **5**, 1430–1438 (1991).
1076 18. Almada, A. E., Wu, X., Kriz, A. J., Burge, C. B. & Sharp, P. A. Promoter directionality is controlled by
1077 U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363 (2013).
1078 19. Mimoso, C. A. & Adelman, K. U1 snRNP increases RNA Pol II elongation rate to enable synthesis
1079 of long genes. *Mol Cell* **83**, 1264-1279.e10 (2023).
1080 20. So, B. R. *et al.* A Complex of U1 snRNP with Cleavage and Polyadenylation Factors Controls
1081 Telescripting, Regulating mRNA Transcription in Human Cells. *Mol Cell* **76**, 590-599.e4 (2019).
1082 21. Mangilet, A. F. *et al.* The Arabidopsis U1 snRNP regulates mRNA 3'-end processing. *Nat. Plants* **10**,
1083 1514–1531 (2024).
1084 22. Carrillo-Barral, N., Rodríguez-Gacio, M. D. C. & Matilla, A. J. Delay of Germination-1 (DOG1): A
1085 Key to Understanding Seed Dormancy. *Plants (Basel)* **9**, 480 (2020).
1086 23. Auge, G. A. *et al.* Secondary dormancy dynamics depends on primary dormancy status in
1087 *Arabidopsis thaliana*. *Seed Sci. Res.* **25**, 230–246 (2015).
1088 24. Graeber, K. *et al. DELAY OF GERMINATION 1* mediates a conserved coat-dormancy mechanism for
1089 the temperature- and gibberellin-dependent control of seed germination. *Proc. Natl. Acad. Sci.*
1090 *U.S.A.* **111**, (2014).
1091 25. Footitt, S., Douterelo-Soler, I., Clay, H. & Finch-Savage, W. E. Dormancy cycling in *Arabidopsis*
1092 seeds is controlled by seasonally distinct hormone-signaling pathways. *Proc. Natl. Acad. Sci.*
1093 *U.S.A.* **108**, 20236–20241 (2011).

1094 26. Alonso-Blanco, C., Bentsink, L., Hanhart, C. J., Vries, H. B. & Koornneef, M. Analysis of Natural
1095 Allelic Variation at Seed Dormancy Loci of *Arabidopsis thaliana*. *Genetics* **164**, 711–729 (2003).
1096 27. Bentsink, L., Jowett, J., Hanhart, C. J. & Koornneef, M. Cloning of DOG1, a quantitative trait locus
1097 controlling seed dormancy in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A* **103**, 17042–17047 (2006).
1098 28. Fedak, H. *et al.* Control of seed dormancy in Arabidopsis by a cis-acting noncoding antisense
1099 transcript. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7846–E7855 (2016).
1100 29. Nakabayashi, K., Bartsch, M., Ding, J. & Soppe, W. J. J. Seed Dormancy in Arabidopsis Requires
1101 Self-Binding Ability of DOG1 Protein and the Presence of Multiple Isoforms Generated by
1102 Alternative Splicing. *PLoS Genet* **11**, e1005737 (2015).
1103 30. Cyrek, M. *et al.* Seed Dormancy in Arabidopsis Is Controlled by Alternative Polyadenylation of
1104 *DOG1*. *Plant Physiol.* **170**, 947–955 (2016).
1105 31. Krzyszton, M. *et al.* Dormancy heterogeneity among Arabidopsis thaliana seeds is linked to
1106 individual seed size. *Plant Communications* 100732 (2023) doi:10.1016/j.xplc.2023.100732.
1107 32. Kowalczyk, J. *et al.* Alternative Polyadenylation of the Sense Transcript Controls Antisense
1108 Transcription of DELAY OF GERMINATION 1 in Arabidopsis. *Molecular Plant* **10**, 1349–1352
1109 (2017).
1110 33. Kendall, S. L. *et al.* Induction of Dormancy in *Arabidopsis* Summer Annuals Requires Parallel
1111 Regulation of *DOG1* and Hormone Metabolism by Low Temperature and CBF Transcription
1112 Factors. *Plant Cell* **23**, 2568–2580 (2011).
1113 34. Rodríguez-Gacio, M. D. C., Matilla-Vázquez, M. A. & Matilla, A. J. Seed dormancy and ABA
1114 signaling: The breakthrough goes on. *Plant Signaling & Behavior* **4**, 1035–1048 (2009).
1115 35. Park, S.-Y. *et al.* Abscisic Acid Inhibits Type 2C Protein Phosphatases via the PYR/PYL Family of
1116 START Proteins. *Science* **324**, 1068–1071 (2009).
1117 36. Baek, W., Lim, C. W., Luan, S. & Lee, S. C. The RING finger E3 ligases PIR1 and PIR2 mediate PP2CA
1118 degradation to enhance abscisic acid response in Arabidopsis. *The Plant Journal* **100**, 473–486
1119 (2019).

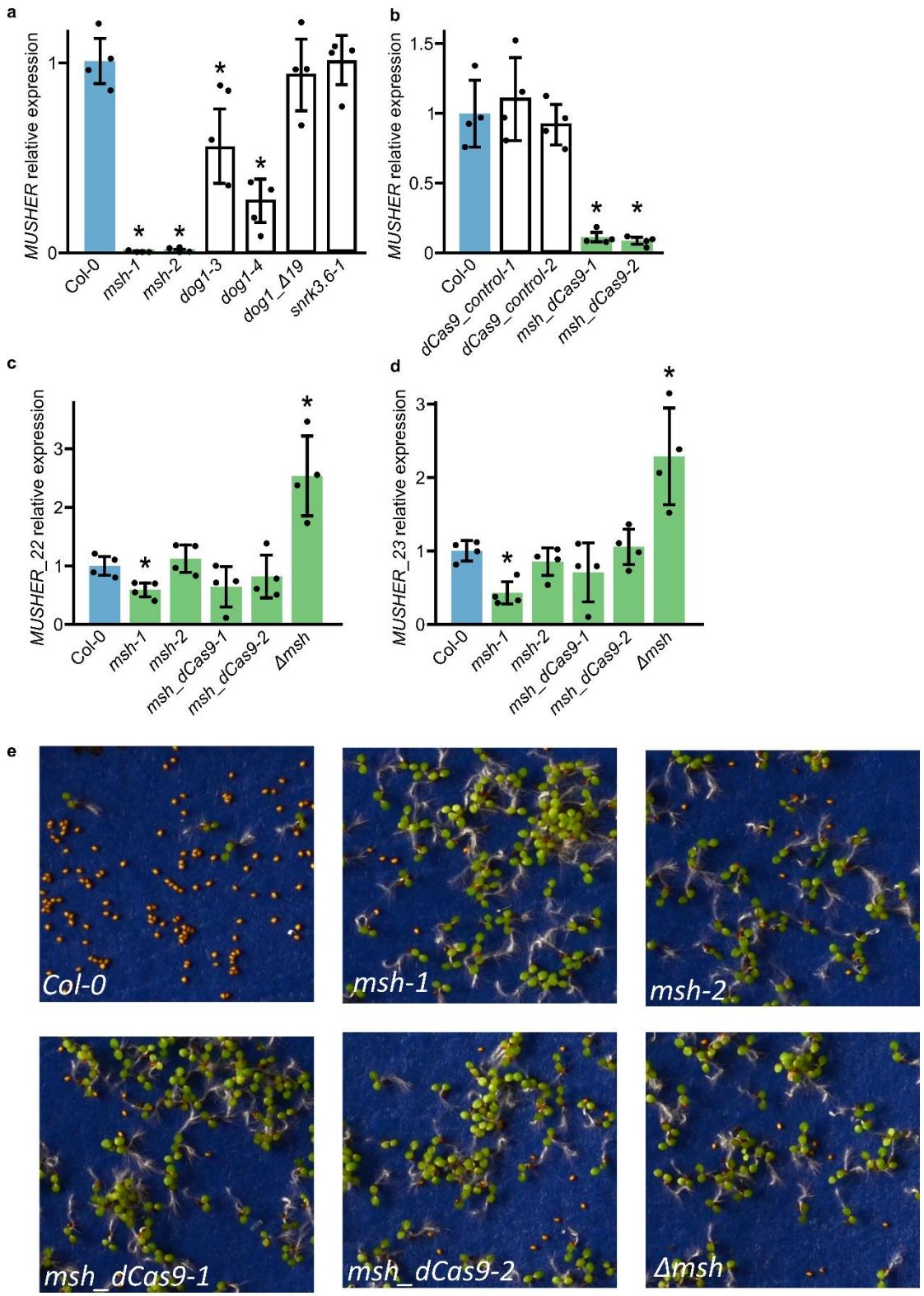
1120 37. Née, G. *et al.* DELAY OF GERMINATION1 requires PP2C phosphatases of the ABA signalling
1121 pathway to control seed dormancy. *Nat Commun* **8**, 72 (2017).
1122 38. Nishimura, N. *et al.* Control of seed dormancy and germination by DOG1-AHG1 PP2C
1123 phosphatase complex via binding to heme. *Nat Commun* **9**, 2132 (2018).
1124 39. Zhang, S. *et al.* New insights into Arabidopsis transcriptome complexity revealed by direct
1125 sequencing of native RNAs. *Nucleic Acids Research* **48**, 7700–7711 (2020).
1126 40. Thomas, Q. A. *et al.* Transcript isoform sequencing reveals widespread promoter-proximal
1127 transcriptional termination in Arabidopsis. *Nat Commun* **11**, 2589 (2020).
1128 41. Thieffry, A. *et al.* Characterization of Arabidopsis thaliana Promoter Bidirectionality and Antisense
1129 RNAs by Inactivation of Nuclear RNA Decay Pathways. *Plant Cell* **32**, 1845–1867 (2020).
1130 42. Čermák, T. *et al.* A Multipurpose Toolkit to Enable Advanced Genome Engineering in Plants. *Plant*
1131 *Cell* **29**, 1196–1217 (2017).
1132 43. Jupe, F. *et al.* The complex architecture and epigenomic impact of plant T-DNA insertions. *PLoS*
1133 *Genet* **15**, e1007819 (2019).
1134 44. Xu, R. *et al.* The 73 kD Subunit of the cleavage and polyadenylation specificity factor (CPSF)
1135 complex affects reproductive development in Arabidopsis. *Plant Mol Biol* **61**, 799–815 (2006).
1136 45. Funck, D., Clauß, K., Frommer, W. B. & Hellmann, H. A. The Arabidopsis CstF64-like RSR1/ESP1
1137 protein participates in glucose signaling and flowering time control. *Front Plant Sci* **3**, (2012).
1138 46. Cortazar, M. A. *et al.* Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation
1139 Facilitates Termination by a “Sitting Duck Torpedo” Mechanism. *Molecular Cell* **76**, 896-908.e4
1140 (2019).
1141 47. Laugesen, A., Højfeldt, J. W. & Helin, K. Molecular Mechanisms Directing PRC2 Recruitment and
1142 H3K27 Methylation. *Mol Cell* **74**, 8–18 (2019).
1143 48. Blackledge, N. P. & Klose, R. J. The molecular principles of gene regulation by Polycomb
1144 repressive complexes. *Nat Rev Mol Cell Biol* **22**, 815–833 (2021).

1145 49. Krzyszton, M. *et al.* Single seeds exhibit transcriptional heterogeneity during secondary dormancy
1146 induction. *Plant Physiology* **190**, 211–225 (2022).
1147 50. Buijs, G. A Perspective on Secondary Seed Dormancy in Arabidopsis thaliana. *Plants* **9**, 749
1148 (2020).
1149 51. Ibarra, S. E. *et al.* Molecular mechanisms underlying the entrance in secondary dormancy of
1150 *Arabidopsis* seeds: Mechanisms underlying secondary dormancy. *Plant Cell Environ.* **39**, 213–221
1151 (2016).
1152 52. Footitt, S., Huang, Z., Clay, H. A., Mead, A. & Finch-Savage, W. E. Temperature, light and nitrate
1153 sensing coordinate A rabidopsis seed dormancy cycling, resulting in winter and summer annual
1154 phenotypes. *The Plant Journal* **74**, 1003–1015 (2013).
1155 53. Wrona, M. *et al.* The BRAHMA-associated SWI/SNF chromatin remodeling complex controls
1156 *Arabidopsis* seed quality and physiology. *Plant Physiology* **197**, kiae642 (2024).
1157 54. Wang, H. *et al.* H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. *Nature*
1158 **615**, 339–348 (2023).
1159 55. Chu, C., Quinn, J. & Chang, H. Y. Chromatin isolation by RNA purification (ChIRP). *J Vis Exp* 3912
1160 (2012) doi:10.3791/3912.
1161 56. Déjardin, J. & Kingston, R. E. Purification of Proteins Associated with Specific Genomic Loci. *Cell*
1162 **136**, 175–186 (2009).
1163 57. Sajeev, N., Koornneef, M. & Bentsink, L. A commitment for *life:* Decades of unraveling the
1164 molecular mechanisms behind seed dormancy and germination. *The Plant Cell* koad328 (2024)
1165 doi:10.1093/plcell/koad328.
1166 58. Chen, M. & Penfield, S. Feedback regulation of COOLAIR expression controls seed dormancy and
1167 flowering time. *Science* **360**, 1014–1017 (2018).
1168 59. Nielsen, M. *et al. COOLAIR* and PRC2 function in parallel to silence *FLC* during
1169 vernalization. *Proc. Natl. Acad. Sci. U.S.A.* **121**, (2024).

1170 60. Jin, Y., Ivanov, M., Dittrich, A. N., Nelson, A. D. & Marquardt, S. LNCRNA FLAIL affects alternative
1171 splicing and represses flowering in *Arabidopsis*. *The EMBO Journal* **42**, e110921 (2023).
1172 61. Née, G. *et al.* DELAY OF GERMINATION1 requires PP2C phosphatases of the ABA signalling
1173 pathway to control seed dormancy. *Nat Commun* **8**, 72 (2017).
1174 62. Krzyszton, M. *et al.* Single seeds exhibit transcriptional heterogeneity during secondary dormancy
1175 induction. *Plant Physiol.* **190**, 211–225 (2022).
1176 63. Nakabayashi, K. *et al.* The Time Required for Dormancy Release in *Arabidopsis* Is Determined by
1177 DELAY OF GERMINATION1 Protein Levels in Freshly Harvested Seeds. *Plant Cell* **24**, 2826–2838
1178 (2012).
1179 64. Thatcher, L. F. *et al.* The Arabidopsis RNA Polymerase II Carboxyl Terminal Domain (CTD)
1180 Phosphatase-Like1 (CPL1) is a biotic stress susceptibility gene. *Sci Rep* **8**, 13454 (2018).
1181 65. Guan, Q., Yue, X., Zeng, H. & Zhu, J. The Protein Phosphatase RCF2 and Its Interacting Partner
1182 NAC019 Are Critical for Heat Stress–Responsive Gene Regulation and Thermotolerance in
1183 *Arabidopsis*. *Plant Cell* **26**, 438–453 (2014).
1184 66. Qi, Y. *et al.* The over-expression of a chrysanthemum gene encoding an RNA polymerase II CTD
1185 phosphatase-like 1 enzyme enhances tolerance to heat stress. *Hortic Res* **5**, 37 (2018).
1186 67. Krzyszton *et al.* Single seeds exhibit transcriptional heterogeneity during secondary dormancy
1187 induction. *Accepted in Plant Physiology* (2022).
1188 68. Ali, F., Qanmber, G., Li, F. & Wang, Z. Updated role of ABA in seed maturation, dormancy, and
1189 germination. *J Adv Res.* **35**, 199–214 (2022).
1190 69. Kim, D.-H. & Sung, S. Vernalization-Triggered Intragenic Chromatin Loop Formation by Long
1191 Noncoding RNAs. *Developmental Cell* **40**, 302-312.e4 (2017).
1192 70. Guo, Y. *et al.* Long noncoding RNAs heat shock RNA omega nucleates TBPH and promotes
1193 intestinal stem cell differentiation upon heat shock. *iScience* **27**, 109732 (2024).

1194 71. Ji, Q., Zong, X., Mao, Y. & Qian, S.-B. A heat shock–responsive lncRNA *Heat* acts as a HSF1-
1195 directed transcriptional brake via m^6 A modification. *Proc. Natl. Acad. Sci. U.S.A.* **118**,
1196 e2102175118 (2021).
1197 72. Guo, C.-J. *et al.* Distinct Processing of lncRNAs Contributes to Non-conserved Functions in Stem
1198 Cells. *Cell* **181**, 621-636.e22 (2020).
1199 73. Bridges, M. C., Daulagala, A. C. & Kourtidis, A. LNCcation: lncRNA localization and function. *J Cell*
1200 *Biol* **220**, e202009045 (2021).
1201 74. Shukla, C. J. *et al.* High-throughput identification of RNA nuclear enrichment sequences. *EMBO J*
1202 **37**, e98452 (2018).
1203 75. Lubelsky, Y. & Ulitsky, I. Sequences enriched in Alu repeats drive nuclear localization of long RNAs
1204 in human cells. *Nature* **555**, 107–111 (2018).
1205 76. Grosch, M. *et al.* Nucleus size and DNA accessibility are linked to the regulation of paraspeckle
1206 formation in cellular differentiation. *BMC Biol* **18**, 42 (2020).
1207 77. Strehle, M. & Guttman, M. Xist drives spatial compartmentalization of DNA and protein to
1208 orchestrate initiation and maintenance of X inactivation. *Current Opinion in Cell Biology* **64**, 139–
1209 147 (2020).
1210 78. Zhao, X. *et al.* Global identification of Arabidopsis lncRNAs reveals the regulation of MAF4 by a
1211 natural antisense RNA. *Nat Commun* **9**, 5056 (2018).
1212 79. Csorba, T., Questa, J. I., Sun, Q. & Dean, C. Antisense COOLAIR mediates the coordinated
1213 switching of chromatin states at FLC during vernalization. *Proc. Natl. Acad. Sci. U.S.A.* **111**,
1214 16160–16165 (2014).
1215 80. Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J. & Dean, C. R-loop stabilization represses
1216 antisense transcription at the Arabidopsis FLC locus. *Science* **340**, 619–621 (2013).
1217 81. Xu, W. *et al.* The R-Loop Atlas of Arabidopsis Development and Responses to Environmental
1218 Stimuli. *Plant Cell* **32**, 888–903 (2020).

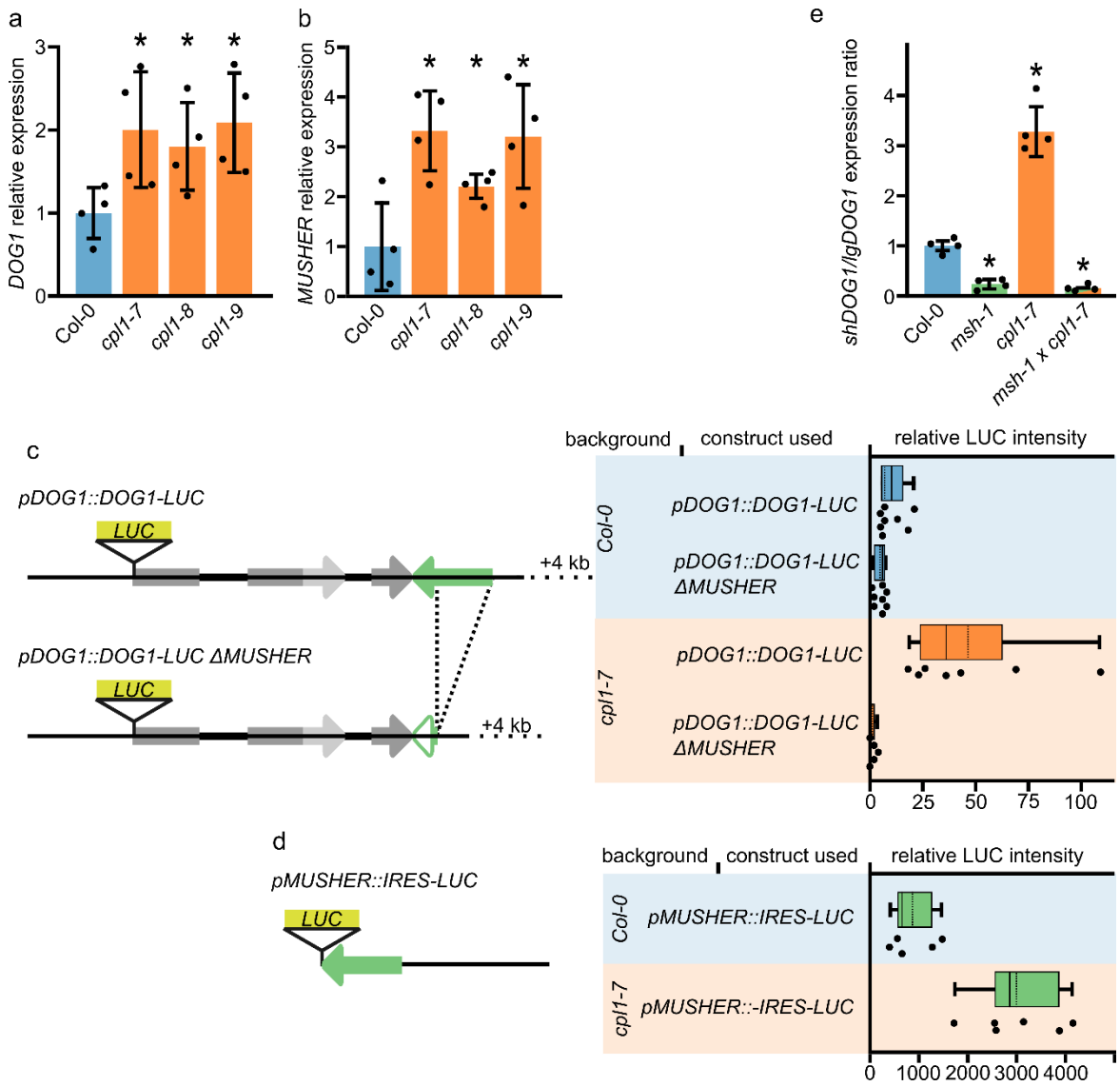
1219 82. Manavella, P. A. *et al.* Fast-Forward Genetics Identifies Plant CPL Phosphatases as Regulators of
1220 miRNA Processing Factor HYL1. *Cell* **151**, 859–870 (2012).
1221 83. Simpson, G. G., Dijkwel, P. P., Quesada, V., Henderson, I. & Dean, C. FY Is an RNA 3' End-
1222 Processing Factor that Interacts with FCA to Control the Arabidopsis Floral Transition. *Cell* **113**,
1223 777–787 (2003).
1224 84. Xing, D., Zhao, H., Xu, R. & Li, Q. Q. Arabidopsis PCFS4, a homologue of yeast polyadenylation
1225 factor Pcf11p, regulates FCA alternative processing and promotes flowering time. *The Plant*
1226 *Journal* **54**, 899–910 (2008).
1227 85. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–
1228 359 (2012).
1229 86. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
1230 87. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for
1231 exploring deep-sequencing data. *Nucleic Acids Research* **42**, W187–W191 (2014).
1232 88. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
1233 89. Smith, T., Heger, A. & Sudbery, I. UMI-tools: Modelling sequencing errors in Unique Molecular
1234 Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2016).
1235 90. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features.
1236 *Bioinformatics* **26**, 841–842 (2010).
1237 91. Warkocki, Z. *et al.* Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s. *Cell* **174**,
1238 1537-1548.e29 (2018).
1239 92. Alpern, D. *et al.* BRB-seq: ultra-affordable high-throughput transcriptomics enabled by bulk RNA
1240 barcoding and sequencing. *Genome Biol* **20**, 71 (2019).
1241 93. Sofroniew, N. *et al.* napari: a multi-dimensional image viewer for Python. Zenodo
1242 https://doi.org/10.5281/ZENODO.6597502 (2022).
1243 94. Bokota, G. *et al.* PartSeg: a tool for quantitative feature extraction from 3D microscopy images for
1244 dummies. *BMC Bioinformatics* **22**, 72 (2021).



Supplementary Fig. 1| *MUSHER* expression is partially dependent on *DOG1*.

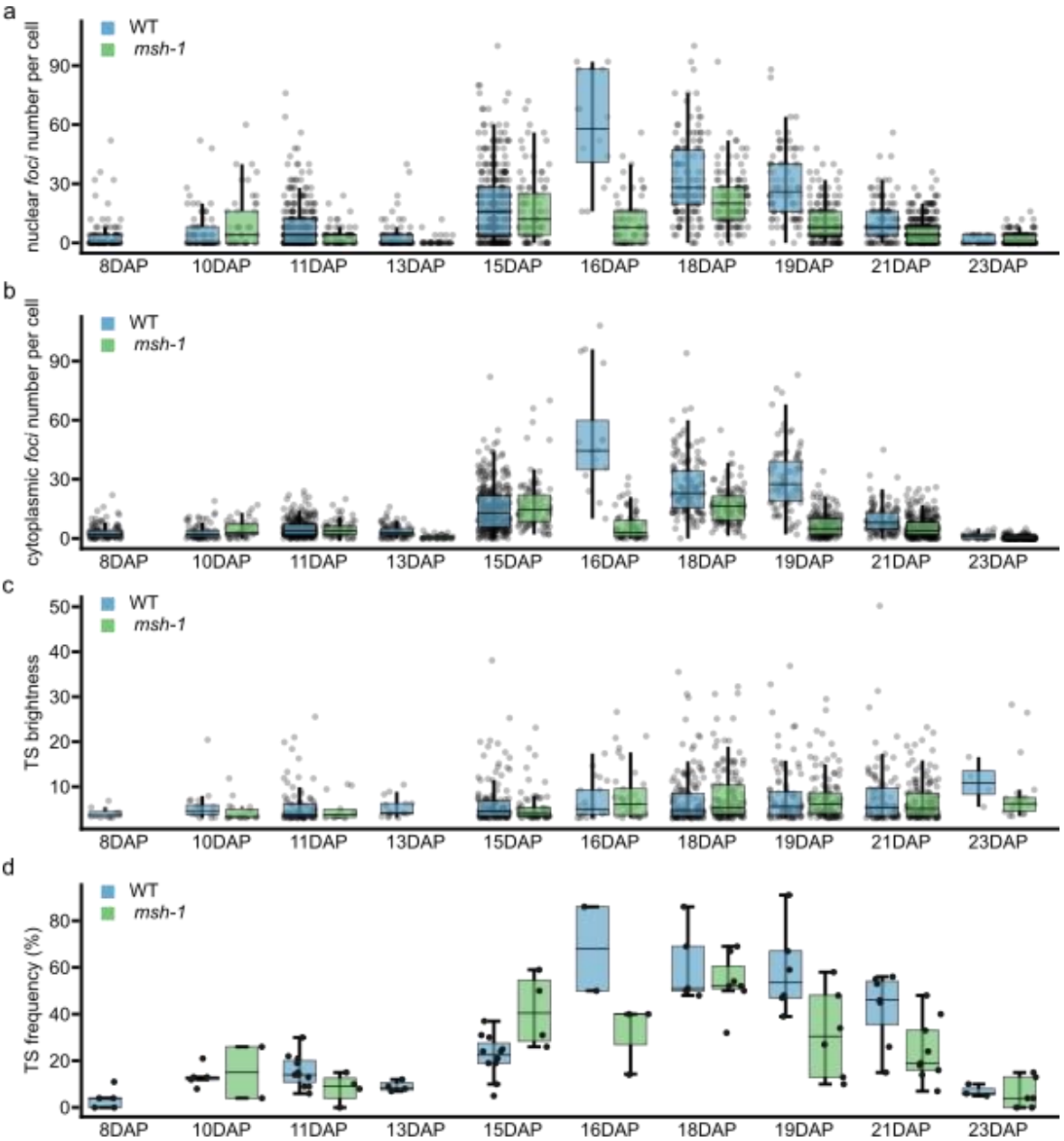
a *MUSHER* relative expression level in developing seeds of WT and indicated mutants, normalised to *UBC21* and relative to WT. *p*-value was calculated in comparison with WT. *p < 0.05. Error bars represent the standard deviation of four biological replicates. **b** *MUSHER* relative expression level in freshly harvested seeds of WT and dCas9 mutants, normalised to *UBC21* and relative to WT. *p*-value was calculated in comparison with WT. *p < 0.05. Error bars represent the standard deviation of four biological replicates. New transcripts around *MUSHER locus* were quantified using *MUSHER_2* and

MUSHER_3 primers. Transcript relative expression level in WT and indicated mutants seeds, normalised to *UBC21* and relative to WT. *p*-value was calculated in comparison with WT. *p < 0.05. Error bars represent the standard deviation of four biological replicates. **e** Representative photographs of *Col-0*, *msh-1*, *msh-2*, *msh_dCas9-1*, *msh_dCs9-2* and *Δmsh* corresponding to germination tests on Fig. 1F.



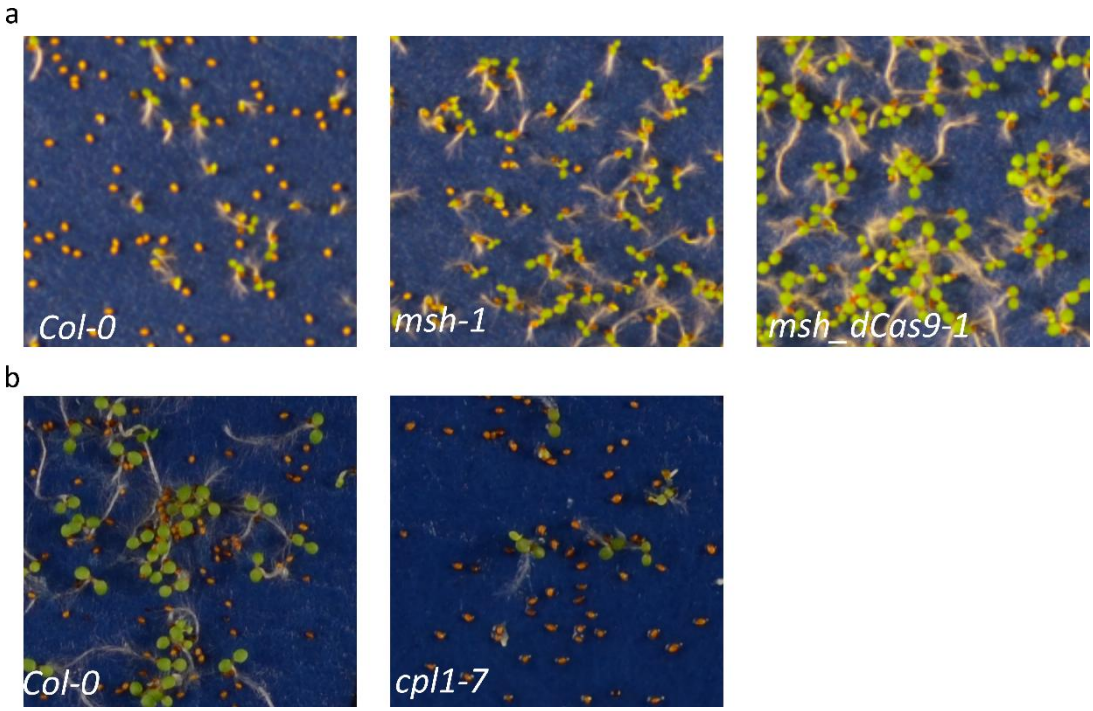
Supplementary Fig. 2| *MUSHER* overexpression activates *DOG1*.

a *DOG1* and **b** *MUSHER* relative expression level in seeds of WT and *cpl1-7* mutant, normalised to *UBC21* and relative to WT. *p*-value was calculated in comparison with WT. **p*-value < 0.05. Error bars represent the standard deviation of four biological replicates. **c** Luciferase (LUC) gene reporter assay was used to measure *MUSHER's* ability to activate the *DOG1* gene expression. Schematic diagram showing regions included in *LUC* gene fusion constructs used for transient transformation of Arabidopsis WT and *cpl1-7* mutant seedlings. In boxplots, middle lines show the medians (solid line) and mean (dotted line); box limits indicate the 25th and 75th percentiles; whiskers extend 1.5 times the interquartile ranges. * *p*-value < 0.05. **d** Luciferase activity assay on transgenic seeds containing the *LUC* gene fused with the *MUSHER* promoter. In boxplots, middle lines show the median (solid line) and mean (dotted line); box limits indicate the 25th and 75th percentiles; whiskers extend 1.5 times the interquartile ranges. * *p*-value < 0.01. **e** The ratio of *shDOG1* and *lgDOG1* isoforms in freshly harvested seeds of WT and *msh-1, cpl1-7* and double *msh-1 cpl1-7* mutants, normalised to *UBC21* and relative to WT. *p*-value from the two-tailed t-test, **p*-value < 0.05. Error bars show standard deviations.

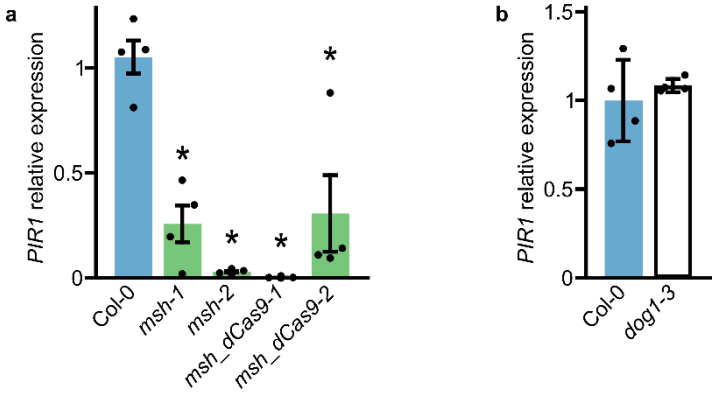


Supplementary Fig. 3| *MUSHER* promotes *DOG1* mRNA abundance in the nucleus and cytoplasm but not burst frequency or intensity.

a Nuclear *DOG1* mRNA foci number per cell during silique development in WT and *msh-1.* **b** Cytoplasmic *DOG1* mRNA foci number per cell during silique development in WT and *msh-1.* **c** Intensity of *foci* corresponding to *DOG1* transcription sites (TS) during silique development in WT and *msh-1. Foci* with a fold-change ≥3 times the average intensity of all nuclear foci, were considered as TS. **d** Frequency of *DOG1* TS during silique development in WT and *msh-1.* For all box plots, the middle line represents the median; the upper and lower lines are the first and third quartiles (Q1 and Q3); whiskers extend 1.5 times the interquartile ranges. The x-axis represents days after pollination (DAP).

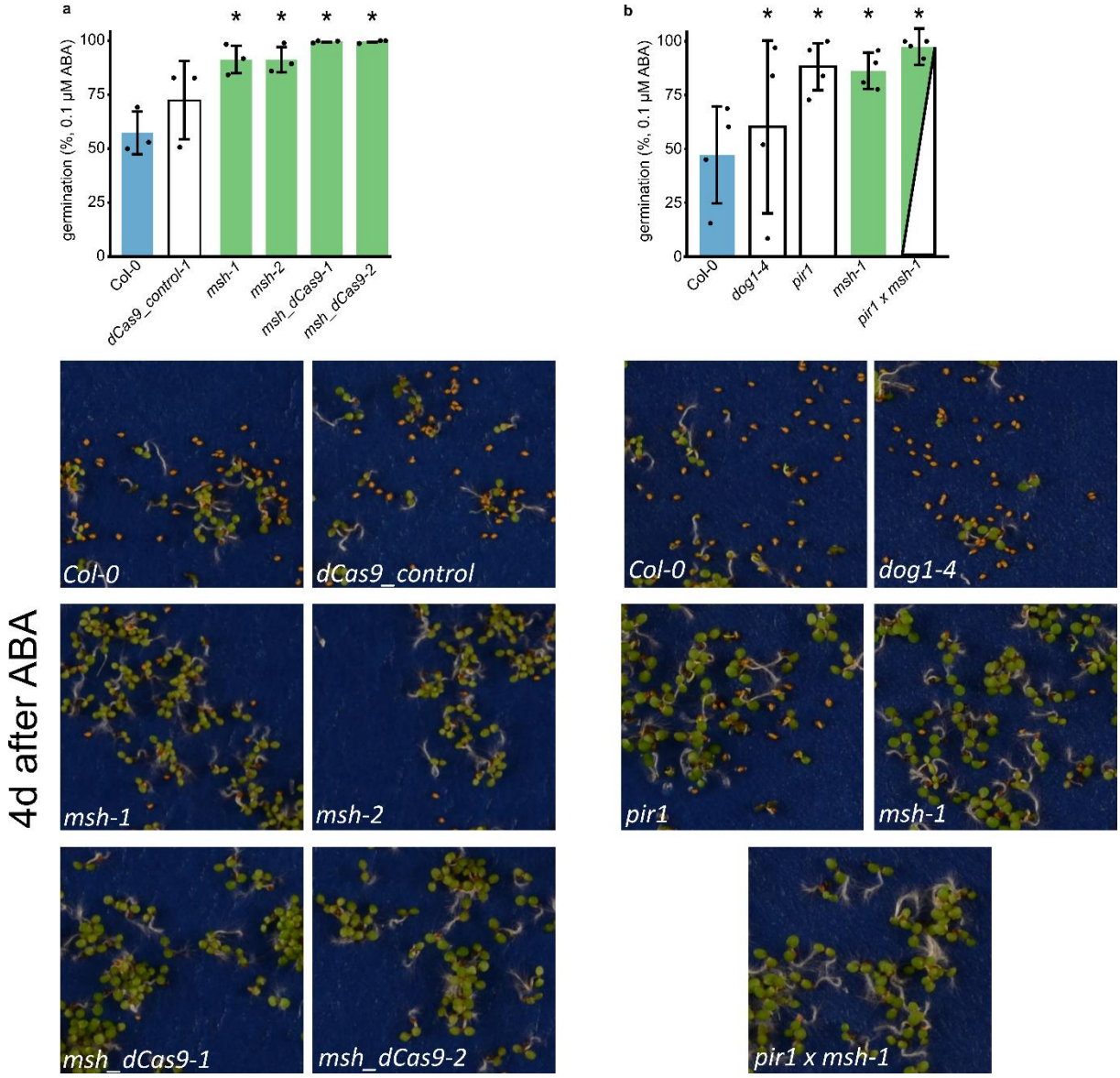


Supplementary Fig. 4| Lower secondary dormancy of *MUSHER* mutants and higher secondary dormancy in *cpl1-7*. **a** Representative photographs of WT, *msh-1* and *msh_dCas9-1* seeds corresponding to germination tests on Fig. 3a. **b** Representative photographs of *Col-0* and *cpl1*-7 germination tests corresponding to Fig. 3b.

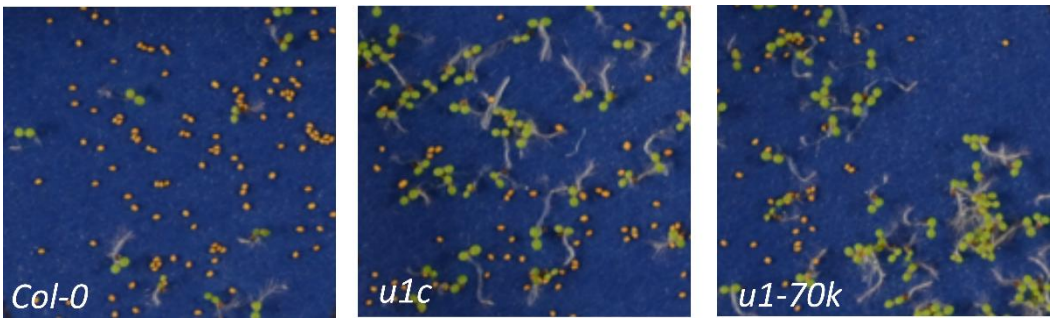


Supplementary Fig. 5 | *MUSHER* mutants showed decreased *PIR1* expression, which remained unaffected in the *dog1-3* mutant.

PIR1 relative expression level in seeds of *msh-1*, *msh-2, msh_dCas9-1, msh_dCas9-2* **(a)** and *dog1-3* **(b)** mutant subjected to secondary dormancy (2 days), normalised to *UBC21* and relative to WT. *p*-value was calculated in comparison with WT. *p-value < 0.05. Error bars represent the standard deviation of four biological replicates.



Supplementary Fig. 6| *MUSHER* mutants are less sensitive to exogenous ABA. The germination rate of WT, *dCas9-control-1*, *msh-1*, *msh-2, msh_dCas9-1, msh_dCas9-2* **(a)** and WT, *dog1-4, pir1, msh-1* and double *pir1 x msh-1* **(b)** mutant seeds on blue paper supplemented with 0,1uM ABA. *P-value<0.05 from two-tailed Student t-test. Error bars show standard deviations from three biological replicates.
Below each chart are representative photographs corresponding to the germination tests above.



Supplementary Fig. 7| Lower secondary dormancy of *u1c and u1-70k* mutants. Representative photographs corresponding to the germination tests shown on Fig. 6e