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 target gene[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

Characterisation of *MUSHER* lncRNA as a regulator of *DOG1* gene expression and seed germination.

To identify new lncRNA controlling dormancy, 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* (Supplementary Fig. 1b) showed reduced *MUSHER* levels in seeds (Supplementary Fig. 1a). Unexpectedly, *Δmsh* despite deletion of mapped *MUSHER* start site showed ectopic expression 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. 1f). 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. 1g), 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, dog*1-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.

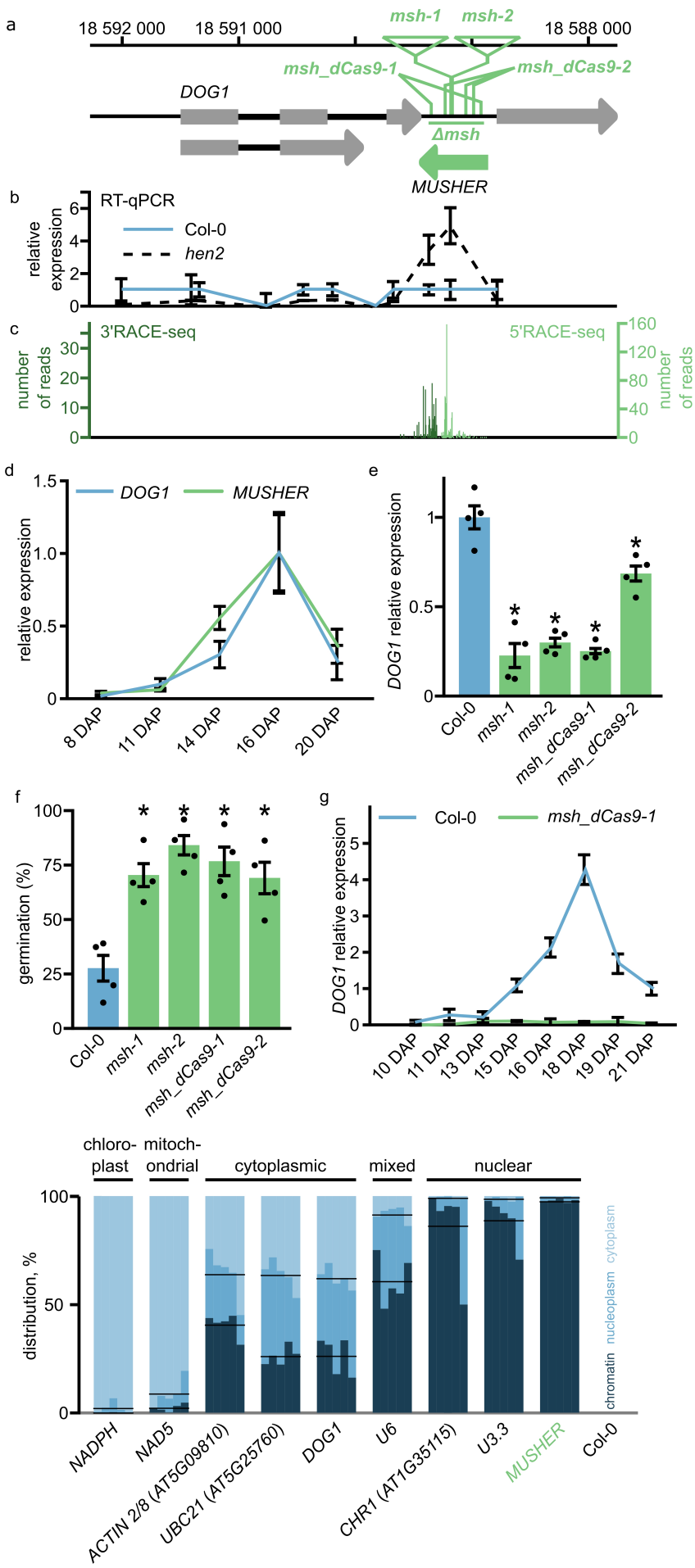
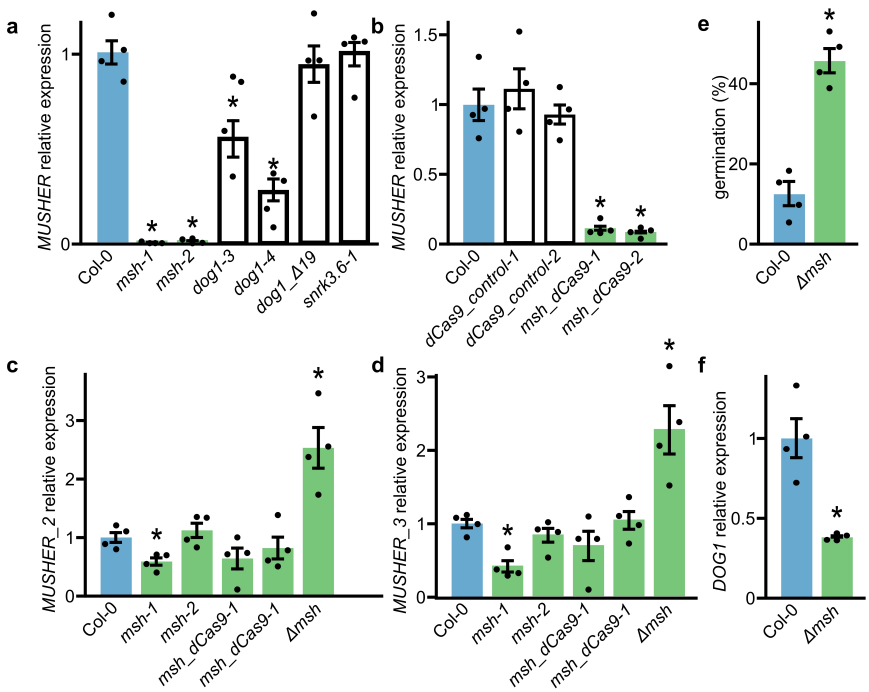


Fig. 1| *MUSHER* lncRNA acts as a positive regulator of seed dormancy and *DOG1* expression.

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 deleted region of MUSHER 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 of tested seeds 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 seed dormancy is established during seed maturation, as a result of the expression of *DOG1* and other genes. 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* chromatin association suggests that it may control *DOG1* expression by binding to its chromatin.



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

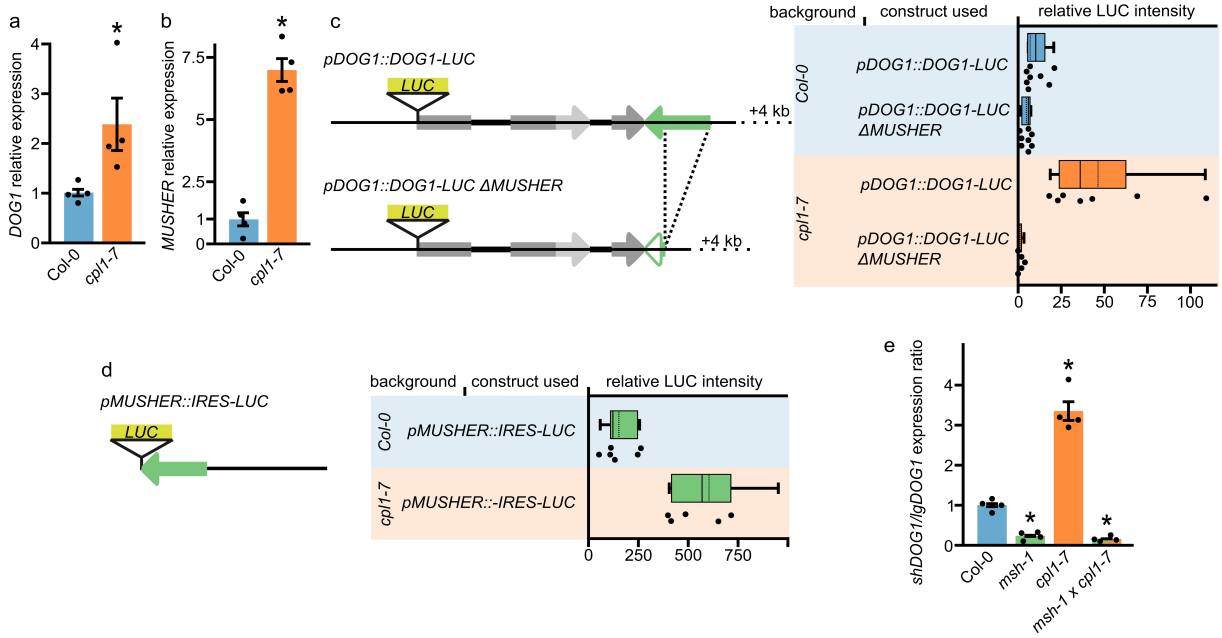
a *MUSHER* relative expression level in developing seeds of WT and indicated mutants, normalised to *UBC21* and relative to WT. *p*-value was calculated in comparison with WT. *p < 0.05. Error bars represent the standard deviation of four biological replicates. **b** *MUSHER* relative expression level 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. *MUSHER*-associated 2nd (c) and 3rd (d) 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 Freshly harvested seeds of WT and Δ*musher* mutant were scored for germination. **e** Germination of tested seeds 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). **f** *DOG1* relative expression levels in freshly harvested seeds of Δ*musher* mutant, relative to WT, normalised to *UBC21*. *p < 0.05 from two-tailed Student's *t*-test.

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 of *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, in the *cpl1-7* mutant, the *MUSHER* transcript level is elevated (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] (*pDOG1::LUC-DOG1ΔMUSHER*). Transient expression in Arabidopsis seedlings showed that *MUSHER* deletion suppressed *DOG1* upregulation in the *cpl1* plants (Supplementary Fig. 2c).



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.

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.

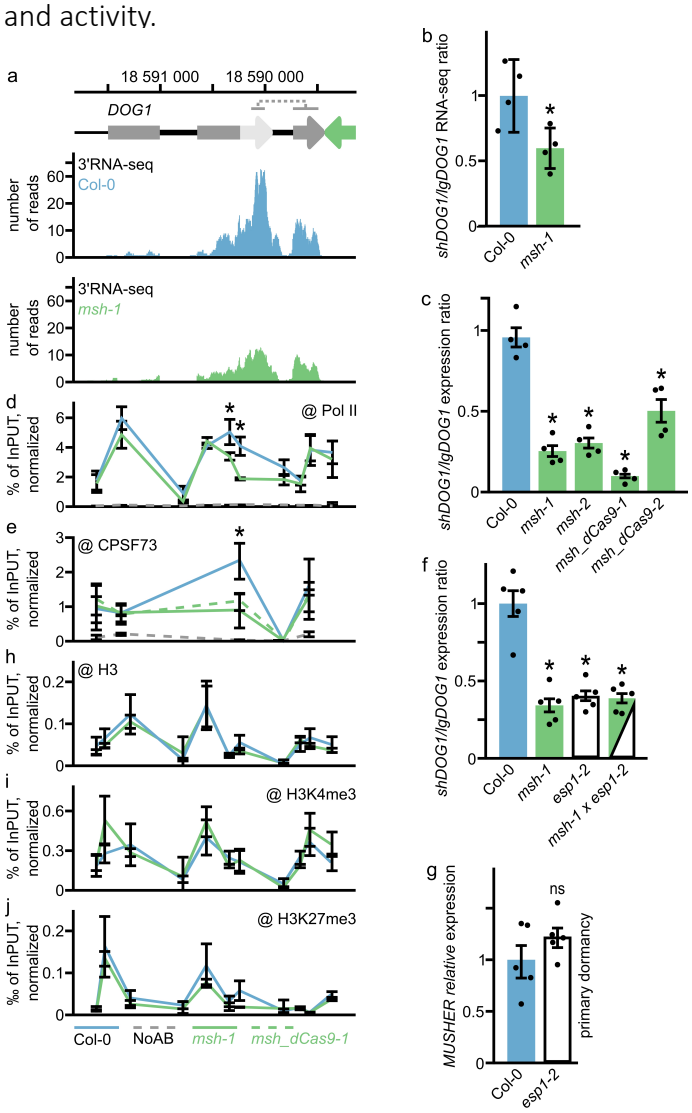


Fig. 2| *MUSHER* enhances *DOG1* proximal polyA site selection.

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). **f** Pol II occupancy was measured by ChIP-qPCR in WT and

msh-1 mutant. (n=3). **g** *MUSHER* relative expression level in seeds of WT and *esp1-2*, normalised to *UBC21* and relative to WT (n=5, ns- not significant). **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

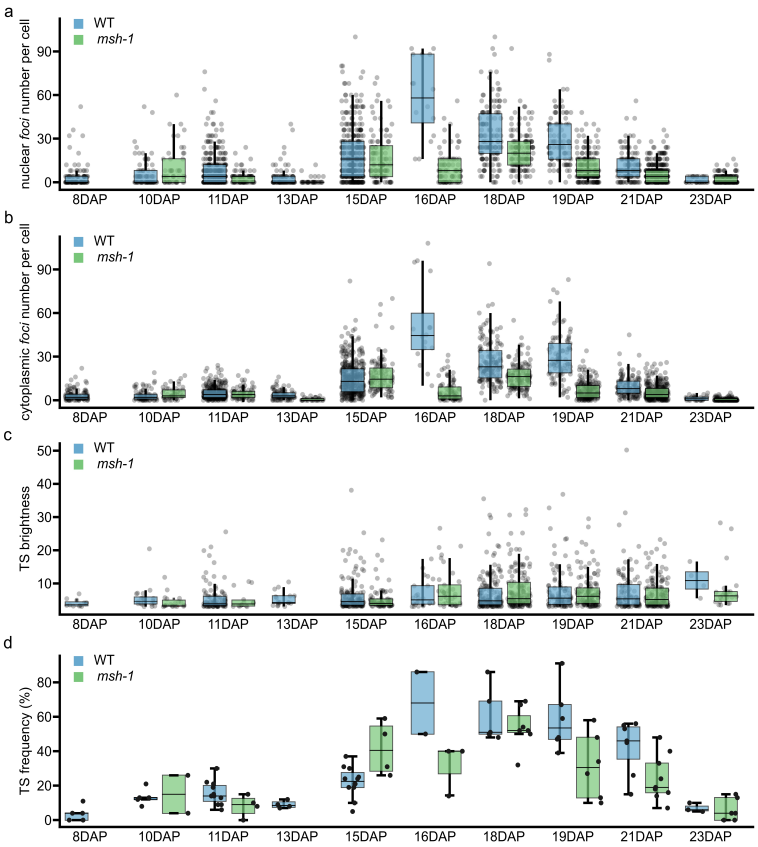
We used 3'RNA-seq to confirm that *shDOG1* is decreased in *msh-1* seeds, while *lgDOG1* remains unchanged when compared to WT (Fig. 2a). Calculation of 3'RNA-seq read density assigned to *DOG1* isoforms revealed a significant reduction of *shDOG1*/*lgDOG1* ratio in *msh-1* (Fig. 2b). Subsequent RT-qPCR 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). Because *cpsf73* mutants are sterile[44], for genetic analysis, we used a mutant in an auxiliary component of Cleavage Stimulatory Factor complex encoding ESP1 (CSTF64L) protein[45]. Analysis of the *msh-1 esp1-2* double mutant showed no additivity in *DOG1* termination site selection defects (Figs. 2e,f). This is supported by no change of *MUSHER* expression in the *esp1* mutant (Fig. 2g). Our results implied that *MUSHER* increases *DOG1* expression by enhancing CPSF complex recruitment to the proximal termination site, as supported by findings in *cpl1-7* mutant showing *MUSHER* overexpression and enhanced 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. 2f). 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 by controlling transcription levels. To test this hypothesis, we applied single-molecule RNA FISH (smFISH) to measure the level of *DOG1* mRNA transcripts and analyse transcription kinetics by quantifying the intensity of *DOG1* transcription[7]. The brightness of dots identified as transcription sites is generally considered to correlate with the number of RNA molecules present at transcription sites, while the frequency of transcription sites reflects the proportion of cells exhibiting detectable transcription activity[7]. We collected embryos from WT and *msh-1* siliques at ten-time points, corresponding to the experiments showing dynamics of *DOG1* and *MUSHER* expression during seed maturation (Fig. 1g). *DOG1* smFISH showed a consistent and significant reduction in the number of both nuclear and cytoplasmic *foci* during seed development in *msh-1* seeds when compared to WT (Supplementary Figs.3a,b). We observed an increase in *DOG1* transcript number in WT siliques between the 16th and 21st day, which was diminished in *msh-1* (Supplementary Figs.3a,b). Notably, the *msh* mutant showed no significant change in the intensity of nuclear dots identified as transcription sites (Supplementary Fig. 3c)[7]

379 and only minor changes in the frequency of cells with active *DOG1* transcription
380 (Supplementary Fig. 3d). We therefore conclude that in agreement with Pol II ChIP (Fig. 2f),
381 smFISH indicated no major changes in *DOG1* transcription initiation in *msh-1* seeds.



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).

406
407 Our data suggest that *MUSHER* promotes *DOG1* proximal termination by locally affecting
408 cleavage and polyadenylation complexes and Pol II occupancy at the proximal transcription
409 termination site without affecting *DOG1* transcription levels. All modes of RNA Pol II activity,
410 including transcription termination, are accompanied and modulated by histone
411 posttranslational modifications. On the contrary, some lncRNAs may recruit specific histone
412 modelling complexes, including PRC2 involved in H3K27me3 deposition[47,48]. To address the
413 effects of *MUSHER* on the *DOG1* chromatin landscape, we analysed selected active (H3K4me3)
414 or repressive (H3K27me3) chromatin histone marks along the *DOG1* locus. Using ChIP-qPCR, we
415 detected no significant difference in both tested histone modification levels in the *msh-1*
416 mutant throughout the *DOG1* locus (Figs. 2h,j) compared to WT. This result showed that defects
417 at the proximal transcription termination site (TTS) are unrelated to chromatin modifications
418 and agree with no changes in Pol II recruitment for most of the *DOG1* locus. In summary, we
419 propose a model in which *MUSHER* lncRNA enhances *DOG1* proximal termination rather than
420 control Pol II recruitment.
421

422 ## *MUSHER* enhances secondary dormancy in response to high-temperature

423 *DOG1* is a known regulator of both primary and secondary dormancy[49,50]. Based on the weak
424 dormancy in *musher* mutants and the requirement for high *DOG1* expression during primary
425 seed dormancy establishment (Fig.1), we hypothesised that *MUSHER* may also be involved in
426 secondary dormancy regulation. We used a well-established secondary dormancy induction
427 protocol that consists of prolonged incubation in the darkness at high temperatures followed

by germination analysis in permissive conditions[49]. In agreement with defects in primary dormancy, *msh-1* and *msh_dCas9-1* showed weaker secondary seed dormancy, and *cpl1-7*, a mutant with high *MUSHER* expression, showed stronger secondary dormancy when compared to WT (Figs.3a,b). Consistent with these phenotypes, we observed the down- and up-regulation of the *shDOG1* transcript in *musher* and *cpl1-7* mutants, respectively (Figs.3c,d).

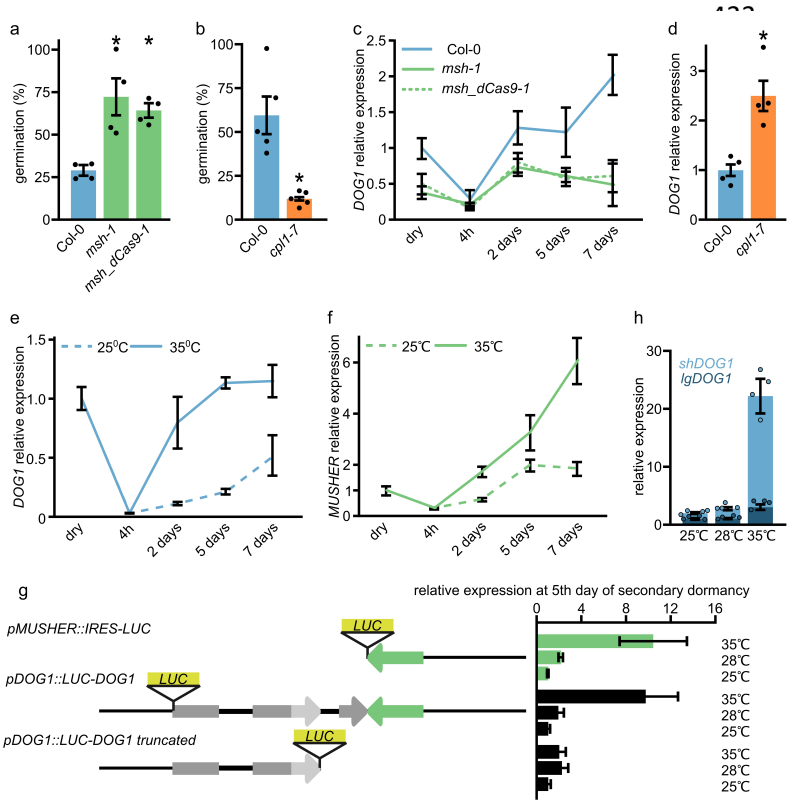


Fig. 3| *MUSHER* activates *DOG1* expression during secondary dormancy induction in response to high temperature.

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 second day of secondary dormancy induction and remained steady 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).

Next, we tested if *MUSHER* could pass the temperature input for *DOG1* regulation during secondary dormancy induction. We subjected luciferase reporter expressing seeds (Supplementary Fig. 2) to secondary dormancy induction at 25, 28, and 35°C. We noted that *MUSHER* and *DOG1* reporters showed similar gradual increases in expression in response to increasing temperatures (Fig. 3g). Interestingly, *DOG1*-expressing construct devoid of *MUSHER* region (*pDOG1::DOG1-LUC truncated*) showed no induction at 35°C (Fig. 3g). 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 the findings that *MUSHER* enhances the selection of the proximal termination site, we noticed a prominent increase in the level of the short but much smaller increase in the long *DOG1* mRNA at high temperatures during secondary dormancy induction (Fig. 3h).

In summary, our data are consistent with a model in which *MUSHER* lncRNA acts as an activator of *DOG1* 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* 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 reads showed a strong reduction in reads corresponding to both *PIR1* and an upstream localised *AT2G35320* gene (Fig. 4b). RT-qPCR confirmed the strong downregulation of *PIR1* expression in *musher* mutant for mature (Fig. 4c) and secondary dormancy-induced seeds (Fig. 4d). Interestingly, *PIR1* expression remained unchanged in the *dog1* mutant, indicating that *PIR1* is controlled by *MUSHER* independently from its effect on *DOG1* (Figs. 4c,d). Consistent with the conclusion that *MUSHER* activates *PIR1*, we observed that increased *MUSHER* expression 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 transcription levels rather than termination of *PIR1* gene (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). As H3K4me3 levels are highly correlated with Pol II transcriptional activity[54], this agrees with the reduced expression of the *PIR1* gene.

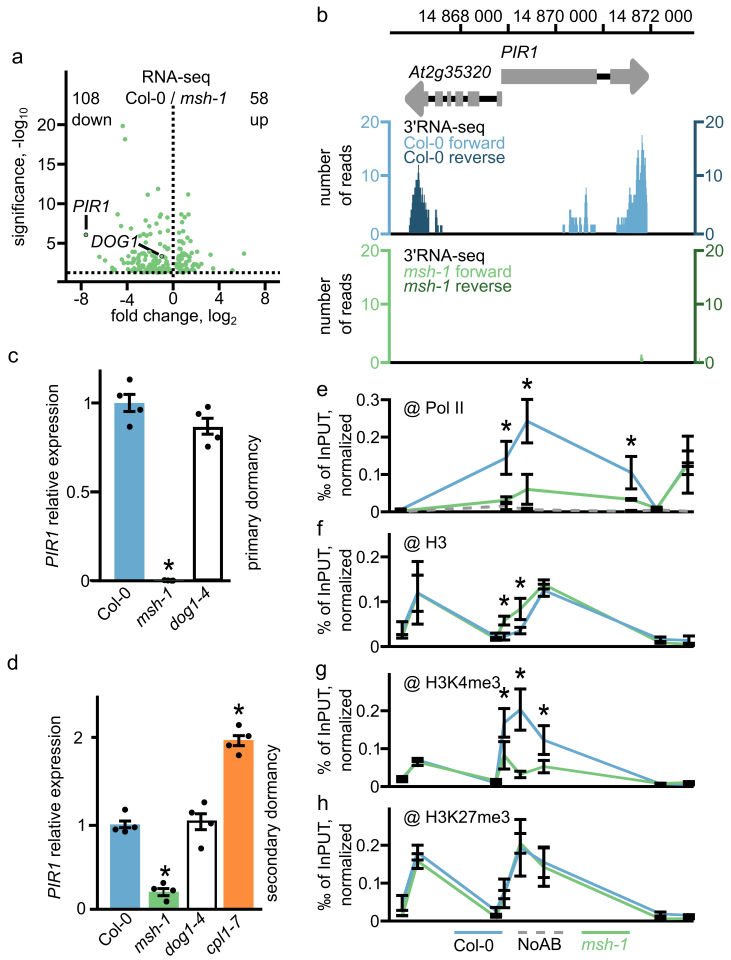


Fig. 4| *MUSHER* enhances *PIR1* expression independently of its role in *DOG1* regulation.

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 suggested that *MUSHER* regulates transcriptional activity at the *PIR1* gene, whereas at *DOG1*, it specifically promotes 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) and regulates gene expression at two *loci*. 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).

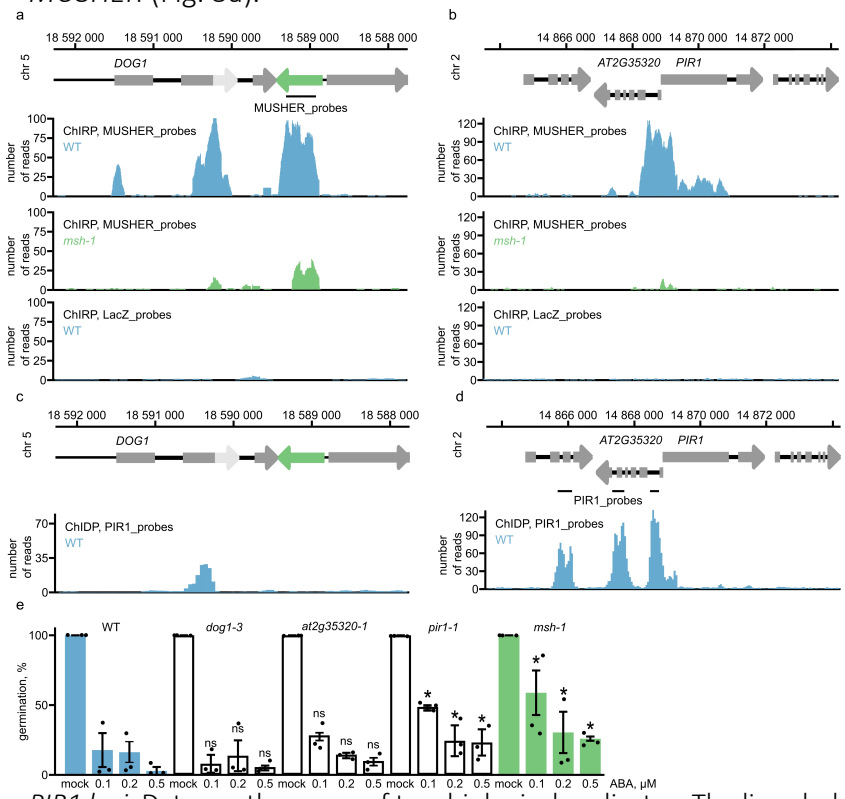


Fig. 5| *MUSHER* lncRNA integrates ABA and DOG1 pathways by interacting with *PIR1* TSS and *DOG1* proximal termination site.

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*, *pir1-1* 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 the *DOG1 locus* is localised next to *MUSHER* on chromosome 5, the *PIR1 locus* is localised on chromosome 2. This encouraged us to test the physical proximity of *PIR1* and *DOG1 locus* chromatin.

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

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

U1 snRNP mediates *MUSHER* association to chromatin

MUSHER binds to and regulates *DOG1* and *PIR1* gene expression, with most *MUSHER* transcripts bound to chromatin (Fig. 1h). This raised the question of whether *MUSHER* association with chromatin is required for its function and what is the mechanism of *MUSHER* binding to chromatin.

In addition to its well-known splicing function, U1 snRNP has been shown in mammals to bind various lncRNAs, facilitating their chromatin tethering[15]. Interestingly, we identified four putative U1 binding sites within the *MUSHER* sequence even though we failed to detect *MUSHER* splicing (Supplementary Table 3 and Fig. 1c). Using probes specific for *MUSHER*, but not *LacZ* probes, we selectively recovered U1, but not U4 or U6 RNA using formaldehyde-assisted RNA ImmunoPrecipitation (RIP) (Fig. 6a). Supporting the specificity of the approach, U1 recovery was strongly diminished when we repeated the experiment using *MUSHER* probes in *msh-1* lines (Fig. 6b). Next, we validated the *MUSHER*-U1 association by RNA Immunoprecipitation (RIP) using the U1 snRNP component U1-70K. We used a recently reported *u1-70k* complementing line expressing GFP tagged U1-70K driven by endogenous U1-70K promoter (*pU1-70K::U1-70K-GFPu1-70k*)[21]. RT-qPCR showed that U1-70K immunoprecipitation recovered spliced control RNA, failed to recover unspliced control mRNA, and, importantly, retrieved a significant amount of *MUSHER* lncRNA (Fig. 6c).

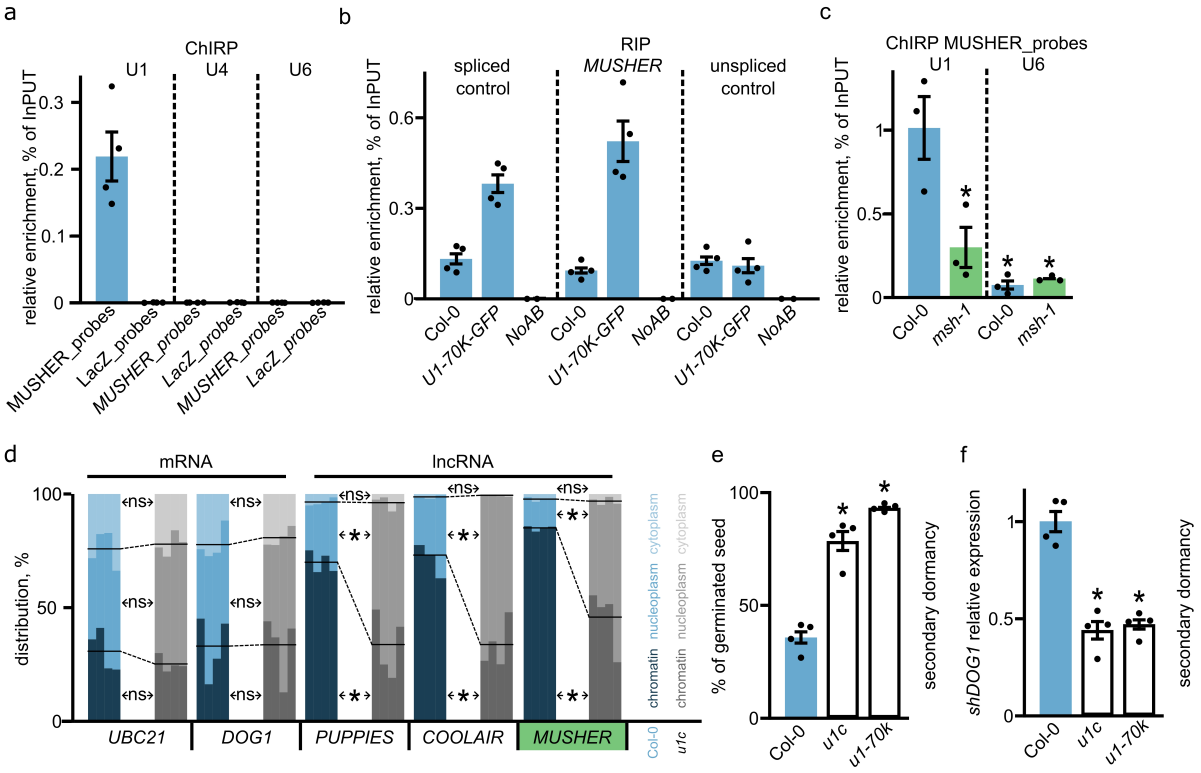


Fig. 6| U1 is required for chromatin tethering of *MUSHER* and other lncRNA.

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. In addition, neither *UBC21* nor *DOG1* mRNA distribution among the cellular fractions analysed was altered in the *u1c* mutant (Fig. 6d). Notably, *MUSHER* chromatin association was reduced in *u1c* mutant, a result that is expected given the presence of four U1 binding sites within the *MUSHER* transcriptional unit. Next, we asked whether the U1 requirement for *MUSHER* chromatin retention 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 in plants.

Finally, we asked if *MUSHER* chromatin association is required for its function. Both *u1c* and another amiRNA mutant in the U1 snRNP component U1-70K, showed 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* shows defects in both primary and secondary dormancy induction. To synchronise the physiological state of seed , they were matured on mother plants, stored until they lost primary dormancy and then subjected to induce 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). This agrees with the model where U1 is required for *MUSHER* association to chromatin and seed dormancy induction. The reduction in *u1c* and *u1-70k* secondary dormancy was accompanied by decreased *shDOG1* expression, as expected from *musher* mutants' analysis (Fig. 6f). In summary, U1 is required for multiple lncRNA chromatin associations and, at least in the case of *MUSHER,* this association is required for the lncRNA gene expression regulatory 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).

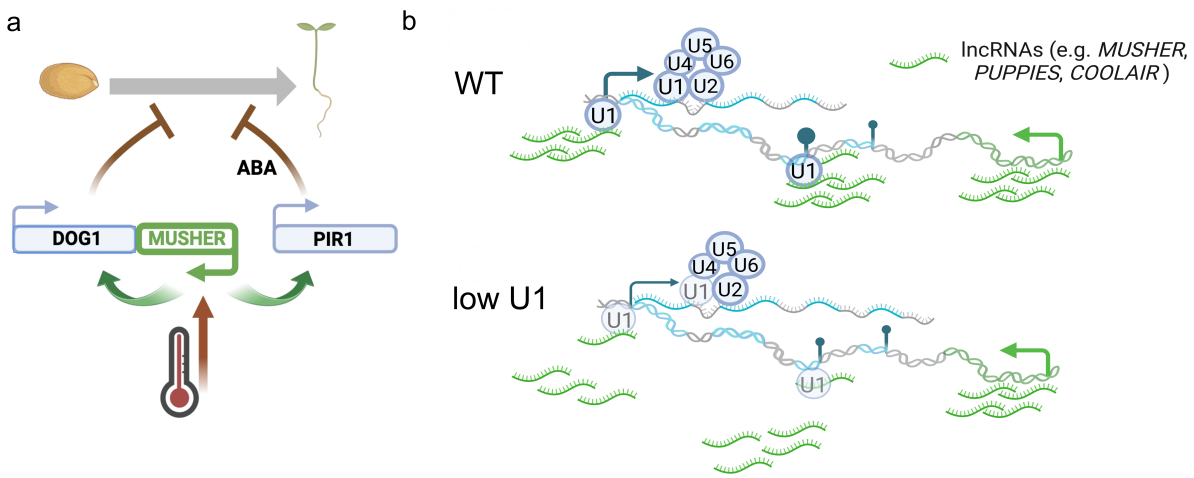


Fig. 7| a *MUSHER* modulates seed dormancy through the transcriptional regulation of DOG1 and ABA pathways. **b** U1 is required for chromatin retention of lncRNA, including *MUSHER* in Arabidopsis.

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

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

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

MUSHER as a Link Between ABA Sensitivity and *DOG1*

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

744 decreased ABA sensitivity. We propose that *MUSHER* lncRNA facilitates the concurrent
745 activation of DOG1 and ABA pathways, integrating them into a shared regulatory mechanism.
746 High temperatures are known to enhance *DOG1* expression and ABA biosynthesis in seeds[51,67].
747 We observed altered ABA sensitivity and reduced secondary dormancy induction in response
748 to high temperature in *msh* mutants. Elevated temperatures contribute to two well-described
749 responses: 1) thermoinhibition, a high-temperature-mediated inhibition of germination in
750 light[68] and 2) secondary dormancy induction in non-dormant imbibed seeds exposed to high
751 temperatures in the darkness[31,49,51]. We hypothesise that *MUSHER*-mediated integration of ABA
752 and DOG1 pathways contributes not only to dormancy but also to other ABA/DOG1-related
753 phenotypes (e.g., drought tolerance or thermoinhibition). We conclude that, similar to other
754 long non-coding RNAs involved in temperature sensing, such as *COOLAIR*[8], *COLDWRAP*[69] and
755 *SVALKA* [6] in plants, and *Hsrω*[70] and *Heat*[71] in animals, *MUSHER* also plays a role in temperature
756 sensing in plants.

757 U1 role in chromatin retention of lncRNA in Arabidopsis.

758 In contrast to mature mRNAs, lncRNAs are often enriched in the nucleus[1,15]. Consistently, many
759 lncRNA have been shown to regulate gene expression by binding to chromatin[1,5,6,39]. The
760 molecular mechanism of human lncRNA retention in the nucleus has only recently started to
761 be explored [15,72,73], and those remain largely unknown in plants. Mechanisms ensuring human
762 lncRNA nuclear retention encompass the presence of specific sequences[74], including *Alu*
763 transposable elements binding the HNRNPK complex[75]. These protein-RNA interactions may
764 contribute to liquid-liquid phase separation (LLPS) that can also be responsible for the nuclear
765 retention of lncRNA. For example, LLPS by *NEAT1* lncRNA results in the formation of nuclear
766 paraspeckles[76] and *XIST* lncRNA was reported to undergo LLPS with different partners localised
767 at inactive X chromosome[77]. However, nuclear retention may not be enough to guide lncRNA
768 directly to chromatin, which has been proposed to be a result of direct RNA-DNA interaction
769 either in the form of R-loops or RNA:DNA triplex formation[15]. In addition, U1 snRNP, involved in
770 5' splice site recognition, has recently been shown to be responsible for the chromatin binding
771 of a large fraction of lncRNAs[15]. In contrast to *Alu* and other sequence elements identified in
772 humans, U1 binding sites are well conserved in plant genomes.

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

786 efficiently. As we did not test the interplay between U1 binding and LLPS, we cannot exclude
787 the possibility that U1 binding is required or otherwise interlinked with the phase separation of
788 lncRNA at chromatin. R-loops have not been reported in seeds, and the one shown in seedlings
789 at *PIR1* locus[81] is unlikely to be formed by *MUSHER* since *MUSHER* is very lowly expressed in
790 seedlings. Therefore, while we cannot exclude this possibility, *MUSHER* is unlikely to form R-
791 loops. As concluded for human U1 lncRNA targets[15], U1 also do not form direct R-loops with
792 the target DNA. This suggests that R-loops probably do not contribute to U1-mediated lncRNA
793 retention. The fact that the tested lncRNA have different modes of action implies that the U1
794 role in chromatin retention is not limited to a specific lncRNA activity. This, together with the
795 described role in chromatin retention of human and mouse lncRNA, strongly supports the idea
796 that the U1 role in chromatin retention of lncRNA is evolutionarily conserved in plants.

797 Methods

798 Plant material and growth conditions

799 WT used in this work is Columbia (Col-0) *Arabidopsis thaliana* accession. Plants were grown on soil in a
800 greenhouse under long-day conditions (16 h light/8 h dark) at 22°C/18°C. All mutants and transgenic
801 lines were created in Col-0 background and were homozygous for the indicated genotype. *MUSHER* lines
802 *msh-1*(SM_29297) and *msh-2*(SM_29310) were obtained from the NASC centre. Other mutants used in
803 this study have been described previously*: cpl1-7*[82], *cpl1-8* (GK-165H09)[32], *cpl1-9* (GK-849H10)[32], *dog1-*
804 *3* (SALK_000867)[27], *dog1-4* (SM_3_20886)[30], *fy-2* (SAIL_ 657D4)[83], *esp1-2* (SALK_078793)[84], *pir1*
805 (SALK_078237)[36], *amiR-u1-70k* and *amiR-u1-c*[21].

806 Germination tests

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

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

818 Cloning, Plant Transformation and CRISPR-dCas9 mutant generation

819 The generation of the dCas9 transgenic plants was performed according to Čermák et al.[42]. The specific
820 sgRNA to Arabidopsis TAIR10 genome were designed using the CRISPR-P 2.0
821 (http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR) to select the canonical NGG PAM motif. sgRNAs were
822 cloned to the dCas9 vector by Golden Gate assembly (*dCas9_control-1*, *dCas9_control-2*, *msh_dCas9-1*
823 and *msh_dCas9-2*) according to the protocols 3B and 5[42]. Primers used for cloning are listed in Table S4.
824 All constructs were verified by Sanger sequencing. Each construct was transformed into Agrobacterium
825 strain GV3101 and then transformed into Col-0 or an appropriate mutant background using the floral
826 dip method. T1 transformant plants were identified by selecting the resistance to glufosinate. The seeds
827 from the selected T1 plants were used for phenotypic analysis. At least two independent transgenic lines

were tested and propagated. As a control for potential indirect effects from dCas9 expression, we also created two dCas9 lines with guides recognising different intergenic regions localised 100 kb from *DOG1 locus*: *dCas9_control_1* and *dCas9_control_2*, which showed no influence on *MUSHER* transcription.

Luciferase measurement

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

RNA extraction

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

RT–qPCR

Trace DNA in the RNA samples was removed using a TURBO DNA-free kit (Thermo Fisher Scientific) according to the manufacturer's protocol. The DNA removal was confirmed by PCR using primers to the *PP2A* gene (AT1G13320). The RNA sample quality and quantity were checked by agarose gel electrophoresis and using a Nanodrop 2000 spectrophotometer. Reverse transcription using 2.5 µg of RNA was performed with a Superscript III cDNA Synthesis kit (Thermo Fisher Scientific) according to the manufacturer's protocol. An equal mixture of oligo(dT) and hexamer primers was used for cDNA synthesis according to the manufacturer's protocol. 5x diluted cDNA was used as a template in the qPCR reaction (LightCycler 480; Roche) 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,

871 1 mM PMSF, 1x Complete protease inhibitors (Roche)). Then antibodies against RNA polymerase II
872 subunit B1 (Agrisera, AS11 1804), H3 (Abcam, ab1791), H3K4me3 (Millipore, 07-473), H3K27me3
873 (Abcam, ab6002) or CPSF73 (PhytoAB, PHY1043A) were added followed by addition of the mixture of
874 Dynabeads Protein A and Protein G beads (Thermo Fisher Scientific). Next, samples were incubated
875 overnight on a rotator at 4°C. Two washing steps with low salt buffer (150 mM NaCl, 1 mM EDTA, 10
876 mM HEPES-KOH pH 7.4, 0.1% Triton X-100) and two washing steps using high salt buffer (500m M NaCl,
877 1 mM EDTA, 10 mM HEPES-KOH pH 7.4, 0.1% Triton X-100) were performed. DNA was released from
878 beads and reverse-crosslinked by incubation at 95°C with 10% Chelex 100 suspension for 15 min. After
879 Proteinase K treatment (55°C) for 90 min and repeated incubation at 95°C for 10 min, samples were
880 centrifuged for 5 min at RT at 10,000 g. Recovered DNA was used in qPCR reaction with primers listed
881 in Table S4.
882

883 RNA ImmunoPrecipitation (RIP)

884 For U1-70K-GFP RIP, we used 100 mg of fresh siliques (13-20 days after pollination), which were ground
885 in liquid nitrogen and resuspended in 20 ml of crosslink buffer (10 mM HEPES-KOH pH 7.4, 50 mM NaCl,
886 100 mM sucrose, 1% formaldehyde). After 10 min incubation on a rotator at 4°C, the crosslinking agent
887 was quenched by adding 1.25 ml 2 M glycine. After another 5 min incubation on a rotator, samples were
888 centrifuged at 4°C and 2,000 g for 15 min. The pellets were resuspended in 20 ml of cold Honda Buffer
889 (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,
890 0.5% Triton X-100, 10 mM β-mercaptoethanol, 1mM PMSF, 1x Complete protease inhibitors (Roche),
891 and 5 U/ml murine RNase inhibitor (NEB)) and filtered through Miracloth. Samples were centrifuged for
892 15 min at 2,000 g at 4°C. The pellet was resuspended in Nuclear Lysis Buffer (50 mM HEPES-KOH pH 7.4,
893 10 mM $MgCl_2$, 1% Triton X-100, 1x Complete protease inhibitors (Roche), 20 U/ml murine RNase
894 inhibitor (NEB)), followed by 20 min incubation at 37°C with TURBO DNAse (Thermo Fisher Scientific).
895 After adding SDS up to 1%, samples were sonicated using Bioruptor (Diagenode). Next, samples were
896 diluted 10 times with RIP Dilution Buffer (0.5 mM EDTA, 10 mM HEPES-KOH pH 7.4, 150 mM NaCl,
897 20 U/ml murine RNase inhibitor (NEB)) and incubated with GFP Agarose beads (Chromotek) for 1h on a
898 rotator at 4°C. After three washing steps with RIP Wash Buffer (150 mM NaCl, 1 mM EDTA, 10 mM
899 HEPES-KOH pH 7.4, 0.05% NP-40, 20 U/ml murine RNase inhibitor (NEB)), RNA was eluted (150 mM
900 NaCl, 1 mM EDTA, 10 mM HEPES-KOH pH 7.4, 1% SDS, 20 U/ml murine RNase inhibitor (NEB)) and
901 reverse-crosslinked by incubation at 95°C for 15 min. After 1h Proteinase K treatment (55°C), RNA was
902 isolated with Trisure (Bioline) according to producer protocol. After cDNA synthesis with SuperScript III
903 (Thermo Fisher Scientific), samples were used in qPCR reaction with appropriate primers listed in Table
904 S4.

905 Nuclei purification for ChIRP, RIRP and ChIDP

906 To uncover RNA-RNA (RNA Isolation by RNA Purification), RNA-DNA (Chromatin Isolation by RNA
907 Purification) and DNA-DNA (Chromatin Isolation by DNA Purification Assay) interactions, we used
908 400 mg of fresh siliques (13-20 days after pollination), which were ground in liquid nitrogen and
909 resuspended in 20 ml of crosslink buffer (10 mM HEPES-KOH pH 7.4, 50 mM NaCl, 100 mM sucrose, 1%
910 formaldehyde (for RNA-RNA and RNA-DNA) or 0.5% glutaraldehyde (for ChIDP)). After 10 min incubation
911 on a rotator at 4°C, the crosslinking reaction was quenched by adding 1.25 ml 2 M glycine. After an
912 additional 5 min incubation on a rotator, samples were centrifuged at 4°C and 2,000 g for 15 min. The
913 pellets were resuspended in 20 ml of cold Honda Buffer (20 mM HEPES-KOH pH 7.4, 0.44M sucrose,
914 1.25% Ficoll, 2.5% Dextran T40, 10 mM $MgCl_2$, 5 mM DTT, 0.5% Triton X-100, 10 mM
915 β-mercaptoethanol, 1 mM PMSF, 1x Complete protease inhibitors (Roche) and 5 U/ml murine RNase
916 inhibitor (NEB)) and filtered through Miracloth. Subsequently, samples were centrifuged for 15 min at

917 2,000 g at 4°C. According to downstream assays, the nuclei pellets were resuspended in appropriate
918 buffers (Element Zero Biolabs kit).

919 Chromatin Isolation by RNA Purification Assay (ChIRP)

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

923 For *MUSHER* probes RIRP and ChIRP with *MUSHER* and *lacZ* probes, nuclei pellets were resuspended in
924 1.1 ml of MagIC Lysis Buffer FA (Element Zero Biolabs kit), supplemented with 1x Complete protease
925 inhibitors (Roche) and 5 mM DTT. Next, samples were sonicated using a Bioruptor (Diagenode) and
926 diluted two times with MagIC Lysate Dilution Buffer FA, with 1% of the volume saved as an input sample.
927 MagIC beads were pre-cleared according to the producers' protocol and added to the samples that were
928 next incubated at 40°C for 30 minutes with occasional shaking. This was followed by three washes using
929 MagIC Wash Buffer I FA and nucleic acids elution by incubation at 92°C for 5 min with 10 mM Tris-HCl
930 pH 7.5. After Proteinase K treatment (55°C, 90 min) and phenol:chloroform:isoamyl alcohol extraction,
931 DNA and RNA were precipitated by adding 1 µl glycogen, 10% of the volume of 3 M LiCl, and an equal
932 volume of isopropanol. Samples were centrifuged for 30 min at 4°C at 20,000 g. Pellets were washed
933 with cold 80% ethanol, resuspended in water, and samples were split in half. The presence of specific
934 DNA fragments interacting directly or indirectly with *MUSHER* probes was validated by qPCR analysis. To
935 identify RNAs enriched with *MUSHER* probes, half of the sample was incubated with TURBO DNAse
936 (Thermo Fisher Scientific), extracted with phenol:chloroform: isoamyl alcohol as above, and used for
937 cDNA synthesis and qPCR analysis. For ChIRP library preparation, DNA fragments were eluted from
938 MagIC beads with RNase If (NEB). Libraries were prepared with the NEBNext® Ultra™ DNA Library Prep
939 Kit for Illumina using NEBNext Multiplex Dual Index Primers. Paired-end sequencing was performed on
940 the Illumina NovaSeq 6000.

941 Chromatin Isolation by DNA Purification Assay (ChIDP)

942 For ChIDP, the nuclei pellet was resuspended in 1.1 ml of MagIC Lysis Buffer FA (Element Zero kit)
943 supplemented with 1x Complete protease inhibitors (Roche) and 5 mM DTT. Samples were sonicated
944 using Bioruptor (Diagenode), diluted with MagIC Lysate Dilution Buffer FA, and split into halves (treated
945 with RNaseA at 37°C for 20 min and non-treated). MagIC beads recognising the *PIR1 locus* were mixed
946 with the lysates that were next incubated at 81°C for 5 min with occasional shaking to denature DNA
947 partially. Subsequently, samples were transferred to 40°C for 20 min to enable binding of the probes.
948 This was followed by six washing steps with the MagIC Lysate Wash buffer I and II, and DNA was eluted
949 with 10 mM Tris-HCl pH 7.5 at 92°C for 5 min. After adding an equal volume of MagIC Lysate Wash buffer
950 I to eluted samples, proteinase K treatment (55°C, 120 min) was performed. To extract DNA fragments,
951 an equal volume of phenol:chloroform:isoamyl alcohol was added. The upper phase was mixed with 1 µl
952 of glycogen, 10% of the volume of 3 M LiCl and an equal amount of isopropanol. After centrifugation for
953 30 min at 20000g at 4°C, the pellet was washed with 80% EtOH and resuspended in water. Libraries were
954 prepared with the NEBNext® Ultra™ DNA Library Prep Kit for Illumina using NEBNext Multiplex Dual
955 Index Primers. Paired-end sequencing was performed on the Illumina NovaSeq 6000.

956 ChIRP-seq and ChIDP-seq analysis

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

961 5'RACE-seq

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

970 3'RACE-seq

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

980 3'RNA-seq

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

988 Fractionation

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

RNA quality was tested by agarose gel electrophoresis and quantified with a Nanodrop 2000 spectrophotometer. cDNA synthesis was performed using a 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 took 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).

1045 2. Dror, I. *et al.* XIST directly regulates X-linked and autosomal genes in naive human pluripotent
1046 cells. *Cell* **187**, 110-129.e31 (2024).
1047 3. Li, J. *et al.* Long noncoding RNA XIST: Mechanisms for X chromosome inactivation, roles in sex-
1048 biased diseases, and therapeutic opportunities. *Genes Dis* **9**, 1478–1492 (2022).
1049 4. Arun, G., Aggarwal, D. & Spector, D. L. MALAT1 Long Non-Coding RNA: Functional Implications.
1050 *Noncoding RNA* **6**, 22 (2020).
1051 5. Ariel, F. *et al.* R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Molecular Cell*
1052 **77**, 1055-1065.e4 (2020).
1053 6. Kindgren, P., Ard, R., Ivanov, M. & Marquardt, S. Transcriptional read-through of the long non-
1054 coding RNA SVALKA governs plant cold acclimation. *Nat Commun* **9**, 4561 (2018).
1055 7. Montez, M. *et al.* Promoter-pervasive transcription causes RNA polymerase II pausing to boost
1056 DOG1 expression in response to salt. *EMBO J* **42**, e112443 (2023).
1057 8. Zhao, Y. *et al.* Natural temperature fluctuations promote COOLAIR regulation of FLC. *Genes Dev*
1058 **35**, 888–898 (2021).
1059 9. Swiezewski, S., Liu, F., Magusin, A. & Dean, C. Cold-induced silencing by long antisense transcripts
1060 of an Arabidopsis Polycomb target. *Nature* **462**, 799–802 (2009).
1061 10. Rosa, S., Duncan, S. & Dean, C. Mutually exclusive sense–antisense transcription at FLC facilitates
1062 environmentally induced gene repression. *Nat Commun* **7**, 13031 (2016).
1063 11. Yang, J. *et al.* Long noncoding RNA ARTA controls ABA response through MYB7 nuclear trafficking
1064 in Arabidopsis. *Developmental Cell* **58**, 1206-1217.e4 (2023).
1065 12. Yatusevich, R. *et al.* Antisense transcription represses Arabidopsis seed dormancy QTL DOG1 to
1066 regulate drought tolerance. *EMBO Rep.* **18**, 2186–2196 (2017).
1067 13. Wang, Y. *et al.* The noncoding RNA *HIDDEN TREASURE 1* promotes phytochrome B-dependent
1068 seed germination by repressing abscisic acid biosynthesis. *The Plant Cell* **35**, 700–716 (2023).
1069 14. Yang, M. *et al.* In vivo single-molecule analysis reveals COOLAIR RNA structural diversity. *Nature*
1070 **609**, 394–399 (2022).

1071 15. Yin, Y. *et al.* U1 snRNP regulates chromatin retention of noncoding RNAs. *Nature* **580**, 147–150
1072 (2020).
1073 16. Kondo, Y., Oubridge, C., van Roon, A.-M. M. & Nagai, K. Crystal structure of human U1 snRNP, a
1074 small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. *Elife*
1075 **4**, e04986 (2015).
1076 17. Goguel, V., Liao, X. L., Rymond, B. C. & Rosbash, M. U1 snRNP can influence 3'-splice site
1077 selection as well as 5'-splice site selection. *Genes Dev* **5**, 1430–1438 (1991).
1078 18. Almada, A. E., Wu, X., Kriz, A. J., Burge, C. B. & Sharp, P. A. Promoter directionality is controlled by
1079 U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363 (2013).
1080 19. Mimoso, C. A. & Adelman, K. U1 snRNP increases RNA Pol II elongation rate to enable synthesis
1081 of long genes. *Mol Cell* **83**, 1264-1279.e10 (2023).
1082 20. So, B. R. *et al.* A Complex of U1 snRNP with Cleavage and Polyadenylation Factors Controls
1083 Telescripting, Regulating mRNA Transcription in Human Cells. *Mol Cell* **76**, 590-599.e4 (2019).
1084 21. Mangilet, A. F. *et al.* The Arabidopsis U1 snRNP regulates mRNA 3'-end processing. *Nat. Plants* **10**,
1085 1514–1531 (2024).
1086 22. Carrillo-Barral, N., Rodríguez-Gacio, M. D. C. & Matilla, A. J. Delay of Germination-1 (DOG1): A
1087 Key to Understanding Seed Dormancy. *Plants (Basel)* **9**, 480 (2020).
1088 23. Auge, G. A. *et al.* Secondary dormancy dynamics depends on primary dormancy status in
1089 *Arabidopsis thaliana*. *Seed Sci. Res.* **25**, 230–246 (2015).
1090 24. Graeber, K. *et al. DELAY OF GERMINATION 1* mediates a conserved coat-dormancy mechanism for
1091 the temperature- and gibberellin-dependent control of seed germination. *Proc. Natl. Acad. Sci.*
1092 *U.S.A.* **111**, (2014).
1093 25. Footitt, S., Douterelo-Soler, I., Clay, H. & Finch-Savage, W. E. Dormancy cycling in *Arabidopsis*
1094 seeds is controlled by seasonally distinct hormone-signaling pathways. *Proc. Natl. Acad. Sci.*
1095 *U.S.A.* **108**, 20236–20241 (2011).

1096 26. Alonso-Blanco, C., Bentsink, L., Hanhart, C. J., Vries, H. B. & Koornneef, M. Analysis of Natural
1097 Allelic Variation at Seed Dormancy Loci of *Arabidopsis thaliana*. *Genetics* **164**, 711–729 (2003).
1098 27. Bentsink, L., Jowett, J., Hanhart, C. J. & Koornneef, M. Cloning of DOG1, a quantitative trait locus
1099 controlling seed dormancy in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A* **103**, 17042–17047 (2006).
1100 28. Fedak, H. *et al.* Control of seed dormancy in Arabidopsis by a cis-acting noncoding antisense
1101 transcript. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7846–E7855 (2016).
1102 29. Nakabayashi, K., Bartsch, M., Ding, J. & Soppe, W. J. J. Seed Dormancy in Arabidopsis Requires
1103 Self-Binding Ability of DOG1 Protein and the Presence of Multiple Isoforms Generated by
1104 Alternative Splicing. *PLoS Genet* **11**, e1005737 (2015).
1105 30. Cyrek, M. *et al.* Seed Dormancy in Arabidopsis Is Controlled by Alternative Polyadenylation of
1106 *DOG1*. *Plant Physiol.* **170**, 947–955 (2016).
1107 31. Krzyszton, M. *et al.* Dormancy heterogeneity among Arabidopsis thaliana seeds is linked to
1108 individual seed size. *Plant Communications* 100732 (2023) doi:10.1016/j.xplc.2023.100732.
1109 32. Kowalczyk, J. *et al.* Alternative Polyadenylation of the Sense Transcript Controls Antisense
1110 Transcription of DELAY OF GERMINATION 1 in Arabidopsis. *Molecular Plant* **10**, 1349–1352
1111 (2017).
1112 33. Kendall, S. L. *et al.* Induction of Dormancy in *Arabidopsis* Summer Annuals Requires Parallel
1113 Regulation of *DOG1* and Hormone Metabolism by Low Temperature and CBF Transcription
1114 Factors. *Plant Cell* **23**, 2568–2580 (2011).
1115 34. Rodríguez-Gacio, M. D. C., Matilla-Vázquez, M. A. & Matilla, A. J. Seed dormancy and ABA
1116 signaling: The breakthrough goes on. *Plant Signaling & Behavior* **4**, 1035–1048 (2009).
1117 35. Park, S.-Y. *et al.* Abscisic Acid Inhibits Type 2C Protein Phosphatases via the PYR/PYL Family of
1118 START Proteins. *Science* **324**, 1068–1071 (2009).
1119 36. Baek, W., Lim, C. W., Luan, S. & Lee, S. C. The RING finger E3 ligases PIR1 and PIR2 mediate PP2CA
1120 degradation to enhance abscisic acid response in Arabidopsis. *The Plant Journal* **100**, 473–486
1121 (2019).

1122 37. Née, G. *et al.* DELAY OF GERMINATION1 requires PP2C phosphatases of the ABA signalling
1123 pathway to control seed dormancy. *Nat Commun* **8**, 72 (2017).
1124 38. Nishimura, N. *et al.* Control of seed dormancy and germination by DOG1-AHG1 PP2C
1125 phosphatase complex via binding to heme. *Nat Commun* **9**, 2132 (2018).
1126 39. Zhang, S. *et al.* New insights into Arabidopsis transcriptome complexity revealed by direct
1127 sequencing of native RNAs. *Nucleic Acids Research* **48**, 7700–7711 (2020).
1128 40. Thomas, Q. A. *et al.* Transcript isoform sequencing reveals widespread promoter-proximal
1129 transcriptional termination in Arabidopsis. *Nat Commun* **11**, 2589 (2020).
1130 41. Thieffry, A. *et al.* Characterization of Arabidopsis thaliana Promoter Bidirectionality and Antisense
1131 RNAs by Inactivation of Nuclear RNA Decay Pathways. *Plant Cell* **32**, 1845–1867 (2020).
1132 42. Čermák, T. *et al.* A Multipurpose Toolkit to Enable Advanced Genome Engineering in Plants. *Plant*
1133 *Cell* **29**, 1196–1217 (2017).
1134 43. Jupe, F. *et al.* The complex architecture and epigenomic impact of plant T-DNA insertions. *PLoS*
1135 *Genet* **15**, e1007819 (2019).
1136 44. Xu, R. *et al.* The 73 kD Subunit of the cleavage and polyadenylation specificity factor (CPSF)
1137 complex affects reproductive development in Arabidopsis. *Plant Mol Biol* **61**, 799–815 (2006).
1138 45. Funck, D., Clauß, K., Frommer, W. B. & Hellmann, H. A. The Arabidopsis CstF64-like RSR1/ESP1
1139 protein participates in glucose signaling and flowering time control. *Front Plant Sci* **3**, (2012).
1140 46. Cortazar, M. A. *et al.* Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation
1141 Facilitates Termination by a “Sitting Duck Torpedo” Mechanism. *Molecular Cell* **76**, 896-908.e4
1142 (2019).
1143 47. Laugesen, A., Højfeldt, J. W. & Helin, K. Molecular Mechanisms Directing PRC2 Recruitment and
1144 H3K27 Methylation. *Mol Cell* **74**, 8–18 (2019).
1145 48. Blackledge, N. P. & Klose, R. J. The molecular principles of gene regulation by Polycomb
1146 repressive complexes. *Nat Rev Mol Cell Biol* **22**, 815–833 (2021).

1147 49. Krzyszton, M. *et al.* Single seeds exhibit transcriptional heterogeneity during secondary dormancy
1148 induction. *Plant Physiology* **190**, 211–225 (2022).
1149 50. Buijs, G. A Perspective on Secondary Seed Dormancy in Arabidopsis thaliana. *Plants* **9**, 749
1150 (2020).
1151 51. Ibarra, S. E. *et al.* Molecular mechanisms underlying the entrance in secondary dormancy of
1152 *Arabidopsis* seeds: Mechanisms underlying secondary dormancy. *Plant Cell Environ.* **39**, 213–221
1153 (2016).
1154 52. Footitt, S., Huang, Z., Clay, H. A., Mead, A. & Finch-Savage, W. E. Temperature, light and nitrate
1155 sensing coordinate A rabidopsis seed dormancy cycling, resulting in winter and summer annual
1156 phenotypes. *The Plant Journal* **74**, 1003–1015 (2013).
1157 53. Wrona, M. *et al.* The BRAHMA-associated SWI/SNF chromatin remodeling complex controls
1158 *Arabidopsis* seed quality and physiology. *Plant Physiology* **197**, kiae642 (2024).
1159 54. Wang, H. *et al.* H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. *Nature*
1160 **615**, 339–348 (2023).
1161 55. Chu, C., Quinn, J. & Chang, H. Y. Chromatin isolation by RNA purification (ChIRP). *J Vis Exp* 3912
1162 (2012) doi:10.3791/3912.
1163 56. Déjardin, J. & Kingston, R. E. Purification of Proteins Associated with Specific Genomic Loci. *Cell*
1164 **136**, 175–186 (2009).
1165 57. Sajeev, N., Koornneef, M. & Bentsink, L. A commitment for *life:* Decades of unraveling the
1166 molecular mechanisms behind seed dormancy and germination. *The Plant Cell* koad328 (2024)
1167 doi:10.1093/plcell/koad328.
1168 58. Chen, M. & Penfield, S. Feedback regulation of COOLAIR expression controls seed dormancy and
1169 flowering time. *Science* **360**, 1014–1017 (2018).
1170 59. Nielsen, M. *et al. COOLAIR* and PRC2 function in parallel to silence *FLC* during
1171 vernalization. *Proc. Natl. Acad. Sci. U.S.A.* **121**, (2024).

1172 60. Jin, Y., Ivanov, M., Dittrich, A. N., Nelson, A. D. & Marquardt, S. LNCRNA FLAIL affects alternative
1173 splicing and represses flowering in *Arabidopsis*. *The EMBO Journal* **42**, e110921 (2023).
1174 61. Née, G. *et al.* DELAY OF GERMINATION1 requires PP2C phosphatases of the ABA signalling
1175 pathway to control seed dormancy. *Nat Commun* **8**, 72 (2017).
1176 62. Krzyszton, M. *et al.* Single seeds exhibit transcriptional heterogeneity during secondary dormancy
1177 induction. *Plant Physiol.* **190**, 211–225 (2022).
1178 63. Nakabayashi, K. *et al.* The Time Required for Dormancy Release in *Arabidopsis* Is Determined by
1179 DELAY OF GERMINATION1 Protein Levels in Freshly Harvested Seeds. *Plant Cell* **24**, 2826–2838
1180 (2012).
1181 64. Thatcher, L. F. *et al.* The Arabidopsis RNA Polymerase II Carboxyl Terminal Domain (CTD)
1182 Phosphatase-Like1 (CPL1) is a biotic stress susceptibility gene. *Sci Rep* **8**, 13454 (2018).
1183 65. Guan, Q., Yue, X., Zeng, H. & Zhu, J. The Protein Phosphatase RCF2 and Its Interacting Partner
1184 NAC019 Are Critical for Heat Stress–Responsive Gene Regulation and Thermotolerance in
1185 *Arabidopsis*. *Plant Cell* **26**, 438–453 (2014).
1186 66. Qi, Y. *et al.* The over-expression of a chrysanthemum gene encoding an RNA polymerase II CTD
1187 phosphatase-like 1 enzyme enhances tolerance to heat stress. *Hortic Res* **5**, 37 (2018).
1188 67. Krzyszton *et al.* Single seeds exhibit transcriptional heterogeneity during secondary dormancy
1189 induction. *Accepted in Plant Physiology* (2022).
1190 68. Ali, F., Qanmber, G., Li, F. & Wang, Z. Updated role of ABA in seed maturation, dormancy, and
1191 germination. *J Adv Res.* **35**, 199–214 (2022).
1192 69. Kim, D.-H. & Sung, S. Vernalization-Triggered Intragenic Chromatin Loop Formation by Long
1193 Noncoding RNAs. *Developmental Cell* **40**, 302-312.e4 (2017).
1194 70. Guo, Y. *et al.* Long noncoding RNAs heat shock RNA omega nucleates TBPH and promotes
1195 intestinal stem cell differentiation upon heat shock. *iScience* **27**, 109732 (2024).

1196 71. Ji, Q., Zong, X., Mao, Y. & Qian, S.-B. A heat shock–responsive lncRNA *Heat* acts as a HSF1-
1197 directed transcriptional brake via m^6 A modification. *Proc. Natl. Acad. Sci. U.S.A.* **118**,
1198 e2102175118 (2021).
1199 72. Guo, C.-J. *et al.* Distinct Processing of lncRNAs Contributes to Non-conserved Functions in Stem
1200 Cells. *Cell* **181**, 621-636.e22 (2020).
1201 73. Bridges, M. C., Daulagala, A. C. & Kourtidis, A. LNCcation: lncRNA localization and function. *J Cell*
1202 *Biol* **220**, e202009045 (2021).
1203 74. Shukla, C. J. *et al.* High-throughput identification of RNA nuclear enrichment sequences. *EMBO J*
1204 **37**, e98452 (2018).
1205 75. Lubelsky, Y. & Ulitsky, I. Sequences enriched in Alu repeats drive nuclear localization of long RNAs
1206 in human cells. *Nature* **555**, 107–111 (2018).
1207 76. Grosch, M. *et al.* Nucleus size and DNA accessibility are linked to the regulation of paraspeckle
1208 formation in cellular differentiation. *BMC Biol* **18**, 42 (2020).
1209 77. Strehle, M. & Guttman, M. Xist drives spatial compartmentalization of DNA and protein to
1210 orchestrate initiation and maintenance of X inactivation. *Current Opinion in Cell Biology* **64**, 139–
1211 147 (2020).
1212 78. Zhao, X. *et al.* Global identification of Arabidopsis lncRNAs reveals the regulation of MAF4 by a
1213 natural antisense RNA. *Nat Commun* **9**, 5056 (2018).
1214 79. Csorba, T., Questa, J. I., Sun, Q. & Dean, C. Antisense COOLAIR mediates the coordinated
1215 switching of chromatin states at FLC during vernalization. *Proc. Natl. Acad. Sci. U.S.A.* **111**,
1216 16160–16165 (2014).
1217 80. Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J. & Dean, C. R-loop stabilization represses
1218 antisense transcription at the Arabidopsis FLC locus. *Science* **340**, 619–621 (2013).
1219 81. Xu, W. *et al.* The R-Loop Atlas of Arabidopsis Development and Responses to Environmental
1220 Stimuli. *Plant Cell* **32**, 888–903 (2020).

1221 82. Manavella, P. A. *et al.* Fast-Forward Genetics Identifies Plant CPL Phosphatases as Regulators of
1222 miRNA Processing Factor HYL1. *Cell* **151**, 859–870 (2012).
1223 83. Simpson, G. G., Dijkwel, P. P., Quesada, V., Henderson, I. & Dean, C. FY Is an RNA 3' End-
1224 Processing Factor that Interacts with FCA to Control the Arabidopsis Floral Transition. *Cell* **113**,
1225 777–787 (2003).
1226 84. Xing, D., Zhao, H., Xu, R. & Li, Q. Q. Arabidopsis PCFS4, a homologue of yeast polyadenylation
1227 factor Pcf11p, regulates FCA alternative processing and promotes flowering time. *The Plant*
1228 *Journal* **54**, 899–910 (2008).
1229 85. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–
1230 359 (2012).
1231 86. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
1232 87. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for
1233 exploring deep-sequencing data. *Nucleic Acids Research* **42**, W187–W191 (2014).
1234 88. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
1235 89. Smith, T., Heger, A. & Sudbery, I. UMI-tools: Modelling sequencing errors in Unique Molecular
1236 Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2016).
1237 90. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features.
1238 *Bioinformatics* **26**, 841–842 (2010).
1239 91. Warkocki, Z. *et al.* Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s. *Cell* **174**,
1240 1537-1548.e29 (2018).
1241 92. Alpern, D. *et al.* BRB-seq: ultra-affordable high-throughput transcriptomics enabled by bulk RNA
1242 barcoding and sequencing. *Genome Biol* **20**, 71 (2019).
1243 93. Sofroniew, N. *et al.* napari: a multi-dimensional image viewer for Python. Zenodo
1244 https://doi.org/10.5281/ZENODO.6597502 (2022).
1245 94. Bokota, G. *et al.* PartSeg: a tool for quantitative feature extraction from 3D microscopy images for
1246 dummies. *BMC Bioinformatics* **22**, 72 (2021).

1247