Research

New Phytologist

The CARBON CATABOLITE REPRESSION 4A-mediated RNA deadenylation pathway acts on the transposon RNAs that are not regulated by small RNAs

Ling Wang[1,2]*, Hui Li[1,2]*, Zhen Lei[1,2], Dong-Hoon Jeong[3,4] and Jungnam Cho[5]

[1]National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, 200032, China; [2]University of Chinese Academy of Science, Beijing, 100049, China; [3]Department of Life Science, Hallym University, Chuncheon, 24252, Korea; [4]Multidisciplinary Genome Institute, Hallym University, Chuncheon, 24252, Korea; [5]Department of Biosciences, Durham University, Durham, DH1 3LE, UK

Author for correspondence:
Jungnam Cho
Email: jungnam.cho@durham.ac.uk

Received: *24 August 2023*
Accepted: *13 November 2023*

New Phytologist (2024) **241:** 1636–1645
doi: 10.1111/nph.19435

Key words: *Arabidopsis thaliana*, long terminal repeat retrotransposon, mRNA poly(A) tail, RNA deadenylation, RNA decay.

Summary

- Transposable elements (TEs) are mobile genetic elements that can impair the host genome stability and integrity. It has been well documented that activated transposons in plants are suppressed by small interfering (si) RNAs. However, transposon repression by the cytoplasmic RNA surveillance system is unknown.
- Here, we show that mRNA deadenylation is critical for controlling transposons in *Arabidopsis*. Trimming of poly(A) tail is a rate-limiting step that precedes the RNA decay and is primarily mediated by the CARBON CATABOLITE REPRESSION 4 (CCR4)-NEGATIVE ON TATA-LESS (NOT) complex.
- We found that the loss of *CCR4a* leads to strong derepression and mobilization of TEs in *Arabidopsis*. Intriguingly, CCR4a regulates a largely distinct set of TEs from those controlled by RNA-dependent RNA Polymerase 6 (RDR6), a key enzyme that produces cytoplasmic siRNAs. This indicates that the cytoplasmic RNA quality control mechanism targets the TEs that are poorly recognized by the previously well-characterized RDR6-mediated pathway, and thereby augments the host genome stability.
- Our study suggests a hitherto unknown mechanism for transposon repression mediated by RNA deadenylation and unveils a complex nature of the host's strategy to maintain the genome integrity.

Introduction

Transposable elements (TEs) are mobile genetic elements that pose a significant threat to the host genome stability and integrity. It is well documented that transposons are subject to epigenetic silencing that is mediated by a so-called RNA-directed DNA methylation (RdDM; Matzke & Mosher, 2014). TEs that escape such transcriptional suppression or are newly introduced to the host genome, thus are not yet epigenetically silenced, are recognized by the RNA-DEPENDENT RNA POLYMERASE 6 (RDR6)-SUPPRESSOR OF GENE SILENCING 3 (SGS3) complex, which generates 21/22-nucleotide (nt) small interfering (si) RNAs and initiates epigenetic silencing (Nuthikattu *et al.*, 2013; Creasey *et al.*, 2014; Panda *et al.*, 2016; Lee *et al.*, 2020). Accumulating evidence suggests that the incompleteness of mRNA (e.g. tail-less or truncated mRNA) is crucial for specific targeting of RDR6 (Luo & Chen, 2007; Creasey *et al.*, 2014; Baeg *et al.*, 2017). In our previous study, we showed that the suboptimal codon usage of transposons causes ribosome stalling and RNA cleavage, which accounts for their frequent targeting to the RDR6-mediated siRNA biogenesis pathway (Kim *et al.*, 2021). In addition, the ribosome-stalled transcripts are preferentially guided to cytoplasmic compartments where SGS3 and RDR6 are localized (Kim *et al.*, 2021; Han *et al.*, 2023; Tan *et al.*, 2023). It is also worth noting that the 21/22-nt siRNAs are associated with only around one-third of active and transcribed TEs in *Arabidopsis*, and the posttranscriptional suppression of transposon that is independent of siRNAs is largely unknown.

Aberrancy of mRNA caused by premature translation termination and ribosome arrest is monitored and resolved by the RNA surveillance pathways (Kervestin & Jacobson, 2012; Lykke-Andersen & Jensen, 2015; Zhang & Guo, 2017; Kurosaki *et al.*, 2019; D'Orazio & Green, 2021). mRNA deadenylation is a primary and rate-limiting step of RNA decay and is catalyzed by multiple deadenylase complexes (Passmore & Coller, 2022): the POLY(A) NUCLEASE 2 (PAN2)-PAN3 complex acts at an earlier phase of deadenylation in metazoa and yeast, and degrades the poly(A) tail to 50–110 nt (Chen & Shyu, 2011; Jonas *et al.*, 2014; Schäfer *et al.*, 2014, 2019; Tang *et al.*, 2019; Passmore & Coller, 2022). However, orthologs of PAN2-PAN3 have

*These authors contributed equally to this work.

not been identified in flowering plants (Pavlopoulou *et al.*, 2013; Chantarachot & Bailey-Serres, 2018). The CARBON CATABOLITE REPRESSION 4 (CCR4)-NEGATIVE ON TATALESS (NOT) complex catalyzes more rapid deadenylation by two catalytic components, CCR4 and CCR4-ASSOCIATED FACTOR 1 (CAF1; Chen & Shyu, 2011; Passmore & Coller, 2022). In *Arabidopsis*, CCR4a, CCR4b, CAF1a, and CAF1b exhibit the catalytic activity of deadenylation while showing some level of target specificity (Liang *et al.*, 2009; Walley *et al.*, 2010; Suzuki *et al.*, 2015; Arae *et al.*, 2019). A third class of deadenylase is POLY(A)-SPECIFIC RIBONUCLEASE (PARN), the orthologs of which have been identified in vertebrates and plants (Pavlopoulou *et al.*, 2013; Passmore & Coller, 2022). Loss of *PARN* genes in *Arabidopsis* causes embryo lethality (Chiba *et al.*, 2004; Nishimura *et al.*, 2005). Overall, a diverse array of mRNA deadenylases has evolved in different species and controls gene expression in rather sequence-specific manner.

Despite the inherent aberrancy of transposon transcripts, their regulation by the cellular RNA surveillance system has been poorly reported. In *Drosophila*, for instance, it was suggested that the mutants defective in *CCR4* accumulate TE transcripts in the chromatin-associated RNA fraction, and the CCR4-NOT complex interacts with the piRNA pathway components in the nucleus, indicating a co-transcriptional suppression of transposons (Kordyukova *et al.*, 2020). In addition, a CCR4-NOT complex component NOT1 was identified in a genetic screen for RdDM regulators in *Arabidopsis* (Zhou *et al.*, 2020). However, it is important to note that it is still uncertain whether TE suppression by the CCR4-NOT complex requires the catalytic activity of RNA deadenylation.

In this study, we investigated the mutants for RNA deadenylases in *Arabidopsis* and assessed the transposon RNA levels. Intriguingly, we found that RNA deadenylases suppress a set of transposons that are not usually regulated by the RDR6-mediated pathway. Oxford Nanopore direct RNA sequencing (ONT-DRS) revealed that CCR4a shortens the poly(A) tails, destabilizes the transcripts, and reduces the steady-state mRNA levels of transposons. Moreover, we also carried out whole-genome resequencing and droplet digital PCR (ddPCR) experiments to interrogate the mobilization of TEs and observed an increased mobility of transposons in the deadenylase mutants. Our study unveils a previously unknown cellular mechanism that degrades transposon RNAs through an evolutionarily conserved RNA surveillance system.

Materials and Methods

Plant materials and growth condition

All *Arabidopsis* (*Arabidopsis thaliana* L. Heynh) plants used in this study are in the Col-0 background. The *ccr4a-1* (SAIL_802_A10), *ccr4b-1* (SALK_151541C), *caf1a-1* (SALK_070336), and *caf1b-3* (SALK_044043) mutants were obtained from the Arabidopsis Biological Resources Center (https://abrc.osu.edu/). The *caf1a-1 caf1b-3* double mutant was identified from the F2 segregation population derived from crosses. *De novo ddm1* mutants were generated by the CRISPR-Cas9 system containing three sgRNAs. The sgRNA sequences were designed by an online web tool (https://chopchop.cbu.uib.no/), and the sgRNA secondary structure was predicted in the UNAFold web server (www.unafold.org/mfold/). The synthesized oligonucleotides were annealed and inserted into the digested entry vectors pENTR_L4_R1, pENTR_L1_L2, and pENTR_R2_L3 at the BbsI (NEB) sites. The entry vectors were subsequently transferred to a destination vector pFG7m34GW (Shimada *et al.*, 2010), carrying the FastGreen fluorescent seed selection marker and *proUBQ10*-driven Cas9 cassettes, using the Gateway LR reaction (Thermo Fisher Scientific, Waltham, MA, USA). *Agrobacterium tumefaciens*-mediated floral dip method was used to transform the *DDM1*-targeting pFG7m34GW vector into the indicated mutant background. Editing events were confirmed by Sanger sequencing. We were unable to identify any mutations at the sgRNA3-targeted regions, and the editing events at the regions targeted by sgRNA1 and sgRNA2 are summarized in Supporting Information Fig. S1. T-DNA was segregated out at T3 generation, and unless otherwise stated, plants at T4 generation were used in this study. Sequences of sgRNAs are listed in Table S1.

Seeds were sterilized using 75% ethanol, sown on Murashige and Skoog (MS) media (0.43 $g\,l^{-1}$ MS salts (pH 5.8), 3 $g\,l^{-1}$ sucrose, 0.8% (w/v) phytoagar) and stratified at 4°C for 3 d. Plants were grown at 22°C under long-day condition (16 h : 8 h, light : dark).

RT-qPCR

Total RNA was isolated from 10-d-old seedlings using the TRIzol extraction method (Tiangen, Beijing, China) and reverse-transcribed using ReverTra Ace qPCR RT Master Mix with gDNA Remover (Toyobo, Osaka, Japan). To quantify the relative abundance of transcripts, quantitative PCR was carried out using a ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad). *Actin2* (AT3G18780) was used as an internal control for normalization. Gene expression levels were determined by the $\Delta\Delta C_t$ method. Sequences of primers are listed in Table S1.

RNA-Seq

For RNA-Seq library construction, total RNA was isolated from 10-d-old seedlings using the TRIzol Reagent (Invitrogen), and poly(A)-RNA was purified from 3 μg of total RNA using poly-T oligo-attached magnetic beads. Library was prepared using the NEBNext Ultra Directional RNA Library Prep Kit (NEB) following the manufacturer's instructions. Sequencing was performed on an Illumina NovaSeq 6000 platform, and 150-bp paired-end (PE150) reads were generated. RNA-Seq dataset is summarized in Table S2.

For RNA-Seq data analysis, the raw sequences were trimmed by TRIMMOMATIC (v.0.39; Bolger *et al.*, 2014) to remove reads containing adapter and low-quality sequences with the parameters: LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50. Trimmed reads were then aligned to *Arabidopsis*

reference genome (TAIR10) using HISAT2 (v.2.2.1; Kim *et al.*, 2015) with settings: --rna-strandness RF --fr. Read count and FPKM values of genes and TEs were calculated by STRINGTIE (v.2.1.7; Pertea *et al.*, 2015). The R package DESEQ2 (Love *et al.*, 2014) was used for the differential expression analysis.

Oxford Nanopore direct RNA sequencing (ONT-DRS)

Total RNA was isolated from 10-d-old seedlings by TRIzol (Qiagen), and poly(A)-RNA was purified using Dynabeads mRNA Purification Kit (Invitrogen) following the manufacturer's instructions. The quality and quantity of mRNA were assessed using the NanoDrop 2000 spectrophotometer and Qubit. Library was prepared using direct RNA sequencing kit (SQK-RNA002; Oxford Nanopore Technologies, Oxord, UK), loaded onto an R9.4 Flow Cell (Flow cell type FLO-MIN106), and sequenced on a GridION device for 72 h. ONT-DRS dataset is summarized in Table S3.

The raw nanopore signals were converted to base sequences by GUPPY (v.6.1.5) using the high-accuracy basecalling model. Since transposons are not properly annotated in the reference assembly of *Arabidopsis* and therefore often omitted in the downstream analysis, we generated a custom transcript assembly by merging the reference transcript assembly and all the *de novo* assembled transcripts derived from the RNA-Seq data generated in this study using STRINGTIE (v.2.1.7; Pertea *et al.*, 2015). Then, the nanopore reads with a mean quality score > 7 were mapped to the custom transcriptome using MINIMAP2 (v.2.24-r1122; Li, 2018) with the following parameters: -ax map-ont -L -p 0 -N 10. Poly(A) tail length was detected by NANOPOLISH (v.0.13.3; Workman *et al.*, 2019). Transcripts with > 15 reads were used to obtain the median poly(A) tail length. The reads with poly(A) tail were re-aligned to TAIR10 genome with the following parameters: -ax splice -k14 -uf and visualized by the python genome package BUGV (Jia *et al.*, 2022).

mRNA half-life

Four-day-old etiolated *Arabidopsis* seedlings were immersed in cordycepin solution (1 mM PIPES (pH 6.25), 15 mM sucrose, 1 mM KCl, 1 mM sodium citrate, and 1 μM cordycepin) and harvested at 0, 0.25, 0.5, 1, 2, and 4 h for three biological replicates. RNA extraction and RNA-Seq were performed as described previously. mRNA half-lives were calculated as follows: decay rate $K_i = -\log_e (F_i/F_0)/T_i$, in which F_i is the FPKM at time i, and T_i is the time of cordycepin treatment. K_i was calculated from each time point, and the half-life is $\log_e(2)/K_a$, in which K_a is the average decay rate measured for all time points.

Whole-genome resequencing

Genomic DNA was extracted using the CTAB method. One microgram genomic DNA was randomly fragmented by ultrasonicator (Covaris, Woburn, MA, USA). An average size of 200–400-bp DNA fragments was selected by Agencourt AMPure XP-Medium kit. The fragments were then end-repaired, 3′ adenylated and ligated with adaptors. The purified double-stranded products were heat denatured to single-stranded DNA and then circularized. The single-stranded circular DNA was sequenced by a DNBSEQ-T7 generating 150-bp paired-end reads. Whole-genome resequencing dataset is summarized in Table S4.

Paired-end short-read whole-genome sequencing data were mapped to TAIR10 and processed following the SPLITREADER pipeline (Baduel *et al.*, 2021). Briefly, discordantly mapping and nonmapping reads were recovered, and then, the reads were remapped to the TE pools and the genome. Insertions supported by at least three reads (DP filter = 3) were filtered and only nonreference insertions were considered.

Droplet digital PCR

Droplet digital PCR was performed on TargetingOne® Digital PCR System (TargetingOne, Beijing, China) following the manufacturer's instruction. Briefly, Genomic DNA was extracted using a N96 DNAsecure Plant Kit (Tiangen). One hundred nanogram of genomic DNA was digested using AluI (NEB) for 4 h at 37°C. The digested DNA was quantified using the Qubit4 DNA quantification system (Thermo Fisher Scientific) and diluted to 0.15 ng μl^{-1}. The reaction mixture containing 2× ddPCR Supermix (Bio-Rad), 0.8 μM primer, 0.25 μM probe, and 0.6 ng of cleaved sample DNA was thoroughly mixed and added into the droplet generation chip. Then, 180 μl of droplet generation oil was added to the mixture in the reaction mix inlet. Subsequently, the generated droplets were transferred into an 8-strip PCR tube and used for PCR reaction that was performed on a PTC-200 Thermal Cycler. FAM (488 nm) and VIC (532 nm) fluorescence signals were detected through the separate channels on the Chip Reader. Finally, the data were subjected to Poisson distribution analysis using the Chip Reader R1 software to obtain the target DNA copy numbers. Sequences of primers and probes are listed in Table S1.

Results

mRNA deadenylases suppress transposons

We previously showed that *Arabidopsis* TE RNAs often undergo ribosome stalling and RNA cleavage, which are required for the RDR6-mediated siRNA biogenesis (Kim *et al.*, 2021). However, a substantial fraction of TEs with signatures of ribosome stalling and RNA cleavage is not associated with siRNAs. Since such aberrancy of RNA is monitored and resolved by RNA surveillance and decay pathways (Harigaya & Parker, 2010; Chen & Shyu, 2011; Graille & Séraphin, 2012; Shoemaker & Green, 2012; D'Orazio & Green, 2021), we reasoned that transposon RNAs might be controlled by the RNA degradation pathways. To test this possibility, we first identified the *Arabidopsis* mutants for RNA deadenylases (*ccr4a-1* and *ccr4b-1* single mutants, and *caf1a-1 caf1b-3* double mutant) and induced *de novo* mutations in *DDM1* using CRISPR-Cas9 to release transposons from epigenetic silencing (Fig. S1). It is worth noting that pre-existing *ddm1* mutants contain many newly inserted transposons that could be

unevenly segregated in genetic crosses with other mutants and therefore may lead to erroneous assessment of transposon expression. For this reason, we generated *de novo* mutants of *DDM1* and used the plant materials collected in the same generation (See Materials and Methods section). RNA-Seq was then carried out in two independent *ddm1* mutant alleles of each RNA deadenylase mutant (Fig. S2). Our transcriptome analysis identified hundreds of genes that are up- or downregulated; however, transposons exhibited a strikingly different pattern that most differentially expressed transposons are upregulated in the deadenylase mutants (Fig. 1a–c). We then compared the upregulated transposons in these double and triple mutants and found that a large fraction of TEs is commonly upregulated, while *CCR4a* displays the greatest impact on transposon RNA levels (Figs 1d,e, S3). These data imply that the mRNA deadenylation pathway is involved in transposon repression.

Differential TE control by RDR6 and CCR4a

It is well documented that some transposons give rise to 21/22-nt siRNAs that can target transposon RNAs for cleavage (Nuthikattu *et al.*, 2013; Creasey *et al.*, 2014). This specific class of siRNAs is also known as epigenetically activated siRNAs (easiRNAs) and is generated by the RDR6-DICER LIKE 2 and 4 (DCL2/4) module (Gasciolli *et al.*, 2005; Nuthikattu *et al.*, 2013). Since the cleaved transcript products are eliminated by the RNA decay pathways, we suspected that the observed derepression of transposons in the deadenylase mutants might be merely a consequence of compromised RNA decay of the easiRNA-cleaved transcripts. However, the transposons regulated by CCR4a marginally overlapped with those targeted by RDR6 (Fig. 2a), suggesting that the reactivation of these TEs is more strongly associated with the loss of *CCR4a* than *RDR6*. Transposon classification analysis further supports this conclusion; the *RDR6*-regulated transposons are strongly enriched

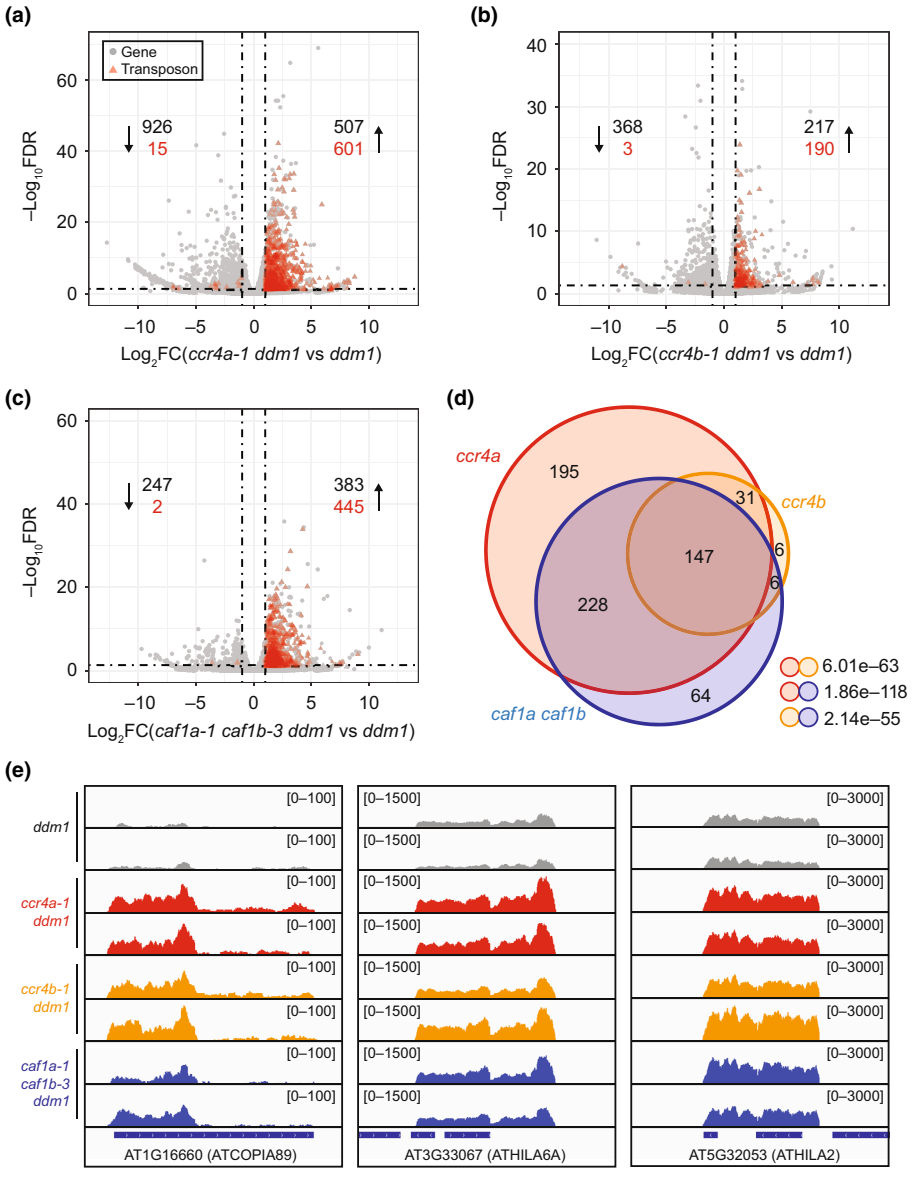

Fig. 1 Loss of mRNA deadenylases leads to transposon derepression. (a–c) Volcano plots shown for *ccr4a-1 ddm1* (a), *ccr4b-1 ddm1* (b), and *caf1a-1 caf1b-3 ddm1* (c) in comparison with the *ddm1* single mutant of *Arabidopsis thaliana*. Mutations for *DDM1* were generated *de novo* by CRISPR-Cas9 and two independent *ddm1* mutant lines were used. Differential expression was defined by the $\log_2$-fold change > 1 or < −1 and FDR values < 0.05. Grey dots and red triangles represent genes and transposons, respectively. Numbers indicate differentially expressed genes and transposons, and up- or downregulation was expressed by arrows. (d) Overlap of transposons upregulated by the mutations of *CCR4a*, *CCR4b*, and both *CAF1a* and *CAF1b*. (e) Genome browser snapshots for representative transposon loci showing the increased RNA levels in the mRNA deadenylase mutants. Numbers in parentheses indicate read coverage and two independent *ddm1* mutant lines are displayed in separate tracks.

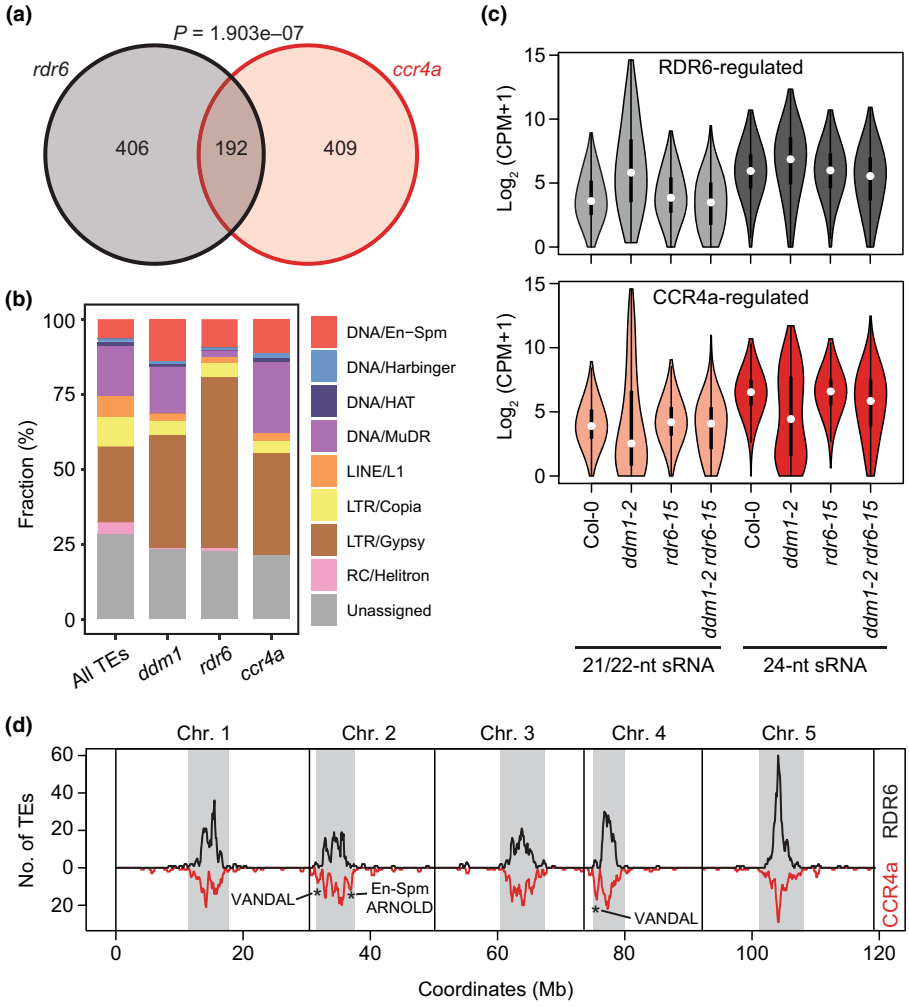

Fig. 2 CCR4a regulates distinct set of transposons from those controlled by RNA Polymerase 6 (RDR6). (a) Overlap of transposons regulated by RDR6 and CCR4a. RDR6-regulated transposons were retrieved from the previous study (Kim *et al*., 2021) and identified by the reduced 21/22-nt siRNA levels in the *rdr6 ddm1* double mutant as compared to the *ddm1* single mutant of *Arabidopsis thaliana*. CCR4a-regulated transposons are those that are upregulated in *ccr4a-1 ddm1* compared with *ddm1* with the $\log_2$-fold change > 1 and FDR values < 0.05. (b) Fraction of transposon families in all annotated transposons, derepressed in *ddm1*, regulated by RDR6, and upregulated in the *ccr4a* mutant. Reactivated transposons in the *ddm1* mutant were identified from a public dataset (GSE52952) by filtering those with the $\log_2$-fold change > 1 and FDR < 0.05. (c) Levels of 21/22- and 24-nt sRNAs in the transposons regulated by RDR6 and CCR4a. White circle, median; black rectangle, upper and lower quartile. The sequencing datasets were obtained from GSE52952. (d) Chromosomal distribution of RDR6- and CCR4a-regulated transposons. Numbers of transposable elements (TEs) in 500-kb overlapping windows sliding in steps of 100 kb are shown. Regions overrepresented in the CCR4a-regulated transposons are marked by asterisks, and representative transposon families corresponding to the region are also indicated. Pericentromeric regions are expressed as grey boxes.

with the LTR/Gypsy family (hypergeometric test, $P = 2.57e{-}36$), and in the *ccr4a* mutants, DNA/MuDR DNA transposon family is strongly overrepresented (hypergeometric test, $P = 5.34e{-}12$; Fig. 2b). To further confirm the divergence of the RDR6- and CCR4a-regulated transposons, we compared the 21/22- and 24-nt siRNA levels. As shown in Fig. 2(c), the 21/22-nt siRNAs of RDR6-controlled TEs were greatly increased in *ddm1*, whereas CCR4a-regulated transposons exhibited a significant reduction in both classes of siRNAs in *ddm1*. In addition, the transposons regulated by RDR6 and CCR4a were mapped across the *Arabidopsis* chromosomes. The RDR6 target transposons were mostly found in the centromeric region, and the transposons regulated by CCR4a were also mapped to the pericentromeric and euchromatic regions in addition to centromeres (Figs 2d, S4). Collectively, loss of mRNA deadenylases is associated with increased RNA levels of TEs that are largely independent to easiRNA and RDR6 control.

CCR4a shortens poly(A) length and destabilizes TE RNAs

We next wanted to assess the poly(A) tail lengths of TE RNAs in the deadenylase mutant. For this, we took advantage of ONT-DRS, which allows for the tail length measurement of native RNA. Transposon transcripts identified by ONT-DRS reproducibly showed a strongly increased levels in the *ccr4a* mutant (Fig. S5), verifying our observation shown in Fig. 1. The ONT-DRS data from *ddm1-L2* revealed that the tail lengths peak at 20, 40–50, and 70–80 nt, which are distanced by *c*. 25 nt (Fig. 3a,b). A similar pattern was also observed in previous studies (Parker *et al*., 2020; Jia *et al*., 2022), suggesting a robust estimation of poly(A) tail length by ONT-DRS. Importantly, the *ccr4a-1 ddm1-L2* mutant displayed a longer tail length distribution compared with *ddm1-L2* in both genes and transposons (Fig. 3a,b), confirming that CCR4a is a key cellular factor shortening the poly(A) tail. We then retrieved the transposon transcripts from our ONT-DRS dataset and analyzed their tail lengths. As shown in Fig. 3(c,d), TE RNAs possess longer poly(A) tails compared with non-TE transcripts, which is consistent with a previous study (Li *et al*., 2021), and the loss of *CCR4a* led to a lengthening of their mRNA tails as did in genic transcripts.

It has been previously reported that highly expressed and stable genes are featured with short steady-state poly(A) tail length (Jia *et al*., 2022; Passmore & Coller, 2022). Nonetheless, lengthening of poly(A) tail contributes to active translation and RNA stability in humans and plants (Suzuki *et al*., 2015; Eisen *et al*., 2020). For example, the poly(A) tail length of a CACTA-like transposon

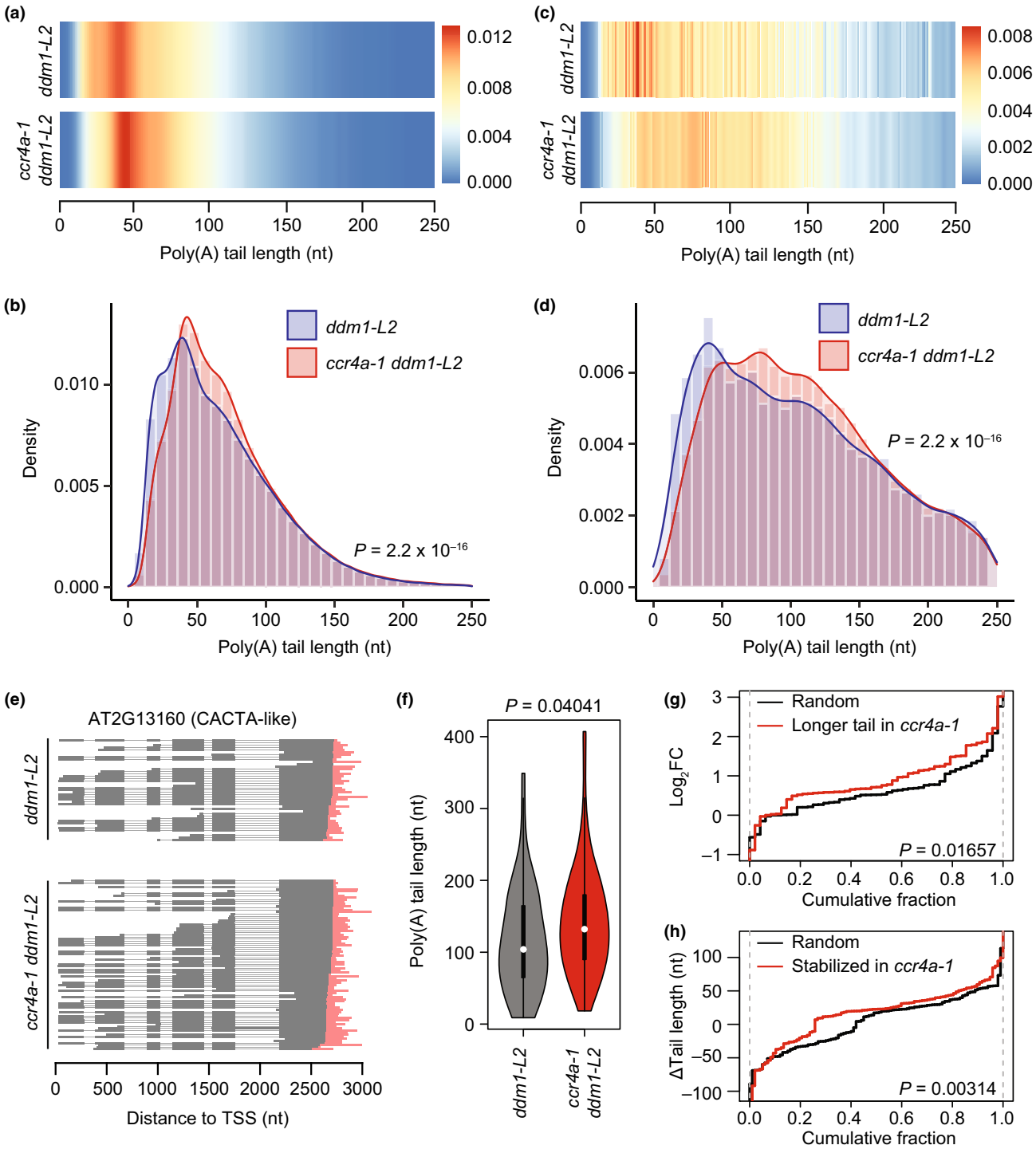

Fig. 3 Longer transposon RNA tails are associated with increased expression. (a–d) mRNA tail lengths of all transcripts (a, b) and transposon RNAs (c, d) in *ddm1-L2* and *ccr4a-1 ddm1-L2* of *Arabidopsis thaliana*, shown as heatmap (a, c; color key shows density of distribution) and density plot (b, d). Poly(A) length was measured by Oxford Nanopore direct RNA sequencing. *P*-value was obtained by the one-sided Wilcoxon rank sum test. (e, f) A CACTA-like transposable element (TE) exhibiting higher RNA levels (e) and longer tail length (f) in *ccr4a-1 ddm1-L2* double mutant. In (e), each line represents individual transcript detected by Oxford Nanopore direct RNA sequencing, and poly(A) tail is shown in red line. In (f), mRNA tail lengths of the CACTA-like element in *ddm1-L2* and *ccr4a-1 ddm1-L2* are compared. White circle, median; black rectangle, upper and lower quartile. *P*-value was obtained by the one-sided Wilcoxon rank sum test. (g) Fold changes of CCR4a-regulated transposons ($n = 48$, $\log_2$-fold change in *ddm1-L2* vs *ccr4a-1 ddm1-L2*) compared with randomly chosen transposons ($n = 48$). CCR4a-regulated transposons are those with longer tails by at least 10 nt in the *ccr4a-1 ddm1-L2* mutant. *P*-value was obtained by the one-sided Wilcoxon rank sum test. (h) Poly(A) tail length difference in transcripts stabilized by the loss of *CCR4a* in *ddm1-L2* and *ccr4a-1 ddm1-L2*. mRNA half-lives were determined for *ddm1-L2* and *ccr4a-1 ddm1-L2* by the transcription arrest assay followed by RNA-Seq. Transcripts with longer half-lives in *ccr4a-1 ddm1-L2* by at least 0.5 h were selected ($n = 97$) and compared against randomly chosen transcripts ($n = 100$). Tail length difference was calculated by subtracting the tail lengths in *ddm1-L2* from those in *ccr4a-1 ddm1-L2*. *P*-value was obtained by the one-sided Wilcoxon rank sum test.

became longer, and its RNA level was increased in the *ccr4a-1 ddm1-L2* mutant compared with *ddm1-L2* (Fig. 3e,f). We further tested other TE transcripts that have longer tails in the *ccr4a-1 ddm1-L2* mutant for their levels and found that almost 90% of these TEs are increased in transcript levels in *ccr4a-1 ddm1-L2* (Fig. 3g). Moreover, a transcription arrest RNA-Seq was carried out to determine the RNA stability in *ddm1-L2* and *ccr4a-1 ddm1-L2*. For this, seedlings were treated with cordycepin, a transcription elongation inhibitor, and then serially harvested at different time points for RNA-Seq, and mRNA half-lives were determined (See Materials and Methods section for details). Genes that became more stabilized in *ccr4a-1 ddm1-L2* as compared to *ddm1-L2* (increased half-lives by at least 0.5 h) exhibited longer tail lengths when *CCR4a* is mutated (Fig. 3h). These data together suggest that CCR4a shortens the poly(A) tail and destabilizes transposon RNAs.

Given that shortening of poly(A) tail is often coupled to translation repression (Tang *et al.*, 2019; Passmore & Coller, 2022) and weak translation leads to RNA localization to cytoplasmic RNA granules, which contain RNA deadenylases and degrading enzymes (Wheeler *et al.*, 2017; Chantarachot & Bailey-Serres, 2018; Arae *et al.*, 2019; Kim *et al.*, 2021), the TE transcripts controlled by RNA deadenylases might be more strongly enriched in RNA granules and actively degraded. Indeed, in the comparison between the RDR6- and CCR4a-regulated transposons, we were able to observe that the transcripts controlled by CCR4a exhibit weaker translational activity and stronger RNA granule enrichment (Fig. S6). Overall, TE transcripts regulated by CCR4a might be guided to the RNA turnover pathway in specific cytoplasmic compartments, further differentiating them from the RDR6-regulated transposons.

CCR4a suppresses transposon mobilization

We have so far demonstrated that loss of RNA deadenylases results in the increased RNA levels of transposons. This led us to test whether transposons mobilize more strongly in the deadenylase mutants. To test this idea, we carried out a whole-genome resequencing experiment to interrogate transposon proliferation. Ten individual plants from each genotype (*ddm1*, *ccr4a-1 ddm1*, *ccr4b-1 ddm1*, and *caf1a-1 caf1b-3 ddm1*) were randomly chosen and analyzed for nonreference and neo-insertions of transposons using the SPLITREADER pipeline (Baduel *et al.*, 2021). Intriguingly, all three RNA deadenylase mutants showed an increased number of transposon insertions compared with the *ddm1* single mutant (Fig. 4a). New insertions were observed for TEs that were previously shown for mobility in natural *Arabidopsis* population (Quadrana *et al.*, 2016) and *ddm1* mutant (Tsukahara *et al.*, 2009). We also found that the transposons that exhibited mobilization are of largely different types from those transcriptionally activated in the RNA deadenylase mutants; for instance, the LTR/Gypsy type was among the most actively induced in *ddm1*, but it was the LTR/Copia family that was most proliferative in the *ddm1* mutant (Fig. 4b). This indicates that an additional layer of regulation exists to control transposon mobilization. Moreover, a

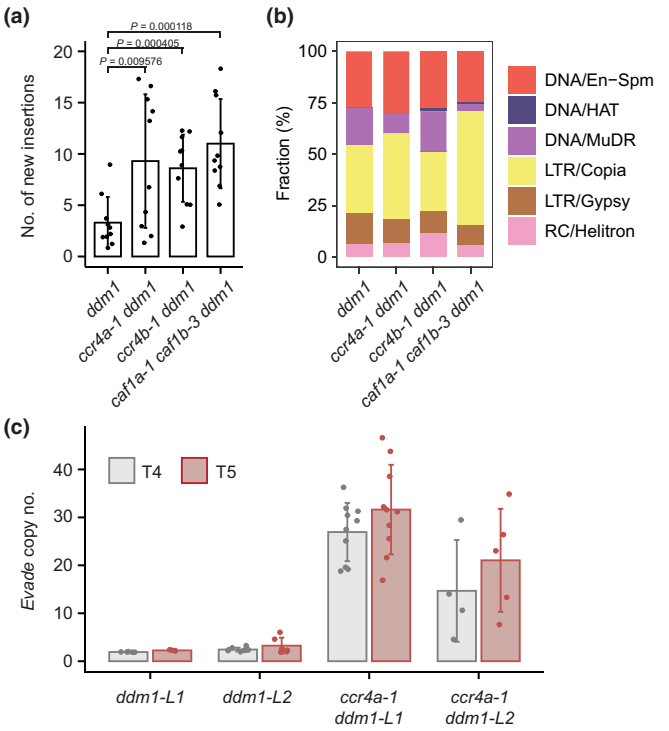

Fig. 4 CCR4a suppresses transposon mobilization. (a) Number of new insertions of transposable elements (TEs) detected in *ddm1*, *ccr4a-1 ddm1*, *ccr4b-1 ddm1*, and *caf1a-1 caf1b-3 ddm1* of *Arabidopsis thaliana*. Ten individual plants were randomly chosen, and whole-genome resequencing was performed for each individual plant independently. Data are presented in mean ± SD. *P*-value was obtained by the one-sided Student's *t*-test. (b) Percentage of TE families that were detected for neo-insertions. (c) Droplet digital PCR experiment determining the copy number of *Evade* retroelement in *ddm1* and *ccr4a-1 ddm1*. Plants were randomly chosen from a pool of selfed plants of an identical genotype and extracted for DNA individually. The experiment was performed at T4 and T5 generations of the *DDM1*-targeting CRISPR-Cas9 transformation. Data are presented in mean ± SD.

ddPCR experiment was carried out to validate the transposition of a representative LTR/Copia element in the *Arabidopsis* genome known as *Evade*. ddPCR is an experimental method that can quantitatively measure DNA copies and is particularly useful for assessing transposon copy number (Fan & Cho, 2021; Fan *et al.*, 2022). In this experiment, we used *ddm1* and *ccr4a-1 ddm1* at T4 and T5 generations. As shown in Fig. 4(c), the copy number of *Evade* was greatly increased in *ccr4a-1 ddm1* as compared with *ddm1*, further supporting our conclusion that RNA deadenylation represses transposition. In short, the RNA deadenylases destabilize transposon RNAs and inhibit the subsequent step of mobilization in *Arabidopsis*.

Discussion

In this study, we showed that RNA deadenylation is a critical cellular mechanism that augments the host's general suppression of TEs distinct from and in addition to the easiRNA-mediated pathway. This suggests a previously unknown complexity of transposon control, which complementarily contributes to the

tight repression of transposons that are of various types and sequence features. Such divergent TE suppression mechanisms possibly involve RNA localization to different cytoplasmic compartments (Fig. S6), in which the key regulators (RDR6 and CCR4a) are localized.

The easiRNA-mediated TE repression of the host can be seen as an efficient and persistent way of controlling transposons because it can switch on the epigenetic silencing that can be maintained through cell divisions and generations and target other TE transcripts with similar sequences (Nuthikattu *et al.*, 2013; Creasey *et al.*, 2014). On the contrary, RNA decay is merely degeneration of transcripts and does not generate any signals or biomolecules that can be amplified and transmitted. This partly accounts for why the RDR6-mediated pathway primarily acts on young transposons, particularly those that are structurally intact and long (Panda *et al.*, 2016), while non-TE transcripts are predominantly controlled by RNA decay (Gazzani *et al.*, 2004; Thran *et al.*, 2012; Branscheid *et al.*, 2015; De Alba *et al.*, 2015). In this regard, the CCR4a-targeted TEs might be older in age compared with those regulated by RDR6 and have likely undergone more evolutionary sequence degeneration, which makes them less harmful to the host genome and thus less demanded for the easiRNA-mediated epigenetic silencing. Further investigations into the sequence and structural features determining the target specificity of RDR6 and CCR4a will be a compelling follow-up study.

Our work mainly focused on the cytoplasmic role of RNA deadenylases; however, a previous report suggested a nuclear role of the CCR4-NOT complex in *Arabidopsis* as one of the essential elements for RdDM (Zhou *et al.*, 2020). This is reminiscent of what is known in *Drosophila* that CCR4 co-transcriptionally represses TEs in association with Piwi (Kordyukova *et al.*, 2020). These together suggest that the RNA deadenylase complex controls transposons in the nucleus; however, it has not yet been elucidated if the nuclear function of the CCR4-NOT complex requires the RNA deadenylation activity. In this study, we directly demonstrated using ONT-DRS that CCR4a shortens the poly(A) tail lengths of active TEs in *Arabidopsis*.

In summary, the shortening of TE RNA poly(A) tail length by RNA deadenylases and thereby RNA destabilization is a critical cytoplasmic mechanism suppressing transposon activity. This work unveils a hidden complexity of transposon regulation, which helps broaden our understanding of the host's defense against endogenous parasitic DNA.

Acknowledgements

We thank Dr Hongxia Wang and Dr Jun Yang for their technical assistance with Oxford Nanopore sequencing. This work was supported by the National Natural Science Foundation of China (32111540256 and 32270569) and the General Program of the Natural Science Foundation of Shanghai (22ZR1469100).

Competing interests

None declared.

Author contributions

JC conceived the idea and designed the experiments. HL and ZL conducted the experiments. LW, HL, D-HJ and JC analyzed the data. LW and JC drafted and wrote the manuscript. JC revised the manuscript. LW and HL contributed equally to this work.

ORCID

Jungnam Cho https://orcid.org/0000-0002-4078-7763

Data availability

All the sequencing data generated in this study is available in the SRA repository under the accession ID PRJNA940263.

References

Arae T, Morita K, Imahori R, Suzuki Y, Yasuda S, Sato T, Yamaguchi J, Chiba Y. 2019. Identification of Arabidopsis CCR4-NOT complexes with Pumilio RNA-binding proteins, APUM5 and APUM2. *Plant & Cell Physiology* **60**: 2015–2025.

Baduel P, Quadrana L, Colot V. 2021. Efficient detection of transposable element insertion polymorphisms between genomes using short-read sequencing data. *Methods in Molecular Biology* **2250**: 157–169.

Baeg K, Iwakawa H, Tomari Y. 2017. The poly(A) tail blocks RDR6 from converting self mRNAs into substrates for gene silencing. *Nature Plants* **3**: 17036.

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

Branscheid A, Marchais A, Schott G, Lange H, Gagliardi D, Andersen SU, Voinnet O, Brodersen P. 2015. SKI2 mediates degradation of RISC 5′-cleavage fragments and prevents secondary siRNA production from miRNA targets in Arabidopsis. *Nucleic Acids Research* **43**: 10975–10988.

Chantarachot T, Bailey-Serres J. 2018. Polysomes, stress granules, and processing bodies: a dynamic triumvirate controlling cytoplasmic mRNA fate and function. *Plant Physiology* **176**: 254–269.

Chen C-YA, Shyu A-B. 2011. Mechanisms of deadenylation-dependent decay. *Wiley Interdisciplinary Reviews: RNA* **2**: 167–183.

Chiba Y, Johnson MA, Lidder P, Vogel JT, van Erp H, Green PJ. 2004. AtPARN is an essential poly(A) ribonuclease in Arabidopsis. *Gene* **328**: 95–102.

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

De Alba AEM, Moreno AB, Gabriel M, Mallory AC, Christ A, Bounon R, Balzergue S, Aubourg S, Gautheret D, Crespi MD *et al.* 2015. In plants, decapping prevents RDR6-dependent production of small interfering RNAs from endogenous mRNAs. *Nucleic Acids Research* **43**: 2902–2913.

D'Orazio KN, Green R. 2021. Ribosome states signal RNA quality control. *Molecular Cell* **81**: 1372–1383.

Eisen TJ, Eichhorn SW, Subtelny AO, Lin KS, McGeary SE, Gupta S, Bartel DP. 2020. The dynamics of cytoplasmic mRNA metabolism. *Molecular Cell* **77**: 786–799.e10.

Fan W, Cho J. 2021. Quantitative measurement of transposon copy number using the droplet digital PCR. *Methods in Molecular Biology* **2250**: 171–176.

Fan W, Wang L, Chu J, Li H, Kim EY, Cho J. 2022. Tracing mobile DNAs: from molecular to population scales. *Frontiers in Plant Science* **13**: 837378.

Gasciolli V, Mallory AC, Bartel DP, Vaucheret H. 2005. Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing *trans*-acting siRNAs. *Current Biology* **15**: 1494–1500.

Gazzani S, Lawrenson T, Woodward C, Headon D, Sablowski R. 2004. A link between mRNA turnover and RNA interference in Arabidopsis. *Science* **306**: 1046–1048.

Graille M, Séraphin B. 2012. Surveillance pathways rescuing eukaryotic ribosomes lost in translation. *Nature Reviews Molecular Cell Biology* **13**: 727–735.

Han Y, Zhang X, Du R, Shan X, Xie D. 2023. The phase separation of SGS3 regulates antiviral immunity and fertility in Arabidopsis. *Science China Life Sciences* **66**: 1938–1941.

Harigaya Y, Parker R. 2010. No-go decay: a quality control mechanism for RNA in translation. *Wiley Interdisciplinary Reviews: RNA* **1**: 132–141.

Jia J, Lu W, Liu B, Fang H, Yu Y, Mo W, Zhang H, Jin X, Shu Y, Long Y *et al.* 2022. An atlas of plant full-length RNA reveals tissue-specific and monocots-dicots conserved regulation of poly(A) tail length. *Nature Plants* **8**: 1118–1126.

Jonas S, Christie M, Peter D, Bhandari D, Loh B, Huntzinger E, Weichenrieder O, Izaurralde E. 2014. An asymmetric PAN3 dimer recruits a single PAN2 exonuclease to mediate mRNA deadenylation and decay. *Nature Structural & Molecular Biology* **21**: 599–608.

Kervestin S, Jacobson A. 2012. NMD: a multifaceted response to premature translational termination. *Nature Reviews. Molecular Cell Biology* **13**: 700–712.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**: 357–360.

Kim EY, Wang L, Lei Z, Li H, Fan W, Cho J. 2021. Ribosome stalling and SGS3 phase separation prime the epigenetic silencing of transposons. *Nature Plants* **7**: 303–309.

Kordyukova M, Sokolova O, Morgunova V, Ryazansky S, Akulenko N, Glukhov S, Kalmykova A. 2020. Nuclear CCR4-NOT mediates the degradation of telomeric and transposon transcripts at chromatin in the *Drosophila germline*. *Nucleic Acids Research* **48**: 141–156.

Kurosaki T, Popp MW, Maquat LE. 2019. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nature Reviews. Molecular Cell Biology* **20**: 406–420.

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

Li H. 2018. MINIMAP2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**: 3094–3100.

Li Q, Chen S, Leung AW-S, Liu Y, Xin Y, Zhang L, Lam H-M, Luo R, Zhang S. 2021. DNA methylation affects pre-mRNA transcriptional initiation and processing in Arabidopsis. *bioRxiv*. doi: 10.1101/2021.04.29.441938.

Liang W, Li C, Liu F, Jiang H, Li S, Sun J, Wu X, Li C. 2009. The Arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. *Cell Research* **19**: 307–316.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESEQ2. *Genome Biology* **15**: 550.

Luo Z, Chen Z. 2007. Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in Arabidopsis. *Plant Cell* **19**: 943–958.

Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nature Reviews. Molecular Cell Biology* **16**: 665–677.

Matzke MA, Mosher RA. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics* **15**: 394–408.

Nishimura N, Kitahata N, Seki M, Narusaka Y, Narusaka M, Kuromori T, Asami T, Shinozaki K, Hirayama T. 2005. Analysis of ABA hypersensitive germination2 revealed the pivotal functions of PARN in stress response in Arabidopsis. *The Plant Journal* **44**: 972–984.

Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN, Slotkin RK. 2013. The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs. *Plant Physiology* **162**: 116–131.

Panda K, Ji L, Neumann DA, Daron J, Schmitz RJ, Slotkin RK. 2016. Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biology* **17**: 170.

Parker MT, Knop K, Sherwood AV, Schurch NJ, Mackinnon K, Gould PD, Hall AJ, Barton GJ, Simpson GG. 2020. Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m(6)A modification. *eLife* **9**: e49658.

Passmore LA, Coller J. 2022. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nature Reviews. Molecular Cell Biology* **23**: 93–106.

Pavlopoulou A, Vlachakis D, Balatsos NAA, Kossida S. 2013. A comprehensive phylogenetic analysis of deadenylases. *Evolutionary Bioinformatics Online* **9**: 491–497.

Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. 2015. STRINGTIE enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology* **33**: 290–295.

Quadrana L, Silveira AB, Mayhew GF, LeBlanc C, Martienssen RA, Jeddeloh JA, Colot V. 2016. The *Arabidopsis thaliana* mobilome and its impact at the species level. *eLife* **5**: 1–25.

Schäfer IB, Rode M, Bonneau F, Schüssler S, Conti E. 2014. The structure of the Pan2–Pan3 core complex reveals cross-talk between deadenylase and pseudokinase. *Nature Structural & Molecular Biology* **21**: 591–598.

Schäfer IB, Yamashita M, Schuller JM, Schüssler S, Reichelt P, Strauss M, Conti E. 2019. Molecular basis for poly(A) RNP architecture and recognition by the Pan2–Pan3 deadenylase. *Cell* **177**: 1619–1631.

Shimada TL, Shimada T, Hara-Nishimura I. 2010. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *The Plant Journal* **61**: 519–528.

Shoemaker CJ, Green R. 2012. Translation drives mRNA quality control. *Nature Structural & Molecular Biology* **19**: 594–601.

Suzuki Y, Arae T, Green PJ, Yamaguchi J, Chiba Y. 2015. AtCCR4a and AtCCR4b are involved in determining the poly(A) length of granule-bound starch synthase 1 transcript and modulating sucrose and starch metabolism in *Arabidopsis thaliana*. *Plant & Cell Physiology* **56**: 863–874.

Tan H, Luo W, Yan W, Liu J, Aizezi Y, Cui R, Tian R, Ma J, Guo H. 2023. Phase separation of SGS3 drives siRNA body formation and promotes endogenous gene silencing. *Cell Reports* **42**: 111985.

Tang TTL, Stowell JAW, Hill CH, Passmore LA. 2019. The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. *Nature Structural & Molecular Biology* **26**: 433–442.

Thran M, Link K, Sonnewald U. 2012. The Arabidopsis DCP2 gene is required for proper mRNA turnover and prevents transgene silencing in Arabidopsis. *The Plant Journal* **72**: 368–377.

Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. 2009. Bursts of retrotransposition reproduced in Arabidopsis. *Nature* **461**: 423–426.

Walley JW, Kelley DR, Nestorova G, Hirschberg DL, Dehesh K. 2010. Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. *Plant Physiology* **152**: 866–875.

Wheeler JR, Mitchell SF, Khong A, Matheny T, Parker R, Jain S. 2017. The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Molecular Cell* **68**: 808–820.

Workman RE, Tang AD, Tang PS, Jain M, Tyson JR, Razaghi R, Zuzarte PC, Gilpatrick T, Payne A, Quick J *et al.* 2019. Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nature Methods* **16**: 1297–1305.

Zhang X, Guo H. 2017. mRNA decay in plants: both quantity and quality matter. *Current Opinion in Plant Biology* **35**: 138–144.

Zhou H-R, Lin R-N, Huang H-W, Li L, Cai T, Zhu J-K, Chen S, He X-J. 2020. The CCR4-NOT complex component NOT1 regulates RNA-directed DNA methylation and transcriptional silencing by facilitating Pol IV-dependent siRNA production. *The Plant Journal* **103**: 1503–1515.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Identification of *de novo* mutated *DDM1* alleles.

Fig. S2 Similarity between two ddm1 alleles in RNA-Seq.

Fig. S3 RT-qPCR validation.

Fig. S4 TE fractions in euchromatic and heterochromatin regions.

Fig. S5 Reactivation of transposons in the *ccr4a* mutant.

Fig. S6 Translational efficiency and RNA granule enrichment.

Table S1 Sequences of primers used in this study.

Table S2 Summary of RNA-Seq.

Table S3 Summary of Oxford Nanopore direct RNA sequencing.

Table S4 Summary of whole-genome resequencing.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.