The TUTase URT1 connects decapping activators and prevents the accumulation of excessively deadenylated mRNAs to avoid siRNA biogenesis

Hélène Scheer1,5, Caroline de Almeida1,5, Emilie Ferrier1, Quentin Simonnot1, Laure Poirier1, David Pflieger1, François M. Sement1, Sandrine Koechler1, Christina Piermaria2, Paweł Krawczyk3,4, Seweryn Mroczek3,4, Johana Chicher2, Lauriane Kuhn2, Andrzej Dziembowski3,4, Philippe Hammann2, Hélène Zuber1✉ & Dominique Gagliardi1✉

Uridylation is a widespread modification destabilizing eukaryotic mRNAs. Yet, molecular mechanisms underlying TUTase-mediated mRNA degradation remain mostly unresolved. Here, we report that the Arabidopsis TUTase URT1 participates in a molecular network connecting several translational repressors/decapping activators. URT1 directly interacts with DECAPPING 5 (DCPS), the Arabidopsis ortholog of human LSM14 and yeast Scd6, and this interaction connects URT1 to additional decay factors like DDX6/Dhh1-like RNA helicases. Nanopore direct RNA sequencing reveals a global role of URT1 in shaping poly(A) tail length, notably by preventing the accumulation of excessively deadenylated mRNAs. Based on in vitro and in planta data, we propose a model that explains how URT1 could reduce the accumulation of oligo(A)-tailed mRNAs both by favoring their degradation and because 3’ terminal uridines intrinsically hinder deadenylation. Importantly, preventing the accumulation of excessively deadenylated mRNAs avoids the biogenesis of illegitimate siRNAs that silence endogenous mRNAs and perturb Arabidopsis growth and development.
Uridylation targets most classes of eukaryotic RNAs, from small and large noncoding RNAs (ncRNAs) to mRNAs. Uridylation of ncRNAs can promote maturation, control stability, or abrogate activity, depending on the type of ncRNAs and its cellular context. For mRNAs, the prevalent role of uridylation is to trigger 5′-3′ and 3′-5′ degradation, which is mostly achieved by the multifunctional Carbon Catabolite Repression-Negative On TATA-less (CCR4-NOT) complex. Two of its core subunits are the deadenylases CCR4 and CCR4-associated factor 1 (CAF1). CCR4 is proposed to shorten poly(A) tails of all mRNAs while CAF1 deadenylates mRNAs with lower rates of translation elongation and poor poly(A) binding protein (PABP) occupancy. Importantly, specificity factors bridge target mRNAs to the CCR4-NOT complex, thereby leading to translational repression and decay of specific mRNAs. Those specificity factors include the RNA-induced silencing complex (RISC) and several RNA binding proteins (RBPs), such as Tristetraprolin (TTP), Pumilio/lem-3 mRNA binding factor (PUF) proteins, and the YTS21-B homology (YTH) domain-containing family proteins YTHDF2 and Meiotic mRNA inter- ception protein 1 (Mmi1). In addition, the CCR4-NOT complex interacts with central regulators of translation and decapping, such as the GRB10-interacting GYF (glycine-tyrosine-phenylalanine domain) proteins (GIGYF) and the DExD/H-box RNA helicases DDX6, Dhh1, or Maternal expression at 31B (PAT1), and Like SM homolog 14 (LSM14), also called Suppress of clathrin dependence (Sced6) in yeast. The binding of DDX6 to CNOT1 or to decapping activators is proposed to be mutually exclusive. A possible scenario is that a succession of interactions allows DDX6 to hand over deadenylated mRNAs to the decapping machinery.

Once the CCR4-NOT complex has been recruited and the poly(A) tail has been shortened, oligo(A) tails of less than ca 25 As are frequently uridylated by terminal uridylyltransferases (TUTases) such as TUT4/7 in mammals or UTR:RNA URIDYLLYTRANSFERASE (URT1) in Arabidopsis thaliana (hereafter Arabidopsis). URT1-mediated uridylation can restore a binding site for a PABP, but its impact on mRNA stability is yet unsolved. In fission yeast and human cultured cells, uridylation of short oligo(A) tails is proposed to favor the binding of the LSM1-7 complex, which recuits the decapping complex through the interaction with PAT1, and ultimately results in the degradation of the decapped mRNA by the 5′-3′ exoribonuclease 1 XRNI (XRNI in plants). Alternatively, U-tails can directly attract Dis3-like protein 2 (Dis3L2) or the RNA exosome to trigger 3′-5′ exoribonucleolytic decay of mRNAs.

In this study, we show that the Arabidopsis TUTase URT1 is integrated in an interaction network comprising the deadenylation complex CCR4-NOT and other translation repressors/decapping activators, including DCP5, the plant ortholog of the translational inhibitor/decapping activator Scd6 or LSM14A. Our interactomic and functional analysis data support a model explaining how URT1-mediated uridylation prevents the accumulation of excessively deadenylated mRNAs. We also show that in absence of URT1-mediated uridylation, excessively deadenylated mRNAs can become a source of spurious siRNAs that silence endogenous mRNAs, with a negative impact on plant fitness.

Results
Conservation of an intrinsically disordered region (IDR) across plant URT1 orthologs. The 764-long amino acid sequence of the Arabidopsis TUTase URT1 encoded by AT2G45620 can be divided in two regions discriminated by compositional biases and the presence of known domains (Fig. 1a). URT1’s C-terminal region contains the catalytic core domain (CCD), the typical signature of terminal nucleotidyltransferase family members. The CCD is composed of a nucleotidyltransferase domain (amino acids 434–567, Superfamily domain SCOP 81302, E value = 3.11 × 10⁻³⁵) followed by a PAP-associated domain or PAP/OAS1 substrate-binding domain (amino acids 571–741, Superfamily domain SCOP 81631, E value = 6.8 × 10⁻⁴⁸) (Fig. 1a). By contrast, the N-terminal region of URT1 (amino acids 1–433) is devoid of known domains and is characterized by a significant enrichment for P/Q/N/G (p value = 2.4 × 10⁻¹⁸) compared to the Arabidopsis proteome (Fig. 1a). Moreover, the whole N-terminal half of URT1 is predicted as a large IDR (Fig. 1b).

To test for the possible conservation of this IDR amongst plant URT1 orthologs, we analyzed 87 sequences (Supplementary Data 1) that were recently compiled to determine the evolutionary history of TUTases in Archaeplasta (i.e., all plants). These URT1 sequences originate from 72 species representing major groups of Archaeplasta: glaucophytes, rhodophytes (red algae), chlorophyte and streptophyte algae, bryophytes (liverworts, hornworts, and mosses), lycophytes and pteridophytes (e.g., ferns), gymnosperms (e.g., conifers and Gingko), and angiosperms (flowering plants). IDRs can tolerate mutations that do not affect their overall function. Indeed, the primary sequence of URT1’s IDR is highly variable between species (Fig. 1d and Supplementary Fig. 1a). Yet, two SLiMs named hereafter M1 and M2 are conserved in land plants, i.e., from bryophytes (including mosses) to flowering plants (Fig. 1d, e and Supplementary Fig. 1a). A PPGF motif is also conserved in most land plant URT1s (Supplementary Fig. 1a). Its conservation is underestimated by sequence alignment partly because of its varying positions in the IDR (Fig. 1e). Yet, a systematic search revealed that URT1 orthologs of flowering plants of at least one URT1 ortholog with M1, M2, and PPGF motifs indicates key functions under selective pressure.

Short linear motifs (SLiMs) are conserved in plant URT1 orthologs. IDRs can tolerate mutations that do not affect their overall function. Indeed, the primary sequence of URT1’s IDR is highly variable between species (Fig. 1d and Supplementary Fig. 1a). Yet, two SLiMs named hereafter M1 and M2 are conserved in land plants, i.e., from bryophytes (including mosses) to flowering plants (Fig. 1d, e and Supplementary Fig. 1a). A PPGF motif is also conserved in most land plant URT1s (Supplementary Fig. 1a). Its conservation is underestimated by sequence alignment partly because of its varying positions in the IDR (Fig. 1e). Yet, a systematic search revealed that URT1 orthologs of flowering plants of at least one URT1 ortholog with M1, M2, and PPGF motifs indicates key functions under selective pressure.

URT1 co-purifies with translational repressors/decapping activators. The conservation of a large IDR containing SLiMs supports the possibility that URT1 interacts with one or several partners. To identify this interaction network, proteins co-purifying with URT1 (tagged with myc or YFP) expressed in urt1 mutants were identified by LC-MS/MS analyses. To obtain a global view of URT1 RNP context, cellular extracts were cross-linked with formaldehyde before immunoprecipitation (IP) (Fig. 2a). The comparison of eight URT1 samples (representing four biological replicates) to seven control samples revealed 62 proteins significantly enriched in URT1 IPs (Fig. 2a).
and Supplementary Data 2). The most enriched molecular functions associated to URT1 co-purifying proteins are mRNA binding (GO:0003729) and RNA binding (GO:0003723) (Benjamini–Hochberg corrected p values of $4.8 \times 10^{-29}$ and $6.1 \times 10^{-18}$, respectively). Both categories are consistent with the known involvement of URT1 in mRNA metabolism.11,12.

The most significantly enriched protein co-purifying with URT1 is ESSENTIAL FOR POTEXVIRUS ACCUMULATION 1 (EXA1) encoded by AT5G42950 and also named GYN4, PSIG1, and MUSE1.25–28. The two closest EXA1 homologs (AT1G27430 and AT1G24300) are also significantly enriched in URT1 IPs (noted GYF protein in Fig. 2a and Supplementary Data 2a). EXA1...
Fig. 2 URT1 co-purifies with translational repressors/decapping activators. Enrichment of proteins co-purified with myc and YFP-tagged URT1 with formaldehyde crosslink (a) or without (b). The dashed line indicates the threshold above which proteins are significantly enriched (adjusted p value < 0.05, quasi-likelihood negative binomial generalized log-linear model). c Common domain organization of Arabidopsis (At) DCP5 and its human (Hs) and yeast (Sc) orthologs LSM14A and Scd6, respectively. d In vitro GST pull-down assay showing a direct URT1-DCP5 interaction. Pull-downs were performed in presence of RNase A with the recombinant proteins 6His-GST, 6His-GST-URT1, 6His-MBP-DCP5, 6His-GST-m1URT1, and 6His-MBP-ΔLSmDCP5. Pull-down and input fractions were analyzed by SDS-PAGE and SYPRO Ruby staining. e Diagram illustrating the point mutations in m1URT1 construct. f Volcano plot showing proteins differentially enriched (log2 fold change > 0) or depleted (log2 fold change < 0) in myc-tagged m1URT1 versus myc-tagged URT1 IPs. The dashed line indicates the significant threshold (adjusted p value < 0.05, quasi-likelihood negative binomial generalized log-linear model). Dot color code is as in a. The source data are available in Supplementary Data 2, at [https://www.ebi.ac.uk/pride/archive/projects/PXD018672] and at [https://doi.org/10.17632/ybcvvmtn9.3].
is a GYF domain-containing protein, orthologous to human and Drosophila GRB10-interacting GYF domain proteins (GIGYF). GIGYF proteins interact with the deadenyl complex CCR4-NOT and several translational repressors or decapping activators such as the S‘ cap-binding protein elf4E-Homologous Protein (4EHP), the RNA helicase DDx6/Me31B, and the decapping activator Pat120,21. Interestingly, orthologs of all known GIGYF interactors are also significantly enriched in Urt1 IPs alongside Eax1. They include subunits of the CCR4-NOT complex like Cnot1, Cnot10, and Cnot11 (AT1G02080, AT5G35430, and AT5G18420, respectively), the 4EHP ortholog called new cap-binding protein (nCBP) (AT5G3180), and the DDx6-like RNA helicases Rh6, Rh8, and Rh12 (AT2G45810, AT4G00660, and AT3G61240, respectively). These results raise the possibility that the chain of interactions described for GIGYF is conserved for Eax1 in Arabidopsis and that Urt1 is connected to these factors, including the CCR4-NOT complex.

Another translational repressor/decapping activator highly enriched in Urt1 IPs is Decapping5 (DCP5) encoded by AT1G26110 (Fig. 2a and Supplementary Data 2a). The known interactants of the human DCP5 ortholog LSM14 are Edc4, the DDX6-like RNA helicases, and the elf4E-binding protein 4E-T.29 There is no 4E-T ortholog in Arabidopsis. However, both Varicose (Vcr, ortholog to EDC4 and encoded by AT3G13000), and the aforementioned DDx6-like RNA helicases (Rh6, Rh8, and Rh12) are enriched in Urt1 IPs. Of note, Decapping 5-Like (DCP5-L, AT5G45330) and Varicose-Related (Vcr, AT3G13290), homologs of Dcp5 and Vcs, respectively, are also significantly enriched in Urt1 IPs (Fig. 2a and Supplementary Data 2).

Repeating the IP experiments without formaldehyde crosslink revealed a much simpler interactome with Dcp5, Rh6, Rh8, Rh12, and the translation initiation factor elf4f (AT3G60240) among the most enriched proteins (Fig. 2b and Supplementary Data 2b). A DcP5-elf4f interaction has not yet been reported, but the yeast DCP5 ortholog Scd6 does interact with elf4f via its C-terminal RGG repeats30, which are also present in DCP5 and LSM14A (Fig. 2c).

Altogether, our IP results indicate that Urt1 is integrated into interaction networks connecting translational repressors and decapping activators, including DCP5.

The Slm1 M1 mediates a direct interaction between DCP5 and Urt1. DCP5 is the most enriched protein in mycUrt1 IPs without crosslinking (Fig. 2b and Supplementary Data 2b) and reciprocally, Urt1 co-purifies with DCP5 tagged with GFP and expressed at endogenous levels (Supplementary Fig. 2a). DCP5 contains a Lsm domain, and the Lsm domain-containing proteins Lsm14, Scd6, and Ed3 interact with helical leucine-rich motifs (HLMs)31, which resemble Urt1’s M1 motif (Fig. 1d). We therefore suspected a direct DCP5-Urt1 interaction. Indeed, in vitro pull-down experiments in presence of RNAse A confirmed a direct interaction between 6His-Gst-Urt1 and 6His-Mbp-DCP5 (Fig. 2d). This direct interaction requires the M1 motif because a mutated version of Urt1, in which leucines 21 and 25 in M1 are mutated into asparagines (6His-Gst-m1Urt1) (Fig. 2e) failed to pull down DCP5 (Fig. 2d). Furthermore, 6His-Gst-Urt1 cannot pull down 6His-MBP-LSmDCP5, indicating that the Lsm domain of DCP5 is necessary for the Urt1-DCP5 interaction (Fig. 2d). Additional assays confirmed that 6His-Gst-Urt11–40 (i.e., the 40 first aminoacids of Urt1 containing the M1 motif but not the M2 motif and fused downstream of 6 histidines and GST) is sufficient to pull down 6His-Mbp-DCP5 LSm (i.e., DCP5’s Lsm domain fused downstream of 6 histidines and MBP). As expected, this interaction is abolished by mutating the M1 motif (Supplementary Fig. 2b). By contrast, mutating the M2 motif has no impact on the Urt1-DCP5 pull-down assays (Supplementary Fig. 2c). Altogether, these data reveal that Urt1 can directly bind to DCP5 via an interaction between Urt1’s conserved M1 motif and DCP5’s Lsm domain.

To test how M1 impacts the Urt1 interactome in planta, Urt1-myc or m1Urt1-myc were expressed in urt1-1 mutant plants and used as baits in IPs following formaldehyde crosslink. Interestingly, the six proteins most significantly depleted by mutating the M1 motif are Dcp5, Dcp5L, Rh6, Rh8, Vcs, and Rh12 (Fig. 2f and Supplementary Data 2c). We conclude from these experiments that the conserved Slm1 M1 connects Urt1 to DCP5 (and possibly DCP5L), which recruits additional translational repressors or decapping activators such as Vcs or the Dhhl1/DDx6-like RNA helicases, Rh6, Rh8, and Rh12.

Ectopic expression of Urt1 remodels poly(A) tail profiles. To investigate the molecular function of Urt1-mediated mRNA uridylation and test the potential role of the M1 and M2 motifs, we first determined how the ectopic expression of Urt1-myc or m1m2Urt1-myc affects the expression of a GFP reporter mRNA co-expressed in Nicotiana benthamiana leaves. To prevent transgene-induced silencing, the silencing suppressor P19 was co-expressed in all experiments and will not be further mentioned. The GFP reporter was co-expressed without Urt1 (ctrl) or with either one of two catalytically impaired versions of Urt1, Urt1D491/3A, or Urt1P618L. (see Fig. 3a for a schematic representation of the different Urt1 versions). Urt1D491/3A is fully inactivated by the mutations of catalytic residues13 whereas the uridylating activity of Urt1P618L mutant is strongly affected, but not abrogated32.

Both active and inactive Urt1 versions were expressed as full-length proteins (Fig. 3b). Yet, the inactive versions of Urt1 are systematically more expressed as compared with active ones, which is important for the interpretation of the results presented hereafter. GFP expression levels were monitored by UV illumination of infiltrated leaf patches. Leaf patches expressing catalytic inactive versions of Urt1 showed GFP levels similar to controls. By contrast, the GFP fluorescence was systematically decreased upon expression of active versions of Urt1 with either wild-type or mutated M1 and M2 motifs (Fig. 3c). Thus, GFP repression requires Urt1’s activity but obviously neither the M1 nor the M2 motif. We do not know at present whether GFP repression is a direct consequence of GFP mRNA uridylation, or due to an indirect effect of Urt1 ectopic overexpression. Of note, GFP repression is not the consequence of a lower expression of the silencing suppressor P19 because similar results were obtained with N. benthamiana plants silenced for RDR6, a key component of transgene-induced post-transcriptional silencing (PTGS) (Supplementary Fig. 3a).

Remarkably, the decrease in GFP expression was not due to lower amounts of GFP mRNAs, because similar steady-state levels of GFP mRNAs were detected in patches expressing active or inactive Urt1 (Fig. 3d). Yet, expression of the inactive Urt1D491/3A resulted in a slight, but systematic, shift in the migration of GFP mRNAs detected by northern blots (Fig. 3d). To investigate the reason for this small size shift and to determine the molecular impact of ectopic Urt1 expression on GFP mRNA tails, those mRNAs were detected by 3’RACE-seq. A primer was ligated to RNA 3’ extremities and used to initiate cDNA synthesis. The 3’ region of GFP reporter mRNAs (including the poly(A) tail and eventually non-A nucleotides) was PCR-amplified and GFP amplicons from six independent biological replicates were sequenced in two MiSeq runs (Supplementary Data 3a, b). In control patches, GFP mRNAs had a mean uridylation level of 3.2% and uridylation...
Fig. 3 Ectopic expression of URT1 remodels poly(A) tail profiles. a Domain organization and mutations of different URT1-myc versions transiently expressed in *N. benthamiana* leaf patches. Legend as in Fig. 1a. b Western blot analysis of URT1-myc expression. c Representative *N. benthamiana* leaf under UV light to detect the expression of the GFP reporter co-expressed with the different URT1-myc versions (top). Quantification of GFP fluorescence of the different patches relative to control (ctrl) for 11 independent replicates (bottom). d Northern blot analysis of the steady-state level of GFP mRNAs. The arrows indicate unspliced and mature forms of GFP mRNAs. e Uridylation percentage of GFP reporter mRNAs for six biological replicates. f, g Distribution profiles of GFP mRNA poly(A) tail sizes. The percentages of sequences were calculated for six biological replicates for uridylated (f) and homopolymeric poly(A) tails (non-uridylated) (g) tails from 1 to 90 nucleotides. The percentages were calculated using the total number of sequences with tails from 1 to 90 nucleotides. Individual points are color-coded for each replicate and the average of all replicates is indicated as a gray area. The pie charts represent the average proportion of tails longer than 90 nucleotides. Letters in (c–g) represent significant statistical *p* value (two-tailed Wilcoxon rank-sum test, *n* = 11 (c) and *n* = 6 (e–g)). Exact *p* values are indicated in Supplementary Data 3e. The source data are available in Supplementary Data 3, at [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148409] and at [https://doi.org/10.17632/ybcvvmtc9n.3].
occurred mostly on GFP mRNAs with 10–25 As (ctrl in Fig. 3e, f). This mRNA population with oligo(A) tails of 10–25 As will be later referred to as oligoadenylated mRNAs. The GFP mRNA uridylation pattern resembles the ones typically observed for Arabidopsis mRNAs10–12 and we propose that this basal uridylation level is performed by the N. benthamiana URT1 ortholog. As expected, GFP mRNA uridylation levels increased upon ectopic expression of URT1-myc and reached up to 15% (Fig. 3e). Interestingly, the size distribution profile for uridylated tails was markedly modified by the ectopic expression of URT1-myc, which resulted in the uridylation of large poly(A) tails up to 90 As (Fig. 3f). Moreover, the number of uridylated tails longer than 90 As was also significantly higher upon ectopic expression of URT1 as compared to the control samples or leaf patches expressing inactive URT1 (pie charts in Fig. 3f). Thus, URT1 can uridylate long poly(A) tails when ectopically expressed, demonstrating that deadenylation is not a pre-requisite for uridylation.

Strikingly, URT1 ectopic expression also affected the size distribution of homopolymeric poly(A) tails (called non-uridylated poly(A) tails hereafter and in all figures). Indeed, non-uridylated poly(A) tails in samples expressing wild-type URT1 showed a clear decrease of the 16–20 peak and an accumulation of longer poly(A) tails as compared to control samples (Fig. 3g and Supplementary Fig. 3b). This accumulation is due to the uridylation activity of URT1, as it is not observed for URT1D491/3A (Fig. 3g), despite the higher expression levels of the inactive protein (Fig. 3b).

Another somehow counterintuitive observation was that overexpression of inactive URT1D491/3A resulted in increased GFP mRNA uridylation as compared to control samples (Fig. 3e). This increased uridylation corresponds to the specific accumulation of uridylated oligoadenylated GFP mRNAs as compared to control samples (Fig. 3f). We hypothesized that inactive URT1D491/3A overexpressed to high levels (Fig. 3b) may sequester decay factors, thereby hindering the degradation of uridylated oligoadenylated mRNAs. The resolution of the URT1 interactome in Arabidopsis hinted that the M1 motif could participate in sequestering decay factors. Therefore, we tested whether the accumulation of oligoadenylated GFP mRNAs upon ectopic overexpression of URT1D491/3A requires the M1 motif. To do so, we co-expressed the reporter GFP mRNAs with URT1, URT1D491/3A, m1URT1D491/3A, m2URT1D491/3A or m1m2URT1D491/3A (see Fig. 4a for a schematic representation of the different URT1 versions). As previously noted, URT1 inactive versions were expressed more highly than active URT1 (Fig. 4b). We then analyzed GFP mRNA tailing profiles by 3′RACE-seq. In line with our previous results, the ectopic expression of URT1 resulted in much longer poly(A) tails for both uridylated and non-uridylated GFP mRNAs (Fig. 4c, d, respectively), and uridylated oligoadenylated GFP mRNAs accumulated upon URT1D491/3A expression (Fig. 4c). Interestingly, the accumulation of uridylated oligoadenylated GFP mRNAs was reduced by mutating the M1, but not the M2 motif (Fig. 4c). This observation supports the idea that overexpressed URT1D491/3A sequesters a factor(s) involved in the turnover of uridylated oligoadenylated mRNAs through an interaction involving the M1 motif. Indeed, the comparison of MS data for seven URT1D491/3A versus eight m1-URT1D491/3A IPs revealed that URT1D491/3A co-purifies with N. benthamiana decay factors including DCP5 homologs and that this interaction requires the M1 motif (Supplementary Fig. 3c, d).

Finally, we noted that poly(A) tails interspersed with non-A ribonucleotides, subsequently called A-rich tails, accumulated in URT1D491/3A samples. This accumulation is strictly dependent on the presence of the M1 motif in overexpressed URT1D491/3A (Fig. 4e and Supplementary Fig. 3e). Of note, only few A-rich tails are detected in control or URT1 samples, and they could have been considered as possible experimental artefacts. However, their dramatic accumulation in URT1D491/3A provides compelling evidence that these A-rich tails are produced in vivo. Either A-rich tails are constitutively produced but do not accumulate in wild-type plants, or their production is induced by overexpressing M1-containing URT1D491/3A. Their simultaneous accumulation with uridylated and oligoadenylated mRNAs, triggered by overexpressing URT1D491/3A versions that contains the M1 motif, suggests that these tails mark mRNAs undergoing degradation.

Taken altogether, those data confirm that URT1 connects decay factors through the M1 motif and, more importantly, reveal that the ectopic overexpression of URT1 remodels the poly(A) tails of GFP reporter mRNAs toward larger sizes.

**Effects of URT1 ectopic expression on tailing of endogenous PR2 mRNAs.** We abstained from tethering URT1 to the GFP reporter mRNA because URT1 is a distributive TUTase11 and tethering would likely entail the synthesis of longer poly(U) tails as compared to the wild-type situation. In our experimental design, the number of uridines added to GFP mRNAs is similar between control and URT1 samples (Supplementary Fig. 3f). Because ectopically expressed URT1 is not tethered to the reporter mRNAs, it potentially uridylates also endogenous mRNAs. To test this possibility, endogenous mRNAs encoding PATHOGENESIS-RELATED PROTEIN 2 (PR2), were analyzed by 3′RACE-seq. PR2 mRNAs were chosen because the agroinfiltration procedure triggers PR2 expression (Supplementary Fig. 4a). Unlike for GFP mRNAs, overexpression of URT1D491/3A led to increased levels of PR2 mRNAs (Supplementary Fig. 4a). Either overexpression of URT1D491/3A induces PR2 transcription at a higher level, or URT1D491/3A overexpression impairs PR2 mRNA turnover. In line with the GFP mRNA results, URT1 ectopic expression increased PR2 uridylation levels, resulted in the uridylation of longer poly(A) tails, and led to the accumulation of PR2 mRNAs with longer poly(A) tails as compared to the control samples (Supplementary Fig. 4b–d). Moreover, URT1D491/3A overexpression led to the accumulation of oligoadenylated uridylated PR2 mRNAs, as well as to the accumulation of PR2 mRNAs with A-rich tails (Supplementary Fig. 4b, c, e). Altogether, these data indicate that URT1 ectopic expression has a similar impact on tail sizes on both the reporter GFP mRNAs and the endogenous PR2 mRNAs.

**URT1-mediated uridylation shapes poly(A) tails in Arabidopsis.** If the shift toward larger poly(A) tail sizes induced by URT1 overexpression reflects bona fide molecular functions of this TUTase, the opposite effect should be observed in urt1 mutants. We therefore compared global poly(A) tail profiles between three biological replicates of wild-type and urt1 plants using nanopore direct RNA sequencing (DRS). Although DRS does not yet allow to detect mRNA uridylation, it is well suited for measuring poly(A) tail sizes, including long ones. The nanopore DRS analysis revealed the impact of URT1 on poly(A) tail size distributions with a clear shift of the distribution toward short oligo(A) tails in urt1 samples (Fig. 5a). This accumulation of deadenylated mRNAs is observed for both bulk and intergenic poly(A) tail size distributions, thereby reflecting a robust effect.

We then used the 3′RACE-seq method to precisely compare polyadenylation and uridylation profiles for 22 mRNAs analyzed in two biological replicates of wild-type and urt1 plants. Those mRNAs were selected because they have various uridylation levels ranging from 1 to 24% (Fig. 5b), and quite distinct poly(A) tail profiles (for instance compare AT2G30570 and AT4G28240, the...
Fig. 4 Accumulation of oligo(A) uridylated and A-rich-tailed GFP mRNAs upon overexpression of URT1\textsuperscript{D491/3A} requires the M1 motif. a Domain organization and mutations of URT1-myc versions. b Western blot analysis of URT1-myc expression. c–e Distribution profiles of GFP mRNA poly(A) tail sizes. The percentages of sequences were calculated for three biological replicates for uridylated (c), non-uridylated (d) and A-rich (e) tails from 1 to 90 nucleotides. The percentages were calculated using the total number of sequences with tails from 1 to 90 nucleotides. Individual points are color-coded for each replicate and the average of all replicates is indicated as a gray area. The source data are available in Supplementary Data 3, at [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148409] and at [https://doi.org/10.17632/ybcvmtcn9.3].
first and last genes shown in Fig. 5b, respectively). In agreement with our previous studies, mRNA uridylation levels drop in urt1 mutants (bar plots in Fig. 5b), and uridylation tags mostly oligoadenylated mRNAs, i.e., mRNAs with tails of <25As (Supplementary Fig. 5). Yet, longer poly(A) tails can also get uridylated, albeit to low levels (Supplementary Fig. 5). In line with the DRS data, the profiles for poly(A) tails are modified upon loss of URT1: a slight shift toward smaller poly(A) tails is frequently observed.
observed in urt1 (Fig. 5b). The changes in poly(A) tail profiles can be split in two effects. First, the urt1 mutant accumulates oligoadenylated mRNAs with oligo(A) tails of 10–25 As for most of the analyzed mRNAs. This effect is particularly obvious for AT4G38770 mRNAs, whose poly(A) tail profile is strongly impacted by loss of URT1, despite its apparently low uridylation frequency in wild type (Fig. 5b). Second, excessively deadenylated mRNAs (with tails of <10 As) accumulate in urt1 mutants. Interestingly, the accumulation of mRNAs with tails <10 is more often observed for highly uridylated mRNAs (Fig. 5b). Hence, URT1 prevents the accumulation of excessively deadenylated mRNAs.

Two hypotheses could explain that poly(A) tails are shortened in urt1 mutants (and conversely, shifted toward longer sizes upon URT1 overexpression in N. benthamiana leaves). First, uridylation by URT1 could impede deadenylation, thereby slowing down the production of oligoadenylated mRNAs. Second, URT1 could trigger the degradation of oligoadenylated mRNAs, thereby shifting the distribution toward longer poly(A) tails. Importantly, those possibilities are not mutually exclusive.

We first used in vitro assays to test whether uridylation could impede deadenylation. The Arabidopsis genome encodes two CCR4 and 11 CAF1 homologs. CAF1s are classified in three groups based on phylogenetic analyses. CAF1 from group A (CAF1a and CAF1b) and group C (CAF1h–k) have been proposed to interact with NOT1, the scaffold protein of the CCR4-NOT complex. We purified recombinant CCR4s and the CAF1 proteins from group A and C and tested their deadenylase activity in vitro. For unknown reasons, only CAF1b had a robust deadenylase activity under the various biochemical conditions tested. Therefore, CAF1b was used to test the intrinsic influence of uridylation on its deadenylation activity. A catalytic mutant CAF1bD42A was used as a negative control. CAF1b and CAF1bD42A were incubated with radiolabeled RNA substrates containing either 14 3′ terminal As, 13 As and 1 U, 12 As and 2 Us, or 14 Us. As expected, CAF1b fastly degraded the oligo(A) tail whereas it was markedly inhibited by 14 Us (Fig. 5c).

Interestingly, the presence of a single 3′ terminal uridine delayed the degradation of the oligo(A) tail (Fig. 5c). Two 3′ terminal uridines even further impeded CAF1b activity (Fig. 5c). Therefore CAF1b activity is slowed down by the presence of a limited number of 3′ terminal uridines, but not fully inhibited. Whether this intrinsic feature is maintained upon plant CCR4-NOT complex assembly is unknown yet. Of note, the CAF1 activity within a fully reconstituted S. pombe CCR4-NOT complex is also slowed down by 3′ terminal uridines, albeit to a lesser extent than by guanosines and cytidines. The fact that uridylation can intrinsically impede deadenylation explains, at least partly, why overexpression of URT1 in N. benthamiana leaves results in the accumulation of long poly(A) tails (Figs. 3 and 4) and why loss of URT1 in Arabidopsis results in the accumulation of excessively deadenylated mRNAs, especially of mRNAs with a high uridylation level (Fig. 5b).

The second possibility to explain the change of poly(A) tail profiles in urt1 mutants is that URT1-mediated uridylation triggers the degradation of its main targets, i.e., deadenylated mRNAs. This possibility is in line with the primordial role of mRNA uridylation, mostly investigated in S. pombe and mammals. We have previously shown in Arabidopsis that 90% of uncapped LOM1 mRNAs are uridylated, whereas only 1% of capped LOM1 mRNAs have terminal uridines. Those uncapped and uridylated mRNAs are unstable and were detected only in a xrn4 mutant, in which cytosolic 5′-3′ RNA degradation is compromised. Those data already suggested a link between uridylation and decay in Arabidopsis, at least for LOM1 mRNAs. Yet, a global relationship between uridylation and decay rates was not yet reported in plants. To this end, we first generated TAIL-seq libraries for three biological replicates of Col-0 plants to rank mRNAs according to their uridylation levels (Supplementary Data 5). This dataset was then compared with three independent datasets reporting transcriptome-wide mRNA half-lives in Arabidopsis. Despite the poor correlation between the three datasets, we found for each comparison that the higher the mRNA uridylation, mostly investigated in Arabidopsis growth and development, at least when plants are grown in standard conditions. Yet, introgressing the urt1 mutation into an xrn4 background, lacking the main cytosolic 5′-3′ exoribonuclease, had a detrimental impact on development (Fig. 6a, b and Supplementary Fig. 6). 6.5-week-old urt1-1 xrn4-3 plants failed to set new leaves (Fig. 6a), and 9.5-week-old urt1-1 xrn4-3 plants had severely impaired statures as compared to control plants or single mutants (Fig. 6a). Moreover, urt1-1 xrn4-3 double mutants failed to develop inflorescences when grown under 12/12 (day/night) conditions (Fig. 6b).
The accumulation of RNA decay intermediates such as uncapped and excessively deadenylated mRNAs can be deleterious in plants because such aberrant mRNAs can erroneously trigger the biogenesis of siRNAs. Because some illegitimate siRNAs are produced in an xrn4 mutant, we checked whether urt1-1 xrn4-3 growth and developmental defects are linked to an increased biogenesis of spurious siRNAs. We first analyzed small RNA libraries from 24-day old in vitro grown seedlings before the onset of visible phenotypes and indeed detected an increased accumulation of 21-nt siRNAs originating from mRNA loci in urt1-1 xrn4-3 (Fig. 6c). Overall, siRNAs derived from 2659 mRNAs significantly accumulated in urt1-1 xrn4-3 (Fig. 6d and Supplementary Data 6). Preventing siRNA production by mutating DCL2 and DCL4 in urt1-1 xrn4-3 abrogated the growth and developmental defects associated to the urt1-1 xrn4-3 mutation (Fig. 6e). This result demonstrates the causality between urt1-1 xrn4-3 phenotype and the production of spurious siRNAs.

Finally, we checked whether the mRNAs that are more prone to trigger the synthesis of spurious siRNAs in urt1-1 xrn4-3 are highly uridylated in wild-type plants. We therefore compared the TAIL-seq results to the small RNA-seq data and indeed, mRNAs that produce spurious siRNAs in urt1-1 xrn4-3 have a significantly higher propensity to uridylation in wild-type plants (Fig. 6f).

Altogether, our data reveal that URT1-mediated uridylation prevents the accumulation of excessively deadenylated mRNAs, and by doing so, avoids the production of spurious siRNAs that can target endogenous mRNAs.

Discussion
Uridylation is now recognized as an integral step of mRNA degradation in eukaryotes. Yet, the full range of its molecular functions in assisting mRNA decay remains to be defined. Based
on the URT1 interactome and the functional analysis of URT1-mediated uridylation presented here, we propose a model integrating the dual function of URT1 in preventing excessive deadenylation and favoring the turnover of deadenylated mRNAs through the direct recruitment of decapping activators (Fig. 7). By preventing the accumulation of excessively deadenylated mRNAs, URT1-mediated uridylation protects endogenous mRNAs from triggering the synthesis of spurious siRNAs in Arabidopsis.

The composite domain organization of most TUTases is proposed to be key for the recruitment of factors that assist TUTases for the recognition of specific RNA substrates or channel the downstream molecular effects of uridylation43,44,45. However, only a few interactants of cytosolic TUTases have been identified to date. In mammals, TUT4/7 contacts the RBP Lin28, which binds Group II let-7 miRNA precursors. The presence or absence of Lin28 toggles TUT4/7 into a processive or a more distributive mode, promoting either degradation or maturation of let-7 precursors, respectively46–48. In Drosophila, the TUTase Tailor binds the 3′–5′ exoribonuclease Dis3l2 to form the terminal RNA uridylation-mediated processing (TRUMP) complex, which degrades a variety of structured mRNAs in the cytoplasm49. In this study, we show that Arabidopsis URT1 co-purifies with several translational repressors/decapping activators, orthologs of which are known to form an intricate and dynamic interaction network in animals13,20,22,50,51. However, whether this network also comprises a TUTase in animals is not yet known.

A key factor at the heart of this dynamic network connecting translational repressors and decapping activators is the CCR4-NOT complex. Although not all components of CCR4-NOT were enriched in URT1 IPs, the detection of the CCR4-NOT complex scaffold subunit NOT1 alongside the other CCR4-NOT subunits CNOT10 and CNOT11 strongly supports a connection between the CCR4-NOT complex and URT1. The prime candidate for connecting URT1 to the CCR4-NOT complex is EXA1, a GYF domain containing protein, homologous to the human and fly GIGYF proteins20,21. GIGYF are translation repressors and decapping activators, that interact with different translation repressors via multiple interfaces, among them the GYF domain20,21. GIGYF interactants include the RNA helicase DDX6/Me31B, PAT1, the 5′ cap-binding protein 4EHP, and the CCR4-NOT complex51,52. Interestingly, orthologs of all these factors are detected in URT1 IPs, supporting the hypothesis that such an interaction network is conserved in Arabidopsis. In line with our data, a two-hybrid screen using Arabidopsis EXA1 as a bait (called GYN4 in this study) retrieved CNOT4 subunits26, providing independent support for a physical association between EXA1 and the CCR4-NOT complex. Interestingly, the GYF domain of EXA1 recognizes a PPGF sequence26, and such a motif is conserved in URT1 of flowering plants. A CCR4-NOT/EXA1/URT1 network may explain, at least in part, how URT1 selects its targets and why deadenylated mRNAs are preferentially uridy- lated. The 3′ extremity of long poly(A) tails that are either protected by PABPs or being shortened by CCR4-NOT would be poorly accessible to URT1, even though URT1 and CCR4-NOT are connected. However, once poly(A) tails get short enough to loosen their association with the last remaining PABP and when CCR4-NOT’s activity is more distributive, the tethering of URT1 to CCR4-NOT via EXA1 could facilitate the uridylation of oligo (A)-tailed mRNAs.

In S. pombe and mammalian cells, uridylation is proposed to favor decapping of deadenylated mRNA by promoting the binding of the LSM1-7 complex5,8,53, which recruits the decapping complex through a connection with Pat134–37. The LSM1-7 complex preferentially binds short oligo(A) tails of <10 As23. Our previous TAIL-seq analyses12 and the 3′RACE-seq data presented here show that the majority of uridylated oligo(A) tails are longer than 10 As and comprise mostly 10–25 As. Such oligo(A) tails are likely suboptimal targets for the LSM1-7 complex. Interestingly, the direct recruitment of decapping activators by URT1 could bypass the requirement for LSM1-7 binding. We demonstrated here that the conserved M1 motif of URT1 directly binds to the decapping activator DCP5. DCP5 interacts with additional decapping factors, like DCP1 and DCP229,30. For yet unknown reasons, neither DCP1 nor DCP2 were significantly enriched in URT1 IPs. By contrast, we detected VCS, an ortholog of human EDC4, which interacts with LSM14 via a FFD motif29, perfectly conserved in Arabidopsis DCP5. Moreover, a LSM14-DCP6 interaction via LSM14′s PFD and TFG motifs (both conserved in DCP5) is required to expose the FFD motif for EDC4 recruitment30. Our URT1 IP data support the idea that a DCP5/ RH6,8,12/VCS connection also exists in Arabidopsis. In addition, the most straightforward interpretation of the poly(A) profiles observed upon overexpression of different URT1 versions in N. benthamiana is that the interaction of URT1’s M1 motif with DCP5 promotes the degradation of oligoadenylated mRNAs (and possibly mRNAs with heteropolymeric A-rich tails). We therefore propose that the conserved M1 motif in the N-terminal IDR of URT1 contacts DCP5, which then recruits the RH6,8,12-VCS decapping activators. In addition, URT1 contacts the GIGYF-like EXA1, likely via URT1′s PPGF motif. EXA1 recruits further translational repressors and decapping activators like nCBP and PAT1. Altogether these interaction networks would facilitate decapping of mRNAs with oligo(A) tails larger than those typically required for LSM1-7 recruitment.

But why has a bypass of the LSM1-7 recruitment been selected during plant evolution? A likely reason is that in plants, excessive deadenylation could trigger the biogenesis of spurious siRNAs targeting endogenous mRNAs to PTGS. Indeed, the RNA-dependent RNA polymerase RDR6, the key enzyme converting “aberrant” RNA into double stranded RNA that will be diced into siRNAs, intrinsically favors fully deadenylated mRNAs over polyadenylated mRNAs as templates59. We therefore propose that a key role for URT1-mediated uridylation is to avoid the accumulation of excessively deadenylated mRNAs that otherwise erroneously enter the RNA silencing pathway. In line with this hypothesis, a connection between URT1 and RNA silencing was recently suggested by the identification of URT1 as a silencing suppressor of transgenes60. Our results indicate two additive modes of action to explain how URT1 limits the accumulation of excessively deadenylated mRNAs (Fig. 7). First, the direct URT1-DCP5 interaction mediates a molecular connection between a TUTase and decapping activators, thereby facilitating the 5′–3′ removal of oligoadenylated mRNAs with oligo(A) tails in the 10–25 A range. Second, uridylation per se can participate in preventing excessive deadenylation by slowing down deadenylation, at least the CAF1 activity tested in this study. Finally, we have previously shown that uridylation by URT1 repairs deadenylated mRNAs to restore an extension sufficient for the binding of a PABP12. Although it is yet unknown how the binding of PABP to uridylated oligo(A) tails influences deadenylation or translation, binding of a PABP may also protect mRNA 3′ extremity from being accessible to RDR6. Altogether, these data illustrate the complexity of uridylation-mediated processes. Although facilitating degradation emerges as the prototypical function of mRNA uridylation, the underlying molecular mechanisms are complex and may differ across eukaryotes. This diversity is yet to be fully explored.

Methods

Plant material. All Arabidopsis thaliana plants used in this study are of Columbia (Col-0) accession. T-DNA mutants were described previously: urt1-1 (SALK_087647C)11, urt1-2 (WiscDsloxHS208_08D)11, xrn4-3 (SALK_014209D)61,
xrn4-5 (SAIL_681E01) 96, del2-1 (SALK_064627) 97, del4-2 (GABI_160G05) 98, and dcp5-1 (SALK_008881) 99. The plant material used for RNA-seq and small RNA-seq corresponds to Arabidopsis plantlets grown for 24 days in vitro on Murashige & Skoog media with 0.8% agar and 12-h light/12-h darkness cycles (22/18 °C). For other analyses using Arabidopsis, flowers were harvested from plants grown on soil with 16-h light/8-h darkness cycles (21/18 °C). Agroinfiltration experiments were performed in leaves of Nicotiana benthamiana plants grown during 4 weeks on soil with 16-h light/8-h darkness cycles (22/18 °C).
Characterization of URT1 sequence and phylogeny. The characterization of URT1 sequences and phylogenetic analyses are detailed in Supplementary Method. All protein sequences used for Fig. 1 and Supplementary Fig. 1 are provided in Supplementary Data 2.

Oligonucleotides and plasmids. Oligonucleotides and expression plasmids used for cloning are listed in Supplementary Data 7. For plant transformation, the URT1 sequence was PCR-amplified from genomic DNA and includes 5’ UTR (URT1-myc constructs) or 3’ UTR (myc-URT1 and YFP-URT1 constructs). For bacterial expression, URT1, DCP5, and CAF1b vectors were PCR-amplified from CDNA templates.

Co-immunopurifications (IP). Details about samples and replicates of co-immunopurification (IP) experiments are provided in Supplementary Data 2e. For IPs without crosslinking on Arabidopsis samples, 300 mg of flower buds or seedlings were ground in 1.5 ml of ice-cold lysis buffer (50 mM Tris-HCl pH 8, 50 mM NaCl, 1% Triton X-100, protease inhibitors (cComplete, EDTA-free Protease Inhibitor Cocktail, Roche)). After cell debris removal by centrifugation (twice 10 min at 16,000 g, 4 °C), supernatants were incubated for 1 h at 4 °C. The membrane was washed 50% with 0.375% formaldehyde (Thermo Fisher Scientific). Beads loaded on magnetized µMACS separation columns equilibrated with 0.05% Tween 20, 50 mM Tris-HCl pH 7.5, 1% Triton X-100, 0.375% formaldehyde and transferred out the 50% less abundant proteins that could adversely affect the dispersion estimation. The size factor used to scale samples were calculated according to the method and statistics were performed using the Proteome Discoverer software (v2.3, Thermo Scientific) with the following parameters: Sequest and MS algorithms including a FDR at 1%, “Top 3 Average” method, no imputation.

In vitro pull-down assays. Recombinant 6His-GST, 6His-GST-URT1, 6His-GST-m1URT1, 6His-GST-m1URT1-D491/3A, 6His-MBP-DCP5, 6His-MBP-UTR1, 6His-ALSMDCP5, and 6His-MBP-DCP5-LSM were expressed into Escherichia coli BL21 DE3 using plasmids listed in Supplementary 7. Conditions for protein expression and purification are detailed in Supplementary Methods. Total protein was incubated in a final volume of 50 ml of 20 mM MOPS pH 7.2, 100 mM KCl, 15% glycerol, and 0.1% Tween 20 with 100 ng/ml of RNase A for 10 min at 4 °C under rotation. Eighty microliters of glutathione sepharose resin (GE healthcare) were added to each reaction and incubated under rotation for 1 h at 4 °C. The resin was sedimented at 500 g for 5 min and washed five times with 500 µl of the same buffer without RNase A. The elution was performed by adding 60 µl of elution buffer 20 mM MOPS pH 7.2, 100 mM KCl, 15% glycerol and 0.1% Tween 20, 10 mM reduced glutathione (Sigma-Aldrich) and incubated 5 min at 4 °C before elution by centrifugation at 500 g for 5 min. Eluted proteins were concentrated with a 10,000 MWCO centrifugal filter unit using SYN Bio (Bio-Rad) and an Attan DIGE imager (Amersham Biosciences) or Fusion FX camera system (Vilber) was used for visualization.

In vitro activity assays. Recombinant 6His-GST-CAF1b and 6His-GST-CAF1bD42A proteins were expressed in E. coli BL21 using plasmids listed in Supplementary 7. Conditions for protein expression and purification are detailed in Supplementary Methods. The deadenylation test was performed in biological triplicates in 20 mM MOPS pH 7.2, 5 mM MgCl2, 50 mM KCl, 7% glycerol, and 0.1% Tween 20. The oligoribonucleotides CACCAACCACU-A14, CACCAACCACU-A12U2, and CACCAACCACA-U14 were used as substrates. Purified 6His-GST-CAF1b and 6His-GST-CAF1bD42A proteins (30 nM) were incubated with 17.5 nM of radiolabelled substrate for 1 h at 25 °C. Aliquots were taken at different time points and separated on a 17% polyacrylamide/7 M urea gel before autoradiography. The amount of intact substrate RNA was estimated as detailed in Supplementary Methods.

Agroinfiltration experiments in N. benthamiana. Agrobacterium tumefaciens GV3101 (pMP90) were transformed with plant expression plasmids listed in Supplementary 7 and inoculated in 10 ml of LB for 20 h at 28 °C. Pre-cultures were then centrifuged at 5000 g for 15 min. The pellets were resuspended at an OD600 of 1 in 5 ml of agroinfiltration buffer (10 mM MgCl2 and 250 µM of 3′,5′-Dimethoxy-4′-hydroxyacetophenone (Sigma-Aldrich)). The cell suspensions containing the P19, URT1, and GFP constructs were mixed to a 1:1:1 ratio and infiltrated into N. benthamiana leaves using needleless syringes. The plants were harvested 4 days after infiltration for RNA and protein extraction. Pictures of the infiltrated leaves were taken under UVI illumination at 365 nm using a UVP Blak-Ray B-100Y UV lamp (Thermo Fisher Scientific) to detect the expression of the GFP reporter. The intensity of the GFP fluorescence was quantified using ImageJ (see details in Supplementary Methods).

Western blot analysis. N. benthamiana infiltrated leaf patches (four different leaf samples) were grounded in SDS-urea extraction buffer (62.5 mM Tris pH 6.8, 4 M urea, 3% SDS, 10% glycerol, 0.01% ammonium thiocyanate). The samples were separated by SDS-PAGE and electrotransferred to a 0.45 µm Immobilon-P PVDF membrane (Millipore). The membrane was incubated overnight at 4 °C with DCP5 antibodies (provided by Rémy Merret and Cécile Bouquet-Antonioli, used at 1/10,000 dilution), URT1 antibodies (described in ref. 12 and used at 1/1,000 dilution), MBP antibodies (Invitrogen, used at 1/5000 dilution), monoclonal c-myc antibody (Roche, used at 1/10,000 dilution). Polyclonal and monoclonal antibodies were detected by goat anti-rabbit or anti-mouse IgG coupled to peroxidase (Invi- trodetection) used at 1/10,000 dilution, respectively, and visualized using a DAB+ Blotting Substrate (Roche). Pictures were taken with a Fusion FX camera system. The PVDF membranes were stained with 0.1% Coomassie Brilliant Blue R-250, 75% acetic acid, 50% methanol) to monitor loading.

Northern blot analysis. For each sample, four infiltrated patches pooled from different leaves were harvested from N. benthamiana four days after agroinfiltration. Total RNA was extracted using TRI-Reagent (Molecular Research Center), adjusted using the Benjamini–Hochberg method from stats R package. The gene ontology analysis for URT1 co-purifying proteins was performed using the Functional Annotation tool (http://www.broadinstitute.org/geneontology) with the following parameters: Sequest and MS algorithms including a FDR at 1%, “Top 3 Average” method, no imputation.
samples. After destaining, membranes were incubated with PerfectHyb hybridization buffer (Sigma-Aldrich) for 30 min at 65 °C and hybridized overnight at 65 °C. Signals of 10 µl were detected using DR2 (Nikon) and exposed to Phosphor screen and visualized with an Amersham Typhoon IP Biomolecular Imager (GE Healthcare Life Sciences).

3' RACE-seq library preparation and data processing. Total RNA was extracted using Tri-Reagent (Molecular Research Center) from N. benthamiana infiltrated leaves or Arabidopsis flowers. Ten pmol of a 5′-ribodylenated DNA oligonucleotide (3′-Adapt RACEseq, Supplementary Data 7) were ligated to 5 µg of total RNA using 10 U of T4 ssRNA Ligase 1 (NEB) in a final volume of 50 µl at 1 h at 37 °C and 1X T4 of RNA Ligase Reaction Buffer (NEB, 50 mM Tris-HCl pH 7.5, 10 mM MgCl2, 1 mM DTT). The ligation products were purified from reagents and non-ligated adapter molecules with Nucleospin RNA Clean-up columns (Macherey Nagel). cDNA synthesis was performed in 20 µl reaction that contains 2-3 µg of purified ligated RNA, 50 pmol of the 3′-RT oligonucleotide (Supplementary Data 7), 10 nmol of dNTP, 0.1 µl of DTT, 40 U of RNaseOUT (Invitrogen), 200 U of SuperScript IV reverse transcriptase (Invitrogen), and 1X of SuperScript IV RT buffer (Invitrogen). Reactions were incubated at 50 °C for 10 min, and then at 80 °C for 10 min in an incubator to inactivate the reverse transcriptase. Two nested PCR amplifications round PCRs 1–20, 30–30 cycles, respectively, were then performed. PC1 was run using 0.5–2 µl of cDNA, 10 pmol of gene-specific primer (Supplementary Data 7), 10 pmol of RACEseq_rev1 primer (Supplementary Data 7), 10 nmol of dNTP, 1 U of GoTaq DNA Polymerase (Promega), and 1X of Green GoTaq Reaction Buffer (Promega) in a 200 µl reaction. The conditions for PCR1 were as follows: a step at 94 °C for 30 s; 30 cycles at 94 °C for 20 s, 50 °C for 20 s, and 72 °C for 30 s; a final step at 72 °C for 30 s. PC2 was performed using 1 µl of PCR1 product, 10 pmol of gene-specific primer (Supplementary Data 7) and 10 pmol of a TruSeq DNA RNA adapter (RII, Supplementary Data 7), 10 nmol of dNTP, 1 U of GoTaq DNA Polymerase (Promega), and 1X of Green GoTaq Reaction Buffer (Promega) in a 200 µl reaction. The conditions for PCR2 were as follows: a step at 94 °C for 1 min; 20–30 cycles at 94 °C for 30 s, 56 °C for 20 s, and 72 °C for 30 s; a final step at 72 °C for 30 s. All PC2 products were purified using one volume of AMPure XP beads (Agenourt). Library was pooled-end sequenced with MiSeq (v3 chemistry) with 41 × 111 bp cycle settings. The library preparations were sequenced on a HiSeq 2500 (HiSeq high-output mode, 150 bp) (Illumina) in a single read using the Illumina R1.1.1.0.20 and CASAVA pipeline v1.8.2. Details about further data processing are described in Supplementary Methods. The numbers of sequences obtained at each processing step of TAIL-seq libraries and the calculated uridylation percentages are provided in Supplementary Data 5.

Small RNA-seq library preparation and data processing. Total RNA was extracted using Tri-Reagent (Molecular Research Center) from WT, urt1-1, smrt-3, and urt1-1 smrt-3 24-day-old seedlings by two biological replicates each, and subsequently treated with DNase I (Thermo Fisher Scientific). Small RNA libraries were prepared and sequenced at Fasters (http://www.fasters.com). Libraries were generated from 3 µg of RNA-treated using the Illumina TruSeq Small RNA protocol after size selection of 18–30 nt RNA fragments on a denaturing polyacrylamide gel. Libraries were sequenced on a HiSeq 2500 (HiSeq high-output mode, 150 bp) (Illumina). The base calling and alignment of RNA reads were performed using the topHat and clipAlign tool and the Trimmomatic v 0.34 tool. The numbers of sequences obtained for small RNA-seq libraries and the list of mRNA loci that show differential small RNA accumulation are provided in Supplementary Data 6.

Statistics and reproducibility. Statistical analyses were performed using R 3.6.1, and the following R packages: edgeR 3.26.5, stats 3.6.1, multcomp 0.1-8. For all analyses, a p value of 0.05 was defined as the threshold of significance. For each figure, the exact value of n and the test used for the statistical analysis is featured in the figure or in the corresponding legend. Fold change and p values in Figs. 3a, b, and Supplementary Figs. 3b, c were computed using the Negative binomial generalized log-linear model implemented in the edgeR package67. Statistical significance shown in Fig. 3 and Supplementary Figs. 3b, 4b was obtained using Pairwise Wilcoxon Rank-Sum tests with data considered as paired (two-tailed). Statistical significance shown in Figs. 3d and 6f was obtained using Pairwise Wilcoxon Rank-Sum tests (two-tailed) with data considered as unpaired. Statistical significance shown Supplementary Fig. 3d was obtained using two-tailed t-test (Proteome Discover software v2.3, Thermo Scientific). For box plot analyses in Figs. 3 and 6, the upper and lower edges correspond to the first and third quartiles, respectively. The median is indicated by a horizontal bar and whiskers show data range except far-outliers. Differential statistical analysis shown in Fig. 5d was performed using the edgeR package and its implemented negative binomial generalized log-linear model. Half-lives shown in Fig. 5c and Supplementary Fig. 5b were calculated using a quasi-Poisson regression model of the R stats package. All p values were adjusted using the Benjamini–Hochberg method. For each graph, the number of independent replicates is indicated directly on the figures. For Figs. 2d, 3b, 3d, 4b, 5c and Supplementary Figures 2a–c, 4a, representative images of 2, 2, 3, 3, 3, 4, 2, and 2 independent experiments, respectively, are shown.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability. NGS datasets generated during this study have been deposited in NCBI’s Gene Expression Omnibus and are accessible through GEO Series accession number GSE148449. GEO Series accession numbers for individual datasets are GSE148406 for 3′ RACE-seq in Arabidopsis, GSE148409 for 3′ RACE-seq in N. benthamiana, GSE148417 for TAIL-seq, and GSE148427 for small RNA seq.
Mass spectrometry proteomics raw data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository\(^3\) with dataset identifiers PXD018672 and PXD002676 for Arabidopsis and N. benthamiana, respectively.

Raw data for Nanopore DRS have been deposited at ENA with the accession number PRJEB40438.

Source data for all figures in the paper, including raw data underlying graphs and uncropped versions of gels or blots presented in the figures, are available as Mendeley data: https://doi.org/10.17632/bcv3cm9.3.

The raw intensity files (.cif files) used to test the TAILseeker3 software (results shown in Supplementary Fig. 3) have not been deposited in a public repository because of their large size but are available from the corresponding author on request.

Web links for associated raw data are indicated in each figure legend. All biological materials used in this study, plant lines, and plasmids are available from the authors or the indicated sources.

**Code availability**

Bioinformatic pipelines including python and bash source code for 3\(^\prime\)UTR and TAIL-seq analyses are available as Mendeley data: https://doi.org/10.17632/v8d9bd692c.1.

Received: 12 May 2020; Accepted: 19 January 2021; Published online: 26 February 2021

**References**

1. De Almeida, C. et al. RNA uridylation and decay in plants. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 373, 20180163 https://doi.org/10.1098/rstb.2018.0163 (2018).

2. Zgukóvá, D. & Vaháčová, Š. The role of 3′ end uridylation in RNA metabolism and cellular physiology. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 373, 20180171 https://doi.org/10.1098/rstb.2018.0171 (2018).

3. Scheer, H., Zuber, H., De Almeida, C. & Gagliardi, D. Uridylation earmarks mRNAs for degradation… and more. *Trends Genet.* 32, 607–619 (2016).

4. De Almeida, C., Scheer, H., Zuber, H. & Gagliardi, D. RNA uridylation: a key post-transcriptional modification shaping the coding and non-coding transcriptome. *WIREs RNA* e1440 https://doi.org/10.1002/wrna.1440 (2018).

5. Rissland, O. S. & Norbury, C. J. Decapping is preceded by 3′UTR and 3′end trimming of oligoadenylated and polyadenylated RNAs. *RNA* 13, 998–1016 (2007).

6. Malecki, M. et al. The exoribonuclease DIs3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* 32, 1842–1854 (2013).

7. Hashimoto, M. et al. EXA1, a GYF domain protein, is responsible for loss-of-susceptibility to plantago asiatica mosaic virus in Arabidopsis thaliana. *Plant J.* 88, 120–131 (2016).

8. Kofer, M., Motzny, K. & Freund, C. GYF domain proteomes reveals interaction sites in known and novel target proteins. *Mol. Cell. Proteom.* 4, 1797–1811 (2005).

9. Matsui, H. et al. The GYF domain protein PSIG1 dampens the induction of cell death during plant-pathogen interactions. *PLoS Genet.* 13, e1007037 (2017).

10. Wu, Z. et al. Regulation of plant immune receptor accumulation through translational repression by a glycine-tyrosine-phenylalanine (GYF) domain protein. *eLife* 6, e26364. https://doi.org/10.7554/eLife.26364 (2017).

11. Brandmann, T. et al. Molecular architecture of LSM14 interactions involved in the assembly of mRNA silencing complexes. *EMBO J.* 37, e97869 (2018).

12. Rajyaguru, P., She, M. & Parker, R. Scd6 targets eLF4G to repress translation: RGG motif proteins as a class of eLF4G-binding proteins. *Mol. Cell* 45, 244–254 (2012).

13. Fromm, S. A. et al. The structural basis of Edc3- and Scd6-mediated activation of the Dcp1/Dcp2 mRNA decay complex. *EMBO J.* 31, 279–290 (2012).

14. Wang, X. et al. Synergistic and independent actions of multiple terminal nucleotidyl transferases in the 3′ tailing of small RNAs in Arabidopsis. *PLoS Genet.* 11, e1005091 (2015).

15. Arae, T. et al. Identification of Arabidopsis CCR4-NOT complexes with Pumilio RNA binding proteins, APUM5 and APUM2. *Plant Cell Physiol.* https://doi.org/10.1093/pcp/pcz089 (2019).

16. Tang, T. T., Stowell, J. A., Hill, C. H. & Passmore, L. A. The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Csf1 deadenylylases. *Nat. Struct. Mol. Biol.* 26, 433–442 (2019).

17. Narsai, R. et al. Genome-wide analysis of mRNA decay rates and their gene expression dynamics of the Arabidopsis transcriptome. *Plant Cell* 19, 3413–3436 (2007).

18. Sorensen, R. S., Deshotel, M. J., Johnson, K., Adler, F. R. & Sieburth, L. E. Arabidopsis mRNA decay landscape arises from specialized RNA decay substrates, decapping-mediated feedback, and redundancy. *Proc. Natl Acad. Sci. USA.* 115, E1485–E1494 (2018).

19. Szabo, E. et al. Metabolic labeling of RNAs uncovers hidden features and dynamics of the Arabidopsis transcriptome. *Plant Cell* 32, 871–887 (2020).

20. Gregory, B. D. et al. A link between RNA metabolism and silencing affecting Arabidopsis development. *Dev. Cell* 44, 854–866 (2008).

21. Martinez de Alba, A. E. et al. In plants, decapping prevents RDR6-dependent production of small interfering RNAs from endogenous mRNAs. *Nucleic Acids Res.* 43, 2902–2913 (2015).

22. Moreno, A. B. et al. Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-transcriptional gene silencing in plants. *Nucleic Acids Res.* 41, 4699–4708 (2013).

23. Parent, J.-S. et al. Post-transcriptional gene silencing triggered by sense transgenes involves uncapped antisense RNA and differs from silencing intentionally triggered by antisense transgenes. *Nucleic Acids Res.* 43, 8464–8475 (2015).

24. Tsuzuki, M., Motomura, K., Kumakura, N. & Takeda, A. Interconnections between mRNA degradation and RDR-dependent siRNA production in mRNA turnover in plants. *J. Plant Res.* 130, 211–226 (2017).

25. Zhang, X. et al. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. *Science* 348, 120–123 (2015).

26. Kwak, J. E. & Wickens, M. A family of poly(U) polymerases. *RNA* 13, 860–867 (2007).

27. Warocki, Z., Liudkovska, V., Gewartowska, O., Mroczek, S. & Dziembowski, A. Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 373, 20180162 https://doi.org/10.1098/rstb.2018.0162 (2018).
46. Hagan, J. P., Piskounova, E. & Gregory, R. I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat. Struct. Mol. Biol. 16, 1021–1025 (2009).

47. Heo, I. et al. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell 151, 521–532 (2012).

48. Thornton, J. E., Chang, H.-M., Piskounova, E. & Gregory, R. I. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA 18, 1875–1885 (2012).

49. Reimão-Pinto, M. M. et al. Molecular basis for cytoplasmic RNA surveillance by uridylation-triggered decay in Drosophila. EMBO J. 35, 2417–2434 (2016).

50. Jonas, S. & Iazaurralde, E. The role of disordered protein regions in the recruitment of translation repressors. EMBO J. 29, 2192–2204 (2010).

51. Peter, D. et al. GIGYF1/2 proteins use auxiliary sequences to selectively bind in mammalian development. Mol. Cell 48, 79–92 (2012).

52. Morita, M. et al. A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. Mol. Cell. Biol. 32, 3585–3593 (2012).

53. Montemayor, E. J. et al. Molecular basis for the distinct cellular functions of the Lsm1–7 and Lsm2–8 complexes. RNA 26, 1400–1413 (2020).

54. Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M. & Séraphin, B. A Sm-like protein complex that participates in mRNA degradation. EMBO J. 19, 1661–1671 (2000).

55. Tharun, S. & Parker, R. Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. Mol. Cell 8, 1075–1083 (2001).

56. Tharun, S. et al. Yeast Sm-like proteins function in mRNA decapping and decay. Nature 404, 515–518 (2000).

57. Lobel, J. H., Tibble, R. W. & Gross, J. D. Pat1 activates late steps in mRNA decay by multiple mechanisms. Proc. Natl Acad. Sci. USA 116, 23512–23517 (2019).

58. Xu, J. & Chua, N.-H. Arabidopsis decapping 5 is required for mRNA decapping, P-body formation, and translational repression during postembryonic development. Plant Cell 21, 3270–3279 (2009).

59. Baeg, K., Iwakawa, H.-O. & Tomari, Y. The pol(A) tail blocks RDR6 from converting self mRNAs into substrates for gene silencing. Nat. Plants 3, 17036 (2017).

60. Li, T. et al. A genetics screen highlights emerging roles for CPL3, RST1 and URT1 in RNA metabolism and silencing. Nat. Plants 5, 539–550 (2019).

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

62. Souret, F. F., Kastenmayer, J. P. & Green, P. J. ADRXN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. Mol. Cell 15, 173–183 (2004).

63. Xie, Z. et al. Genetic and functional diversification of small RNA pathways in plants. PLoS Biol. 2, e104 (2004).

64. Nakazawa, Y., Hiraguri, A., Moriyama, H. & Fukuhara, T. The dsRNA-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA 18, 1875–1885 (2012).

65. Reimão-Pinto, M. M. et al. Molecular basis for cytoplasmic RNA surveillance by uridylation-triggered decay in Drosophila. EMBO J. 35, 2417–2434 (2016).

66. Nakazawa, Y., Hiraguri, A., Moriyama, H. & Fukuhara, T. The dsRNA-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA 18, 1875–1885 (2012).

67. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010).

68. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DEXSeq. Genome Biol. 15, 550 (2014).

69. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nat. Biotechnol. 26, 190–190 (2008).

70. Bilska, A. et al. Immunoglobulin expression and the humoral immune response is regulated by the non-canonical poly(A) polymerase TENT5C. Nat. Commun. 11, 2032 (2020).

71. Gwaterowska, O. et al. Cytoplasmic polyadenylation by TENT5A is required for proper bone formation. bioRxiv. https://doi.org/10.1101/2020.08.18.256115 (2020).

72. Edgar, R. Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30, 207–210 (2002).

73. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 47, D442–D450 (2019).

74. Walsh, I., Martin, A. J. M., Di Domenico, T. & Tosatto, S. C. E. ESpritz: accurate and fast prediction of protein disorder. Bioinformatics 28, 503–509 (2012).

75. Iversen, D., Walsh, I., Minervini, G. & Tosatto, S. C. E. FELTS: fast estimator of latent local structure. Bioinformatics 33, 1889–1891 (2017).

76. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. 38, W529–W533 (2010).