HAL
open science

Interplay between uridylation and deadenylation during mRNA degradation in Arabidopsis thaliana

Caroline de Almeida

To cite this version:

Caroline de Almeida. Interplay between uridylation and deadenylation during mRNA degradation in Arabidopsis thaliana. Vegetal Biology. Université de Strasbourg, 2019. English. NNT : 2019STRAJ126 . tel-03191334

HAL Id: tel-03191334

https://theses.hal.science/tel-03191334v1

Submitted on 7 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

Institut de Biologie Moléculaire des Plantes – CNRS – UPR2357

THÈSE présentée par :

Caroline DE ALMEIDA

soutenue le : **30 septembre 2019**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : Aspects Moléculaires et Cellulaires de la Biologie

Interplay between uridylation and deadenylation during mRNA degradation in *Arabidopsis thaliana*

Directeur de thèse :	Dr GAGLIARDI Dominique
Rapporteur Externe :	Pr VANACOVA Štěpánka
Rapporteur Externe :	Pr WACHTER Andreas
Examinateur Interne :	Dr MAUXION Fabienne
Membre Invité :	Dr ZUBER Hélène

Acknowledgments

First, I would like to thank Štěpánka Vaňáčová, Andreas Wachter and Fabienne Mauxion for having accepted to evaluate my work. Merci également au LabEx NetRNA d'avoir financé ma thèse.

J'aimerais remercier mon directeur de thèse, Dominique Gagliardi (« Gag »), de m'avoir accueilli au sein de son équipe et de m'avoir soutenu tout au long de ma thèse. Je te remercie pour ton expertise, tes précieux conseils et les (longues… ;-)) discussions scientifiques qui m'ont beaucoup apporté et aidé à développer mon esprit critique et scientifique.

Un grand merci à Hélène Zuber pour sa gentillesse et sa patience durant ces 4 années. Merci pour ton soutien moral et intellectuel, tu as toujours été là pour répondre à mes innombrables questions et tu as toujours su m'encourager quand ça n'allait pas trop. Et merci aussi pour ta patience et tes bons conseils lors de l'écriture de cette thèse. Tu es une personne et une scientifique remarquable, je suis confiante que tu iras loin dans ta vie personnelle et professionnelle !

Ein besonderer Dank gilt Heike für ihre Betreuung und ihre guten Ratschläge. Ich danke dir aufrichtig für deinen konstruktiven Beitrag und deine unermüdliche Unterstützung, die mir bei der Umsetzung dieses Manuskripts sehr geholfen haben. Hättest du mich damals nicht als Praktikantin genommen, wäre ich wohl jetzt nicht hier. Du bist eine sehr aufmerksame und hilfsbereite Person, die ich sehr schätze und nicht so schnell vergessen werde.

Merci à Marlène et Hélène d'avoir été là ! Vous êtes des amies précieuses que je ne suis pas prête à oublier ! Merci Marlène pour ton grain de folie, ta bonne humeur et ton honnêteté, et pour tous ces moments passés au labo et surtout en dehors du labo avec Hélène ;-) ! Merci Hélène, je n'oublierais jamais ces bons moments de rire qu'on a passés ensemble au labo ! Merci pour ton soutien et ta présence dans les moments un peu moins joyeux, et toute ton aide durant la thèse qui n'aurais pas été la même sans toi. J'espère que tu auras tout ce que tu mérites dans ta vie personnelle et professionnelle, tu as toutes les capacités pour aller très loin.

Merci aux autres membres actuels ou anciens de l'équipe, Damien, Clara, Natka, Pierre, Simon, Anne-Caroline, Marie, Aude et Gaétan pour votre aide et pour votre bonne humeur. Merci à Laure pour ta gentillesse, tu es toujours prête à aider et à l'écoute. Merci à mes stagiaires Florence, Marine et Quentin, je vous souhaite beaucoup de réussite pour la suite.

Merci à l'équipe de la plateforme de séquençage pour votre aide et conseils, merci à tous les membres de la plateforme de production de plante, Michèle de la laverie, Laurence du magasin, et les membres de la plateforme de protéomique de l'IBMC pour votre expertise et précieuse aide.

Merci à tous les autres formidables personnes que j'ai pu connaitre à l'IBMP, Marion, Marco, Thibaut, Magdalène, Lucie, Chloé, Adrien, Valentin, Habib, Arnaud, Herrade (et Thibaud), Xue, Cédric, Nicolas, Marcel, Vianney, Mathilde, Anthony, Elodie, Sandrine, Thalia, Clémentine, Kamel, Esther … et tous les autres. Aux « gurlzzz et boyyzz » de la salle de rédaction, merci pour votre soutien, votre gentillesse et les bons moments passés ensemble. Ne lâchez rien vous allez tous y arrivez !

Merci à mes bons amis Déborah, Simon, Nicolas, Yannick et Pierre, qui ont toujours été là pour moi, et qui (surtout à la fin) m'ont toujours supporté malgré ma mauvaise humeur « occasionnelle » ;-). Merci à ceux que j'ai connu à l'université, Léa et Yacine, et particulièrement mes amis et anciens camarades de licence Urša, Nadjwa, Carole, Théo, Daniel, Charles, et surtout mon cousin Gilles et Nicolas. J'ai passé de superbes moments avec vous et je ne vous oublierais jamais.

Merci dem Natascha, Jill, Annick, Laurent, Isabelle an Caroline, dass dier ëmmer Kontakt mat mir gehalen hutt obwuel ech säit e puer Joren elo net méi am Land sinn. Merci fir är Ënnerstëtzung an är positiv Wieder.

Mon plus grand merci ira à ma famille qui m'a toujours soutenue et encouragé. Merci à mes parents et mes sœurs pour votre patience tout on long de ma scolarité, et particulièrement durant ces 4 dernières années ! Merci de m'avoir appris à être persévérante et patiente et à toujours essayer de donner le meilleur de moi-même. Un grand merci à Florent d'avoir toujours cru en moi, tu as toujours été à l'écoute et tu as toujours su m'encourager. Merci pour toutes ces découvertes musicales et culturelles que nous avons pu faire ces dernières années. Merci à toi et ma famille d'être toujours à mes cotées, sans vous je ne serai sans doute pas arrivée jusqu'ici.

Table of contents

Introduction **1**

1. Post-transcriptional RNA modifications. **2**

1.1. Uridylation, a widespread modification of non-coding and coding RNAs. 3

1.1.1. Key features of terminal uridylyltransferases. 3

1.1.2. The N-terminal IDR of Arabidopsis URT1 is required for the binding of partner proteins. 5

1.1.3. Molecular functions of RNA uridylation. 6

2. Essential protective elements of eukaryotic mRNAs. **61**

2.1. Co-transcriptional addition of the 5' cap and 3' polyA-tails. 61

2.1.1. The 5’ cap structure. 61

2.1.2. The 3’ polyA tail. 62

2.2. The cap binding complex and polyA binding proteins. 63

2.2.1. The nuclear cap binding complex CBC. 64

2.2.2. Nuclear polyA binding proteins PABPN. 65

2.2.3. The cytoplasmic cap binding complex eIF4F. 66

2.2.4. Cytoplasmic polyA binding proteins (PABPC). 67

2.2.5. eIF4F and PABPC promote translation initiation. 67

3. Deadenylation, the rate-limiting step of bulk mRNA decay. **69**

3.1. The polyA ribonuclease PARN. 69

3.2. CCR4-like Nocturnin (NOC). 71

3.3. PolyA specific ribonuclease complex PAN2/3. 71

3.4. The main deadenylase complex CCR4/NOT. 72

3.4.1. Architecture of the CCR4/NOT complex. 72

3.4.2. Functions of the CCR4/NOT complex. 73

3.4.3. Recruitment of CCR4/NOT to mRNA. 74

3.4.4. Deadenylation by the CCR4/NOT complex. 75

3.4.5. CCR4/NOT complexes in plants. 77

4. The main factors involved in bulk mRNA decay. **79**

4.1. The 3’ to 5’ degradation pathway. 79

4.2. The 5’ to 3’ degradation pathway. 81

4.2.1. Deadenylation-independent decapping of mRNAs. 81

4.2.2. Deadenylation-dependent mRNA decapping. 82

Thesis objectives **85**

Results **89**

1. Optimisation of TAIL-seq and 3'RACE-seq methods to analyse mRNA polyA tail length and 3' modifications. 90

1.1. PolyA tails size estimation by Tailseeker and basecalling analysis pipelines. 91

1.2. First analysis of the new Arabidopsis TAIL-seq dataset. 92

1.3. 3'RACE-seq enables the deep analysis of target mRNA 3' ends. 93

2. Assessing the role of URT1 by transient expression assays in *Nicotiana benthamiana* leaves. 95

2.1. Overexpression of AtURT1 in *Nicotiana benthamiana* changes the polyA tail distribution of reporter *GFP* mRNAs. 95

2.1.1. Overexpression of active URT1 increases the uridylation level of the reporter GFP mRNA. 96

2.1.2. Expression of URT1-myc changes the polyA tail size distribution of reporter *GFP* mRNA. 97

2.1.3. The polyA size distribution is modulated by URT1-mediated uridylation. 98

2.2. Functional implication of the intrinsically disordered region of URT1 in the turnover of short polyA tails. 99

2.2.1. The removal of the IDR reduces uridylation levels of reporter *GFP* mRNA. 100

2.2.2. The N-terminal IDR is important for the regulation of polyA tail sizes by URT1. 101

2.2.3. *GFP* mRNAs with A-rich tails and G/A or U/A extensions accumulate when inactive forms of URT1-myc are overexpressed. 101

2.2.4. Uridylation by URT1 does not destabilise the *GFP* reporter mRNA. 102

2.3. Overexpression of active and inactive versions of URT1 impacts tail size and uridylation levels of endogenous *PR2* mRNAs in *Nicotiana benthamiana*. 103

2.3.1. Expression of URT1 leads to accumulation of *NbPR2* transcripts with longer polyA tails. 103

2.3.2. Expression of URT1 increases the uridylation level of endogenous *PR2* transcript. 104

2.3.3. Overexpression of URT1$^{D491/3A}$-myc results in increased levels of *PR2* mRNAs with A-rich tails. 104

2.3.4. NbPR2 mRNAs accumulate in URT1$^{D491/3A}$-myc samples. 105

2.4. The M1 motif is required for the accumulation of mRNAs with short polyA tails upon overexpression of URT1$^{D491/3A}$-myc. 106

2.4.1. The M1 motif is implicated in the regulation of polyA tail length. 106

2.4.2. The M1 motif is required to suppress the overaccumulation of uridylated short polyA tails induced by overexpression of inactive URT1-myc. 107

2.4.3. Overexpression of URT1$^{D491/3A}$M1-myc leads to the accumulation of A-rich polyA tails. 107

3. Terminal uridines impede the activity of CAF1b deadenylase. 109

4. Expression of URT1 leads to longer polyA tail profiles in *Arabidopsis thaliana*. 112

4.1. Uridylation frequencies vary between individual mRNAs. 112

4.2. mRNAs with higher uridylation frequencies have longer polyA tails. 113

5. The expression of HESO1-myc changes the polyA size distributions and decreases the steady-state level of reporter *GFP* mRNAs. 115

5.1. Expression of HESO1-myc does not increase mRNA uridylation levels but impacts the size distribution of polyA tails. 116
5.2. Expression of HESO1-myc affects polyA tail size distributions similar to the expression of URT1-myc. 117
5.3. Expression of HESO1-myc affects the steady-state of reporter *GFP* mRNAs. 118

Discussion 119

1. URT1-mediated uridylation controls the extent of mRNA deadenylation. 120

1.1. URT1 can uridylate longer polyA tails. 121
1.2. URT1 influences the turnover of mRNAs with short polyA tails. 122

2. The polyA size distribution is shaped by the binding of PABPs. 123

2.1. Highly uridylated mRNAs show a dispersed and phased polyA tail size distribution *in vivo*. 124

3. HESO1 and URT1 may have overlapping and specific functions in the turnover of mRNAs. 127

4. General Conclusion 129

Materials and Methods 131

Materials. 132

1. Plant materials. 132
2. Bacterial strains. 132
3. Vectors. 133
4. Antibodies. 134
5. Primers. 134

Table 1: Primers used for Gateway cloning. 134
Table 2: Primers used for 3'RACE-seq and TAIL-seq. 135
Table 3: Primers used for mutagenesis. 137

Methods. 138

1. Gateway cloning method. 138

1.1. PCR amplification of sequence of interest. 138
1.2. Gateway BP and LR reactions. 139
1.3. TOP10F' bacterial transformation. 139
1.4. Site-directed mutagenesis. 139

2. *N. benthamiana* leaf infiltration. 140

2.1. Agrobacterial transformation of GV3101 cells. 140
2.2. Agroinfiltration of *N. benthamiana* leaves. 140

3. Western blot analysis of protein extractions from *N. benthamiana* leaves. 141

3.1. Protein extractions. 141

3.2. SDS-PAGE. 141
3.3. Western blot. 141
3.4. Coomassie staining of polyacrylamide gels or PVDF membranes. 142
4. Northern blot analysis of total RNA extracts. 142
4.1. RNA extraction. 142
4.2. Agarose gel separation of total RNAs. 143
4.3. RNA transfer on Hybond N+ membrane 143
4.4. Methylene blue coloration. 143
4.5. Preparation of radiolabeled hybridisation probes. 144
4.6. Hybridisation with radiolabeled probes. 144
5. Protein production and purification. 144
5.1. Bacterial transformation of BL21 cells. 144
5.2. Batch production of fusion proteins. 145
5.3. Sypro Ruby staining. 146
6. *In vitro* deadenylation test. 146
6.1. 5' end radiolabelling of RNA substrates. 146
6.2. Deadenylation test. 147
6.3. UREA PAGE separation of substrates. 147
7. 3'RACE-seq library preparation. 147
7.1. 3' adapter ligation and reverse transcription. 147
7.2. Nested PCR amplification of 3'RACE-seq libraries. 148
7.3. Library purifications on AMPure beads. 148
8. TAIL-seq library reparation 149
8.1. DNase treatment and ribodepletion. 149
8.2. 3' adapter ligation and RNase T1 treatment. 150
8.3. Biotin pull-down, 5' phosphorylation and gel purification. 150
8.4. 5' adapter ligation and reverse transcription. 151
8.5. PCR amplification of librairies. 151
9. RACE-seq data analysis 152
10. TAIL-seq data analysis 153
10.1. Base calling-based pipeline 153
10.2. PolyA length estimation using Tailseeker software 153
10.3. Spike-in analysis 154

Supplementary data 155

Bibliography 188

Résumé en français 213

List of Figures

Introduction **1**

Figure 1: Domain architectures of TUTases across organisms. 4

Figure 2: Polyadenylation of mRNAs in humans. 65

Figure 3: Simplified scheme of translation initiation in eukaryotes. 67

Figure 4: The CCR4/NOT core subunits in human and yeast. 73

Figure 5: Functions associated to the CCR4/NOT complexes in nucleus and cytoplasm. 74

Figure 6: Recruitment of CCR4/NOT by specific RNA binding proteins to the mRNA 3' ends. 75

Figure 7: Model of mRNA deadenylation by CCR4/NOT. 76

Figure 8: mRNA de-circularisation by deadenylation. 77

Figure 9: Phylogenetic relationship between *Arabidopsis thaliana* CAF1 proteins. 77

Figure 10: Plant CCR4/NOT complex. 78

Results **89**

Figure 11: Schematic description of the experimental procedures of the TAIL-seq and 3'RACE-seq methods. 90

Figure 12: Estimation of polyA tail sizes using the Tailseeker or basecalling pipelines on the TAIL-seq dataset. 91

Figure 13: Analysis of the TAIL-seq wild-type dataset 92

Figure 14: Comparison between the 3'RACE-seq datasets and the TAIL-seq dataset. 93

Figure 15: Transient expression of URT1-myc and URT1$^{D491/3A}$-myc in *N. benthamiana*. 95

Figure 16: Uridylation levels of reporter *GFP* mRNAs. 97

Figure 17: Size repartition of the polyA tails of reporter *GFP* mRNAs. 97

Figure 18: Size distribution of all polyA tails of reporter *GFP* mRNAs. 97

Figure 19: Size distribution of unmodified polyA tails and uridylated polyA tails. 98

Figure 20: Expression of additional M1M2-myc, ΔIDR-myc and ΔIDR$^{D491/3A}$-myc constructs in *N. benthamiana*. 99

Figure 21: The N-terminal IDR is required for the accumulation of transcripts with short uridylated tails upon overexpression of URT1$^{D491/3A}$. 100

Figure 22: The N-terminal IDR is important for the regulation of polyA tail sizes by URT1. 101

Figure 23: Expression of URT1$^{D491/3A}$-myc leads to accumulation of short A-rich tails. 102

Figure 24: Northern blot analysis of reporter *GFP* mRNAs. 103

Figure 25: Size distribution of all tails on *NbPR2* mRNA. 103

Figure 26: Level of uridylated tails and tail size distribution of uridylated tails of *NbPR2* transcripts. 104

Figure 27: Size distribution and proportion of A-rich tails of *NbPR2* mRNA. 105

Figure 28: Northern blot analysis of reporter *NbPR2* mRNAs. 106

Figure 29: Transient expression of URT1$^{D491/3A}$-myc mutated in the M1 and/or M2 motifs in *N. benthamiana*. 106

Figure 30: Size distribution of all tails of *GFP* mRNA. 106

Figure 31: Size distribution of uridylated polyA tails of *GFP* mRNA. 107

Figure 32: Size distribution of A-rich tails of *GFP* mRNA. 107

Figure 33: Deadenylation reaction by hisGST-CAF1b. 110

Figure 34: Uridylation level of analysed Arabidopsis mRNAs. 112

Figure 35: Size distribution of uridylated tails of Arabidopsis mRNAs. 113

Figure 36: Size distribution of polyA tails of Arabidopsis mRNAs. 114

Figure 37: Transient expression of URT1-myc and HESO1-myc constructs in *N. benthamiana*. 115

Figure 38: Uridylation level of reporter *GFP* and endogenous *NbPR2* mRNAs. 116

Figure 39: Distribution of uridylated polyA tails of reporter *GFP* and endogenous *NbPR2* mRNAs. . 116

Figure 40: Distribution of all tails for *GFP* and *NbPR2* mRNAs. 117

Figure 41: Northern blot analysis of reporter *GFP* and *NbPR2* mRNAs. 118

Figure 42: Summary chart of the possible functions of URT1 and HESO1. 118

Discussion 119

Figure 43: URT1-mediated uridylation slows down deadenylation. 130

Materials and Methods 131

Figure 44: Flowchart of the 3'RACE-seq data analysis. 152

Supplementary data 155

High-resolution mapping of 3' extremities of RNA exosome substrates by 3' RACE-seq. 156

Table S1: Name of the core subunits of the CCR4/NOT complex in human, yeast, flies and Arabidopsis. 177

Figure S1: Number of reads for all and only uridylated polyA tails for *GFP* mRNAs. 178

Figure S2: A-rich tails with heteropolymeric extensions for *GFP* mRNAs. 179

Figure S3: Number of reads and number of uridines added to *NbPR2* mRNAs. 180

Figure S4: Size distribution of all tails of *NbPR2* mRNAs. 181

Figure S5: Size distribution of uridylated tails of *NbPR2* mRNAs. 182

Figure S6: Size distribution of A-rich tails of *NbPR2* mRNAs. 183

Figure S7: Number of reads for the *GFP* mRNAs in the 3'RACE-seq run3 dataset. 184

Figure S8: *In vitro* buffer conditions tested with the purified deadenylases. 185

Figure S9: Number of uridines added to the reporter *GFP* and *NbPR2* mRNAs. 186

Figure S10: Proportion of A-rich tails of reporter *GFP* and endogenous *NbPR2* mRNAs. 187

Introduction

1. Post-transcriptional RNA modifications.

Gene expression is fundamental for all living organisms. Transcription gives rise to a broad spectrum of RNA molecules that need to be maturated, transported, eventually translated and ultimately degraded. During their life cycle, both coding and non-coding RNAs are subjected to various co-transcriptional and post-transcriptional RNA modifications that orchestrate their metabolism and biological functions. RNAs undergo two types of modifications: chemical modifications of nucleotides within the RNA sequence, and modifications by tailing that consist in the non-templated addition of nucleotides to the RNA 3' extremities (De Almeida et al., 2018; Chang et al., 2014; Komarnitsky et al., 2000; Lim et al., 2018; Song and Yi, 2017; Zuber et al., 2016). For instance, important modifications of RNA polymerase II (Pol II) transcripts are the addition of the m7G Cap and the polyA tail to the 5' and 3' extremities, respectively (see Chapter 2.1).

In the last years, an increasing number of new types of RNA modifications have been found. Over 160 different RNA modifications have been described to date (Boccaletto et al., 2018). Some non-coding RNAs, like ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs), have many modified nucleotides that are required for their accurate maturation and function (Roundtree et al., 2017). An abundant chemical RNA modification is the methylation of adenosines at the nitrogen-6 position into N^6-methyladenosine or m6A, which is detected on long non-coding RNAs but also on messenger RNAs (mRNAs). Some RNA modifications are recognised by “readers”, such as the YTH (for YT521-B Homology) proteins that specifically recognise m6A. Recognition by YTHs can accelerate mRNA degradation through the recruitment of the deadenylase complex CCR4/NOT, but also modulate mRNA translation efficiency or regulate alternative splicing in mammals (Liao et al., 2018). Recently, m6A was shown to promote liquid–liquid phase separation of mRNAs bound by YTH proteins. The partition of m6A-mRNAs into phase-separated compartments such as P-bodies or stress granules participates in regulating the fate of these mRNAs (Ries et al., 2019). Certain chemical internal modifications can be removed by “erasers”, such as the enzymes FTO (fat mass and obesity-associated) or ALKBH5 (AlkB homolog 5) that demethylate m6A to regulate gene expression (Rajecka et al., 2019).

RNA modifications by tailing are catalyzed by terminal nucleotidyltransferases (TNTases) or ribonucleotidyltransferases (rNTases), and consist in the addition of untemplated guanosines, cytidines, adenosines or uridines to the 3’ ends of RNA. TNTases belong to the superfamily of DNA polymerase β-like nucleotidyltransferases which regroups all enzymes that add untemplated

nucleotides to proteins, antibiotics or RNAs (Belliot et al., 2008; Machín et al., 2001; Rohayem et al., 2006; Shirafuji et al., 1979).

The most studied tailing modification is adenylation which occurs in both eukaryotes and prokaryotes (Hajnsdorf and Kaberdin, 2018; Schmidt and Norbury, 2010; Song et al., 2015; Ustyantsev et al., 2017). The co-transcriptional polyadenylation of Pol II transcripts operates only in eukaryotes and is called "canonical" polyadenylation for historical reasons: it has been discovered first. Non-canonical adenylation refers to any untemplated addition of adenosines to the 3' end of both noncoding RNAs and mRNAs that is not catalyzed by the canonical polyA polymerases. The primordial role of non-canonical polyadenylation is to facilitate RNA degradation, because the non-structured tail serves as a landing pad for exoribonucleases (Glaunsinger and Lee, 2010; Lange et al., 2009; Levy and Schuster, 2016; Slomovic et al., 2010; Song et al., 2015; Wei et al., 1975; West et al., 2006; Zimmer et al., 2009). PolyA-mediated RNA degradation operates in bacteria, archaea and organelles. In addition, non-canonical adenylation earmarks the substrates of nuclear RNA surveillance by the eukaryotic exosome (Hilleren et al., 2001; Kilchert et al., 2016; LaCava et al., 2005; Ogami et al., 2018; Vanacova et al., 2007; Vaňáčová et al., 2005; Wyers et al., 2005; Zinder and Lima, 2017). Non-canonical polyadenylation is also observed on deadenylated mRNAs that are re-adenylated in the cytosol mostly during developmental transitions or in special cell types like oocytes or neurons (Charlesworth et al., 2013; Norbury, 2013).

1.1. Uridylation, a widespread modification of non-coding and coding RNAs.

Another pervasive post-transcriptional modification conserved in all eukaryotes investigated to date (except *Saccharomyces cerevisi*ae) is the addition of uridines, or uridylation. Uridylation is catalysed by terminal uridylyltransferases (TUTases), a protein family with various roles in the regulation of gene expression in different eukaryotes.

1.1.1. Key features of terminal uridylyltransferases.

All proteins of the TUTase family have a minimal catalytic core domain (CCD) composed of a polymerase β-like nucleotidyltransferase domain and a polyA polymerase-associated domain that has gained a specificity for uridines during evolution (Martin and Keller, 2007). In addition, most members of the TUTase family possess supplementary domains and regions that influence target

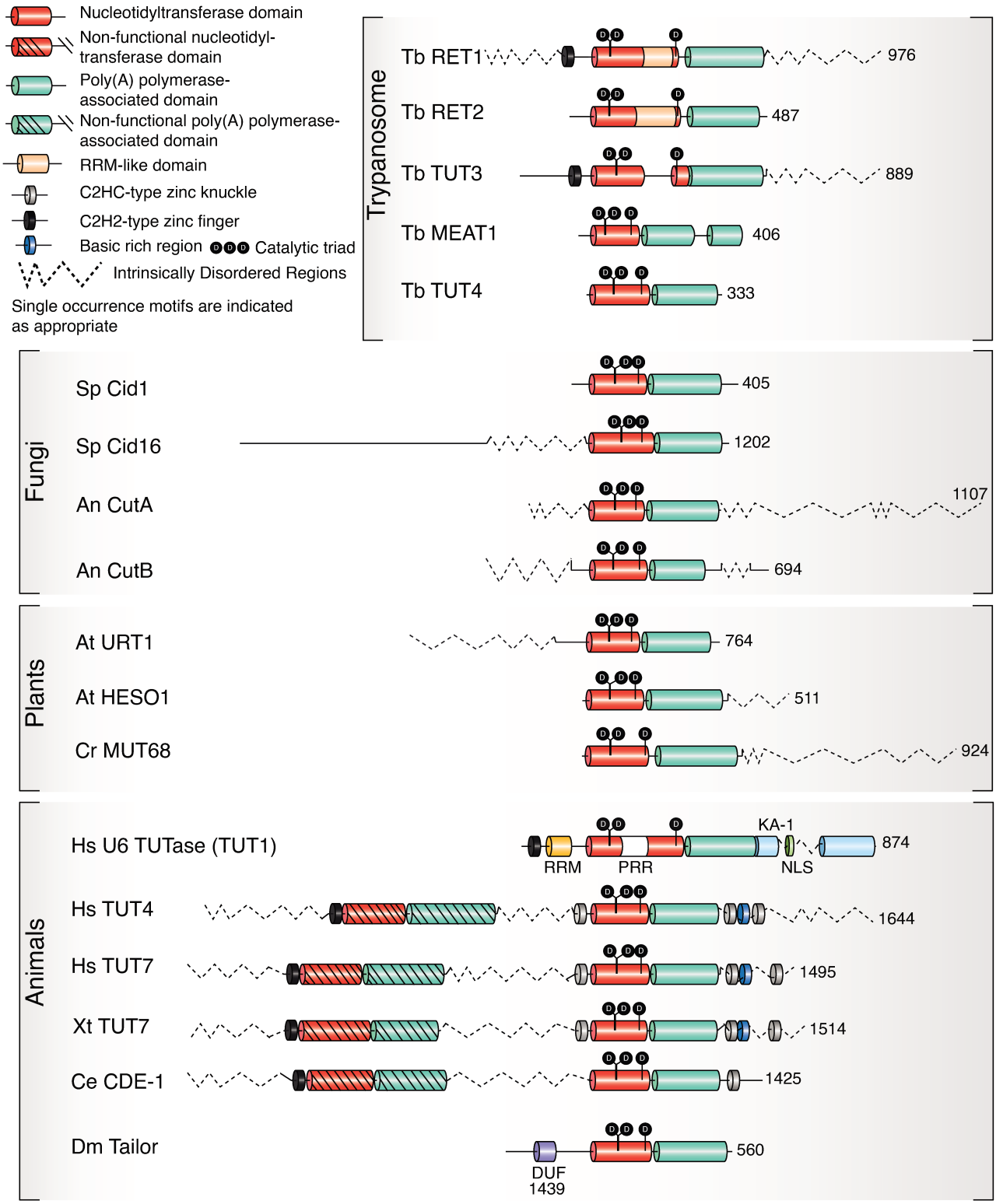


Figure 1: Domain architectures of TUTases across organisms.
Tb, *Trypanosoma brucei*; Sp, *Schizosaccharomyces pombe*; Hs, *Homo sapiens*; An, *Aspergillus nidulans*; At, *Arabidopsis thaliana*; Cr, *Chlamydomonas reinhardtii*; Xt, *Xenopus tropicalis*; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*. RRM, RNA recognition motif; PRR, proline-rich region; KA-1, kinase-associated-1; NTL, nuclear localisation signal; DUF, domain of unknown function. Intrinsically disordered regions (IDR) were predicted using DISOPRED.
De Almeida et al., 2018

recognition, activity and the recruitment of auxiliary factors to the TUTases (De Almeida et al., 2018; Warkocki et al., 2018a; Yashiro and Tomita, 2018; Zigáčková and Vaňáčová, 2018). Hence, most members of the TUTase family display a complex domain organisation (Figure 1).

A TUTase with a minimal CCD configuration is *Schizosaccharomyces pombe* Cid1 (Figure 1). Deprived from any additional domains and motifs, Cid1 binds to its targets through the formation of basic patches at the surface of its three-dimensional structure (Yates et al., 2012). Trypanosome TUT4 and MEAT1 have minimal CCD domain architectures comparable to Cid1 (Figure 1). In contrast to these uncomplex enzymes, trypanosome RET2 has a supplemental RRM (RNA recognition motif)-like domain that assures its binding to the RNA targets. Drosophila Tailor has a supplemental "domain of unknown function", DUF1439, which has been shown to stabilise the interaction between Tailor and the 3'-5' exoribonuclease Dis3L2 (Lin et al., 2017). Particular among the TUTases with a complex domain architecture is the human TUT1 protein which possesses in its N-terminus a C2H2-type zinc finger domain and an RRM domain, a disordered proline rich region (PRR) inserted in the nucleotidyltransferase domain, and an additional C-terminal RNA-binding domain KA-1 with an nuclear localisation signal (NLS) (Yamashita et al., 2017). Other examples for TUTases with a complex domain architecture are human TUT4/7, which among other features contain C2HC-type zinc-knuckle domains located on both sides of the CCD as well as basic rich regions which promote binding to their target RNAs (Figure 1) (Lapointe and Wickens, 2013).

A frequent and often conserved trait of members of the TUTase family is the presence of intrinsically disordered regions (IDR) (Figure 1). IDRs correspond to amino acid sequences that are enriched in charged and polar amino acids, but lack hydrophobic amino acids. The characteristic feature of IDRs is the absence of secondary structures (α-helices or β-sheets). On its own, IDRs have also no defined 3D structure, but they can fold into a stable 3D structure upon binding to other proteins or RNAs (Dyson, 2016; Järvelin et al., 2016; van der Lee et al., 2014). IDRs often contain short linear motifs (SLiMs), which can be target sites for post-translational protein modifications or binding sites for ligands and protein partners (van der Lee et al., 2014). IDRs increase the functional versatility of TUTase proteins. For example, the IDR that is present in the N-terminal region of trypanosome RET1 has been shown to promote its dimerization *in vitro* and is involved in the recruitment of RET1 to the mitochondrial 3' processome (MPsome), a complex ensuring the maturation of guide RNAs. The C-terminal IDR was shown to increase the processivity of RET1 *in vitro*, probably through the stabilisation of its binding to the target RNA (Rajappa-Titu et al., 2016). Human TUT4, TUT7, Xenopus TUT7, and *C. elegans* CDE1 contain long IDRs in their

N-termini (Figure 1) which surround duplicated CCD domains. These duplicated CCDs are predicted to be inactive because they lack aspartate residues required for the binding of metal ions. Yet, the N-terminal part of TUT4 with the inactive CCD is important for cell proliferation by regulating the G1-to-S phase transition. This function does not require the nucleotidyltransferase activity of the C-terminal CCD (Blahna et al., 2011). A recent study indicates that the inactive N-terminal CCDs are critical for the interaction with the Lin28 protein. The binding of Lin28 stabilises the interaction between TUT4/7 and its target pre-let7 miRNA. This increases the processivity of TUT4/7 and leads to the oligouridylation of pre-let7 miRNA (Faehnle et al., 2017). Complementary details about the domain organisation and features of the terminal uridylyltransferases can be found in page 2 to 4 of the review article published in WIREs RNA in 2018 (De Almeida et al., 2018).

1.1.2. The N-terminal IDR of Arabidopsis URT1 is required for the binding of partner proteins.

In plants, the TUTases that have been characterised to date are URT1 and HESO1 in *Arabidopsis thaliana*, and MUT68 in the green algae *Chlamydomonas reinhardtii*. Of note, MUT68 is the orthologue of Arabidopsis HESO1 (De Almeida, C. and Scheer et al., 2018; Ibrahim et al., 2010; Ren et al., 2014; Tu et al., 2015). All three proteins have IDRs. The IDRs of HESO1 and MUT68 are located in their C-terminal parts and their functions have not yet been studied (Figure 1). URT1 possesses a large IDR in its N-terminal part. Although not conserved on the level of the amino acid sequences, the N-terminal regions of all plant URT1 proteins contain IDRs. Moreover, the IDR of Arabidopsis URT1 contains several conserved SLiMs that are strictly conserved among plant URT1 proteins. Emilie Ferrier and Hélène Scheer, former PhD students of our research group, have identified two highly conserved motifs M1 and M2 in the IDR of URT1 (Ferrier, 2013, Scheer, 2018). The function of the M2 motif is still unknown, but *in vitro* pull-down assays indicated that the M1 motif is required for the interaction of URT1 with the decapping activator DCP5 (Scheer, 2018).

Early studies demonstrated that the N-terminal part of URT1 is dispensable for its *in vitro* activity (Ferrier, 2013). However, the presence of the IDR in URT1 homologs and the conservation of the short motifs suggested an important function. Initial results obtained by Emilie Ferrier suggested that the N-terminal region with the IDR is important for the localisation of URT1 in P-bodies (Ferrier, 2013).

P-bodies are cytoplasmic foci composed of aggregates of translationally repressed mRNPs. They are constitutively formed, but highly induced upon stress (Kedersha et al., 2005; Teixeira and

Parker, 2007; Teixeira et al., 2005). They are composed of specific mRNA processing proteins aggregated to their targets, thus embodying distinct cytoplasmic sites for the storage of mRNAs (Kedersha et al., 2005; Stoecklin and Kedersha, 2013). P-bodies contain proteins implicated in mRNA decay including the deadenylation factors CCR4 and PAN3, decapping activators, and 5'-3' exoribonucleases such as Arabidopsis XRN4 (detailed in Chapter 3). Of note, 3'-5' exoribonucleases other than deadenylases are not detected in P-bodies. P-bodies do further contain proteins involved in nonsense mediated decay (NMD) and translation repression (Eulalio et al., 2007; Parker and Sheth, 2007). Based on these observations, it was proposed that P-bodies store translationally repressed mRNAs and are sites of RNA degradation. Recent studies using single-molecule RNA imaging have now shown that RNA degradation does not occur in P-bodies but on polysomes and/or distributed in the cytoplasm (Horvathova et al., 2017; Tutucci et al., 2018). This fits to the observation that decapping does not take place in P-bodies (Schutz et al., 2017). Interestingly, P-bodies sequester predominantly mRNAs encoding regulatory proteins whereas constitutively and heavily translated mRNAs that encode proteins with housekeeping functions are excluded (Hubstenberger et al., 2017). The current hypothesis is that P-bodies inactivate decay factors through molecular crowding and enable the re-mobilisation of deadenylated and repressed mRNAs for translation to quickly adapt protein production to cellular needs (Hubstenberger et al., 2017).

The localisation of URT1 in P-bodies provided early evidence that URT1's function in plants could be linked to the control of mRNA stability and/or translational repression. Interestingly, DCP5, which is recruited to URT1 through the M1 motif in the N-terminal IDR part of the protein, is also implicated in decapping and translational repression, and has been located into P-bodies in *A. thaliana* (Xu and Chua, 2009). These data suggested that the N-terminal IDR of plants is important for both localisation and interactions of URT1. One of the open questions that I address in my thesis is to investigate the role of the N-terminal IDR for URT1's *in vivo* functions.

1.1.3. Molecular functions of RNA uridylation.

Uridylation marks both non-coding and coding RNAs and controls RNA fate in different manners, depending on the targeted transcripts, the activity of the TUTase responsible for the modification, and the co-factors that interact with these TUTases. A frequent downstream consequence of RNA uridylation is the rapid degradation of its RNA targets. Yet, recent studies propose that uridylation is also implicated in translational repression, mRNA storage, RNA maturation and activity, and RNA

sorting. The following review articles I co-authored during my thesis present an overview of the different roles of RNA uridylation.

The first review published in the journal WIREs RNA summarises the biological roles of TUTases in eukaryotes (De Almeida et al., 2018). I am first author of this publication and participated to the writing, the realisation of the figures and the proofreading of the final text.

The second publication summarises how uridylation influences RNA fate specifically in plants (De Almeida, C. and Scheer et al., 2018). I am co-first author of this publication in Phil. Trans. R. Soc. and I participated to the writing, proofreading, evolutionary analyses and illustrations.

The third review has been published in Trends in Genetics and summarises the functions that have been attributed to mRNA uridylation (Scheer et al., 2016). I am second author of this publication and participated to the writing, the realisation of the figures and the proofreading of the review.

Advanced Review

RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome

Caroline De Almeida, Hélène Scheer, Hélène Zuber and Dominique Gagliardi *

RNA uridylation is a potent and widespread posttranscriptional regulator of gene expression. RNA uridylation has been detected in a range of eukaryotes including trypanosomes, animals, plants, and fungi, but with the noticeable exception of budding yeast. Virtually all classes of eukaryotic RNAs can be uridylated and uridylation can also tag viral RNAs. The untemplated addition of a few uridines at the 3′ end of a transcript can have a decisive impact on RNA's fate. In rare instances, uridylation is an intrinsic step in the maturation of noncoding RNAs like for the U6 spliceosomal RNA or mitochondrial guide RNAs in trypanosomes. Uridylation can also switch specific miRNA precursors from a degradative to a processing mode. This switch depends on the number of uridines added which is regulated by the cellular context. Yet, the typical consequence of uridylation on mature noncoding RNAs or their precursors is to accelerate decay. Importantly, mRNAs are also tagged by uridylation. In fact, the advent of novel high throughput sequencing protocols has recently revealed the pervasiveness of mRNA uridylation, from plants to humans. As for noncoding RNAs, the main function to date for mRNA uridylation is to promote degradation. Yet, additional roles begin to be ascribed to U-tailing such as the control of mRNA deadenylation, translation control and possibly storage. All these new findings illustrate that we are just beginning to appreciate the diversity of roles played by RNA uridylation and its full temporal and spatial implication in regulating gene expression. © 2017 Wiley Periodicals, Inc.

How to cite this article:
WIREs RNA 2018, 9:e1440. doi: 10.1002/wrna.1440

INTRODUCTION

RNA uridylation, the untemplated addition of uridines at the 3′ extremity of RNAs, is a widespread posttranscriptional modification that targets both coding and noncoding RNAs. Except for *Saccharomyces cerevisiae*, which seems to have lost the capacity to uridylate RNA, RNA uridylation is detected in various eukaryotic organisms including trypanosomes, fission yeast, plants, insects, nematodes, and humans.[1–6] Because RNA uridylation emerges as a generic feature in RNA metabolism and to understand its impact fully, it is useful to briefly recall some basic principles underlying the production of functional transcripts and their elimination.

Genome expression necessitates the constitutive and regulated production of thousands of coding and noncoding RNAs. Virtually, all of these RNAs are produced as primary transcripts that require further processing and modifications to become functional transcripts. One of the classical steps in RNA processing is the production of mature 5′ and 3′ extremities. This maturation is achieved either by endoribonucleolytic cleavage and/or by exoribonucleolytic trimming of precursor transcripts.

*Correspondence to: dominique.gagliardi@ibmp-cnrs.unistra.fr

Institut de Biologie Moleculaire des Plantes (IBMP), CNRS, University of Strasbourg, Strasbourg, France

Conflict of interest: The authors have declared no conflicts of interest for this article.

Once created, extremities must be protected from the constitutive attacks of scavenging exoribonucleases. This protection is classically achieved by specific terminal features e.g., the m7G 5′ cap or the 3′ polyadenosine tail of eukaryotic messenger RNAs (mRNAs), which are bound by dedicated proteins. Alternatively, nonmodified extremities are simply buried into the ribonucleoproteic particles (RNPs). Stabilization of extremities can also involve a chemical modification such as 2′-O-ribose methylation of the 3′ terminal nucleotide of small RNAs in plants and Piwi-interacting RNAs (piRNAs) in animals.[7] However, the stabilization of 5′ and 3′ extremities has to be overcome at one point because genetic expression must be dynamic. Any coding and noncoding RNA will ultimately be degraded and this degradation can be regulated in response to developmental or environmental stimuli. A plethora of mechanisms is in charge of initiating or facilitating transcript degradation, and RNA uridylation is emerging as such a pervasive mechanism in eukaryotes. In this review, we will focus on how RNA uridylation impacts the processing and stability of coding and noncoding RNAs. The most recent studies clearly support that the overall function of uridylation is to destabilize its target RNA by helping the degradation machinery in overcoming the protection of extremities. However, this destabilizing role coexists with additional functions: uridylation can be required for the processing of functional transcripts, inhibit or promote translation, impede mRNA deadenylation and possibly be involved in mRNA storage. The aim of this review is to expose the fundamental and diverse functions of RNA uridylation. We will describe the core machinery involved in adding uridines to RNA 3′ ends. We will also present the different classes of RNAs that are targeted by uridylation and the main 'readers' that recognize uridylated transcripts to reveal the many roles of this key posttranscriptional modification in shaping the coding and noncoding transcriptome.

KEY FEATURES OF TERMINAL URIDYLYLTRANSFERASES

The central actors of RNA uridylation are of course the terminal uridylyltransferases (TUTases) that catalyze the untemplated addition of uridines at the 3′ end of target transcripts. TUTases belong to the DNA polymerase β (Pol β)-like nucleotidyltransferase superfamily, which include RNA-specific nucleotidyltransferases divided in three subgroups.[2,8–11] Subgroup 1 contains the nuclear poly(A) polymerases ('canonical' poly(A) polymerases) responsible for the polyadenylation of nascent mRNAs and other RNAs transcribed by RNA polymerase II. Subgroup 2 comprises TUTases that are able to recognize a diversity of RNA substrates, from guide RNAs (gRNAs) and mRNAs in trypanosome mitochondria, to miRNAs, other noncoding RNAs and mRNAs. Other various nucleotidyltransferases involved in the posttranscriptional adenylation, cytidylation, guanylation, and the CCA addition to transfer RNAs (tRNAs) in some Archaea as well as the 2′-5′-oligo(A) synthetases are also members of subgroup 2. Finally, Subgroup 3 is represented by the CCA-adding enzymes that mature tRNAs in eukaryotes and in some bacteria.

The Catalytic Domain of TUTases

The core catalytic domain (CCD) of TUTases is defined by a Pol β-like nucleotidyltransferase domain (NTD) and a poly(A) polymerase-associated (PAP-assoc) domain, which has evolved to bind uridine 5'-triphosphate (UTP) rather than adenosine 5'-triphosphate (ATP). The atomic structure of the CCD in complex with nucleotides or as apodomains has been solved for several TUTases: the mitochondrial RET1, RET2, MEAT1 and the cytosolic TUT4 from *Trypanosoma brucei*, as well as Cid1 from *Schizosaccharomyces pombe* and human U6 TUTase (TUT1).[12–17] The juxtaposition of the two domains forming the CCD creates a large cleft, which contains the catalytic and UTP-binding sites. Some TUTases like the human U6 TUTase or the trypanosomal RET1, RET2, and TUT3 have additional sequences inserted in the CCD (Figure 1). A domain which folds like a RNA recognition motif (RRM) despite lacking the typical signature of RRM, is inserted in the CCD of RET1.[15] A similar domain organization exists in RET2 but the orientation of the RRM-like fold is different. This difference could account for the differential recognition of single-stranded versus double-stranded RNA by RET1 and RET2, respectively.[15] RET1 also contains a C2H2 Zinc finger (ZnF) adjacent to the NTD that is essential for the folding and stabilization of the catalytic domain (Figure 1).[15]

UTP specificity involves conserved aspartate and glutamate residues in both mitochondrial and cytosolic TUTases such as RET1 and Cid1, respectively.[14,18,19] In addition, a histidine residue in Cid1 (His336), conserved in some plant and human cytosolic TUTases but absent from trypanosomal TUTases, has been involved in UTP selectivity.[14] A single amino acid substitution at this position is sufficient to switch the specificity of Cid1 from UTP to ATP, and vice versa for human Gld2 (TUT2).[14,20] Those experiments illustrate the ease of switching the nucleotide specificity of a terminal nucleotidyltransferase during evolution to allow for the acquisition of a novel biological function linked to RNA tailing.

 © 2017 Wiley Periodicals, Inc.

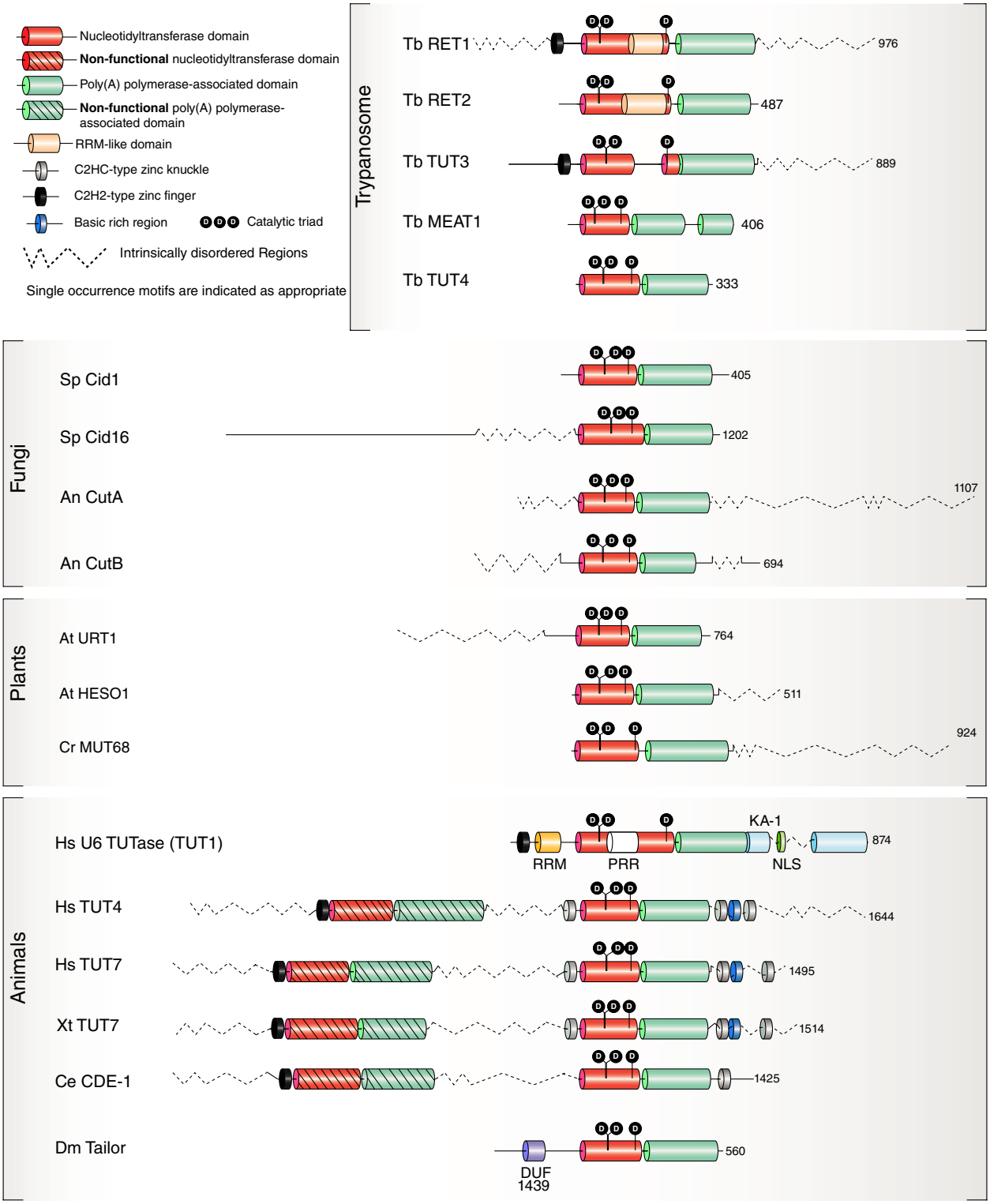


FIGURE 1 | Domain architectures of TUTases are diverse across organisms. Tb, *Trypanosoma brucei*; Sp, *Schizosaccharomyces pombe*; Hs, *Homo sapiens*; An, *Aspergillus nidulans*; At, *Arabidopsis thaliana*; Cr, *Chlamydomonas reinhardtii*; Xt, *Xenopus tropicalis*; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*. RRM, RNA recognition motif; PRR, proline-rich region; KA-1, kinase-associated-1; NLS, nuclear localization signal; DUF, domain of unknown function. Intrinsically disordered regions (IDR) have been predicted using DISOPRED.

 © 2017 Wiley Periodicals, Inc.

Complex and Diverse Domain Architectures of TUTases

The association of the NTD with a PAP-assoc domain constitutes the minimal catalytic module for a TUTase. This minimal configuration actually corresponds to the structural organization of Cid1 from *S. pombe* (Figure 1). Cid1 does not have a dedicated RNA-binding domain adjacent to the catalytic module but rather binds its RNA substrates through interaction with three basic patches distributed at the surface of the enzyme.[13] No interacting partners have been identified to date and Cid1 appears to act as a standalone enzyme. In contrast to Cid1, most other TUTases typically present a multipartite domain architecture with domains or regions mediating protein–protein interactions or protein–RNA binding. This architecture is evolving fast and therefore, TUTases display various domain organizations as illustrated for a selection of characterized TUTases from trypanosomes, fungi, animals, and plants (Figure 1).

A striking example of complex domain architecture is illustrated by the human TUT4/7 or Xenopus TUT7.[8,21] Those TUTases contain a duplicated CCD domain but only the C-terminal one is active. Yet, the inactive CCD is required for structural functionalities, independently of catalyzing RNA tailing.[22,23] Another key feature of HsTUT4/7 and XtTUT7 is the presence of C2H2-type ZnF and C2HC-type ZnF motifs (also known as zinc knuckles), located upstream of the inactive CCD and on both sides of the active one, respectively (Figure 1).[8] The last two C2HC-type ZnF surround a basic amino acid-rich (BR) stretch, conserved from Xenopus to human.[21] BR stretches are known as disordered regions that can promote RNA binding.[24] Those BR stretches belong to one of the three large intrinsically disordered regions (IDRs) predicted in the N-terminal, middle and C-terminal regions of HsTUT4/7 and XtTUT7. Although, the presence of large IDRs is actually common across TUTases, their diversity in size and position largely contributes to the variability of TUTase organization across organisms (Figure 1). Such IDRs could confer specific RNA or protein binding capacity to their respective TUTase. For instance, the N-terminal region of RET1 is dispensable for activity *in vitro* but likely mediates protein–protein interactions crucial for RET1's function *in vivo*.[15] Indeed, RET1 can oligomerize, is integrated into a complex termed the mitochondrial 3′ processome (MPsome) and can also transiently interact with pentatricopeptide repeat (PPR)-containing RNA-binding factors and a mitochondrial poly(A) polymerase.[15,25,26] By contrast, the C-terminal region is not required for incorporation of RET1 into the MPsome but could bind RET1's RNA substrates. Indeed, its deletion decreases the RNA-binding capacity of RET1 thereby strongly reducing processivity and catalytic efficiency.[15] In addition to IDRs, ZnF can also mediate protein–RNA and protein–protein interactions.[27–29]

Deciphering the intricate interaction network of TUTases is key to fully understand the effect of RNA uridylation. Two categories of factors interact with TUTases: the auxiliary factors that assist TUTases in modifying their targets, and some of the 'readers' that recognize the uridylated status of transcripts and translate this information into a biological output (stabilization, decay, translation inhibition, etc.) (Table 1). These factors are detailed next when we survey the diverse roles of RNA uridylation in shaping the transcriptome.

VARIOUS CONSEQUENCES OF URIDYLATING NONCODING RNAs

Processing of Mitochondrial gRNAs in Trypanosomes

Uridylation is fundamental for RNA metabolism in mitochondria of trypanosomes. Both coding and noncoding RNAs encoded in this organelle are uridylated and the majority of mRNAs are massively edited by U insertions and deletions.[58] Editing requires gRNAs that specify the position of editing sites to the RNA editing core complex (RECC). gRNAs are 50–60 nt long RNAs terminating with 15–20 untemplated uridines. The vast majority of guide RNAs are encoded by thousands of minicircles of about 1 kbp, which constitute the mitochondrial genome of trypanosomes together with a few 25 kbp maxicircles encoding multicistronic ribosomal RNAs (rRNAs) and mRNAs.[61]

gRNAs maturation involves the mitochondrial 3' processome (MPsome) constituted by the TUTase RET1 in association with the 3′→5′ exoribonuclease DSS1 and three large proteins with no known motifs.[6,25] The bidirectional transcription of minicircles generates sense and antisense gRNA precursors of about 800–1200 nt and with a 50 nt overlapping region in their 5′ region (Figure 2). The first step in the maturation process of these precursors by the MPsome is the uridylation by RET1, which recruits the 3′→5′ exoribonucleolytic activity of DSS1. The progression of the MPsome is impeded 10–12 nt away from the stable duplex region that is formed by annealing of the complementary sequences created by the bidirectional transcription of the sense and antisense precursors. RET1 then performs a

 © 2017 Wiley Periodicals, Inc.

TABLE 1 | Main Factors Assisting TUTases

Protein/ Complex	Protein Type	Organism(s)	Relevant TUTase(s)	Description	References
Lin28	RNA-binding protein	Mouse, human	TUT4/7	Lin28A binds to pre-let-7 and recruits TUT4/7 to initiate oligo-uridylation of pre-let-7 miRNA and subsequent degradation by Dis3L2.	23,30–34
Trim25	E3 ubiquitin/ISG15 ligase	Mouse, human	TUT4	Trim25 binds to the conserved terminal loop of pre-let-7 and promotes TUT4 Lin28 mediated-uridylation.	35
LSm1-7	RNA-binding protein complex	Human, *S. pombe*	TUT4/7	The LSm1-7 complex preferentially binds to oligoadenylated mRNAs with 3′ terminal uridines, thereby promoting decapping.	36–40
Dis3L2	3′→5′ exoribonuclease	Human, *S. pombe*, Drosophila, mouse	TUT4/7, Tailor	Dis3L2 preferentially degrades uridylated RNAs, including pre-microRNAs, mRNAs and unprocessed and structured noncoding RNAs. In Drosophila, Dis3L2 and Tailor form a cytoplasmic terminal RNA uridylation-mediated processing (TRUMP) complex.	34,41–46
Usb1 (Mpn1)	3′→5′ exoribonuclease/ phosphodiesterase	Human	U6 TUTase (TUT1)	Usb1 nibbles the 3′ extremity of the uridylated U6 snRNA and leaves an extension of five Us and a terminal 2′, 3′ cyclic phosphate that favor the binding of LSm2-8 complex.	47–49
Ago2	Endoribonuclease/ RNA-binding protein	Human	U6 TUTase (TUT1)	U6 TUTase co-purifies with Ago2 and Dis3L2 via RNA-mediated interaction. The three proteins seem to be part of the same complex.	50
LSm2-8	RNA-binding protein complex	Human	U6 TUTase (TUT1)	LSm2-8 binds to U6 snRNA with 3′ five U extension and a terminal 2′, 3′ cyclic phosphate produced by dual action of U6 TUTase/Usb1 exoribonuclease.	51
SART3 (Tip110)	RNA-binding protein	*C. elegans*	USIP-1	USIP-1 forms a complex with SART3 and U6 snRNA and uridylates U6 snRNA to promote its recycling.	52
EGO-1	RNA-directed RNA polymerase	*C. elegans*	CDE-1	Localization of CDE-1 in embryo mitotic chromosomes requires the RdRP EGO-1, which interacts with CDE-1 and the Argonaute protein CSR-1. CDE-1 uridylates CSR-1-bound siRNAs.	53
AGO1	Endoribonuclease/ RNA-binding protein	Arabidopsis	HESO1	AGO1, key factor of the RNA-induced silencing complex, interacts with HESO1 through its Piwi/ Argonaute/Zwille (PAZ) and Piwi domains. HESO1 uridylates AGO1-bound miRNAs to trigger their degradation.	54
RICE1/2	3′ to 5′ exoribonuclease	Arabidopsis	HESO1	RISC-interacting clearing 3′→5′ exoribonucleases 1 and 2 (RICE1/2) are 3′ to 5′ exoribonucleases that initiate degradation of uridylated 5′ RISC-cleaved fragments to recycle RISC.	55
SDN1/2	3′ to 5′ exoribonuclease	Arabidopsis	HESO1	SDN1/2 are 3′→5′ exoribonucleases that trim AGO1-bound small RNAs prior to tailing by HESO1 and degradation.	56
PABP	RNA-binding protein	Arabidopsis	URT1	PABP binds oligo(A/U) tails *in vivo* and determines the size of U extension added by URT1.	57
DSS1	3′ to 5′ exoribonuclease	*T. brucei*	RET1	RET1 is integrated into the MPsome complex composed of the exoribonuclease DSS1 and	25,58

(*continued overleaf*)

 © 2017 Wiley Periodicals, Inc.

TABLE 1 | Continued

Protein/ Complex	Protein Type	Organism(s)	Relevant TUTase(s)	Description	References
				three other proteins with no known domain. The complex is involved in gRNA maturation.	
KPAF1/2	PPR-proteins	*T. brucei*	RET1	The heterodimer composed of KPAF1 and KPAF2 PPR-proteins induces the formation of long A/U tail by RET1 and the KPAP1 poly(A) polymerase.	26
MP81	RNA-binding protein	*T. brucei*	RET2	RET2 is a subunit of the U-insertion subdomain of the RNA editing core complex (RECC) that catalyzes mRNA editing in trypanosome mitochondria. RET2 interacts with MP81 protein resulting in the stabilization of both proteins and enhancement of TUTase activity.	59,60

secondary uridylation step, which does not trigger degradation but could rather promote the MPsome disengagement from the duplex intermediate. The duplex is unwound, the antisense strand degraded and the sense uridylated gRNA integrated into the gRNA-binding complex to direct the editosome to editing sites (Figure 2). The U-tails of gRNAs could promote the interaction with the purine-rich pre-edited mRNA or recruit protein factors.[9] Although further investigation is still required to fully define the biological function of the secondary uridylation step, solving the processing pathway of gRNAs revealed a dual and complex role for uridylation.[25]

Uridylation in U6 snRNA Maturation and Stability

Another well-characterized substrate of RNA uridylation is the human spliceosomal U6 small nuclear RNA (U6 snRNA).[62] U6 snRNA transcription by RNA polymerase III (Pol III) is terminated by a short stretch of four encoded uridines (Figure 3). Interestingly, most human U6 snRNAs end with five Us and a 2′, 3′ cyclic phosphate. Minor forms of U6 snRNA have 3′ oligo(U) tails of up to 20 residues and a 3′ OH extremity. The heterogeneity of 3′ termini of U6 snRNA reveals the combined action of two opposing activities in the 3′ maturation process: U6 snRNA 3′ extremities are uridylated by the U6 TUTase (TUT1)[63–65] and nibbled by a distributive 3′→5′ exoribonuclease Usb1 (Mpn1) (Figure 3).[47–49,66] Usb1 belongs to the LigT-like superfamily of 2H phosphoesterases and catalyzes the formation of a terminal 2′, 3′ cyclic phosphate while removing uridines added by U6 TUTase to leave a five U-tail.[47–49]

Uridylation is rightly considered as an integral step in the 3′ maturation and stabilization of U6 snRNA in humans.[62] Paradoxically, the stabilizing effect of uridylation by the U6 TUTase is in fact due to the action of the 3′→5′ exoribonuclease Usb1. Indeed, the production of a terminal 2′, 3′ cyclic phosphate favors the binding of the stabilizing LSm2-8 complex.[51] In addition, Usb1's action protects U6 snRNA from adenylation by the terminal nucleotidyltransferase Trf4 and subsequent degradation by the nuclear RNA exosome.[47] Although, the recruitment of an exoribonucleolytic activity by U-tails appears as a common process in eukaryotes, the example of human U6 snRNA illustrates that the outcome of such a recruitment is not necessary the destabilization of the target RNA.

Control of miRNA Processing by Uridylation

One of the most spectacular regulatory roles of RNA uridylation in animals is to control the biogenesis of specific miRNAs (Box 1). Uridylation affects miRNA processing and the degradation of miRNA precursors with various consequences on development, diseases and evolution of miRNA families.

Dual Role of Uridylation in let-7 miRNA Biogenesis

The let-7 miRNA family is highly conserved in bilaterian animals. It suppresses cell proliferation and promotes cell differentiation. In humans, 9 out of 12 let-7 precursors are processed as pre-miRNAs with 1 nt 3′ overhang (Group II precursors). The remaining three precursors have a typical 2 nt 3′ overhang (Group I), like canonical miRNA precursors.[68] Group I pre-miRNAs are classically processed by Dicer (Figure 4). By contrast, Group II let-7 pre-miRNAs are poor substrates of Dicer because of their 1 nt 3′ overhang. TUT4 and TUT7 can mono-uridylate those precursors, thereby restoring a full

 © 2017 Wiley Periodicals, Inc.

BOX 1

DIFFERENT miRNA PROCESSING PATHWAYS IN ANIMALS

miRNAs are small RNAs of 21–22 nt that associate with Argonaute (Ago) proteins within RNA-induced silencing complexes (RISCs). RISCs silence target mRNAs by repressing translation and inducing decay.[67] The classical pathway of miRNA processing in animals involves two RNase III-like enzymes, Drosha and Dicer. Drosha cleaves primary miRNA (pri-miRNA) transcripts in the nucleus to generate a pre-miRNA hairpin with a 2 nt 3′ overhang. The pre-miRNA is processed as a duplex RNA by Dicer in the cytoplasm. The miRNA strand of the duplex is retained in complex with a Ago protein to form the core components of RISC while the passenger strand of the duplex is degraded.[67] Alternative processing pathways exist that can bypass either Dicer activity as for miR-451 or Drosha action as for pre-miR-320 and mirtrons. Mirtrons are encoded as short hairpin introns that use splicing and debranching instead of Drosha to generate the pre-miRNA hairpin that is further processed by Dicer.[67] Finally, processing of Group II let-7 miRNAs involves both Drosha and Dicer but mono- and poly-uridylation of pre-let-7 miRNAs either promote or prevent maturation by Dicer, respectively.[67]

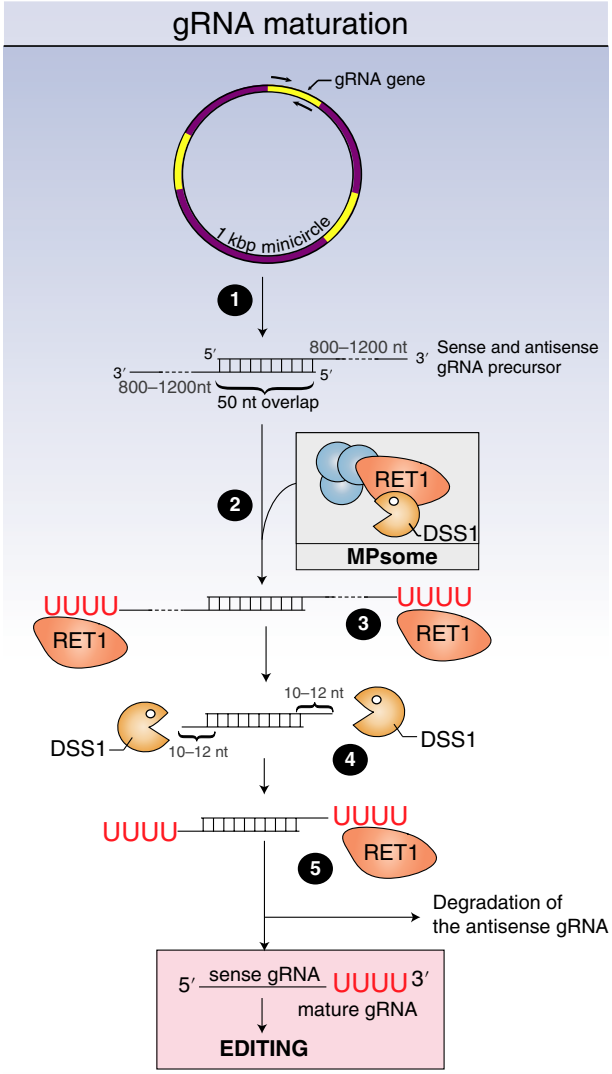


FIGURE 2 | Uridylation and guide RNA (gRNA) maturation in trypanosome mitochondria. gRNAs are processed through a sequential maturation process by the mitochondrial 3′ processome (MPsome), containing the TUTase RET1 in complex with the 3′→5′ exoribonuclease DSS1. (1) gRNAs are generated by bidirectional transcription of minicircles. The sense and antisense gRNA precursors have complementary regions in the 5′ end and form a duplex. After recruitment of the MPsome (2), the precursors undergo a first uridylation step by RET1 (3), leading to the degradation of the precursors by DSS1 (4). Progression of the MPsome is impeded 10–12 nt from the paired region and a second uridylation step by RET1 occurs (5). After antisense gRNA degradation, mature uridylated gRNAs are incorporated into the gRNA-binding complex to direct the editosome to editing sites.

processing capacity for Dicer (Figure 4).[46,68] A recent structure of the catalytic domain of TUT7 engaged in the mono-uridylation of Group II pre-let-7 hairpin revealed a duplex-RNA-binding pocket favoring the addition of a single uridine.[23] Another terminal nucleotidyltransferase, Gld2 (TUT2 or PAPD4), can mono-uridylate and mono-adenylate Group II let-7 pre-miRNAs *in vitro* and promotes let-7 processing *in vivo*.[68] Although it remains to be formally demonstrated whether Gld2's impact on let-7 processing is due to adenylation rather than uridylation,[20] TUT4/7 are definitely crucial to promote Group II pre-let-7 processing, demonstrating that uridylation is required for Group II let-7 biogenesis.[68]

let-7 is expressed in differentiated cells where it negatively regulates several known oncogenes, but is repressed in embryonic stem cells and in several cancers in mammals.[69,70] A prominent factor regulating let-7 accumulation is the RNA-binding protein Lin28.[30–32,41,70–75] The genome of vertebrates encodes two paralogous Lin28 proteins. Both Lin28A and Lin28B downregulate let-7 production by distinct mechanisms including the sequestration of precursors away from nuclear processing factors and uridylation -mediated degradation of cytosolic precursors

 © 2017 Wiley Periodicals, Inc.

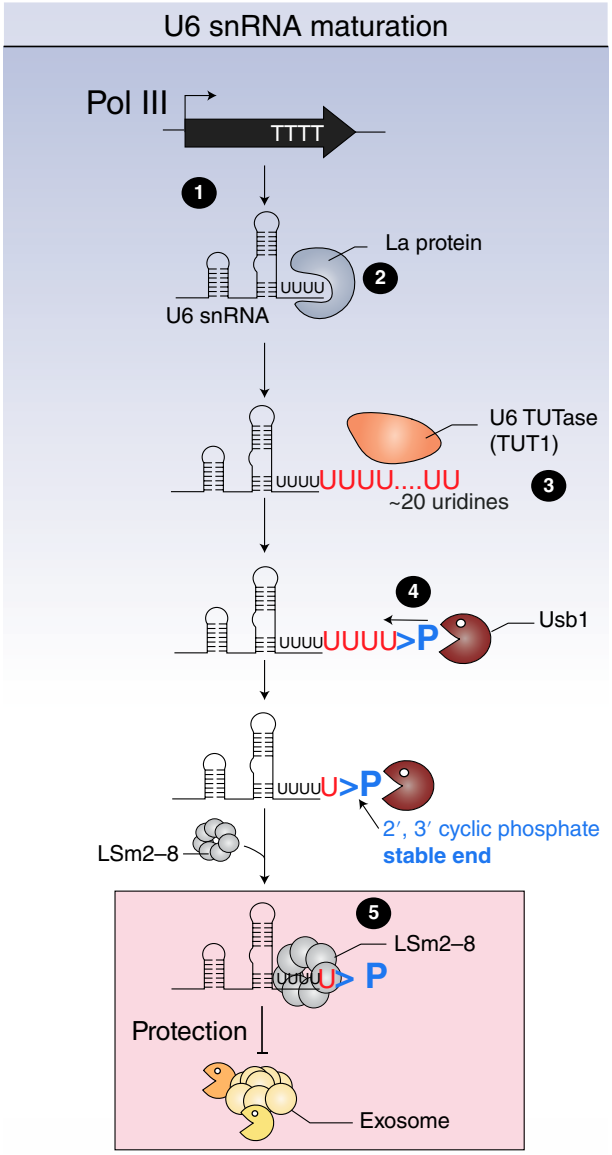


FIGURE 3 | Uridylation is critical for U6 snRNA maturation. U6 snRNAs are transcribed by polymerase III (Pol III) which terminates transcription by a stretch of four encoded uridines (1) that are immediately bound by the La protein (2). U6 snRNAs are then uridylated by U6 TUTase (TUT1) (3) which favors nibbling by the exoribonuclease Usb1 (4). Usb1 is a phosphodiesterase and generates terminal 2′, 3′ cyclic phosphate. The particular 3′ end formed by four encoded uridines and one exogenous uridine with a terminal 2′, 3′ cyclic phosphate facilitates the recruitment of the LSm2-8 complex that prevents degradation by the exosome (5).

(Figure 4).[69] The molecular mechanisms underlying the uridylation-mediated degradation of pre-let-7 miRNAs is thoroughly investigated. The cold-shock domain of the cytosolic Lin28A recognizes the terminal loop of pre-let-7 miRNA and a zinc knuckle stabilizes this interaction by recognizing a conserved GAGG sequence near the 3′ end.[31,76] Recent studies propose that this interaction between the zinc knuckle of Lin28 and GGAG motif of pre-let7 forms a specific surface recognized by the N-terminal half of TUT4/7 containing the inactive CCD domain (also called the Lin28-interacting module or LIM), thereby establishing a stable ternary complex between pre-let-7: Lin28A:TUT4/7.[23] The Lin28A-stabilized interaction between pre-let-7 and TUT4/7 results in the oligo-uridylation of pre-let-7 miRNAs.[30–33] In addition, two C2HC-type zinc knuckles close to the active CCD of TUT4/7 establish uracil-specific interactions, facilitating oligo-uridylation.[23] Finally, the activity of TUT4 is stimulated by the E3 ligase Trim25, which specifically binds the pre-let-7 miRNAs.[35] The oligo-uridylation of pre-let-7 miRNAs inhibits Dicer action and recruits the 3′→5′ exoribonuclease Dis3L2 (Box 2) which degrades let-7 pre-miRNA (Figure 4).[34,41,77] TUT4 and TUT7 redundantly uridylate pre-let-7 miRNAs.[78] However, this redundancy may depend on the cell type because TUT4 knockdown was also shown to fully mimic Lin28A knockdown in human cancer cells expressing Lin28A.[73] Lin28B was also reported to promote uridylation and Dis3L2-mediated degradation of pre-let-7 miRNAs in certain cancer cells.[72]

Altogether, studies of let-7 biogenesis revealed a fascinating regulatory role of RNA

BOX 2

Dis3L2: A 3′→5′ EXORIBONUCLEASE DEGRADING URIDYLATED RNA SUBSTRATES

Dis3L2 is a 3′→5′ exoribonuclease of the RNase II/R family which is localized in the cytosol and conserved in eukaryotes except in *S. cerevisiae*.[4] Mutations in human Dis3L2 are associated with the Perlman syndrome of fetal overgrowth, predisposition to develop Wilms' tumors and a series of additional cancers.[4,41] Dis3L2 acts independently of the exosome and degrades a variety of RNA substrates in the cytosol, ranging from mRNAs to a range of noncoding RNAs including small RNAs.[34,41–43,45,79–81] Of note, uridylation favors Dis3L2-mediated decay. Determination of the structure of mouse Dis3L2 in complex with an oligo(U) RNA revealed extensive uracil-specific interactions, explaining how Dis3L2 preferentially recognizes uridylated RNA substrates.[77]

 © 2017 Wiley Periodicals, Inc.

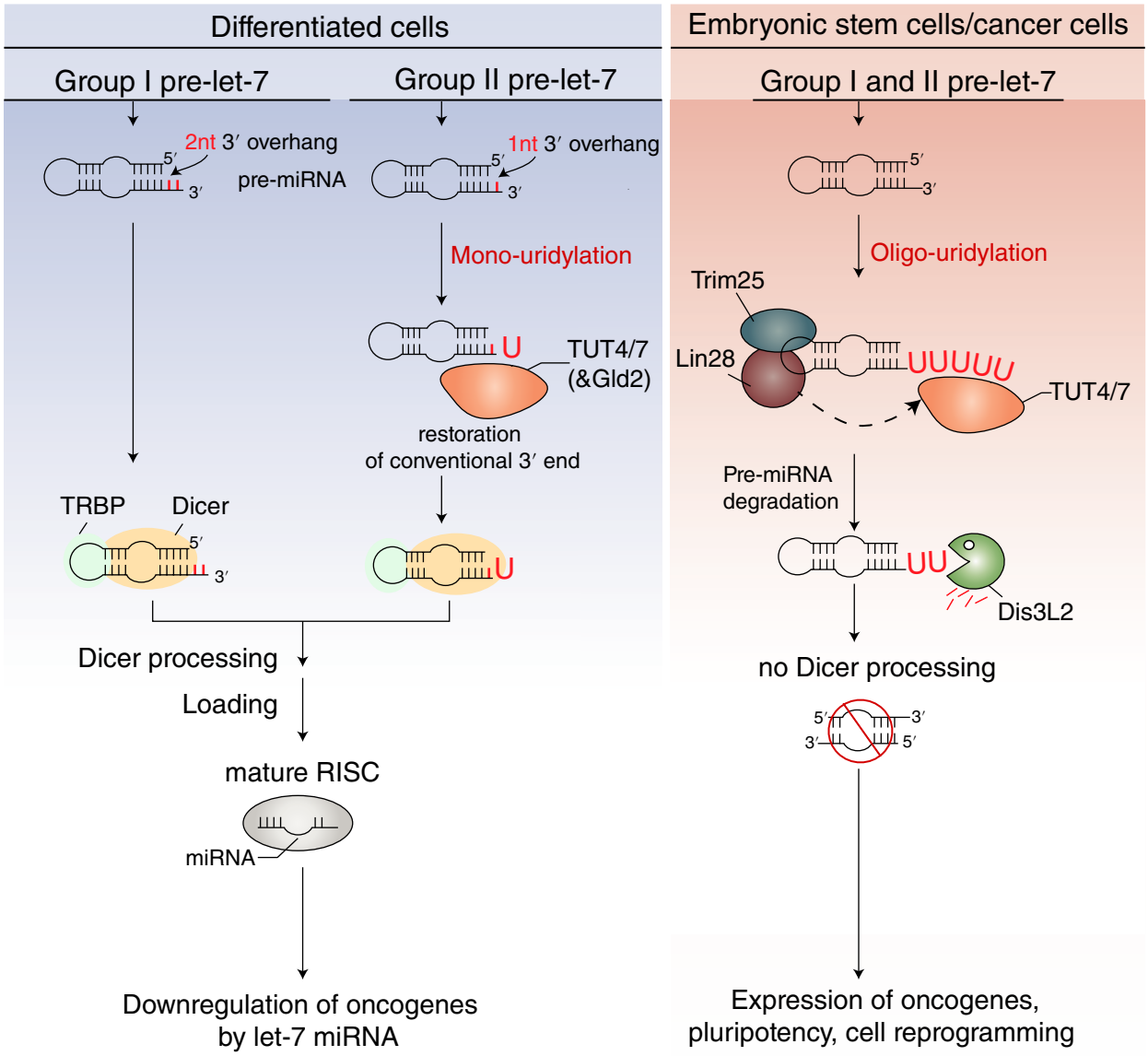


FIGURE 4 | Dual role of uridylation in let-7 miRNA maturation. Group I pre-let-7 miRNAs end with a 2 nt 3′ overhang and are further processed by Dicer to generate mature let-7 which downregulate oncogenes in differentiated cells. Group II pre-let-7 miRNAs end with a 1 nt 3′ overhang and are mono-uridylated by TUT4/7 and Gld2 (or possibly mono-adenylated by Gld2). This single nt addition on the 3′ end restores full Dicer competence to produce let-7 miRNAs. In embryonic stem cells and many cancers, the RNA binding Lin28A (and possibly Lin28B[69]) binds pre-let-7 and, together with the E3 ligase Trim25, recruits TUT4/7 to oligo-uridylate pre-let-7 leading to its degradation by Dis3L2. Prevention of mature let-7 production induces pluripotency, cell reprogramming, and cancers.

uridylation. Mono-uridylation of pre-let-7 miRNAs by TUT4/7 in the absence of Lin28A enhances Dicer processing, thereby downregulating oncogene expression through let-7 action. By contrast, oligo-uridylation by the same set of TUTases triggered by the processivity factors Lin28A/Trim25 leads to the degradation of pre-let-7 miRNAs by Dis3L2, which promotes cell proliferation and limits cell differentiation.

Uridylation Represses the Mirtron Pathway

The stability of mirtron precursors is also controlled by uridylation. Mirtrons are small RNAs defined by their Drosha-independent processing which relies on the splicing machinery to generate short hairpins further processed by Dicer (Figure 5).[67,82] Mirtrons have been identified in flies, worms and humans, and are usually low expressed and not conserved. In *Drosophila melanogaster*, the mirtron precursors are

 © 2017 Wiley Periodicals, Inc.

uridylated by the TUTase Tailor impeding processing by Dicer (Figure 5).[83,84] Strikingly, mirtron hairpins are preferentially uridylated as compared to classical hairpins generated by Drosha.[83–86] This preference is explained by the specificity of Tailor for substrates with a 3′ terminal G.[83,84] Because of their biogenesis by the splicing machinery, all mirtrons end with a 3′ AG. Hence, substrate preference of Tailor may have evolved to suppress mirtron biogenesis to avoid the creation of spurious novel miRNAs.[83,84] Importantly, precursors of conserved miRNAs are specifically depleted for 3′ G whereas ones for nonconserved miRNAs are not. Therefore, Tailor-mediated control of the accumulation of *de novo* created mirtrons likely resulted in a selective pressure that shaped canonical miRNAs in Drosophila.[83,84]

Uridylation of Mature Small RNAs: Decay and More

The untemplated 3′ addition of nucleotides is also common on mature small RNAs. Added nucleotides are mostly adenosines and uridines. Uridylation has been studied for different classes of small silencing RNAs including miRNAs, short interfering RNAs, and piRNAs.[3,87] Small RNA uridylation was reported across diverse organisms, from fission yeast, nematodes, flies, frogs, plants, and mammals. Again, uridylation influences small RNA fate in several ways, the most prominent one being destabilization.

Methylation, Uridylation and Decay of Plant Small RNAs

Uridylation-mediated destabilization of small RNAs is thoroughly investigated in plants. The terminal ribose of all plant small silencing RNAs is 2′-O methylated by the methyltransferase HEN1.[88–90] HEN1 has two double-stranded RNA-binding domains that recognize the duplexes generated by Dicer-like (DCL) enzymes and methylates the 3′ extremity of both strands of the duplex. In Arabidopsis, loss of small RNA methylation in *hen1* mutants results in strong developmental defects due to a decrease in miRNA abundance.[88,89] Of note, siRNA accumulation is also affected in Arabidopsis and in rice *hen1* mutants.[88,89,91] This decrease is accompanied by 3′ end untemplated uridylation and trimming of small RNAs.[88,89] Hence, methylation by HEN1 stabilizes small RNAs by preventing their uridylation and subsequent degradation. Interestingly, the patterns of trimming and uridylation are quite diverse across miRNA families. Some miRNAs are particularly trimmed and tailed while others are not affected by the lack of HEN1-mediated methylation.[92] Yet, particular patterns of trimming/tailing are conserved for the same miRNA between maize, rice and Arabidopsis *hen1* mutants, suggesting that structural elements or sequences conserved between monocotyledons and dicotyledons can influence uridylation and trimming.[92] Of note, certain miRNAs such as miR158 are substantially trimmed and tailed in a wild-type context because of inefficient methylation even in the presence of HEN1.[92]

A genetic suppressor screen in Arabidopsis identified the TUTase HESO1 (HEN1 SUPPRESSOR 1) as the terminal uridylyltransferase responsible for uridylation of small RNAs.[93] In a *heso1 hen1* mutant, unmethylated small RNAs are less uridylated and their global level increases.[93–95] It was therefore concluded that uridylation by HESO1 destabilizes small RNAs.[93,95] Forward and reverse genetic experiments revealed URT1 as another TUTase able to uridylate miRNAs, but not siRNAs, in the *hen1 heso1* genetic context.[96,97] Of note, the biological impact of URT1 in uridylating miRNAs in a wild-type context remains to be determined and HESO1 represents the major TUTase uridylating both siRNAs and miRNAs in *Arabidopsis*. In *Chlamydomonas reinhardtii*, uridylation by the TUTase MUT68 also destabilizes small RNAs.[98]

Both biochemical and sequencing analyses indicate that trimming precedes uridylation by HESO1 (Figure 6(a)). Indeed, the catalytic activity of HESO1 is inhibited by the 2′-O methylation deposited by HEN1 and extensive tailing is observed on trimmed small RNAs.[92,93,95] The 3′→5′ exoribonucleases that trim methylated small RNAs prior to tailing were recently identified as SDN1 and SDN2.[56] SDN1/2 interact with AGO1 as does HESO1,[54] giving a rational explanation as why trimming and tailing depends on AGO.[92] SDN1/2 are proposed to trim small RNAs while loaded on an Ago protein, thereby alleviating the inhibitory effect of HEN1 methylation on HESO1 tailing (Figure 6(a)).[56] Uridylation by HESO1, and possibly URT1, then triggers the degradation of the small RNAs by a yet unidentified enzyme.

Other Small RNAs Protected by Methylation Against Uridylation

The protection of 3′ ends by methylation is not restricted to plant small RNAs. HEN1 homologues are conserved in animals and methylate piRNAs.[99–105] Piwi proteins belong to the Argonaute family and are loaded with piRNAs expressed in germlines to repress transposable elements. In contrast to Arabidopsis HEN1, animals Hen1 homologues lack double stranded RNA-binding domains and rather interact with Piwi proteins to methylate single-stranded piRNAs of various sizes. Methylation by HEN1

 © 2017 Wiley Periodicals, Inc.

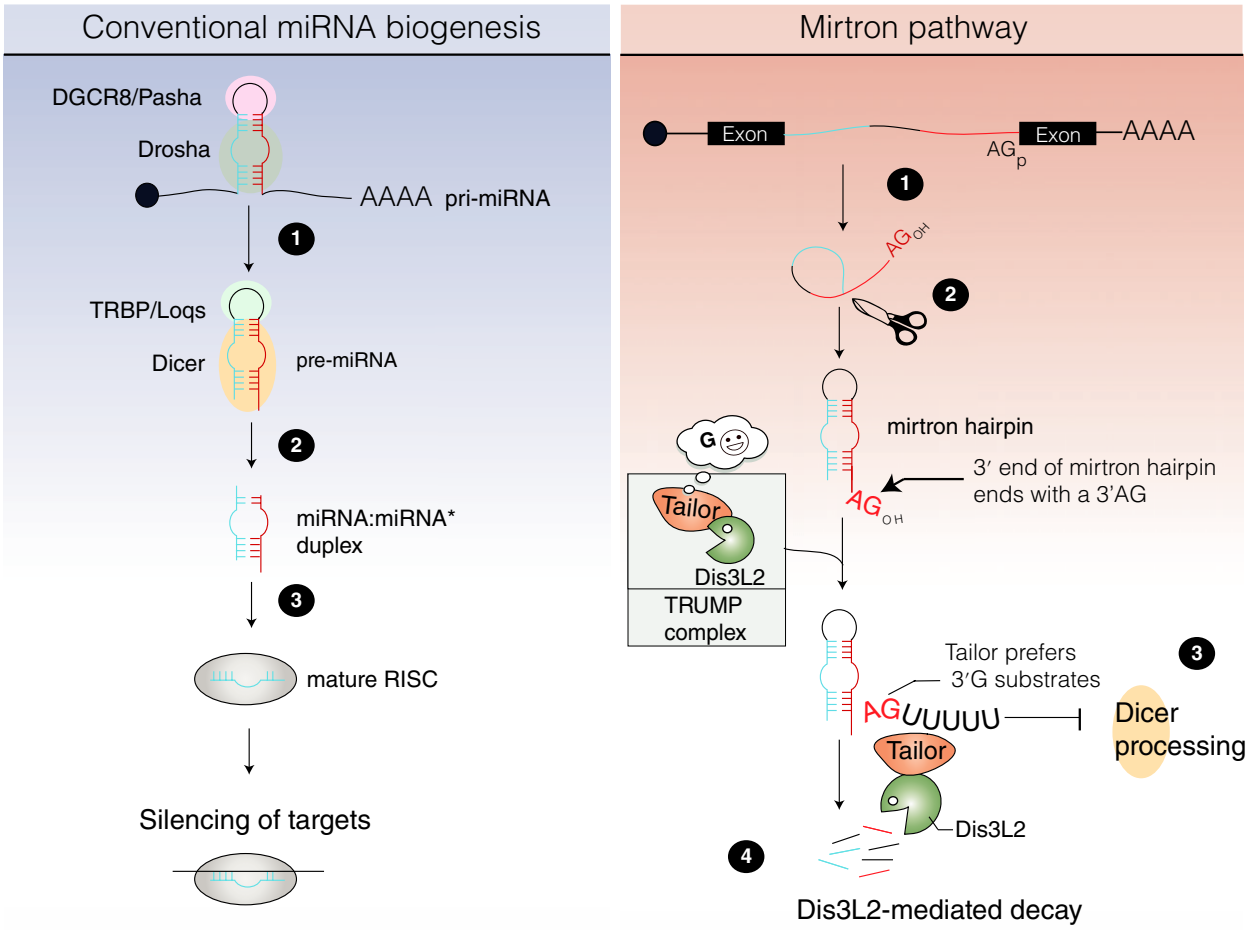


FIGURE 5 | Uridylation restricts mirtron accumulation. Left panel: (1) The conventional processing of animal miRNAs starts with the cleavage of the pri-miRNA by the nuclear Drosha/DGCR8 complex. (2) The generated pre-miRNA hairpin is then further processed by Dicer/TRBP in the cytoplasm. (3) The resulting miRNA:miRNA* duplex is loaded into Argonaute where the mature miRNA is retained, forming the RISC complex. Right panel: As compared to classical processing for animal miRNAs, mirtrons do not rely on Drosha but instead on the splicing (1) and lariat-debranching (2) machinery to generate mirtron hairpins ending with AG. Those hairpins are preferential substrates of the TUTase Tailor, which has a better affinity for substrates ending with a G (3). Tailor together with Dis3L2 forms the TRUMP complex in Drosophila. Uridylation by Tailor promotes degradation by Dis3L2, impeding Dicer processing and preventing the formation of mirtrons (4).

homologues, such as Pimet/Hen1 in Drosophila, Hen1 in zebrafish or HENMT1 in mouse, may prevent uridylation-mediated decay of piRNAs.[25,99,102,104]

Another class of small RNAs that are 3′ methylated is siRNAs in Drosophila.[100] Loss of methylation results in trimming and uridylation of those siRNAs.[106] Only Ago2-bound siRNAs, but not Ago1-bound miRNAs, are methylated in Drosophila and this difference may be linked to their respective mode of target recognition. One of the major differences between siRNAs and miRNAs is the extensive pairing of siRNAs to their targets as compared with the partial pairing of miRNAs, mostly restricted to the seed sequence. Interestingly, increasing the complementarity between a target sequence and Ago1-bound miRNAs is enough to trigger uridylation and trimming.[106] Those experiments explain why siRNAs which trigger the destruction of viral or transposon RNAs using extensive complementarity are methylated to protect their 3′ end from tailing and trimming, whereas miRNAs are not. Indeed, extensive complementarity of a small RNA with its target weakens its interaction with the Piwi/Argonaute/Zwille (PAZ) domain of Ago proteins, thereby allowing accessibility to TUTases and 3′→5′ exoribonucleases.[106]

Finally, methylation can also prevent uridylation of siRNAs in certain trypanosomatids.[107] The core components of RNA silencing are not consistently conserved across kinetoplastids. In *T. brucei* but not *Leishmania* (*Viannia*) sp., an HEN1 homologue methylates siRNAs thereby preventing their trimming and uridylation.[107]

 © 2017 Wiley Periodicals, Inc.

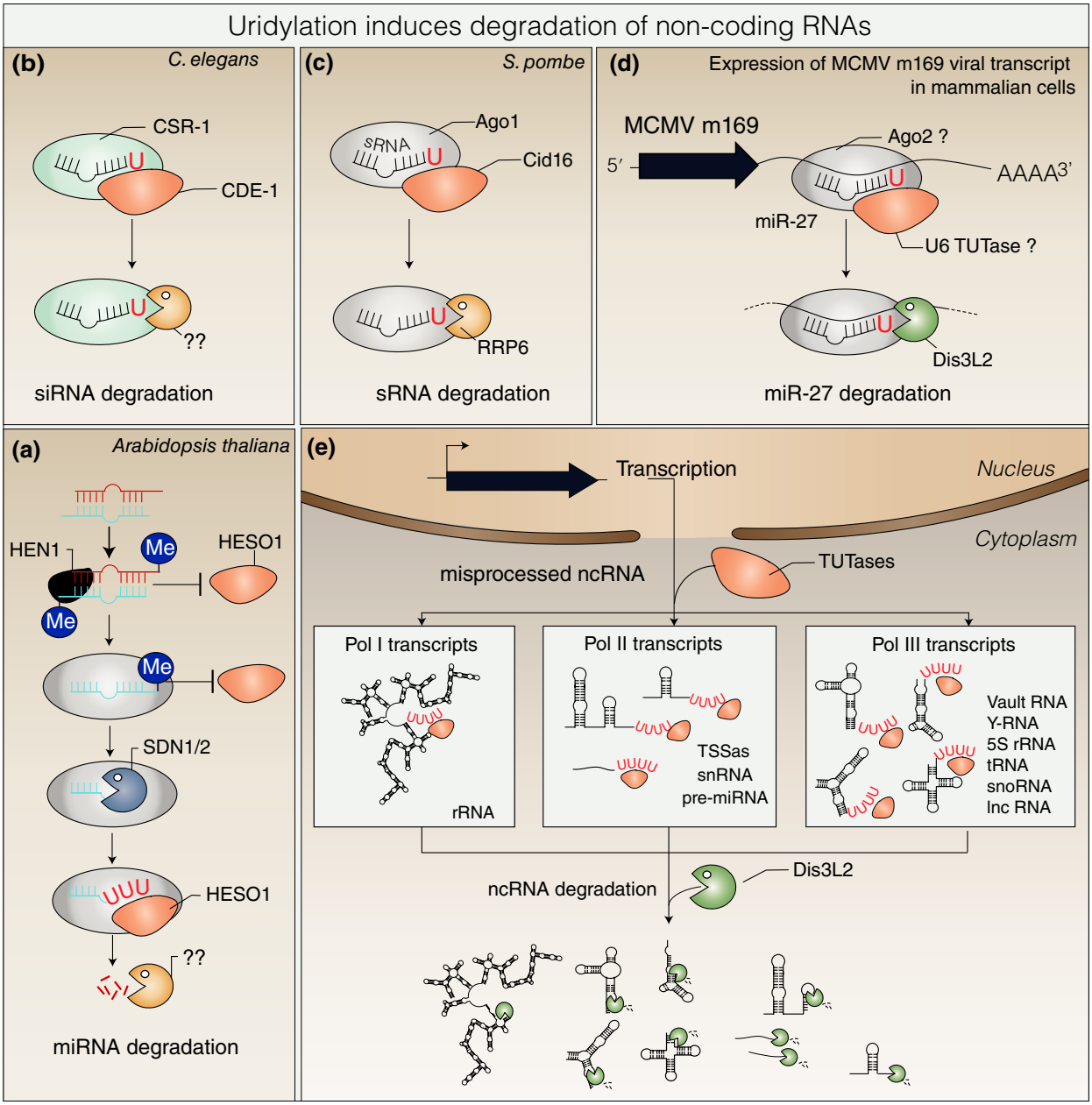


FIGURE 6 | Uridylation induces the degradation of various noncoding RNAs. A few examples were selected to illustrate the destabilizing role of uridylation on noncoding RNAs. (a) Methylation of small RNAs in *Arabidopsis* preventing uridylation by HESO1. SDN1/2 3′→5′ exoribonucleases trim small RNAs, thereby removing the terminal methylated nucleotide. After trimming by SDN1/2, HESO1 can proceed and uridylate small RNAs that are subsequently degraded by a yet unknown ribonuclease. (b) Uridylation of siRNAs by CDE-1 in *C. elegans*. (c) Uridylation of all classes of sRNAs by Cid16 in *S. pombe*. (d) Target RNA-directed miRNA degradation (TDMD) of miR-27 in mammalian cells expressing the mouse cytomegalovirus (MCMV) m169 transcript or in MCMV-infected murine cells. (e) Uridylation marks a plethora of structured and misprocessed noncoding transcripts produced by Pol I, Pol II, and Pol III to target them to cytosolic destruction by the 3′→5′ exoribonuclease Dis3L2. TSSas, transcription start site-associated short RNAs.

Uridylation and Decay of Small RNAs

Clearly, uridylation is not restricted to piRNAs, fly siRNAs, and plant small RNAs but is a widespread process revealed by numerous small RNA deep sequencing analyses in animals and fission yeast.[43,46,78,108,109] Uridylation of miRNAs processed from the 3′ arm of the pre-miRNA is frequent, indicating that mono-uridylation may often occur on pre-miRNAs, prior to loading onto Ago. Yet, clear examples of uridylation-mediated decay of small RNAs associated with Ago

 © 2017 Wiley Periodicals, Inc.

proteins have been reported.[39,50,53,54,92,110] Several TUTases have been involved in small RNA uridylation including CDE-1 in *Caenorhabditis elegans*, Cid16 in *S. pombe*, and TUT4, TUT7 and TUT1 in human cells.[53,78,110–113] CDE-1 is a TUTase that uridylates siRNAs bound by the Ago protein CSR-1 in *C. elegans* (Figure 6(b)).[53] In absence of CDE-1, CSR-1 siRNAs accumulate, leading to defects in chromosome segregation. Accumulated CSR-1 siRNAs also 'leak' into other Ago-mediated pathways, resulting in spurious gene silencing.[53] It was therefore concluded that uridylation by CDE-1 is required to destabilize CSR-1 bound siRNAs to restrict those specific siRNAs to enter other silencing pathways.[53] A related process was demonstrated recently in *S. pombe*.[110] More than 20% of Ago-bound small RNAs in fission yeast have one or two untemplated nucleotides, mostly adenosines but also uridines. Those nucleotides are added by the poly(-A) polymerase Cid14 and the TUTase Cid16, respectively.[110] Both uridylation and adenylation trigger the degradation of Ago-bound small RNAs by RRP6, a catalytic subunit of the exosome (Figure 6(c)).[110] Both Cid14 and Cid16 are essential to eliminate spurious small RNAs to prevent uncontrolled RNA silencing from targeting euchromatic genes.[110]

In humans, uridylation was proposed to decrease the abundance of several miRNAs.[46,78,111,114,115] Uridylation definitely participates in the destabilization of miRNAs in case of high complementarity to their targets.[50,106,114,116–118] This process is referred to as target RNA-directed miRNA degradation (TDMD).[117] TDMD plays a crucial role in the context of mouse cytomegalovirus (MCMV) infection. Binding of miR-27a/b to the abundant MCMV m169 transcript triggers their uridylation and degradation (Figure 6(d)).[116,119] Interestingly, TUT1 (U6 TUTase) co-purifies with Ago2 and with tailed and trimmed isoforms of miR-27 only when TDMD is induced.[50] Dis3L2 also co-purifies with Ago2 and degrades the uridylated miR-27 isoforms.[50]

Altogether, these examples show that uridylation can participate in the degradation of small RNAs in various eukaryotes. Yet, a link between uridylation and decay is not systematic,[78] likely reflecting alternative roles for miRNA mono-uridylation as illustrated below.

Uridylation Controls the Activity of miRNAs

Uridylation is crucial to control the activity of specific miRNAs. The stability of interleukin-6 (IL-6) and other specific cytokine mRNAs is tightly controlled to regulate the inflammatory response. This regulation is partly achieved by the miR26 family, which targets the 3′ UTR of IL-6 transcripts. Uridylation of miR-26 by the murine TUT4 abrogates IL-6 repression by preventing the binding of miR-26 to its targets without affecting miR-26 stability.[120] Knocking out TUT4 in mice does not alter embryogenesis but reduces growth and survival after birth. Genome-wide studies revealed a decreased tailing of some miRNAs but without affecting their abundance.[113] Importantly, TUT4 prevents the miRNA-mediated silencing of IGF-1 transcripts, IGF-1 being essential for early growth and survival. The phenotypes of TUT4-deficient mice were, therefore, partly explained by a decrease in IGF-1 mRNAs and protein due to TUT4 deficiency.[113] Those *in vivo* experiments revealed a general mechanism by which uridylation controls the activity of miRNAs independently of their stability.

Uridylation may Control Export to Exosomes

Uridylation was also proposed to act as a sorting signal to target miRNAs to endosome-derived exosomes.[121] Sequencing small RNA populations from human B cells and their secreted exosomes revealed distinct populations of intracellular and secreted miRNAs discriminated by their 3′ untemplated adenylation and uridylation. Intracellular subsets of miRNAs are preferentially adenylated whereas miRNAs targeted to extracellular vesicles are preferentially uridylated.[121] Although mechanistic insights are still required to fully understand this process, it exemplifies that more functions of miRNA uridylation are likely to be discovered.

Uridylation and Surveillance of Defective Noncoding RNAs

Since the identification of pre-let-7 miRNA as the first Dis3L2 uridylated RNA target,[34,41] genome-wide studies have thoroughly expanded the repertoire of Dis3L2 uridylated substrates in Drosophila, mouse, and human cells.[42–45] The most striking conclusion of these studies is that Dis3L2 and TUTases (TUT4/7 in mammals and Tailor in Drosophila) are the key factors of a RNA surveillance pathway targeting unprocessed, structured noncoding RNAs in the cytosol[42–45] (Figure 6(e)). The majority of Dis3L2 substrates correspond to noncoding RNAs transcribed by Pol III that include unprocessed tRNAs, vault and Y RNAs, an Alu-like element BC200 RNA, 7SL, 7SK, RNase P and RNase MRP RNAs, and 5S rRNA.[42–45] As Pol III terminates transcription following the synthesis of a short stretch of uridines, unprocessed transcripts may be targeted directly by Dis3L2 after export to the cytosol. However, TUT4/7 definitely assists Dis3L2 mediated-degradation by

 © 2017 Wiley Periodicals, Inc.

synthesizing short oligouridine tails, mostly at positions close to stable secondary structures (Figure 6(e)).[42,43,45] An elegant experiment based on the decay rate analysis of randomized terminal sequences confirmed that a short stretch of terminal uridines significantly enhances degradation by the Drosophila Dis3L2.[44]

In addition to many Pol III transcripts, Dis3L2 also degrades transcription start site-associated short RNAs (TSSas) that are uridylated in the cytoplasm (Figure 6(e)).[45] TSSas are generated from bidirectional promoters and stalling of RNA polymerase II followed by premature transcription termination. Similarly, a short transcript originating from the 5′ UTR of ferritin pre-mRNA is among the substrates of Dis3L2.[42] In addition, Dis3L2 targets several pre-miRNAs besides the expected let-7 pre-miRNAs (Figure 6(e)). The two most represented are pre-miR-484 and pre-miR-320, produced as TSS-terminated transcripts.[45] These prematurely terminated transcripts are possibly degraded as other TSSas. Alternatively, Dis3L2 could be involved in regulating the biogenesis of miR-484 and miR-320 miRNAs.[45] Of note, TUT4/7 definitely participates in the surveillance of defective pre-miRNAs. TUT7 recognizes 3′ trimmed pre-miRNAs and oligo-uridylates those defective precursors in the absence of Lin28.[46] In addition, TUT4/7 uridylates Ago-bound defective pre-miRNAs which are then degraded by the catalytic subunits of the RNA exosome, Dis3 and Rrp6.[39] It remains to be determined whether the exosome-bound Dis3 and Rrp6 preferentially act on nuclear substrates whereas Dis3L2 degrades defective uridylated pre-miRNAs in the cytosol, according to their respective main localization.

A major function of Dis3L2, together with TUT4/7, is the degradation of read-through forms of snRNAs (Figure 6(e)).[42–45] Importantly, Dis3L2 does not seem to participate in the production of mature forms of snRNAs but rather eliminates misprocessed precursors.[42] In addition, many of the Dis3L2 substrates originate from pseudogenes,[45] reinforcing the idea that Dis3L2 participates to a cytosolic pathway of RNA surveillance for various noncoding RNAs.

In line with the conserved cooperation between TUTases and Dis3L2 in degrading cytosolic RNAs, a complex between Tailor and Dis3L2 was discovered in Drosophila.[44] This complex was called terminal RNA uridylation-mediated processing (TRUMP), a reference to the Trf4/Air2/Mtr4p polyadenylation (TRAMP) complex which polyadenylates RNAs to facilitate their degradation by the nuclear exosome.[122] Hence, an interesting parallel emerges between noncoding RNA surveillance pathways mediated by TRAMP and the RNA exosome in the nucleus, and by TUTases and Dis3L2 in the cytosol.[42–45]

URIDYLATION OF mRNAs: DECAY AND OTHER CONSEQUENCES

Besides noncoding RNAs, uridylation also tags mRNAs. We and others have recently reviewed mRNA uridylation[5,9,123] and only key and novel aspects are presented here. mRNA uridylation was first reported for nonpolyadenylated RNA species: 5′ RISC-cleaved fragments in Arabidopsis and mouse[124] and the nonpolyadenylated replication-dependent histone mRNA in human cells.[125] Uridylation was thereafter detected for several polyadenylated mRNAs in fungi, plants, and animals.[38,126–129] The advent of a transcriptome-wide method called TAIL-seq designed to detect uridylation (and other untemplated addition of nucleotides) at the 3′ end of mRNAs,[129] has revealed the pervasiveness of mRNA uridylation in human cells and Arabidopsis.[40,57,129] Although the first described function of mRNA uridylation is to favor degradation,[38,40,80,124–127,129,130] the downstream consequences of uridylation emerge as multiple (Figure 7).

Uridylation Facilitates mRNA Decay

RISC-Cleaved mRNAs, Other Truncated mRNAs and Recycling of RISC

Uridylation of the 5′ fragment generated by RISC cleavage is conserved from plants to animals.[124,131] Arabidopsis HESO1 and human TUT2 are involved in uridylating 5′ RISC-cleaved fragments.[54,131] Yet, the full repertoire of TUTases involved in this process, their redundancy versus specificity remains to be fully explored. Uridylation favors the decay of 5′ RISC-cleaved fragments by promoting decapping, a required step prior to elimination by the cytosolic 5′→3′ exoribonuclease XRN.[37,124] The RNA exosome and its cytosolic cofactor, the Ski complex, also contribute to the elimination of 5′ RISC-cleaved fragments.[132,133] Of note, tailing by adenosines in *C. reinhardtii* also promotes the degradation of 5′ RISC-cleaved mRNAs by RRP6, a cofactor of the RNA exosome.[134]

Interestingly, the destabilization of 5′ RISC-cleaved fragments by uridylation was recently proposed to be important for recycling the RISC complex.[55] Two paralogous 3′→5′ exoribonucleases, called RISC-interacting clearing 3′→5′ exoribonucleases 1 and

 © 2017 Wiley Periodicals, Inc.

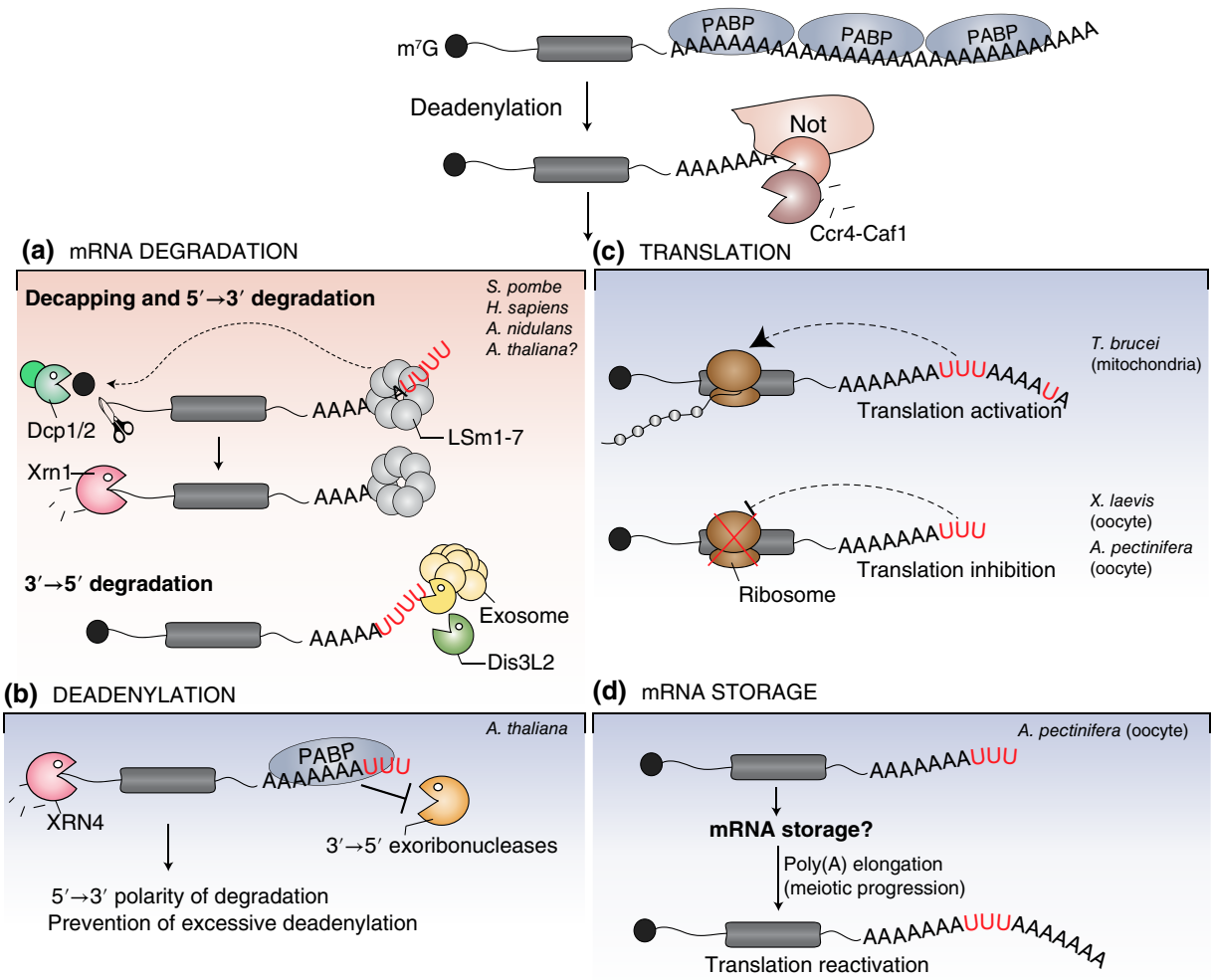


FIGURE 7 | Uridylation plays diverse roles in mRNA metabolism. Uridylation usually occurs after a deadenylation step. (a) The conserved effect of mRNA uridylation is to trigger degradation in eukaryotes. Recognition of uridylated oligoadenylated mRNAs by the LSm1-7 complex induces decapping and subsequent 5′→3′ degradation by XRN1. Alternatively, uridylated mRNAs are degraded from their 3′ end by Dis3L2 or the exosome. (b) In *Arabidopsis*, uridylation prevents excessive deadenylation of mRNAs by restoring an extension of suficient length to allow for PABP binding. (c) Uridylation can also inhibit translation in Xenopus (*X. laevis*) and starish (*Asterina pectinifera*) oocytes or activate translation of mitochondrial mRNAs in trypanosomes. (d) Uridylation could also be involved in mRNA storage in starfish oocytes.

2 (RICE1/2), stimulate the degradation of uridylated 5′ RISC-cleaved fragments in Arabidopsis[55] (Figure 8). RICE1/2 have a DnaQ-like exonuclease fold and forms a donut-shaped homohexamer.[55,135] The active sites are located at the interface formed by hexamer subunits explaining that oligomerization of RICEs is essential for activity. RICEs degrade single strand RNA and associate with AGO1 and AGO10.[55] miRNAs are not the targets of RICEs because downregulation of RICEs reduces miRNA levels with the concomitant accumulation of uridylated 5′ RISC-cleaved fragments. RICE-mediated degradation of uridylated 5′ RISC-cleaved fragments was therefore proposed to maintain functional RISC.[55] It is yet unknown whether RICE homologues would play a similar role in other organisms, including humans.

Small RNA-independent pathways can also generate mRNA fragments and a potential role of uridylation in assisting the elimination of such fragments remains to be explored thoroughly. Of note, mRNA fragments produced during apoptosis are uridylated by TUT4/7 and degraded by Dis3L2, as illustrated for *ACTB* and *EEF1A* mRNAs.[136] The advent of high throughput sequencing-based methods such as TAIL-seq[129] or 3′ RACE-seq[42] will probably reveal other examples of mRNA decay intermediates eliminated through the uridylation-mediated pathway.

 © 2017 Wiley Periodicals, Inc.

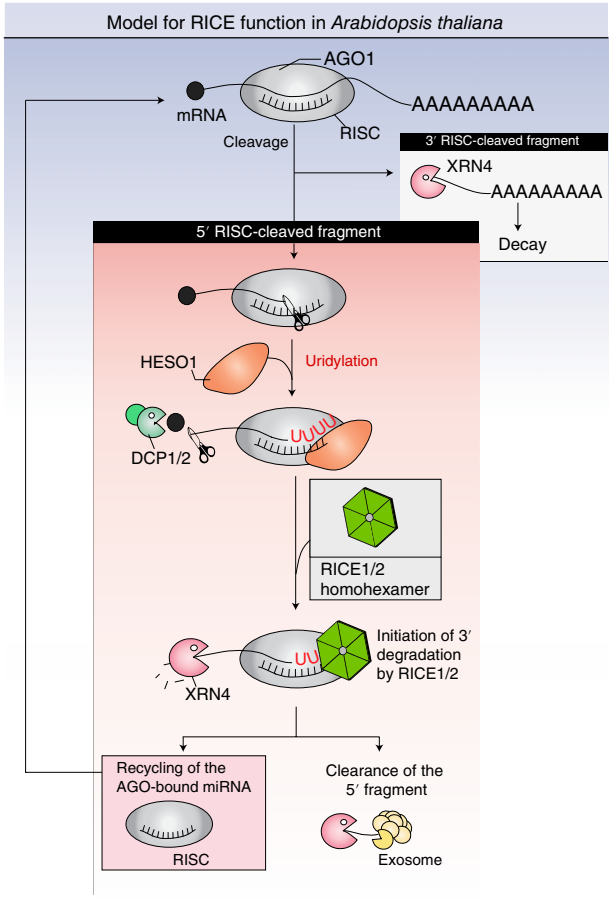


FIGURE 8 | Recycling of RNA-induced silencing complex (RISC) by RICE1/2. MicroRNA is incorporated into the RISC for target recognition. Perfect pairing of plant miRNA with its target supports slicing of the mRNA by the Argonaute protein. This cleavage results in two pieces, known as the 5′ and 3′ RISC-cleaved fragments, that will undergo different decay processes. The 3′ RISC-cleaved fragment is targeted by the 5′→3′ exoribonuclease XRN4. The 5′ RISC-cleaved fragment is uridylated by HESO1 and the degradation is initiated by RICE1/2. Clearance of the 5′ RISC-cleaved fragment is ensured by XRN4 and the exosome.

Replication-Dependent Histone mRNAs in Mammals

Replication-dependent histone mRNAs are not polyadenylated in mammals, but end with a terminal stem-loop (SL) structure, essential for processing, export from the nucleus, translation and stability.[137,138] The SL interacts with the SL-binding protein (SLBP) on the 5′ side and with the exoribonuclease Eri1 (3hExo) on the 3′ side.[139] Mature histone mRNAs end 3 nucleotides downstream of the SL. Nibbling of these terminal nucleotides (likely by Eri1) can be counterbalanced by uridylation that restores the full-length size of histone mRNAs.[140] At the end of the S-phase (or when replication is inhibited), histone mRNAs are rapidly eliminated. Degradation requires translation and the helicase Upf1, otherwise known as the central component of the nonsense-mediated decay pathway.[141] Upf1 interacts with SLBP and somehow favors the recruitment of either TUT4 or TUT7.[40,142,143] Binding of the LSm1-7 complex to the U-tail promotes Eri1 to nibble the SL,[144] and several uridylation/nibbling cycles overcome the protective effect of the SL to allow 3′→5′ degradation by the exosome-associated exonuclease PM/Scl-100 (Rrp6).[145] Alternatively, binding of the LSm1-7 complex activates decapping and subsequent 5′→3′ degradation.[125,146]

Polyadenylated mRNAs

The uridylation of polyadenylated mRNAs is conserved across eukaryotes, from animals to plants and fungi, with the noticeable exception so far of *S. cerevisiae*.[38,126–129a] Uridylation was first shown to induce both 5′→3′ and 3′→5′ degradation of selected model mRNAs in *S. pombe*.[38,80] The addition of uridines 3′ of the poly(A) tail by Cid1 has two effects. Firstly, it can promote the recruitment of the LSm1-7 complex, which in turn activates decapping and subsequent 5′→3′ degradation,[38] in line with a previous observation made using human cell extracts.[37] Secondly, uridylation can attract the exoribonuclease Dis3L2 to digest the mRNA from its 3' end.[80] A related posttranscriptional modification, the addition of CUCU by the TUTases CutA and CutB, plays a similar destabilizing role in *Aspergillus nidulans*.[126,127]

A landmark in the study of mRNA uridylation was the development of TAIL-seq, a high-throughput sequencing method allowing both the determination of poly(A) tail length and the detection of untemplated nucleotides.[40,129] TAIL-seq was decisive to demonstrate and generalize a link between uridylation and degradation of mRNAs. Firstly, TAIL-seq revealed that mRNA uridylation is widespread in human cells,[40,129] a conclusion later extended to Arabidopsis.[57] Secondly, uridylation of human mRNAs by TUT4/7 tags is preceded by deadenylation,[129] confirming at a transcriptome-wide level previous observations made in *A. nidulans* and Arabidopsis for candidate mRNAs.[126–128] Interestingly, uridylation is independent of deadenylation in *S. pombe*,[38] possibly because poly(A) tails are shorter than in plants or animals. Thirdly, knock down of key factors of both 5'→3′ and 3′→5′ RNA degradation pathways resulted in the accumulation of uridylated mRNAs. In addition, this accumulation is further increased by the concomitant depletion of XRN and the exosome, confirming that uridylation facilitates mRNA degradation from both ends.[40] Of note,

 © 2017 Wiley Periodicals, Inc.

depletion of Dis3L2 had only a modest effect on the accumulation of uridylated mRNAs as compared with knock down of the exosome or XRN1.[40] Fourthly and importantly, depletion of TUT4/7 resulted in increasing half-lives by 30% on average for 80% of the mRNAs detected in the study.[40] Altogether, those studies established uridylation as a generic step of mRNA degradation in eukaryotes (Figure 7).

Uridylation Prevents Excessive Deadenylation of Plant mRNAs

In Arabidopsis, uridylated mRNAs accumulate upon impairment of the 5′→3′ RNA decay pathway, indicating that uridylation likely tags plant mRNAs for degradation as in other eukaryotes.[57,128] Yet, URT1, the main TUTase responsible for 80% of mRNA uridylation in Arabidopsis, plays a distinct role in preventing excessive deadenylation of mRNAs. Indeed, *urt1* mutants accumulate excessively deadenylated mRNAs and overexpression of URT1 increases the oligo(A) tail length of deadenylated mRNAs.[57,128] Furthermore, TAIL-seq analysis revealed that URT1-mediated uridylation repairs oligo(A) tails to restore an average extension length of about 16 nucleotides. This length is sufficient for the binding of a poly(A) binding protein (PABP), thereby explaining the protection against excessive deadenylation conferred by URT1-mediated uridylation.[57] Although, URT1-mediated uridylation does not seem to affect the rate of mRNA decay, it participates in establishing the 5′→3′ polarity of mRNA degradation (Figure 7) which could be essential during co-translational decay.[128] A second TUTase, yet to be formally identified, does not prevent excessive deadenylation and may promote RNA decay.[5,57] Intriguingly, *in vitro* assays using human cell extracts also reported that uridylation favors decapping while conferring protection of the 3′ end, likely through binding of the LSm1-7 complex.[37] Favoring the 5′→3′ polarity of mRNA degradation might constitute an important role of uridylation besides facilitating degradation.

Translation Control by Uridylation

In trypanosomes, translation of both edited and never-edited mitochondrial mRNAs require 3′ tailing (Figure 7). Tails consist of a short A-tail extended by long A/U heteropolymers of 200–300 nucleotides. For edited mRNAs, the long A/U extensions are added once editing is completed. The short A-tail is synthesized by the poly(A) polymerase KPAP1 while long A/U extensions are added by the combined action of KPAP1 and the TUTase RET1.[26] The long A/U tails are required to activate translation by recruiting the small subunit of the ribosome.[26]

Conversely, uridylation was proposed to repress translation (Figure 7). Tethering XtTUT7 to reporter mRNAs injected in Xenopus oocytes prevented translation but did not affect mRNA stability.[21] Repression of reporter gene expression was also observed by tethering TUTases to reporter transcripts in human cells.[40] However, in this case, gene repression was linked to transcript destabilization.[40] Therefore, uridylation-mediated translation inhibition could be dependent on the cellular context. In line with this, 96% of *cyclin B* mRNAs stored in starfish oocytes are uridylated and uridylation triggers trimming followed by poly(A) extension only upon meiotic reinitiation by hormonal stimulation (Figure 7).[147] Further work is required to determine whether uridylation could trigger translation inhibition and storage under particular physiological conditions or at certain developmental stages.

URIDYLATION OF VIRAL RNAs

Extensive internal and terminal uridylation has been reported for various viral genomic RNAs and virus-encoded RNAs infecting fungi, plant, and animal cells.[148–152] Uridylation targets positive, negative, or double-stranded RNA viruses that can end with a poly(A) tail, a tRNA-like sequence (TLS) or a non-TLS heteropolymeric sequence (Het).[152] Uridylation of viral RNAs is therefore a widespread process in eukaryotes.[152] Uridylation, together with adenylation, was proposed to repair various truncated viral RNAs such as Beet necrotic yellow vein virus (BNYVV), Sindbis virus (SIN), coxsackievirus B3 (CVB3) and hepatitis C virus (HCV).[148–151] In light of our current knowledge on uridylation-mediated RNA degradation, the potential of uridylation as a restrictive mechanism during viral infections would be worth investigating.

CONCLUSION

The diversity of posttranscriptional regulations mediated by RNA uridylation is yet to be fully explored. Numerous examples have now demonstrated that uridylation can mark virtually all classes of RNAs expressed in eukaryotic cells, including pathogenic RNAs such as viral RNAs. Those substrates can be of all sizes and have various termini, from unstructured poly(A) tails to structured ends. TUTases have evolved to recognize a huge diversity of RNA substrates, either directly or through the assistance of auxiliary factors. The further

 © 2017 Wiley Periodicals, Inc.

identification of such factors assisting TUTases in recognizing specific RNA substrates, and of 'readers' that influence the fate of uridylated transcripts will definitely be key to unravel all regulatory roles due to uridylation.

Uridylation-mediated RNA degradation is definitely among the crucial functions of uridylation. Notably, a uridylation- and Dis3L2-mediated surveillance pathway is key for the degradation of defective noncoding RNAs.[42–45] Uridylated misprocessed transcripts were also detected in human mitochondria raising the possibility that uridylated-mediated RNA surveillance might also operate in this organelle.[153,154] Uridylation is also assisting the degradation of cytosolic mRNAs. The basic molecular mechanisms explaining how uridylation can promote mRNA degradation from both ends have been detailed. Yet, uridylation can influence the process of RNA degradation by additional ways than just accelerating decay. For instance, by establishing the polarity of degradation and possibly its subcellular localization. In addition, the potential of uridylation in regulating translation or mRNA storage is just beginning to be evaluated. Further studies are required to fully elucidate the molecular mechanisms underlying uridylation-mediated regulation of gene expression and to fully appreciate its impact during development or in response to pathogen attacks and diseases.

NOTE

[a] Since the acceptation of this manuscript for publication, a study by Morgan et al. (doi:10.1038/nature23318) showed that uridylation by TUT4/7 is crucial to shape the mouse maternal transcriptome by eliminating mRNAs during oocyte growth.

ACKNOWLEDGMENTS

The activity in our group is currently supported by the Centre National de la Recherche Scientifique (CNRS) and research grants from the French National Research Agency as part of the 'Investments for the Future' program in the frame of the LABEX ANR-10-LABX-0036_NETRNA and ANR-15-CE12-0008-01 to D.G.

REFERENCES

1. Scott DD, Norbury CJ. RNA decay via 3′ uridylation. *Biochim Biophys Acta* 2013, 1829:654–665. https://doi.org/10.1016/j.bbagrm.2013.01.009.
2. Norbury CJ. Cytoplasmic RNA: a case of the tail wagging the dog. *Nat Rev Mol Cell Biol* 2013, 13:643–653. https://doi.org/10.1038/nrm3645.
3. Song JB, Song J, Mo BX, Chen XM. Uridylation and adenylation of RNAs. *Sci China Life Sci* 2015, 58:1057–1066. https://doi.org/10.1007/s11427-015-4954-9.
4. Viegas SC, Silva IJ, Apura P, Matos RG, Arraiano CM. Surprises in the 3′-end: "U" can decide too!. *FEBS J* 2015, 282:3489–3499. https://doi.org/10.1111/febs.13377.
5. Scheer H, Zuber H, De Almeida C, Gagliardi D. Uridylation earmarks mRNAs for degradation… and more. *Trends Genet* 2016, 32:607–619. https://doi.org/10.1016/j.tig.2016.08.003.
6. Aphasizhev R, Suematsu T, Zhang L, Aphasizheva I. Constructive edge of uridylation-induced RNA degradation. *RNA Biol* 2016, 13:1078–1083. https://doi.org/10.1080/15476286.2016.1229736.
7. Ji L, Chen X. Regulation of small RNA stability: methylation and beyond. *Cell Res* 2012, 22:624–636. https://doi.org/10.1038/cr.2012.36.
8. Martin G, Keller W. RNA-specific ribonucleotidyl transferases. *RNA* 2007, 13:1834–1849. https://doi.org/10.1261/rna.652807.
9. Munoz-Tello P, Rajappa L, Coquille S, Thore S. Poly-uridylation in eukaryotes: a 3′-end modification regulating RNA life. *Biomed Res Int* 2015, 2015:1–12. https://doi.org/10.1155/2015/968127.
10. Aphasizhev R, Aphasizheva I. Terminal RNA uridylyltransferases of trypanosomes. *Biochim Biophys Acta* 2008, 1779:270–280. https://doi.org/10.1016/j.bbagrm.2007.12.007.
11. Aphasizhev R. RNA uridylyltransferases. *Cell Mol Life Sci* 2005, 62:2194–2203. https://doi.org/10.1007/s00018-005-5198-9.
12. Munoz-Tello P, Gabus C, Thore S. Functional implications from the Cid1 poly(U) polymerase crystal structure. *Structure* 2012, 20:977–986. https://doi.org/10.1016/j.str.2012.04.006.
13. Yates LA, Fleurdépine S, Rissland OS, De Colibus L, Harlos K, Norbury CJ, Gilbert RJC. Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. *Nat Struct Mol Biol* 2012, 19:782–787. https://doi.org/10.1038/nsmb.2329.
14. Yates LA, Durrant BP, Fleurdépine S, Harlos K, Norbury CJ, Gilbert RJC. Structural plasticity of Cid1 provides a basis for its distributive RNA

 © 2017 Wiley Periodicals, Inc.

terminal uridylyl transferase activity. *Nucleic Acids Res* 2015, 43:2968–2979. https://doi.org/10.1093/nar/gkv122.

15. Rajappa-Titu L, Suematsu T, Munoz-Tello P, Long M, Demir Ö, Cheng KJ, Stagno JR, Luecke H, Amaro RE, Aphasizheva I, et al. RNA editing TUTase 1: structural foundation of substrate recognition, complex interactions and drug targeting. *Nucleic Acids Res* 2016, 44:10862–10878. https://doi.org/10.1093/nar/gkw917.
16. Yamashita S, Takagi Y, Nagaike T, Tomita K. Crystal structures of U6 snRNA-specific terminal uridylyltransferase. *Nat Commun* 2017, 8:15788. https://doi.org/10.1038/ncomms15788.
17. Stagno J, Aphasizheva I, Bruystens J, Luecke Hartmut H, Aphasizhev R. Structure of the mitochondrial editosome-like complex associated TUTase 1 reveals divergent mechanisms of UTP selection and domain organization. *J Mol Biol* 2010, 399:464–475. https://doi.org/10.1016/j.jmb.2010.04.021.
18. Stagno J, Aphasizheva I, Rosengarth A, Luecke H, Aphasizhev R. UTP-bound and Apo structures of a minimal RNA uridylyltransferase. *J Mol Biol* 2007, 366:882–899. https://doi.org/10.1016/j.jmb.2006.11.065.
19. Stagno J, Aphasizheva I, Aphasizhev R, Luecke H. Dual role of the RNA substrate in selectivity and catalysis by terminal uridylyl transferases. *Proc Natl Acad Sci* 2007, 104:14634–14639. https://doi.org/10.1073/pnas.0704259104.
20. Chung C, Jo DHS, Heinemann IU. Nucleotide specificity of the human terminal nucleotidyltransferase Gld2 (TUT2). *RNA* 2016, 2:1239–1249. https://doi.org/10.1261/rna.056077.116.
21. Lapointe CP, Wickens M. The nucleic acid-binding domain and translational repression activity of a Xenopus terminal uridylyl transferase. *J Biol Chem* 2013, 288:20723–20733. https://doi.org/10.1074/jbc.M113.455451.
22. Blahna MT, Jones MR, Quinton LJ, Matsuura KY, Mizgerd JP. Terminal uridyltransferase enzyme Zcchc11 promotes cell proliferation independent of its uridyltransferase activity. *J Biol Chem* 2011, 286:42381–42389. https://doi.org/10.1074/jbc.M111.259689.
23. Faehnle CR, Walleshauser J, Joshua-tor L. Multidomain utilization by TUT4 and TUT7 in control of let-7 biogenesis. *Nat Struct Mol Biol* 2017, 24:658–665. https://doi.org/10.1038/nsmb.3428.
24. Järvelin AI, Noerenberg M, Davis I, Castello A. The new (dis)order in RNA regulation. *Cell Commun Signal* 2016, 14:9. https://doi.org/10.1186/s12964-016-0132-3.
25. Suematsu T, Zhang L, Aphasizheva I, Monti S, Huang L, Wang Q, Costello CE, Aphasizhev R. Antisense transcripts delimit exonucleolytic activity of the mitochondrial 3′ Processome to generate guide RNAs. *Mol Cell* 2016, 61:364–378. https://doi.org/10.1016/j.molcel.2016.01.004.
26. Aphasizheva I, Maslov D, Wang X, Huang L, Aphasizhev R. Pentatricopeptide repeat proteins stimulate mRNA adenylation/uridylation to activate mitochondrial translation in trypanosomes. *Mol Cell* 2011, 42:106–117. https://doi.org/10.1016/j.molcel.2011.02.021.
27. Calabretta S, Richard S. Emerging roles of disordered sequences in RNA-binding proteins. *Trends Biochem Sci* 2015, 40:662–672. https://doi.org/10.1016/j.tibs.2015.08.012.
28. Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, Curk T, Krijgsveld J, Hentze MW. Comprehensive identification of RNA-binding domains in human cells. *Mol Cell* 2016, 63:696–710. https://doi.org/10.1016/j.molcel.2016.06.029.
29. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 2012, 149:1393–1406. https://doi.org/10.1016/j.cell.2012.04.031.
30. Hagan JP, Piskounova E, Gregory RI. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat Struct Mol Biol* 2009, 16:1021–1025. https://doi.org/10.1038/nsmb.1676.
31. Heo I, Joo C, Kim YK, Ha M, Yoon MJ, Cho J, Yeom KH, Han J, Kim VN. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* 2009, 138:696–708. https://doi.org/10.1016/j.cell.2009.08.002.
32. Thornton JE, Chang H-M, Piskounova E, Gregory RI. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA* 2012, 18:1875–1885. https://doi.org/10.1261/rna.034538.112.
33. Wang L, Nam Y, Lee AK, Yu C, Roth K, Chen C, Ransey EM, Sliz P. LIN28 zinc knuckle domain is required and sufficient to induce let-7 oligouridylation. *Cell Rep* 2017, 18:2664–2675. https://doi.org/10.1016/j.celrep.2017.02.044.
34. Ustianenko D, Hrossova D, Potesil D, Chalupnikova K, Hrazdilova K, Pachernik J, Cetkovska K, Uldrijan S, Zdrahal Z, Vanacova S. Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* 2013, 19:1632–1638. https://doi.org/10.1261/rna.040055.113.
35. Choudhury NR, Nowak JS, Zuo J, Rappsilber J, Spoel SH, Michlewski G. Trim25 is an RNA-specific activator of Lin28a/TuT4-mediated uridylation. *Cell*

 © 2017 Wiley Periodicals, Inc.

Rep 2014, 9:1265–1272. https://doi.org/10.1016/j.celrep.2014.10.017.

36. Chowdhury A, Mukhopadhyay J, Tharun S. The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA* 2007, 13:998–1016. https://doi.org/10.1261/rna.502507.
37. Song M-G, Kiledjian M. 3′ terminal oligo U-tract-mediated stimulation of decapping. *RNA* 2007, 13:2356–2365. https://doi.org/10.1261/rna.765807.
38. Rissland OS, Norbury CJ. Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. *Nat Struct Mol Biol* 2009, 16:616–623. https://doi.org/10.1038/nsmb.1601.
39. Liu X, Zheng Q, Vrettos N, Maragkakis M, Alexiou P, Gregory BD, Mourelatos ZA. A microRNA precursor surveillance system in quality control of microRNA synthesis. *Mol Cell* 2014, 55:868–879. https://doi.org/10.1016/j.molcel.2014.07.017.
40. Lim J, Ha M, Chang H, Kwon SC, Simanshu DK, Patel DJ, Kim VN. Uridylation by TUT4 and TUT7 marks mRNA for degradation. *Cell* 2014, 159:1365–1376. https://doi.org/10.1016/j.cell.2014.10.055.
41. Chang H, Triboulet R, Thornton JE, Gregory RI. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28–let-7 pathway. *Nature* 2013, 497:244–248. https://doi.org/10.1038/nature12119.
42. Labno A, Warkocki Z, Kulinski T, Krawczyk PS, Bijata K, Tomecki R, Bijata K, Tomecki R, Dziembowski A. Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. *Nucleic Acids Res* 2016, 44:10437–10453. https://doi.org/10.1093/nar/gkw649.
43. Pirouz M, Du P, Munafo M, Gregory RI. Dis3l2-mediated decay is a quality control pathway for noncoding RNAs. *Cell Rep* 2016, 16:1861–1873. https://doi.org/10.1016/j.celrep.2016.07.025.
44. Reimão-Pinto MM, Manzenreither RA, Burkard TR, Sledz P, Jinek M, Mechtler K, Ameres SL. Molecular basis for cytoplasmic RNA surveillance by uridylation-triggered decay in Drosophila. *EMBO J* 2016, 35:2417–2434. https://doi.org/10.15252/embj.201695164.
45. Ustianenko D, Pasulka J, Feketova Z, Bednarik L, Zigackova D, Fortova A, Zavolan M, Vanacova S. TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. *EMBO J* 2016, 35:2179–2191. https://doi.org/10.15252/embj.201694857.
46. Kim B, Ha M, Loeff L, Chang H, Simanshu DK, Li S, Fareh M, Patel DJ, Joo C, Kim VN. TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. *EMBO J* 2015, 34:1801–1815. https://doi.org/10.15252/embj.201590931.
47. Hilcenko C, Simpson PJ, Finch AJ, Bowler FR, Churcher MJ, Jin L, Packman LC, Shlien A, Campbell P, Kirwan M, et al. Aberrant 3′ oligoadenylation of spliceosomal U6 small nuclear RNA in poikiloderma with neutropenia. *Blood* 2013, 121:1028–1038. https://doi.org/10.1182/blood-2012-10.
48. Mroczek S, Krwawicz J, Kutner J, Lazniewski M, Kuciński I, Ginalski K, Dziembowski A. C16orf57, a gene mutated in poikiloderma with neutropenia, encodes a putative phosphodiesterase responsible for the U6 snRNA 3′ end modification. *Genes Dev* 2012, 26:1911–1925. https://doi.org/10.1101/gad.193169.112.
49. Shchepachev V, Wischnewski H, Missiaglia E, Soneson C, Azzalin CM. Mpn1, mutated in poikiloderma with neutropenia Protein 1, is a conserved 3′-to-5′ RNA exonuclease processing U6 small nuclear RNA. *Cell Rep* 2012, 2:855–865. https://doi.org/10.1016/j.celrep.2012.08.031.
50. Haas G, Cetin S, Messmer M, Chane-Woon-Ming B, Terenzi O, Chicher J, Kuhn L, Hammann P, Pfeffer S. Identification of factors involved in target RNA-directed microRNA degradation. *Nucleic Acids Res* 2016, 44:2873–2887. https://doi.org/10.1093/nar/gkw040.
51. Licht K, Medenbach J, Luhrmann R, Kambach C, Bindereif A. 3′-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. *RNA* 2008, 14:1532–1538. https://doi.org/10.1261/rna.1129608.
52. Rüegger S, Miki TS, Hess D, Großhans H. The ribonucleotidyl transferase USIP-1 acts with SART3 to promote U6 snRNA recycling. *Nucleic Acids Res* 2015, 43:3344–3357. https://doi.org/10.1093/nar/gkv196.
53. van Wolfswinkel JC, Claycomb JM, Batista PJ, Mello CC, Berezikov E, Ketting RF. CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. *Cell* 2009, 139:135–148. https://doi.org/10.1016/j.cell.2009.09.012.
54. Ren G, Xie M, Zhang S, Vinovskis C, Chen X, Yu B. Methylation protects microRNAs from an AGO1-associated activity that uridylates 5′ RNA fragments generated by AGO1 cleavage. *Proc Natl Acad Sci* 2014, 111:6365–6370. https://doi.org/10.1073/pnas.1405083111.
55. Zhang Z, Hu F, Sung MW, Shu C, Castillo-González C, Koiwa H, Tang G, Dickman M, Li P, Zhang X. RISC-interacting clearing 3′-5′ exoribonucleases (RICEs) degrade uridylated cleavage fragments to maintain functional RISC in *Arabidopsis thaliana*. *Elife* 2017, 6:1–29. https://doi.org/10.7554/eLife.24466.

 © 2017 Wiley Periodicals, Inc.

56. Yu Y, Ji L, Le BH, Zhai J, Chen J, Luscher E, Gao L, Liu C, Cao X, Mo B, et al. ARGONAUTE10 promotes the degradation of miR165/6 through the SDN1 and SDN2 exonucleases in Arabidopsis. *PLoS Biol* 2017, 15:1–26. https://doi.org/10.1371/journal.pbio.2001272.
57. Zuber H, Scheer H, Ferrier E, Sement FM, Mercier P, Stupfler B, Gagliardi D. Uridylation and PABP cooperate to repair mRNA deadenylated ends in Arabidopsis. *Cell Rep* 2016, 14:2707–2717. https://doi.org/10.1016/j.celrep.2016.02.060.
58. Aphasizheva I, Aphasizhev R. U-insertion/deletion mRNA-editing holoenzyme: definition in sight. *Trends Parasitol* 2016, 32:144–156. https://doi.org/10.1016/j.pt.2015.10.004.
59. Aphasizhev R, Aphasizheva I, Simpson L. A tale of two TUTases. *Proc Natl Acad Sci U S A* 2003, 100:10617–10622. https://doi.org/10.1073/pnas.1833120100.
60. Ringpis GE, Aphasizheva I, Wang X, Huang L, Lathrop RH, Hatfield GW, Aphasizhev R. Mechanism of U insertion RNA editing in trypanosome mitochondria: the bimodal TUTase activity of the core complex. *J Mol Biol* 2010, 399:680–695. https://doi.org/10.1016/j.jmb.2010.03.050.
61. Aphasizhev R, Aphasizheva I. Mitochondrial RNA processing in trypanosomes. *Res Microbiol* 2011, 162:655–663. https://doi.org/10.1016/j.resmic.2011.04.015.
62. Mroczek S, Dziembowski A. U6 RNA biogenesis and disease association. *WIREs RNA* 2013, 4:581–592. https://doi.org/10.1002/wrna.1181.
63. Trippe R, Sandrock B, Benecke BJ. A highly specific terminal uridylyl transferase modifies the 3′-end of U6 small nuclear RNA. *Nucleic Acids Res* 1998, 26:3119–3126. https://doi.org/10.1093/nar/26.13.3119.
64. Trippe R, Guschina E, Hossbach M, Urlaub H, Lührmann R, Benecke B-J. Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA* 2006, 12:1494–1504. https://doi.org/10.1261/rna.87706.
65. Trippe R, Richly H, Benecke BJ. Biochemical characterization of a U6 small nuclear RNA-specific terminal uridylyltransferase. *Eur J Biochem* 2003, 270:971–980. https://doi.org/10.1046/j.1432-1033.2003.03466.x.
66. Shchepachev V, Azzalin CM. The Mpn1 RNA exonuclease: cellular functions and implication in disease. *FEBS Lett* 2013, 587:1858–1862. https://doi.org/10.1016/j.febslet.2013.05.005.
67. Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* 2014, 15:509–524. https://doi.org/10.1038/nrm3838.
68. Heo I, Ha M, Lim J, Yoon MJ, Park JE, Kwon SC, Chang H, Kim VN. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell* 2012, 151:521–532. https://doi.org/10.1016/j.cell.2012.09.022.
69. Balzeau J, Menezes MR, Cao S, Hagan JP. The LIN28/let-7 pathway in cancer. *Front Genet* 2017, 8:1–16. https://doi.org/10.3389/fgene.2017.00031.
70. Kim SK, Lee H, Han K, Kim SC, Choi Y, Park SW, Bak G, Lee Y, Choi JK, Kim TK, et al. SET7/9 methylation of the pluripotency factor LIN28A is a nucleolar localization mechanism that blocks let-7 biogenesis in human ESCs. *Cell Stem Cell* 2014, 15:735–749. https://doi.org/10.1016/j.stem.2014.10.016.
71. Heo I, Joo C, Cho J, Ha M, Han J, Kim VN. Lin28 mediates the terminal uridylation of let-7 precursor microRNA. *Mol Cell* 2008, 32:276–284. https://doi.org/10.1016/j.molcel.2008.09.014.
72. Suzuki HI, Katsura A, Miyazono K. A role of uridylation pathway for blockade of let-7 microRNA biogenesis by Lin28B. *Cancer Sci* 2015, 106:1174–1181. https://doi.org/10.1111/cas.12721.
73. Piskounova E, Polytarchou C, Thornton JE, Lapierre RJ, Pothoulakis C, Hagan JP, Iliopoulos D, Gregory RI. Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell* 2011, 147:1066–1079. https://doi.org/10.1016/j.cell.2011.10.039.
74. Lehrbach NJ, Armisen J, Lightfoot HL, Murfitt KJ, Bugaut A, Balasubramanian S, Miska EA. LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in *Caenorhabditis elegans*. *Nat Struct Mol Biol* 2009, 16:1016–1020. https://doi.org/10.1038/nsmb.1675.
75. Huang Y. A mirror of two faces: Lin28 as a master regulator of both miRNA and mRNA. *WIREs RNA* 2012, 3:483–494. https://doi.org/10.1002/wrna.1112.
76. Nam Y, Chen C, Gregory RI, Chou JJ, Sliz P. Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell* 2011, 147:1080–1091. https://doi.org/10.1016/j.cell.2011.10.020.
77. Faehnle CR, Walleshauser J, Joshua-Tor L. Mechanism of Dis3l2 substrate recognition in the Lin28–let-7 pathway. *Nature* 2014, 514:252–256. https://doi.org/10.1038/nature13553.
78. Thornton JE, Du P, Jing L, Sjekloca L, Lin S, Grossi E, Sliz P, Zon LI, Gregory RI. Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). *Nucleic Acids Res* 2014, 42:11777–11791. https://doi.org/10.1093/nar/gku805.
79. Lubas M, Damgaard CK, Tomecki R, Cysewski D, Jensen TH, Dziembowski A. Exonuclease hDIS3L2 specifies an exosome-independent 3′-5′ degradation pathway of human cytoplasmic mRNA. *EMBO J* 2013, 32:1855–1868. https://doi.org/10.1038/emboj.2013.135.

 © 2017 Wiley Periodicals, Inc.

80. Malecki M, Viegas SC, Carneiro T, Golik P, Dressaire C, Ferreira MG, Arraiano CM. The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J* 2013, 32:1842–1854. https://doi.org/10.1038/emboj.2013.63.
81. Zhang W, Murphy C, Sieburth LE. Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. *Proc Natl Acad Sci* 2010, 107:15981–15985. https://doi.org/10.1073/pnas.1007060107.
82. Westholm JO, Lai EC. Mirtrons: microRNA biogenesis via splicing. *Biochimie* 2011, 93:1897–1904. https://doi.org/10.1016/j.biochi.2011.06.017.
83. Reimão-Pinto MM, Ignatova V, Burkard TR, Hung JH, Manzenreither RA, Sowemimo I, Herzog VA, Reichholf B, Fariña-Lopez S, Ameres SL. Uridylation of RNA hairpins by Tailor confines the emergence of microRNAs in Drosophila. *Mol Cell* 2015, 59:203–216. https://doi.org/10.1016/j.molcel.2015.05.033.
84. Bortolamiol-Becet D, Hu F, Jee D, Wen J, Okamura K, Lin CJ, Ameres SL, Lai EC. Selective suppression of the splicing-mediated microRNA pathway by the terminal uridyltransferase Tailor. *Mol Cell* 2015, 59:217–228. https://doi.org/10.1016/j.molcel.2015.05.034.
85. Wen J, Ladewig E, Shenker S, Mohammed J, Lai EC. Analysis of nearly one thousand mammalian mirtrons reveals novel features of dicer substrates. *PLoS Comput Biol* 2015, 11:e1004441. https://doi.org/10.1371/journal.pcbi.1004441.
86. Westholm JO, Ladewig E, Okamura K, Robine N, Lai EC. Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs. *RNA* 2012, 18:177–192. https://doi.org/10.1261/rna.030627.111.
87. Ren G, Chen X, Yu B. Small RNAs meet their targets: when methylation defends miRNAs from uridylation. *RNA Biol* 2014, 11:1099–1104. https://doi.org/10.4161/rna.36243.
88. Li J, Yang Z, Yu B, Liu J, Chen X. Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in Arabidopsis. *Curr Biol* 2005, 15:1501–1507. https://doi.org/10.1016/j.cub.2005.07.029.
89. Yu B. Methylation as a crucial step in plant microRNA biogenesis. *Science* 2005, 307:932–935. https://doi.org/10.1126/science.1107130.
90. Yang Z, Ebright YW, Yu B, Chen X. HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2′ OH of the 3′ terminal nucleotide. *Nucleic Acids Res* 2006, 34:667–675. https://doi.org/10.1093/nar/gkj474.
91. Abe M, Yoshikawa T, Nosaka M, Sakakibara H, Sato Y, Nagato Y, Itoh J-i. WAVY LEAF1, an ortholog of Arabidopsis HEN1, regulates shoot development by maintaining microRNA and trans-acting small interfering RNA accumulation in rice. *Plant Physiol* 2010, 154:1335–1346. https://doi.org/10.1104/pp.110.160234.
92. Zhai J, Zhao Y, Simon SA, Huang S, Petsch K, Arikit S, Pillay M, Ji L, Xie M, Cao X, et al. Plant microRNAs display differential 3′ truncation and tailing modifications that are ARGONAUTE1 dependent and conserved across species. *Plant Cell* 2013, 25:2417–2428. https://doi.org/10.1105/tpc.113.114603.
93. Ren G, Chen X, Yu B. Uridylation of miRNAs by HEN1 SUPPRESSOR1 in Arabidopsis. *Curr Biol* 2012, 22:695–700. https://doi.org/10.1016/j.cub.2012.02.052.
94. Zhao Y, Mo B, Chen X. Mechanisms that impact microRNA stability in plants. *RNA Biol* 2012, 9:1218–1223. https://doi.org/10.4161/rna.22034.
95. Zhao Y, Yu Y, Zhai J, Ramachandran V, Theresa T. HESO1, a nucleotidyl transferase in Arabidopsis, uridylates unmethylated miRNAs and siRNAs to trigger their degradation. *Curr Biol* 2012, 22:689–694. https://doi.org/10.1016/j.cub.2012.02.051.HESO1.
96. Wang X, Zhang S, Dou Y, Zhang C, Chen X, Yu B, Ren G. Synergistic and independent actions of multiple terminal nucleotidyl transferases in the 3′ tailing of small RNAs in Arabidopsis. *PLoS Genet* 2015, 11:e1005091. https://doi.org/10.1371/journal.pgen.1005091.
97. Tu B, Liu L, Xu C, Zhai J, Li S, Lopez MA, Zhao Y, Yu Y, Ramachandran V, Ren G, et al. Distinct and cooperative activities of HESO1 and URT1 nucleotidyl transferases in microRNA turnover in Arabidopsis. *PLoS Genet* 2015, 11:e1005119. https://doi.org/10.1371/journal.pgen.1005119.
98. Ibrahim F, Rymarquis LA, Kim E-J, Becker J, Balassa E, Green PJ, Cerutti H. Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. *Proc Natl Acad Sci* 2010, 107:3906–3911. https://doi.org/10.1073/pnas.0912632107.
99. Lim SL, Qu ZP, Kortschak RD, Lawrence DM, Geoghegan J, Hempfling AL, Bergmann M, Goodnow CC, Ormandy CJ, Wong L, et al. HENMT1 and piRNA stability are required for adult male germ cell transposon repression and to define the spermatogenic program in the mouse. *PLoS Genet* 2015, 11:1–30. https://doi.org/10.1371/journal.pgen.1005620.
100. Horwich MD, Li C, Matranga C, Vagin V, Farley G, Wang P, Zamore PD. The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol* 2007, 17:1265–1272. https://doi.org/10.1016/j.cub.2007.06.030.

 © 2017 Wiley Periodicals, Inc.

101. Saito K, Sakaguchi Y, Suzuki T, Siomi H, Siomi MC. Pimet, the Drosophila homolog of HEN1, mediates 2′-O-methylation of Piwi-interacting RNAs at their 3′ ends. *Genes Dev* 2007, 21:1603–1608. https://doi.org/10.1101/gad.1563607.the.

102. Kurth HM, Mochizuki K. 2′-O-methylation stabilizes piwi-associated small RNAs and ensures DNA elimination in Tetrahymena. *RNA* 2009, 15:675–685. https://doi.org/10.1261/rna.1455509.

103. Montgomery TA, Rim YS, Zhang C, Dowen RH, Phillips CM, Fischer SEJ, Ruvkun G. PIWI associated siRNAs and piRNAs specifically require the *Caenorhabditis elegans* HEN1 ortholog henn-1. *PLoS Genet* 2012, 8:e1002616. https://doi.org/10.1371/journal.pgen.1002616.

104. Kamminga LM, Luteijn MJ, den Broeder MJ, Redl S, Kaaij LJT, Roovers EF, Ladurner P, Berezikov E, Ketting RF. Hen1 is required for oocyte development and piRNA stability in zebrafish. *EMBO J* 2010, 29:3688–3700. https://doi.org/10.1038/emboj.2010.233.

105. Kirino Y, Mourelatos Z. The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. *RNA* 2007, 13:1397–1401. https://doi.org/10.1261/rna.659307.

106. Ameres SL, Horwich MD, Hung J-H, Xu J, Ghildiyal M, Weng Z, Zamore PD. Target RNA-directed trimming and tailing of small silencing RNAs. *Science* 2010, 328:1534–1539. https://doi.org/10.1126/science.1187058.

107. Shi H, Barnes RL, Carriero N, Atayde VD, Tschudi C, Ullu E. Role of the *Trypanosoma brucei* HEN1 family methyltransferase in small interfering RNA modification. *Eukaryot Cell* 2014, 13:77–86. https://doi.org/10.1128/EC.00233-13.

108. Newman MA, Mani V, Hammond SM. Deep sequencing of microRNA precursors reveals extensive 3′ end modification. *RNA* 2011, 17:1795–1803. https://doi.org/10.1261/rna.2713611.

109. Gutierrez-Vazquez C, Enright AJ, Rodríguez-Galán A, Perez-García A, Collier P, Jones MR, Benes V, Mizgerd JP, Mittelbrunn M, Ramiro AR, et al. 3′ uridylation controls mature microRNA turnover during CD4 T cell activation. *RNA* 2017, 6:882–891. https://doi.org/10.1261/rna.060095.116.

110. Pisacane P, Halic M. Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi. *Nat Commun* 2017, 8:15332. https://doi.org/10.1038/ncomms15332.

111. Knouf EC, Wyman SK, Tewari M. The human TUT1 nucleotidyl transferase as a global regulator of microRNA abundance. *PLoS One* 2013, 8:e69630. https://doi.org/10.1371/journal.pone.0069630.

112. Wyman SK, Knouf EC, Parkin RK, Fritz BR, Lin DW, Dennis LM, Krouse MA, Webster PJ, Tewari M. Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Res* 2011, 21:1450–1461. https://doi.org/10.1101/gr.118059.110.

113. Jones MR, Blahna MT, Kozlowski E, Matsuura KY, Ferrari JD, Morris SA, Powers JT, Daley GQ, Quinton LJ, Mizgerd JP. Zcchc11 uridylates mature miRNAs to enhance neonatal IGF-1 expression, growth, and survival. *PLoS Genet* 2012, 8:e1003105. https://doi.org/10.1371/journal.pgen.1003105.

114. Baccarini A, Chauhan H, Gardner TJ, Jayaprakash AD, Sachidanandam R, Brown BD. Kinetic analysis reveals the fate of a microRNA following target regulation in mammalian cells. *Curr Biol* 2011, 21:369–376. https://doi.org/10.1016/j.cub.2011.01.067.

115. Chung CZ, Seidl LE, Mann MR, Heinemann IU. Tipping the balance of RNA stability by 3′ editing of the transcriptome. *Biochim Biophys Acta* 2017, in press. https://doi.org/10.1016/j.bbagen.2017.05.003.

116. Marcinowski L, Tanguy M, Krmpotic A, Rädle B, Lisnic VJ, Tuddenham L, Chane-Woon-Ming B, Ruzsics Z, Erhard F, Benkartek C, et al. Degradation of cellular miR-27 by a novel, highly abundant viral transcript is important for efficient virus replication *in vivo*. *PLoS Pathog* 2012, 8:e1002510. https://doi.org/10.1371/journal.ppat.1002510.

117. de la Mata M, Gaidatzis D, Vitanescu M, Stadler MB, Wentzel C, Scheiffele P, Filipowicz W, Grosshans H. Potent degradation of neuronal miRNAs induced by highly complementary targets. *EMBO Rep* 2015, 16:500–511. https://doi.org/10.15252/embr.201540078.

118. Dolken L, Perot J, Cognat V, Alioua A, John M, Soutschek J, Ruzsics Z, Koszinowski U, Voinnet O, Pfeffer S. Mouse cytomegalovirus microRNAs dominate the cellular small RNA profile during lytic infection and show features of posttranscriptional regulation. *J Virol* 2007, 81:13771–13782. https://doi.org/10.1128/JVI.01313-07.

119. Libri V, Helwak A, Miesen P, Santhakumar D, Borger JG, Kudla G, Grey F, Tollervey D, Buck AH. Murine cytomegalovirus encodes a miR-27 inhibitor disguised as a target. *Proc Natl Acad Sci* 2012, 109:279–284. https://doi.org/10.1073/pnas.1114204109.

120. Jones MR, Quinton LJ, Blahna MT, Neilson JR, Fu S, Ivanov AR, Wolf DA, Mizgerd JP. Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat Cell Biol* 2009, 11:1157–1163. https://doi.org/10.1038/ncb1931.

121. Koppers-Lalic D, Hackenberg M, Bijnsdorp IV, van Eijndhoven MAJ, Sadek P, Sie D, Zini N, Middeldorp JM, Ylstra B, de Menezes RX, et al. Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes. *Cell Rep* 2014,

 © 2017 Wiley Periodicals, Inc.

8:1649–1658. https://doi.org/10.1016/j.celrep.2014.08.027.

122. Schmidt K, Butler JS. Nuclear RNA surveillance: role of TRAMP in controlling exosome specificity. *WIREs RNA* 2013, 4:217–231. https://doi.org/10.1002/wrna.1155.
123. Labno A, Tomecki R, Dziembowski A. Cytoplasmic RNA decay pathways – enzymes and mechanisms. *Biochim Biophys Acta* 1863, 2016:3125–3147. https://doi.org/10.1016/j.bbamcr.2016.09.023.
124. Shen B, Goodman HM. Uridine addition after microRNA-directed cleavage. *Science* 2004, 306:997. https://doi.org/10.1126/science.1103521.
125. Mullen TE, Marzluff WF. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. *Genes Dev* 2008, 22:50–65. https://doi.org/10.1101/gad.1622708.
126. Morozov IY, Jones MG, Razak AA, Rigden DJ, Caddick MX. CUCU modification of mRNA promotes decapping and transcript degradation in Aspergillus nidulans. *Mol Cell Biol* 2010, 30:460–469. https://doi.org/10.1128/MCB.00997-09.
127. Morozov IY, Jones MG, Gould PD, Crome V, Wilson JB, Hall AJW, Rigden DJ, Caddick MX. mRNA 3′ tagging is induced by nonsense-mediated decay and promotes ribosome dissociation. *Mol Cell Biol* 2012, 32:2585–2595. https://doi.org/10.1128/MCB.00316-12.
128. Sement FM, Ferrier E, Zuber H, Merret R, Alioua M, Deragon JM, Bousquet-Antonelli C, Lange H, Gagliardi D. Uridylation prevents 3′ trimming of oligoadenylated mRNAs. *Nucleic Acids Res* 2013, 41:7115–7127. https://doi.org/10.1093/nar/gkt465.
129. Chang H, Lim J, Ha M, Kim VN. TAIL-seq: genome-wide determination of poly(A) tail length and 3′ end modifications. *Mol Cell* 2014, 53:1044–1052. https://doi.org/10.1016/j.molcel.2014.02.007.
130. Aphasizheva I, Aphasizhev R. RET1-catalyzed uridylylation shapes the mitochondrial transcriptome in *Trypanosoma brucei*. *Mol Cell Biol* 2010, 30:1555–1567. https://doi.org/10.1128/MCB.01281-09.
131. Xu K, Lin J, Zandi R, Roth JA, Ji L. MicroRNA-mediated target mRNA cleavage and 3′-uridylation in human cells. *Sci Rep* 2016, 6:30242. https://doi.org/10.1038/srep30242.
132. Orban TI, Izaurralde E. Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA* 2005, 11:459–469. https://doi.org/10.1261/rna.7231505.
133. Branscheid A, Marchais A, Schott G, Lange H, Gagliardi D, Andersen SU, Voinnet O, Brodersen P. SKI2 mediates degradation of RISC 5′-cleavage fragments and prevents secondary siRNA production from miRNA targets in Arabidopsis. *Nucleic Acids Res* 2015, 43:10975–10988. https://doi.org/10.1093/nar/gkv1014.
134. Ibrahim F, Rohr J, Jeong W-J, Hesson J, Cerutti H. Untemplated oligoadenylation promotes degradation of RISC-cleaved transcripts. *Science* 2006, 314:1893. https://doi.org/10.1126/science.1135268.
135. Smith DW, Han MR, Park JS, Kim KR, Yeom T, Lee JY, Kim DJ, Bingman CA, Kim HJ, Jo K, et al. Crystal structure of the protein from *Arabidopsis thaliana* gene At5g06450, a putative DnaQ-like exonuclease domain-containing protein with homohexameric assembly. *Proteins* 2013, 81:1669–1675. https://doi.org/10.1002/prot.24315.
136. Thomas MP, Liu X, Whangbo J, McCrossan G, Sanborn KB, Basar E, Walch M, Lieberman J. Apoptosis triggers specific, rapid, and global mRNA decay with 3′ uridylated intermediates degraded by DIS3L2. *Cell Rep* 2015, 11:1079–1089. https://doi.org/10.1016/j.celrep.2015.04.026.
137. Hoefig KP, Heissmeyer V. Degradation of oligouridylated histone mRNAs: see UUUUU and goodbye. *WIREs RNA* 2014, 5:577–589. https://doi.org/10.1002/wrna.1232.
138. Marzluff WF, Wagner EJ, Duronio RJ. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet* 2008, 9:843–854. https://doi.org/10.1038/nrg2438.
139. Tan D, Marzluff WF, Dominski Z, Tong L. Structure of histone mRNA stem-loop, human stem-loop binding protein, and 3′ hExo ternary complex. *Science* 2013, 339:318–321. https://doi.org/10.1126/science.1228705.
140. Welch JD, Slevin MK, Tatomer DC, Duronio RJ, Prins JF, Marzluff WF. EnD-Seq and AppEnD : sequencing 3′ ends to identify nontemplated tails and degradation intermediates. *RNA* 2015, 21:1375–1389. https://doi.org/10.1261/rna.048785.114.6.
141. Kaygun H, Marzluff WF. Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. *Nat Struct Mol Biol* 2005, 12:794–800. https://doi.org/10.1038/nsmb972.
142. Schmidt M-J, West S, Norbury CJ. The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. *RNA* 2011, 17:39–44. https://doi.org/10.1261/rna.2252511.
143. Lackey PE, Welch JD, Marzluff WF. TUT7 catalyzes the uridylation of the 3′ end for rapid degradation of histone mRNA. *RNA* 2016, 22:1673–1688. https://doi.org/10.1261/rna.058107.116.
144. Hoefig KP, Rath N, Heinz GA, Wolf C, Dameris J, Schepers A, Kremmer E, Ansel KM, Heissmeyer V. Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. *Nat Struct Mol Biol* 2013, 20:73–81. https://doi.org/10.1038/nsmb.2450.

 © 2017 Wiley Periodicals, Inc.

145. Slevin MK, Meaux S, Welch JD, Bigler R, Miliani de Marval P, Su W, Rhoads RE, Prins JF, Marzluff WF. Deep sequencing shows multiple oligouridylations are required for 3′ to 5′ degradation of histone mRNAs on polyribosomes. *Mol Cell* 2014, 53:1020–1030. https://doi.org/10.1016/j.molcel.2014.02.027.

146. Su W, Slepenkov SV, Slevin MK, Lyons SM, Ziemniak M, Kowalska J, Darzynkiewicz E, Jemielity J, Marzluff WF, Rhoads RE. mRNAs containing the histone 3′ stem-loop are degraded primarily by decapping mediated by oligouridylation of the 3′ end. *RNA* 2013, 19:1–16. https://doi.org/10.1261/rna.034470.112.

147. Ochi H, Chiba K. Hormonal stimulation of starfish oocytes induces partial degradation of the 3′ termini of cyclin B mRNAs with oligo(U) tails, followed by poly(A) elongation. *RNA* 2016, 22:1–8. https://doi.org/10.1261/rna.054882.115.

148. van Ooij MJM, Polacek C, Glaudemans DHRF, Kuijpers J, van Kuppeveld FJM, Andino R, Agol VI, Melchers WJG. Polyadenylation of genomic RNA and initiation of antigenomic RNA in a positive-strand RNA virus are controlled by the same cis-element. *Nucleic Acids Res* 2006, 34:2953–2965. https://doi.org/10.1093/nar/gkl349.

149. Jupin I, Bouzoubaa S, Richards K, Jonard G, Guilley H. Multiplication of beet necrotic yellow vein virus RNA 3 lacking a 3′ poly(A) tail is accompanied by reappearance of the poly(A) tail and a novel short U-rich tract preceding it. *Virology* 1990, 178:281–284. https://doi.org/10.1016/0042-6822(90)90404-F.

150. Raju R, Hajjou M, Hill KR, Botta V, Botta S. *In vivo* addition of poly(A) tail and AU-rich sequences to the 3′ terminus of the Sindbis virus RNA genome: a novel 3′-end repair pathway. *J Virol* 1999, 73:2410–2419.

151. van Leeuwen HC, Liefhebber JMP, Spaan WJM. Repair and polyadenylation of a naturally occurring hepatitis C virus 3′ nontranslated region-shorter variant in selectable replicon cell lines. *J Virol* 2006, 80:4336–4343. https://doi.org/10.1128/JVI.80.9.4336-4343.2006.

152. Huo Y, Shen J, Wu H, Zhang C, Guo L, Yang J, Li W. Widespread 3′-end uridylation in eukaryotic RNA viruses. *Sci Rep* 2016, 6:25454. https://doi.org/10.1038/srep25454.

153. Szczesny RJ, Borowski LS, Brzezniak LK, Dmochowska A, Gewartowski K, Bartnik E, Stepien PP. Human mitochondrial RNA turnover caught in flagranti: involvement of hSuv3p helicase in RNA surveillance. *Nucleic Acids Res* 2009, 38:279–298. https://doi.org/10.1093/nar/gkp903.

154. Slomovic S, Schuster G. Stable PNPase RNAi silencing : its effect on the processing and adenylation of human mitochondrial RNA. *RNA* 2008, 14:310–323. https://doi.org/10.1261/rna.697308.age.

 © 2017 Wiley Periodicals, Inc.

CellPress

Review

Uridylation Earmarks mRNAs for Degradation… and More

Hélène Scheer,[1,2] Hélène Zuber,[1,2] Caroline De Almeida,[1] and Dominique Gagliardi[1,*]

Groundbreaking discoveries have uncovered the widespread post-transcriptional modifications of all classes of RNA. These studies have led to the emerging notion of an 'epitranscriptome' as a new layer of gene regulation. Diverse modifications control RNA fate, including the 3′ addition of untemplated nucleotides or 3′ tailing. The most exciting recent discoveries in 3′ tailing are related to uridylation. Uridylation targets various noncoding RNAs, from small RNAs and their precursors to rRNAs, and U tails mostly regulate processing or degradation. Interestingly, uridylation is also a pervasive modification of mRNAs. In this review, we discuss how the addition of few uridines to the 3′ end of mRNAs influences mRNA decay. We also consider recent findings that reveal other consequences of uridylation on mRNA fate.

Trends

Uridylation of mRNAs is widespread and conserved among eukaryotes.

Uridylation has a fundamental role in mRNA decay and triggers both 5′–3′ and 3′–5′ degradation.

Uridylation can also 'repair' mRNA extremities as shown for replication-dependent histone mRNAs during S-phase in humans and for deadenylated mRNAs in *Arabidopsis*.

Uridylation may have other alternative functions in different organisms, at specific developmental stages or for particular mRNAs. An alternative consequence of uridylation could be translation regulation.

The Emerging Epitranscriptome

Over 100 post-transcriptional modifications can affect RNA [1,2]. Well-known targets of RNA-modifying activities include rRNAs, tRNAs, or small RNAs. Yet, RNA modifications are not restricted to noncoding RNAs and novel next-generation sequencing strategies have recently revealed the pervasiveness of several mRNA modifications in archaea, bacteria, and eukaryotes [3–7]. RNA modifications can impact function, localization, or stability of transcripts and are an integral part of the regulatory processes that rapidly adjust the transcriptome in response to developmental and environmental cues [8–11]. RNA modifications are established by a variety of enzymes or 'writers' and are recognized by effector RNA-binding proteins or 'readers'. They can be dynamically regulated and the first examples of reversibility involving 'erasers' have been described [8,11]. Hence, RNA modifications share many conceptual similarities with **epigenetic marks** (see Glossary) that modulate chromatin structure and activity. Because of this analogy, the notion of an epitranscriptome is emerging besides the recognized epigenome.

RNA modifications can be divided into two main subclasses: the chemical modification of nucleosides and the tailing of RNA 3′ extremities. Nucleoside modifications are extremely diverse in nature and represent by far the majority of RNA modifications. Modified nucleosides include **N6-methyladenosine**, N1-methyladenosine, **pseudouridine,** or 5-hydroxymethylcytosine [1,2,7,12–14]. The second subclass of RNA modifications, that is, the tailing of 3′ extremities, encompasses adenylation, uridylation, cytidylation, and guanylation. **Noncanonical adenylation** is a widespread, mostly post-transcriptional, modification that triggers the degradation of virtually all classes of noncoding RNAs in all organisms. Noncanonical adenylation also destabilizes mRNAs in bacteria, in most archaea, in chloroplasts, and in plant and human mitochondria [15–17]. Finally, cytoplasmic adenylation activates translation during developmental or physiological transitions including oocytes maturation or synapse function [18,19]. RNA cytidylation and/or guanylation are much less characterized and have been reported only in a few instances, for example, for mRNAs in humans, *Aspergillus nidulans*, and *Arabidopsis* [20–23]. Their precise functions are yet to be determined for mRNAs. By contrast, much progress has

[1]Institut de Biologie Moléculaire des Plantes (IBMP), Centre National de la Recherche Scientifique (CNRS), University of Strasbourg, 67000 Strasbourg, France
[2]These authors contributed equally to this work.

*Correspondence: dominique.gagliardi@ibmp-cnrs.unistra.fr (D. Gagliardi).

© 2016 Elsevier Ltd. All rights reserved.

CellPress

Box 1. Domain Organization of Terminal Uridylyltransferases (TUTases) Uridylating mRNAs

General Domain Organization

As other terminal nucleotidyltransferases belonging to the DNA polymerase β-like nucleotidyltransferase superfamily [24], TUTases contain two archetypical domains:

(i) a Polβ nucleotidyltransferase domain (NTD) with three aspartate/glutamate residues that are indispensable for the chelation of divalent metal ions supporting catalytic activity.

(ii) a poly(A) polymerase-associated domain (PAP), which contains a type II-nucleotide recognition motif (NRM).

Together, NTD and PAP form the core catalytic domain (CCD), which defines the minimal catalytic organization present in all TUTases represented in Figure I. Although the CCD is duplicated in HsTUT7/TUT4 and XtXTUT7, only the C terminal CCD is active [24,25]. The N terminal CCD is inactivated by amino acid substitutions in the NTD catalytic triad. In addition, a histidine, which is indispensable for UTP selectivity, is lacking in the N terminal NRM [81,84,85]. The inactive CCD could still be required for allosteric activation of the protein or to mediate protein–protein interactions, thereby maintaining nucleotidyltransferase-independent functionalities [86]. Besides the CCD domains, HsTUT4, HsTUT7, XtXTUT7, and TbRET1 also possess one C2H2-type zinc finger (ZnF) and, with the exception of TbRET1, two C2HC-type ZnF motifs (also known as zinc knuckle). Such motifs can promote protein–protein interactions and RNA binding [87–89].

Disorder in TUTases

Stretches of basic rich lysine and arginine residues (BR), initially described for XtXTUT7 [81], are also found in human TUT4 and TUT7. Of note, BR regions are known as disordered regions that promote RNA binding [90]. Disorder predictions using DISOPRED [91] reveal that all TUTases have long intrinsically disordered regions (IDRs). IDR-containing proteins are enriched in HeLa cell mRNA interactome and IDRs extensively mediate protein–RNA interactions [92–94]. Therefore, such flexible disordered regions could be involved in RNA substrate binding, especially for TUTases that lack canonical RNA recognition domains. Alternatively, IDRs can mediate protein–protein interactions with effectors involved in RNA substrate recognition or in downstream consequences of uridylation. Lastly, IDRs may influence the localization to P-bodies and stress granules, as shown for decapping factors [95]. In line with this hypothesis, CutA and URT1 are present in P-bodies and stress granules, respectively [35,96].

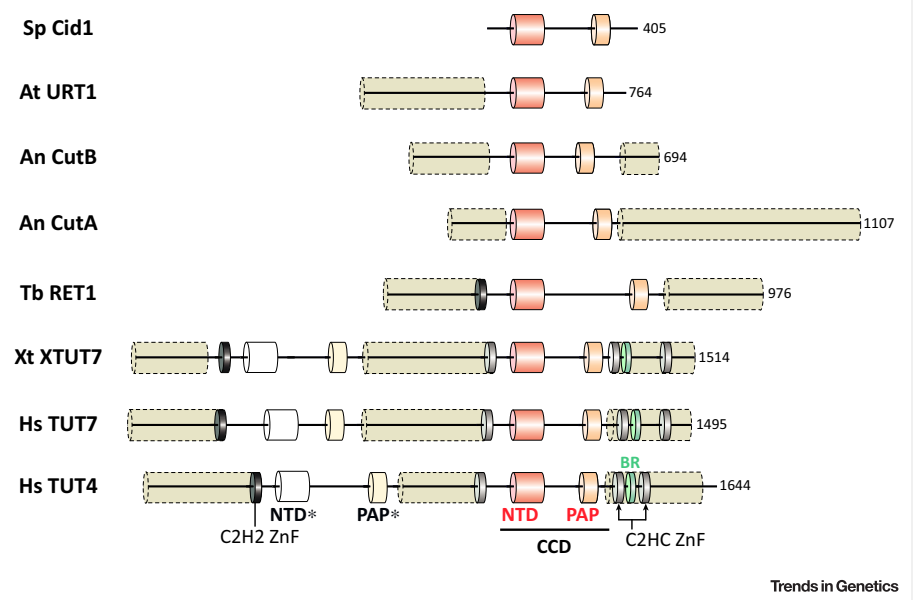


Figure I. Domain Organization of Terminal Uridylyltransferases (TUTases) Uridylating mRNAs. Nucleotidyltransferase domain (NTD) is shown in red, poly(A) polymerase-associated domain (PAP) in orange. Together, the NTD and PAP form the catalytic core domain (CCD). Nonfunctional NTD and PAP domains are marked with an asterisk and are shown in white and pale yellow, respectively. C2H2-type zinc finger (ZnF) domains are in black, C2HC-type in gray, and stretches of basic rich (BR) residues in turquoise blue. Long regions highlighted with beige dashed line cylinders are predicted to be intrinsically disordered using DISOPRED [82]. An, *Aspergillus nidulans*; At, *Arabidopsis thaliana*; Hs: *Homo sapiens*; Sp, *Schizosaccharomyces pombe*; Tb, *Trypanosoma brucei*; Xt, *Xenopus tropicalis*.

been made to understand the impact of uridylation on the transcriptome. The untemplated addition of uridines is catalyzed by terminal uridylyltransferases (TUTases), which belong to the Pol β superfamily and more specifically to the noncanonical terminal nucleotidyltransferases (TNTases) subgroup [24,25]. Besides the characteristic nucleotidyltransferase domain (NTD),

Glossary

Epigenetic marks: DNA and histone modifications that regulate chromatin structure and genome expression but do not alter the genetic sequence.

Intrinsically disordered regions (IDR): protein segments devoid of intrinsically defined 3D structure. IDRs can adopt a precise tridimensional folding upon binding with a target protein or RNA.

N6-methyladenosine (m^6A): an abundant modification present in coding and many noncoding RNAs. Lack of m^6A is embryo-lethal in *Arabidopsis* and leads to apoptosis in mammalian cells. m^6A is involved in the regulation of gene expression by modulating splicing, nuclear export, localization, translation, and stability of mRNA. Importantly, m^6A methylation in mRNAs is reversible.

Noncanonical adenylation: any untemplated addition of adenosines at the 3′ end of noncoding RNAs and mRNAs that is not catalyzed by the canonical poly(A) polymerase, which co-transcriptionally polyadenylates RNA polymerase II transcripts.

Nonsense-mediated decay (NMD): first identified as an RNA surveillance mechanism that insures the degradation of mRNAs with premature termination codons. In fact, NMD factors regulate the stability of numerous transcripts, including RNAs with no obvious coding capacity.

P-bodies and stress granules: two types of dynamic cytoplasmic granules formed by translationally repressed mRNPs. P-bodies are present in nonstressed cells and their formation is further induced upon stress. By contrast, stress granules only accumulate under stress conditions. Archetypical components of P-bodies and stress granules include factors of the mRNA decay machinery and translation initiation factors, respectively.

Pseudouridines: pseudouridines are formed by isomerization of uridines by Ψ synthases. Pseudouridylation stabilizes the structure of noncoding RNAs (like tRNAs or rRNAs) and results in the rapid and regulated rewiring of mRNA coding information by allowing noncanonical base pairing in the ribosome decoding center.

RNA exosome: the eukaryotic RNA exosome complex provides the main 3′–5′ exoribonucleolytic activity in both nuclear and cytoplasmic

TUTases have a fast-evolving, diversified architecture (Box 1). This diversity in noncatalytic domains and the presence of multiple **intrinsically disordered regions** may reflect the variety of RNA substrates recognized by TUTases or specific interaction networks with various cofactors. Those cofactors are involved in RNA substrate recognition or define the multiple downstream consequences of uridylation. The multiplicity of roles played by uridylation in RNA metabolism is particularly well illustrated for mitochondrial RNAs in trypanosomes [26]. In those organelles, short U tails can induce mRNA degradation but long A/U tails promote translation [27]. In addition, uridylation is an intrinsic and necessary step for the processing and function of guide RNAs implicated in U-insertion/deletion mRNA editing [28]. Besides specific roles in mitochondria of kinetoplastid protists, uridylation targets a plethora of noncoding RNAs including miRNAs, siRNAs, Piwi-interacting RNAs, miRNA precursors, rRNAs, and the U6 spliceosomal RNA (Box 2). Uridylation of miRNAs and pre-miRNAs can have opposite consequences, from triggering degradation to favoring maturation or abrogating activity, as reviewed recently [17,19,29–31]. Uridylation was also recently reported to target several RNA viruses, extending the repertoire of RNA substrates recognized by TUTases [32]. Importantly, prominent targets of TUTases are endogenous mRNAs. In fact, with an increasing number of reports in various organisms such as *Schizosaccharomyces pombe*, *A. nidulans*, *Arabidopsis thaliana*, *Trypanosoma brucei*, and humans, the uridylation of polyadenylated mRNAs has recently been recognized as a conserved process in eukaryotic mRNA metabolism [20–23,27,33–38]. The housekeeping function of mRNA uridylation is to promote degradation. In this review, we discuss the recent progress toward understanding the distinct molecular mechanisms by which uridylation can impact mRNA metabolism.

compartments of eukaryotic cells. The RNA exosome has crucial roles in RNA processing, surveillance, and turnover of virtually all classes of RNA.

RNA-induced silencing complex (RISC): complex containing Argonaute proteins and small interfering RNAs that guide the complex to its target transcripts. RISC silences gene expression by translation repression or mRNA degradation.

S-phase: S-phase (synthesis phase) is the phase of the cell cycle during which chromosome replication occurs. Histone mRNA levels increase considerably because the production of new histones is required for nucleosome assembly.

Uridylation Promotes Degradation of Nonpolyadenylated and Cleaved mRNAs

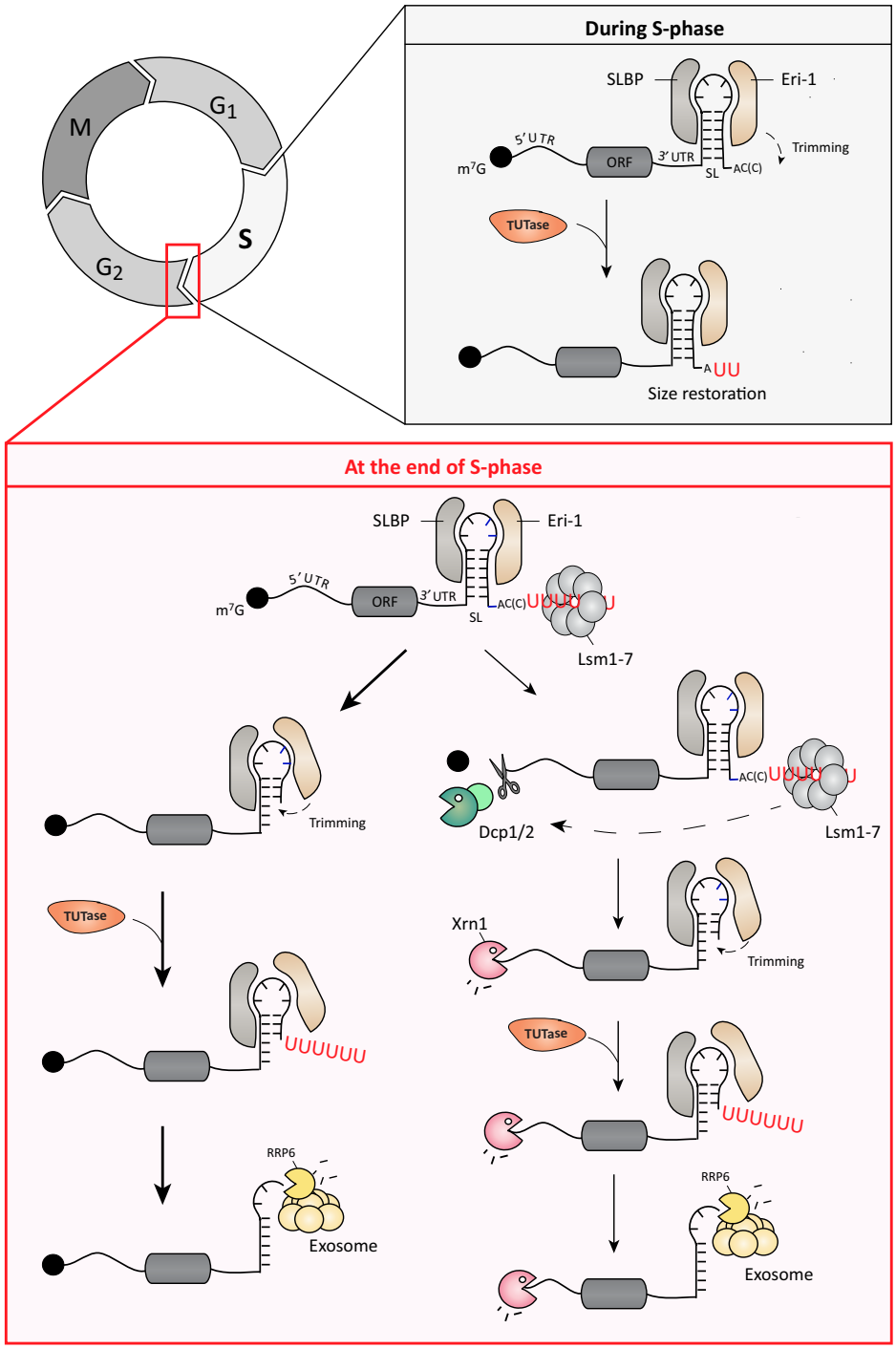
A link between uridylation and the degradation process was first found when it was noticed that the 5′ fragments of mRNAs cleaved by the **RNA-induced silencing complex (RISC)** can be

Box 2. Different Roles of U Tailing in Noncoding RNA Metabolism

Uridylation impacts the fate of noncoding RNAs in various ways, from facilitating maturation and stabilizing processed RNAs to triggering degradation. This versatility of roles is illustrated in the following section.

Uridylation of small interfering RNAs usually leads to their degradation [17,29,31,97,98]. Destabilization is prevented by 2′-O-methylation of the 3′ terminal nucleotide by the methyltransferase HEN1 for siRNAs and miRNAs in plants, Piwi-interacting RNAs in animals as well as siRNAs in *Drosophila* [98–101]. Interestingly, the *Arabidopsis* terminal uridylyltransferases (TUTases) HESO1 and URT1 cooperate or compete for miRNA uridylation, which results in synergistic or opposed impact on stability [70,71]. Uridylation plays also a complex role in animal miRNA biogenesis [55,74–76,102–104]. Mono-uridylation of Group II let7 pre-miRNAs by TUT4/7 produces a 2-nt 3′ overhang, creating an optimal end structure for Dicer processing [55,76]. By contrast, in the presence of the RNA-binding protein Lin28, oligo-uridylation of pre-let7 is favored, which leads to degradation by Dis3L2 [58,74,77]. TUT4/7 can also oligo-uridylate trimmed pre-miRNAs, independently of Lin28, probably leading to the subsequent degradation of nonfunctional pre-miRNAs [17,55]. Indeed, the pervasive uridylation of Ago-bound pre-miRNAs by TUT4/7 contributes to a pre-miRNA surveillance pathway, as shown in mouse embryonic fibroblasts [103]. Mono-uridylation and oligo-uridylation that produces a 3′ overhang different from the canonical 2-nt 3′ overhang optimal for Dicer processing triggers degradation by the exosome. This pre-miRNA surveillance pathway eliminates defective precursors that could compete with functional pre-miRNAs for Ago [103].

Uridylation is also important for the metabolism of other noncoding RNAs. In mitochondria of trypanosomes, U-insertion/deletion mRNA editing is directed by guide RNAs (gRNAs), key actors of the editosome. gRNAs are matured by the mitochondrial 3′ processome, a complex constituted by the TUTase RET1, the 3′–5′ exonuclease DSS1, and three additional subunits. gRNA maturation is initiated by uridylation of long precursors by RET1, which promotes 3′–5′ degradation of 3′ extensions by DSS1. Pausing of DSS1 progression by head-to-head hybridization of precursors triggers secondary uridylation by RET1 [28]. The mature uridylated gRNAs are then incorporated into the editosome [28,36,105]. The maturation of the U6 small nuclear RNA (snRNA), essential component of the spliceosome, also involves uridylation by U6 TUTase (TUT1), which stabilizes U6 snRNA prior to its incorporation into a functional splicing complex [31,106]. As a last example, rRNA maturation intermediates can also be uridylated, presumably to facilitate processing or elimination through the recruitment of 3′–5′ exoribonucleases [107,108].



Trends in Genetics

Figure 1. Uridylation and Degradation of Replication-Dependent Histone mRNAs. Metazoan replication-dependent histone mRNAs end with a stem–loop (SL) structure, instead of the classical poly(A) tail observed for all other eukaryotic mRNAs [43]. This SL is essential for the processing, export from the nucleus, translation, and stability of histone mRNAs. The SL is bound by the SL-binding protein (SLBP) on its 5′ side and by the exoribonuclease Eri-1 (3′hExo) on its 3′ side [45]. Histone mRNAs are stable and actively translated during the S-phase when DNA is replicated. Nibbling of 1–2 nt at the 3′ end (likely by Eri-1) is repaired by uridylation that restores the full length size of mature histone mRNAs [72] (gray panel). At the end of S-phase, histone mRNAs are no longer needed and are rapidly eliminated (red panel). Degradation is initiated by a

(Figure legend continued on the bottom of the next page.)

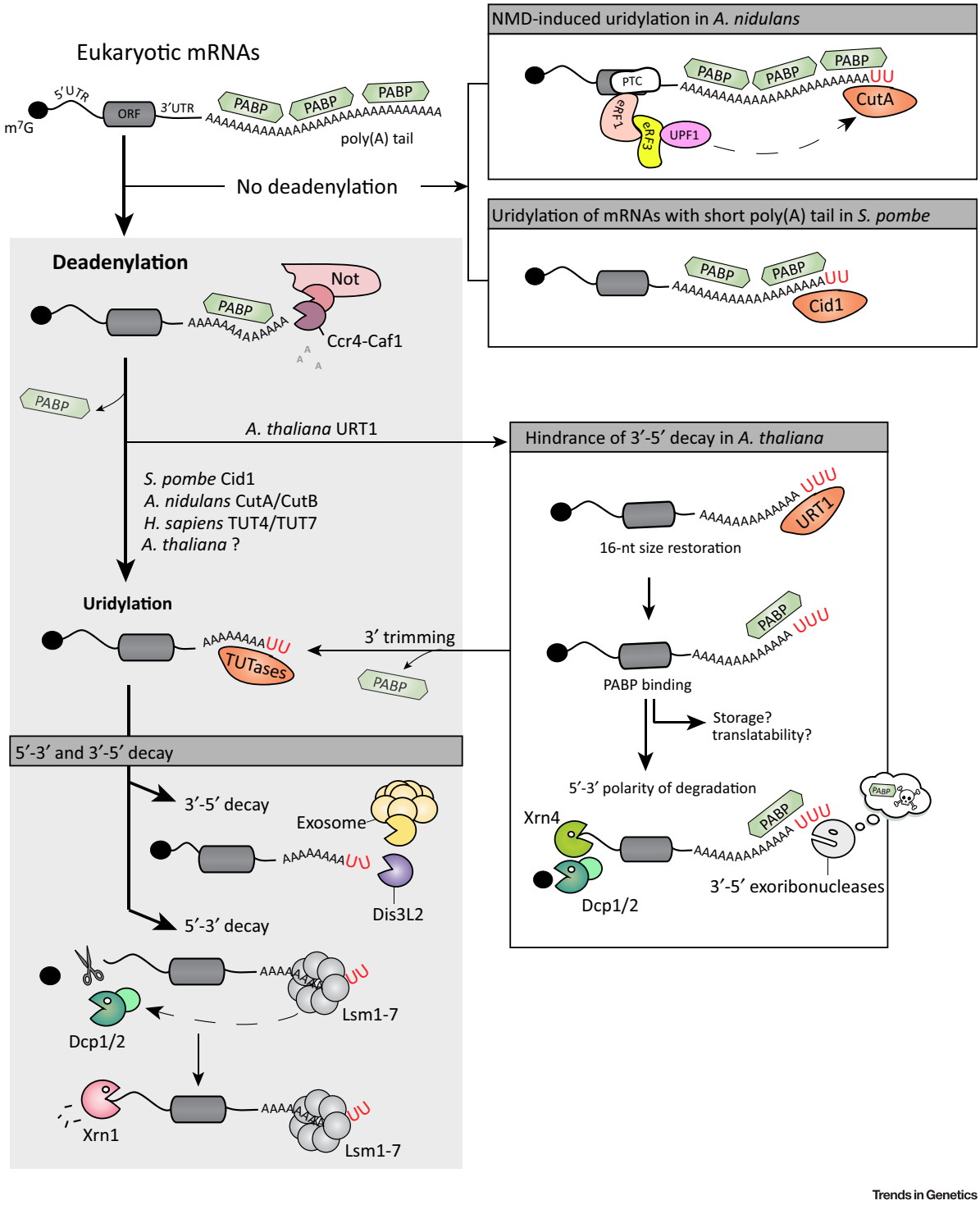
modified at their 3′ end by U-rich short tails [39]. This observation, originally made in *Arabidopsis* and mouse, is also reported in human cells [39,40], indicating an evolutionary-conserved mechanism. Moreover, the addition of 3′ uridines correlates with decapping and 5′ shortening of the cleaved products in *Arabidopsis*, a first hint that uridylation might stimulate the 5′ to 3′ degradation pathway [39]. In line with this, U tracts added at the 3′ end of a generic, capped, nonpolyadenylated RNA sequence recruit decapping factors and promote decapping in mammalian cell extracts [41]. In human cells, TUT2 is implicated in the uridylation of 5′ RISC-cleaved fragments, whereas TUT3 and other TUTases may mediate the uridylation of secondary 3′–5′ decayed fragments [40]. In *Arabidopsis*, the TUTase HESO1 was implicated in uridylating the 5′ fragments produced by AGO1-mediated cleavage of miRNA targets [42]. Residual uridylation persists in *heso1* mutants, indicating that at least another TUTase is able to modify the 3′ end of 5′ RISC-cleaved fragments. Whether uridylation also tags endonucleolytic fragments generated from RISC-independent pathways remains to be addressed.

The next major breakthrough toward the realization that uridylation is an integral step of mRNA decay was the discovery that uridylation elicits the degradation of replication-dependent histone mRNAs in humans (Figure 1). Upon inhibition of DNA replication or at the end of the **S-phase**, histone mRNAs are quickly degraded [43]. This process involves uridylation, which triggers both the 5′–3′ and 3′–5′ decay of histone mRNAs [44] (Figure 1). In metazoans, replication-dependent histone mRNAs are a notable exception among eukaryotic mRNAs because they are not polyadenylated. Instead, these mRNAs end with a conserved stem–loop structure, which is crucial for processing, export, translation, and degradation [43]. The mature form of histone mRNAs ends 2–3 nt 3′ of the stem–loop, and forms a complex with the stem–loop binding protein (SLBP) and the Eri-1 (3′hExo) exoribonuclease, which are bound to the 5′ and 3′ part of the stem–loop, respectively [45]. Interaction of SLBP and translation initiation factors is crucial for efficient histone mRNA translation. A switch from translation to degradation is signaled by the phosphorylation of the RNA helicase UPF1. UPF1, possibly recruited during translation termination, promotes the disruption of the interaction between translation initiation factors and SLBP [46]. As a result, degradation of histone mRNAs is initiated. A major signal triggering the degradation of histone mRNAs is the uridylation of the 3′ terminal stem–loop [44] (Figure 1). Oligouridylation of the 3′-end extremity promotes the binding of the decapping factor: the heptameric Lsm1-7 complex. Lsm1-7 interacts with Eri-1, which attacks the stem–loop in a stepwise manner [47,48]. Cycles of uridylation/nibbling ultimately lead to the destruction of the stem–loop, promoting subsequent degradation of histone mRNAs by RRP6 (PM/Scl-100), one of the two exoribonucleases associated to the **RNA exosome** [48]. Most of 3′ decay intermediates remain capped, suggesting a preponderant 3′–5′ polarity of degradation [48]. However, binding of Lsm1-7 can also promote decapping and subsequent 5′-3′ degradation by the cytosolic 5′–3′ exoribonuclease Xrn1 [44,49]. Several TUTases were proposed to uridylate histone mRNAs based on RNAi experiments and low-throughput sequencing analysis [44,50]. However, studying the impact of siRNA-based knockdown of candidate TUTases using a high-throughput sequencing method recently revealed TUT7 as the major TUTase uridylating both histone mRNA 3′ ends and degradation intermediates in the stem [51].

first round of oligo-uridylation, which triggers exoribonucleolytic decay from 3′ to 5′, from 5′ to 3′, or simultaneously from both ends. The binding of Lsm1-7 complex to the U tail can induce Eri-1 to nibble the SL by 2–4 nt [47,48]. A second round of uridylation favors new 3′–5′ exoribonucleolytic attacks, thereby displacing SLBP and allowing the recruitment of the RRP6 (PM/Scl-100)-associated exosome for further degradation (left side of the red panel). Of note, uridylation and 3′–5′ decay of histone mRNAs can proceed independently of decapping on polysomes [48]. Alternatively, the Lsm1-7 complex can induce decapping of histone mRNAs and subsequent 5′–3′ decay by Xrn1. Histone mRNAs subject to 5′–3′ decay can also simultaneously be degraded by Eri-1 and exosome [44,47,49] (right side of the red panel). ORF, open reading frame; TUTase, terminal uridylyltransferase; UTR, untranslated region.

Key Figure

Uridylation-Mediated Decay of Polyadenylated mRNAs



(See figure legend on the bottom of the next page.)

Uridylation as a New Integral Step of Polyadenylated mRNA Decay

The role of uridylation in the degradation of mammalian cell-cycle-dependent histone mRNAs turned out not to be the only case of uridylation-promoted mRNA decay. It rather embodies the first example of a generic process: uridylation of mRNAs is a global phenomenon and it elicits both the 5′–3′ and 3′–5′ decay of eukaryotic mRNAs (Figure 2, Key Figure). Uridylation of polyadenylated mRNAs was first identified in *S. pombe* for several model transcripts [34]. The TUTase Cid1 catalyzes the addition of mostly one to two uridines at the 3′ end of fission yeast mRNAs. Interestingly, the half-life of the *urg1* mRNA increases when Cid1 is deleted, indicating that uridylation can influence the stability of this mRNA [34]. Moreover, uridylated mRNAs accumulate when mRNA degradation components such as the Ccr4 deadenylase or the Dcp1 and Lsm1 decapping factors are mutated. Of note, uridylation is independent of deadenylation in *S. pombe* (at least for the tested target transcripts) and uridylation and deadenylation act redundantly to promote 5′–3′ degradation [34] (Figure 2). The bypass of the deadenylation step in the general mRNA decay pathway could be specific to *S. pombe* and possibly other organisms for which the average poly(A) tail length is rather short as compared with other eukaryotes [22,52]. The proposed molecular mechanism to explain the stimulation of 5′–3′ decay by uridylation is that the addition of one or two Us on the relatively short poly(A) tails of *S. pombe* mRNAs facilitates the binding of the Lsm1-7 complex. This complex in turn recruits the Dcp1–Dcp2 decapping machinery and decapping triggers the subsequent 5′–3′ degradation by Xrn1 [19,34].

mRNA uridylation in *S. pombe* also triggers 3′–5′ decay by the exoribonuclease Dis3L2, as shown for the *adh1* mRNA [53]. Dis3L2 is a member of the 3′–5′ exoribonuclease II/R family and functions independently of the exosome [53,54]. Dis3L2 is conserved across eukaryotes and was identified in humans as degrading uridylated mRNAs and noncoding RNAs, such as pre-miRNAs and small nuclear RNAs [54–57]. The solution of the structure of the mouse Dis3L2 in complex with an oligo(U)-tailed RNA explained the preferential degradation of uridylated RNAs by revealing extensive uracil-specific interactions spanning 12 Us [58]. U-specific hydrogen bonds exist between Dis3L2 and the uracil base of a U-tailed RNA substrate. Most of these interactions are disrupted when A and C tails are modeled into the Dis3L2 structure, in line with the idea that Dis3L2 targets preferentially uridylated RNA substrates. *In vitro* experiments revealed that two Us are sufficient to confer preferential degradation of an oligo(A) tail by *S. pombe* Dis3L2 [53]. Increasing the size of the U extension enhances the preferential degradation by Dis3L2. *In vivo*, the impact of Dis3L2 deletion on the accumulation of a restricted number of mRNAs and their uridylation is detected only when the 5′–3′ pathway is impaired, because of the redundancy of the 5′–3′ and 3′–5′ RNA decay pathways [53].

All basic components involved in uridylation-mediated mRNA decay in *S. pombe* are present in multicellular eukaryotes. Yet, distinctive features exist. In contrast to fission yeast, deadenylation

Figure 2. Deadenylation-dependent uridylation pathway. The general pathway of mRNA decay is initiated by shortening of the poly(A) tail by two deadenylase activities: the Pan2–Pan3 complex and the multimeric Ccr4–Not complex [83]. Shortening of the poly(A) tail displaces poly(A)-binding proteins (PABPs) until the 3′ extremity of an oligo(A) tail is too short to accommodate a PABP and becomes accessible to terminal uridylyltransferases (TUTases). The addition of untemplated uridines promotes the association of the Lsm1-7 complex at the 3′ end of the mRNA and leads to decapping by Dcp1/2. The unprotected 5′ extremity is subsequently degraded by the 5′–3′ exoribonuclease Xrn1. Alternatively, U tails can directly trigger Dis3L2- or exosome-mediated 3′–5′ degradation [37,53]. In *Arabidopsis thaliana*, URT1-mediated uridylation restores an oligo(A) tail size distribution centered on 16 nt, which allows for PABP binding. Uridylation and PAPB binding hinder 3′–5′ trimming to prevent excessive deadenylation. In addition, slowing down 3′–5′ ribonucleolytic attacks favors 5′–3′ directionality of degradation [23]. Alternatively, binding of PABP to uridylated oligo(A) tails could regulate storage or translatability. However, even if slowed down, deadenylation can still proceed. Below a certain tail size (e.g., <10 As), uridylation can no longer restore the PABP binding site, leading to both 3′–5′ and 5′–3′ degradation. Two deadenylation-independent uridylation instances have been reported. (i) *Aspergillus nidulans* mRNAs with premature termination codon can undergo mRNA 3′ tagging by CutA/CutB, recruited by UPF1. U-rich extensions elicit cap removal and 5′–3′ decay without a prior deadenylation step [21]. (ii) The average size of poly(A) tails in *Schizosaccharomyces pombe* is relatively short as compared to other organisms and mRNA uridylation is independent of and redundant with deadenylation [19]. NMD, nonsense mediated decay; ORF, open reading frame; PTC, premature termination codon; UTR, untranslated region.

precedes uridylation in *A. nidulans*, *Arabidopsis*, and humans [20–23,35,37]. In these organisms, mRNAs are uridylated by the TUTases CutA/CutB, URT1, and TUT4/TUT7, respectively. These TUTases target mRNAs with oligo(A) tails of less than 20 nt. Two features were shown in humans to explain the preferential uridylation of oligo(A) tails: TUT4/7 has an intrinsic preference for short tails, and binding of the cytosolic poly(A)-binding protein (PABPC1) prevents TUT7 action on mRNAs with poly(A) tails longer than 25 As [37]. In *Arabidopsis*, even when URT1 is massively overexpressed, the deadenylation step remains a prerequisite to uridylation [23]. However, specific cases of deadenylation-independent uridylation exist. For instance, the *A. nidulans* CutA and CutB can uridylate poly(A) tails longer than 30 nt for transcripts containing premature stop codon [also known as premature termination codon (PTC)]. The 3′ tailing of these typical **nonsense-mediated decay** (NMD) substrates is dependent on UPF1, a major component of the NMD pathway [21] (Figure 2).

Tailing by CutA promotes decapping and the degradation rate of model mRNA substrates decreases in Δ*cutA* and Δ*cutB* mutants [20,21]. Hence, a clear correlation between mRNA uridylation and degradation exists in *A. nidulans*, at least for the tested mRNAs. Although mRNA uridylation in *Arabidopsis* is also definitely part of the mRNA decay process, uridylation by at least two distinct TUTases has complex consequences on mRNA metabolism as detailed in the following section [23,35] (Figure 2). One of these consequences is likely the stimulation of mRNA degradation, since uridylated mRNAs accumulate when the 5′–3′ RNA degradation pathway is impaired.

The definite proof that uridylation must be considered as an integral step of the general pathway of polyadenylated mRNA decay was obtained recently by determining the global impact of uridylation on both 5′–3′ and 3′–5′ decay pathways in human cell lines [22,37]. A novel deep-sequencing method, called TAIL-seq, was designed to analyze both poly(A) tail size and potential 3′-end modifications at the transcriptome-wide level (Box 3) [22]. TAIL-seq was decisive to unambiguously determine the impact of uridylation on facilitating both 5′–3′ and 3′–5′ decay. First, uridylation was shown at the global scale to preferentially occur on deadenylated transcripts [22,37]. Second, uridylation frequency negatively correlates with global mRNA half-lives [22]. Third, depletion of both TUT4 and TUT7, which redundantly uridylate human mRNAs, eliminates mRNA U tailing and slows down RNA decay [37]. Fourth, depletion of mRNA decay factors Xrn1, Dcp1, and Lsm1 involved in 5′–3′ decay, or RRP41 (a subunit of the exosome) and Dis3L2 involved in the 3′–5′ degradation pathway, leads to the accumulation of uridylated mRNAs [37]. Although this accumulation was rather modest for Dis3L2 depletion [37], the role of uridylation and Dis3L2 in 3′ to 5′ mRNA decay in human cells is further supported by the observation that apoptosis-induced decay of mRNAs involves the uridylation of degradation intermediates by TUT4 and TUT7 and subsequent elimination by Dis3L2 [59]. Taken together, the results obtained for selected mRNA models in several organisms and, more importantly, the global data generated by TAIL-seq analysis in human cell lines revealed a fundamental role of U tails as a mark for mRNA decay in eukaryotes.

A Complex Role for Uridylation in Plant mRNA Turnover

TAIL-seq was also instrumental to define an additional role of mRNA uridylation in plants, besides its role in stimulating degradation. In *Arabidopsis*, at least two TUTases uridylate mRNAs and their activities are not functionally redundant. URT1 is the main TUTase targeting mRNAs after the deadenylation step [23,35]. Interestingly, URT1 extends deadenylated mRNAs with U residues to restore a size distribution of tails centered at 16 nt [23] (Figure 2). Hence, URT1 'repairs' mRNAs' deadenylated tails. Two lines of evidence indicate that this defined size distribution reflects the footprint of PABP: PABP determines the size of U extensions added by URT1 *in vitro* and PABP binds oligo(A/U) tails *in vivo* with a similar size distribution centered

Box 3. High-Throughput Sequencing Methods Dedicated to mRNA 3′ Extension Investigation

The addition of nucleotides at the 3′ end of mRNAs can now be investigated by high-throughput sequencing approaches either at the global scale or for specific targets. Those recent methods have revolutionized the study of mRNA 3′ extensions, which was previously restricted to the analysis by Sanger sequencing of a limited number of clones.

TAIL-seq represents the first method to simultaneously measure poly(A) tail length and 3′ tailing at transcriptome scale [22]. Briefly, RNA samples are depleted for rRNAs, ligated to a biotinylated 3′ adaptor, and fragmented. The 3′ fragments are affinity purified and ligated to a 5′ adaptor prior to cDNA synthesis and library amplification. Then, paired-end sequencing is performed: Read 1 is used for transcript identification, while Read 2 allows the detection of any nucleotides added at the 3′ end of polyadenylated mRNAs as well as the determination of poly(A) sizes. The latter is performed using fluorescence intensity files from the Illumina sequencer rather than a base-call analysis protocol.

Gene-targeted approaches to study 3′ tailing using high-throughput sequencing have also been described. These methods provide an ultradeep analysis of 3′ extremities for transcripts of interest and/or decay intermediates, at a reduced cost as compared to a genome-wide method. EnD-seq [72] and circTAIL-seq [109] illustrate the variety of such protocols that can be developed to address specific questions. EnD-seq was used to analyze 3′ tailing of the nonpolyadenylated histone mRNAs in humans [72]. As for TAIL-seq, an adaptor is ligated at the 3′ end of RNA and an adaptor antisense primer is used to initiate reverse transcription (RT). Alternatively, an RT primer ending with adenosines can be used to enrich for low abundant oligo-uridylated intermediates. No 5′ adaptor ligation is needed, and the Illumina-compatible sequences are incorporated during PCR amplification.

Both TAIL-seq and EnD-seq focus on the 3′ ends of mRNAs. By contrast, circTAIL-seq, a circular RT-PCR protocol adapted for high-throughput sequencing [109], is designed to simultaneously characterize 5′ and 3′ mRNA ends. Transcripts are first circularized using RNA ligase. The method was originally developed for noncapped RNAs. However, RNA used for the circularization step can be appropriately treated with phosphatase/pyrophosphatase combinations to discriminate between capped and uncapped mRNAs. After circularization, RT is performed using gene-specific primers. PCR amplicons are then analyzed by Illumina sequencing. Using circTAIL-seq, 3′ features such as poly(A) tail length and 3′ tailing can be linked to cap status and 5′ position for individual mRNA molecules.

around 16 nt [23]. This observation is coherent with previous reports documenting that the minimal length bound by PABP is 12 As [although 25 As are typically bound by a PABP on a poly (A) tail] and that PABP binds sequences other than homopolymeric poly(A) tails such as AU-rich elements in mRNAs [60–63]. Interestingly, the RNA-recognition motifs of PABP have different respective affinities for homopolymeric A stretches or heteropolymeric sequences [63–67]. Which of the four RNA-recognition motifs present in PABP bind to uridylated oligo(A) tails *in vivo* remains to be determined.

The current data in *Arabidopsis* support a model where URT1 and PABP cooperate to control the extent of deadenylation by hindering 3′ trimming of deadenylated mRNAs. Importantly, URT1-mediated uridylation and PABP slow down deadenylation but do not fully prevent 3′–5′ shortening of oligo(A) tails [23] (Figure 2). The mRNAs with very short oligo (A) tails that are ultimately produced are then uridylated by a TUTase activity other than URT1 [23] (Figure 2). Hence, the consequences of mRNA uridylation depend on the oligo(A) tail length: when oligo(A) size is greater than 13–15, uridylation by URT1 cooperates with PABP to slow down deadenylation, while in the case of shorter oligo(A) tails (<10 As), uridylation by (an) alternative TUTase(s) fails to restore the PABP binding site and presumably facilitates degradation (Figure 2). Indeed, only the shorter oligo(A) tails (<10As) accumulate when 5′–3′ degradation is impaired [23]. These short uridylated tails could be recognized by decay factors such as the Lsm1-7 complex to promote degradation as reported in other eukaryotes (Figure 2). HESO1, the second uridylyltransferase identified in *Arabidopsis* [42,68,69], represents an interesting candidate for this alternative mRNA uridylation activity. Uridylation by HESO1 could favor mRNA degradation as it does for small RNAs and RISC-cleaved transcripts. An overlap of RNA substrates between HESO1 and URT1 is already known for miRNAs. Indeed, URT1 can uridylate miRNAs in the absence of HESO1 and the methyltransferase HEN1, which mediates 2′-O-methylation of the 3′ terminal ribose of plant miRNAs and siRNAs [70,71].

Altogether, the current data indicate a dual role of uridylation in mRNA turnover in *Arabidopsis*. In addition to the canonical role of uridylation in stimulating mRNA degradation, which is likely conserved in plants, URT1-mediated uridylation creates oligo (A/U) tails of sufficient length to allow PABP binding, thereby preventing excessive deadenylation. A similar protective effect was previously observed in mammalian cell extracts, where U tracts added at the 3′ RNA end prevented 3′–5′ exonucleolytic decay, presumably via the binding of Lsm1-7 complex [41]. In addition, an analogous role for uridylation in restoring the normal length of mRNA extension was proposed for human histone mRNAs [72]. At the end of S-phase, uridylation initiates histone mRNA decay as discussed earlier. By contrast, during S-phase, a significant fraction of cytoplasmic histone mRNAs end with one or two uridines, which have replaced the 1–2 nt at the 3′ end of mature histone mRNAs (Figure 1). This uridylation could occur after the nibbling of the mRNA 3′ end by Eri-1 and help to maintain the integrity of histone transcripts [72].

Additional Functions of mRNA Uridylation

Besides mRNA decay, uridylation can also influence translation. Such a link is described for trypanosomal mitochondrial mRNAs. Upon completion of editing, two pentatricopeptide repeat-containing proteins, called kinetoplast polyadenylation/uridylation factors (KPAFs) 1 and 2, modulate the activity of KPAP1 poly(A) polymerase and RET1 TUTase, leading to the synthesis of A/U (70/30% ratio) heteropolymeric tails [27,36,38]. Interestingly, the fully edited RPS12 mRNA with long A/U extension, but not with a short A tail, is enriched in translating mitochondrial ribosomes and KPAF1 inhibition results in an inhibition of protein synthesis [27]. These results indicate a key role of long A/U extensions in recruiting edited mRNAs to mitochondrial ribosomes [27,73]. RET1 also contributes to a second type of mRNA uridylation: RET1 can add short continuous U tails to selected mRNAs, which negatively regulates their steady-state level [36,38]. This example illustrates the ability of one uridylyltransferase to play ambivalent roles in mRNA metabolism, that is, degradation versus promoting translation. Another dual function of uridylation was described in mammals for Group II let7 pre-miRNAs. The absence or presence of the RNA-binding protein Lin28 influences uridylation by TUT4/7 by favoring processing of the let7 miRNAs or degradation of the let7 pre-miRNAs, respectively (Box 2) [55,58,74–78].

Uridylation is also suspected to negatively regulate translation of nucleus-encoded mRNAs. In *A. nidulans*, uridylation by CutA/CutB was proposed to favor polysome dissociation for NMD targets [21]. This hypothesis was based on the observation that mutations in CutA/CutB increased the proportion of a PTC-containing mRNA associated with ribosomes. Because the corresponding protein does not accumulate in *cutA/cutB* mutants, uridylation was suggested to promote ribosome dissociation of the NMD target after translation termination at the PTC [21]. The underlying molecular mechanism is unknown and it may only apply to NMD targets or mRNAs with stalled ribosomes. Yet, this example illustrates the potential of uridylation in promoting polysome dissociation. Decapping and cytoplasmic recapping have recently been proposed as regulating translation, and uridylation may also be involved because uridylated mRNAs are enriched among cytoplasmic capping targets [79,80]. A possible interpretation brought forward to explain the overlap between the data sets is that uridylation induces decapping, which is an obvious prerequisite for the recapping step. Whether a causal relationship exists between uridylation and cytoplasmic capping remains to be experimentally investigated. A clear translation inhibition due to uridylation was observed in *Xenopus* oocytes because tethering of the uridylyltransferase XTUT7 represses the translation of a reporter mRNA [81]. In line with this observation, uridylation was recently proposed to inactivate translation of maternal mRNAs stored in starfish oocytes [82]. This hypothesis awaits further experimental confirmation and future work will tell whether translation inhibition by uridylation exists only at particular developmental stages or operates in diverse cell types.

CellPress

Concluding Remarks

In the last couple of years, uridylation has been recognized as a pervasive and conserved modification of eukaryotic mRNAs. Its key and conserved role in mRNA metabolism is to stimulate degradation. Yet, alternate functions began to be ascribed to uridylation. Although our knowledge on uridylation is rapidly expanding, we are still just beginning to appreciate the various consequences of uridine addition on the decay process itself, on translation inhibition, or on localization of target transcripts (see Outstanding Questions). A complete understanding of these various processes will require the identification of the whole machinery involved, from the whole set of TUTases (writers) to the factors that recognize uridylated mRNAs (readers) and ribonucleolytic activities that could potentially reverse uridylation (erasers). Localization, be it in the cytosol, on polysomes, in **P-bodies, or in stress granules**, will certainly influence the downstream consequences of uridylation. The identification of site-specific interactants of TUTases is almost an entirely open field of investigation. Such interactants could modulate the activity of TUTases, the recognition of their substrates, or the effects of uridylation. The impact of mRNA uridylation on stress-related responses and on developmental transitions is also to be explored. In this context, the advance of high-throughput sequencing methods that are dedicated to 3′-end analysis (Box 3) and their future developments will continue to revolutionize the study of mRNA uridylation and other 3′-end modifications such as cytidylation and guanylation. In the near future, mRNA tailing will be addressed in multiple organisms, developmental stages, growth conditions, and genetic backgrounds to draw a global picture of the fundamental roles of untemplated nucleotide addition in mRNA metabolism.

Outstanding Questions

What are the respective interactants of terminal uridylyltransferases (TUTases)? What is the conservation of such interacting protein networks across organisms?

Do TUTases fulfill alternative functions depending on distinct cytoplasmic localization and protein environment?

What is the full diversity of functions linked to uridylation in mRNA metabolism?

Can uridylation 'repair' the extremities of other mRNAs than replication-dependent histone mRNAs during S-phase in mammals and deadenylated mRNAs in *Arabidopsis*?

What is the impact of uridylation on mRNA translatability across development?

Can uridylation impact mRNA storage?

To what extent, mRNA uridylation impacts developmental transitions, diseases, and stress responses?

What are the terminal nucleotidyltransferases responsible for guanylation and cytidylation and what is the role of these modifications in mRNA metabolism?

Acknowledgments

We apologize to our colleagues whose contributions were omitted due to space limitation. Work in our laboratory is supported by the Centre National de la Recherche Scientifique (CNRS) and research grants from the French National Research Agency as part of the 'Investments for the Future' program in the frame of the LABEX ANR-10-LABX-0036_NETRNA and ANR-15-CE12-0008-01 to D.G.

References

1. Wang, X. and He, C. (2014) Dynamic RNA modifications in posttranscriptional Regulation. *Mol. Cell* 56, 5–12
2. Roundtree, I.A. and He, C. (2016) RNA epigenetics–chemical messages for posttranscriptional gene regulation. *Curr. Opin. Chem. Biol.* 30, 46–51
3. Dominissini, D. *et al.* (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206
4. Edelheit, S. *et al.* (2013) Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5 C within archaeal mRNAs. *PLoS Genet.* 9, e1003602
5. Luo, G-Z. *et al.* (2014) Unique features of the m6A methylome in *Arabidopsis thaliana*. *Nat. Commun.* 5, 5630
6. Deng, X. *et al.* (2015) Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids Res.* 43, 6557–6567
7. Dominissini, D. *et al.* (2016) The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 530, 441–446
8. Fu, Y. *et al.* (2014) Gene expression regulation mediated through reversible m^6A RNA methylation. *Nat. Rev. Genet.* 15, 293–306
9. Chen, K. *et al.* (2016) Nucleic acid modifications in regulation of gene expression. *Cell Chem. Biol.* 23, 74–85
10. Licht, K. and Jantsch, M.F. (2016) Rapid and dynamic transcriptome regulation by RNA editing and RNA modifications. *J. Cell Biol.* 213, 15–22
11. Meyer, K.D. and Jaffrey, S.R. (2014) The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat. Rev. Mol. Cell Biol.* 15, 313–326
12. Delatte, B. *et al.* (2016) Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* 351, 282–285
13. Li, X. *et al.* (2016) Pseudouridine: the fifth RNA nucleotide with renewed interests. *Curr. Opin. Chem. Biol.* 33, 108–116
14. Li, X. *et al.* (2016) Transcriptome-wide mapping reveals reversible and dynamic N1-methyladenosine methylome. *Nat. Chem. Biol.* 12, 311–316
15. Lange, H. *et al.* (2009) Polyadenylation-assisted RNA degradation processes in plants. *Trends Plant Sci.* 14, 497–504
16. Chang, J.H. and Tong, L. (2012) Mitochondrial poly(A) polymerase and polyadenylation. *Biochim. Biophys. Acta* 1819, 992–997
17. Song, J. *et al.* (2015) Uridylation and adenylation of RNAs. *Sci. China Life Sci.* 58, 1057–1066
18. Charlesworth, A. *et al.* (2013) Specificity factors in cytoplasmic polyadenylation. *Wiley Interdiscip. Rev. RNA* 4, 437–461
19. Norbury, C.J. (2013) Cytoplasmic RNA: a case of the tail wagging the dog. *Nat. Rev. Mol. Cell Biol.* 14, 643–653
20. Morozov, I.Y. *et al.* (2010) CUCU modification of mRNA promotes decapping and transcript degradation in *Aspergillus nidulans*. *Mol. Cell. Biol.* 30, 460–469
21. Morozov, I.Y. *et al.* (2012) mRNA 3′ tagging is induced by nonsense-mediated decay and promotes ribosome dissociation. *Mol. Cell. Biol.* 32, 2585–2595
22. Chang, H. *et al.* (2014) TAIL-seq: genome-wide determination of poly(A) tail length and 3′ end modifications. *Mol. Cell* 53, 1044–1052
23. Zuber, H. *et al.* (2016) Uridylation and PABP cooperate to repair mRNA deadenylated ends in *Arabidopsis*. *Cell Rep.* 14, 2707–2717
24. Martin, G. and Keller, W. (2007) RNA-specific ribonucleotidyl transferases. *RNA* 13, 1834–1849
25. Kwak, J.E. and Wickens, M. (2007) A family of poly(U) polymerases. *RNA* 13, 860–867
26. Aphasizhev, R. and Aphasizheva, I. (2011) Mitochondrial RNA processing in trypanosomes. *Res. Microbiol.* 162, 655–663

27. Aphasizheva, I. *et al.* (2011) Pentatricopeptide repeat proteins stimulate mRNA adenylation/uridylation to activate mitochondrial translation in trypanosomes. *Mol. Cell* 42, 106–117
28. Suematsu, T. *et al.* (2016) Antisense transcripts delimit exonucleolytic activity of the mitochondrial 3′ processome to generate guide RNAs. *Mol. Cell* 61, 364–378
29. Lee, M. *et al.* (2014) Emerging roles of RNA modification: m6A and U-Tail. *Cell* 158, 980–987
30. Viegas, S.C. *et al.* (2015) *Surprises in the 3′-end: "U" can decide too! FEBS J.* 282, 3489–3499
31. Munoz-Tello, P. *et al.* (2015) Polyuridylation in eukaryotes: a 3′-end modification regulating RNA Life. *BioMed Res. Int.* 2015, 968127
32. Huo, Y. *et al.* (2016) Widespread 3′-end uridylation in eukaryotic RNA viruses. *Sci. Rep.* 6, 25454
33. Rissland, O.S. *et al.* (2007) Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell. Biol.* 27, 3612–3624
34. Rissland, O.S. and Norbury, C.J. (2009) Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. *Nat. Struct. Mol. Biol.* 16, 616–623
35. Sement, F.M. *et al.* (2013) Uridylation prevents 3′ trimming of oligoadenylated mRNAs. *Nucleic Acids Res.* 41, 7115–7127
36. Aphasizheva, I. and Aphasizhev, R. (2010) RET1-catalyzed uridylylation shapes the mitochondrial transcriptome in *Trypanosoma brucei*. *Mol. Cell. Biol.* 30, 1555–1567
37. Lim, J. *et al.* (2014) Uridylation by TUT4 and TUT7 marks mRNA for degradation. *Cell* 159, 1365–1376
38. Etheridge, R.D. *et al.* (2008) 3′ Adenylation determines mRNA abundance and monitors completion of RNA editing in *T. brucei* mitochondria. *EMBO J.* 27, 1596–1608
39. Shen, B. and Goodman, H.M. (2004) Uridine addition after microRNA-directed cleavage. *Science* 306, 997
40. Xu, K. *et al.* (2016) MicroRNA-mediated target mRNA cleavage and 3′-uridylation in human cells. *Sci. Rep.* 6, 30242
41. Song, M-G. and Kiledjian, M. (2007) 3′ Terminal oligo U-tract-mediated stimulation of decapping. *RNA* 13, 2356–2365
42. Ren, G. *et al.* (2014) Methylation protects microRNAs from an AGO1-associated activity that uridylates 5′ RNA fragments generated by AGO1 cleavage. *Proc. Natl. Acad. Sci. U. S. A.* 111, 6365–6370
43. Marzluff, W.F. *et al.* (2008) Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat. Rev. Genet.* 9, 843–854
44. Mullen, T.E. and Marzluff, W.F. (2008) Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. *Genes Dev.* 22, 50–65
45. Tan, D. *et al.* (2013) Structure of histone mRNA stem-loop, human stem-loop binding protein, and 3′hExo ternary complex. *Science* 339, 318–321
46. Choe, J. *et al.* (2014) The mRNP remodeling mediated by UPF1 promotes rapid degradation of replication-dependent histone mRNA. *Nucleic Acids Res.* 42, 9334–9349
47. Hoefig, K.P. *et al.* (2013) Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. *Nat. Struct. Mol. Biol.* 20, 73–81
48. Slevin, M.K. *et al.* (2014) Deep sequencing shows multiple oligouridylations are required for 3′ to 5′ degradation of histone mRNAs on polyribosomes. *Mol. Cell* 53, 1020–1030
49. Su, W. *et al.* (2013) mRNAs containing the histone 3′ stem-loop are degraded primarily by decapping mediated by oligouridylation of the 3′ end. *RNA* 19, 1–16
50. Schmidt, M-J. *et al.* (2010) The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. *RNA* 17, 39–44
51. Lackey, P.E. *et al.* (2016) TUT7 catalyzes the uridylation of the 3′ end for rapid degradation of histone mRNA. *RNA.* http://dx.doi.org/10.1261/ma.058107.116 (in press)
52. Subtelny, A.O. *et al.* (2014) Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* 508, 66–71
53. Malecki, M. *et al.* (2013) The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* 32, 1842–1854
54. Lubas, M. *et al.* (2013) Exonuclease hDIS3L2 specifies an exosome-independent 3′-5′ degradation pathway of human cytoplasmic mRNA. *EMBO J.* 32, 1855–1868
55. Kim, B. *et al.* (2015) TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. *EMBO J.* 34, 1801–1815
56. Ustianenko, D. *et al.* (2013) Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* 19, 1632–1638
57. Łabno, A. *et al.* (2016) Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. *Nucleic Acids Res.* http://dx.doi.org/10.1093/nar/gkw649 (in press)
58. Faehnle, C.R. *et al.* (2014) Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. *Nature* 514, 252–256
59. Thomas, M.P. *et al.* (2015) Apoptosis triggers specific, rapid, and global mRNA decay with 3′ uridylated intermediates degraded by DIS3L2. *Cell Rep.* 11, 1079–1089
60. Baejen, C. *et al.* (2014) Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition. *Mol. Cell* 55, 745–757
61. Eliseeva, I.A. *et al.* (2013) Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochem. Mosc.* 78, 1377–1391
62. Kini, H.K. *et al.* (2016) Cytoplasmic poly(A) binding protein-1 binds to genomically encoded sequences within mammalian mRNAs. *RNA* 22, 61–74
63. Sladic, R.T. *et al.* (2004) Human PABP binds AU-rich RNA via RNA-binding domains 3 and 4. *Eur. J. Biochem.* 271, 450–457
64. Khanam, T. *et al.* (2006) Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1 + 2 and 3 + 4. *RNA Biol.* 3, 170–177
65. Kühn, U. and Pieler, T. (1996) *Xenopus* poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. *J. Mol. Biol.* 256, 20–30
66. Patel, G.P. and Bag, J. (2006) IMP1 interacts with poly(A)-binding protein (PABP) and the autoregulatory translational control element of PABP-mRNA through the KH III-IV domain. *FEBS J.* 273, 5678–5690
67. Mullin, C. *et al.* (2004) Interaction of rat poly(A)-binding protein with poly(A)- and non-poly(A) sequences is preferentially mediated by RNA recognition motifs 3 + 4. *FEBS Lett.* 576, 437–441
68. Ren, G. *et al.* (2012) Uridylation of miRNAs by HEN1 SUPPRESSOR1 in *Arabidopsis*. *Curr. Biol.* 22, 695–700
69. Zhao, Y. *et al.* (2012) The *Arabidopsis* nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. *Curr. Biol.* 22, 689–694
70. Wang, X. *et al.* (2015) Synergistic and independent actions of multiple terminal nucleotidyl transferases in the 3′ tailing of small RNAs in *Arabidopsis*. *PLoS Genet.* 11, e1005091
71. Tu, B. *et al.* (2015) Distinct and cooperative activities of HESO1 and URT1 nucleotidyl transferases in microRNA turnover in *Arabidopsis*. *PLoS Genet.* 11, e1005119
72. Welch, J.D. *et al.* (2015) EnD-Seq and AppEnD: sequencing 3′ ends to identify nontemplated tails and degradation intermediates. *RNA* 21, 1375–1389
73. Read, L.K. *et al.* (2011) Marked for translation: long A/U tails as an interface between completion of RNA editing and ribosome recruitment. *Mol. Cell* 42, 6–8
74. Heo, I. *et al.* (2009) TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* 138, 696–708
75. Heo, I. *et al.* (2008) Lin28 mediates the terminal uridylation of let-7 precursor microRNA. *Mol. Cell* 32, 276–284
76. Heo, I. *et al.* (2012) Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell* 151, 521–532
77. Choudhury, N.R. *et al.* (2014) Trim25 is an RNA-specific activator of Lin28a/TuT4-mediated uridylation. *Cell Rep.* 9, 1265–1272

78. Hagan, J.P. *et al.* (2009) Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* 16, 1021–1025
79. Mukherjee, C. *et al.* (2012) Identification of cytoplasmic capping targets reveals a role for Cap homeostasis in translation and mRNA stability. *Cell Rep.* 2, 674–684
80. Kiss, D.L. *et al.* (2016) Cap homeostasis is independent of poly(A) tail length. *Nucleic Acids Res.* 44, 304–314
81. Lapointe, C.P. and Wickens, M. (2013) The nucleic acid-binding domain and translational repression activity of a *Xenopus* terminal uridylyl transferase. *J. Biol. Chem.* 288, 20723–20733
82. Ochi, H. and Chiba, K. (2016) Hormonal stimulation of starfish oocytes induces partial degradation of the 3′ termini of cyclin B mRNAs with oligo(U) tails, followed by poly(A) elongation. *RNA* 22, 822–829
83. Wahle, E. and Winkler, G.S. (2013) RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim. Biophys. Acta* 1829, 561–570
84. Lunde, B.M. *et al.* (2012) Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. *Nucleic Acids Res.* 40, 9815–9824
85. Yates, L.A. *et al.* (2012) Structural basis for the activity of a cytoplasmic RNA terminal U-transferase. *Nat. Struct. Mol. Biol.* 19, 782–787
86. Blahna, M.T. *et al.* (2011) Terminal uridyltransferase enzyme Zcchc11 promotes cell proliferation independent of its uridyltransferase activity. *J. Biol. Chem.* 286, 42381–42389
87. Brown, R.S. (2005) Zinc finger proteins: getting a grip on RNA. *Curr. Opin. Struct. Biol.* 15, 94–98
88. Loughlin, F.E. and Mackay, J.P. (2006) Zinc fingers are known as domains for binding DNA and RNA. Do they also mediate protein-protein interactions? *IUBMB Life* 58, 731–733
89. Loughlin, F.E. *et al.* (2012) Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. *Nat. Struct. Mol. Biol.* 19, 84–89
90. Järvelin, A.I. *et al.* (2016) The new (dis)order in RNA regulation. *Cell Commun. Signal.* 14, 9
91. Jones, D.T. and Cozzetto, D. (2015) DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* 31, 857–863
92. Calabretta, S. and Richard, S. (2015) Emerging roles of disordered sequences in RNA-binding proteins. *Trends Biochem. Sci.* 40, 662–672
93. Castello, A. *et al.* (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406
94. Castello, A. *et al.* (2016) Comprehensive identification of RNA-binding domains in human cells. *Mol. Cell* Published online July 19. http://dx.doi.org/10.1016/j.molcel.2016.06.029
95. Jonas, S. and Izaurralde, E. (2013) The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes Dev.* 27, 2628–2641
96. Morozov, I.Y. *et al.* (2010) Distinct roles for Caf1, Ccr4, Edc3 and CutA in the co-ordination of transcript deadenylation, decapping and P-body formation in *Aspergillus nidulans*. *Mol. Microbiol.* 76, 503–516
97. Haas, G. *et al.* (2016) Identification of factors involved in target RNA-directed microRNA degradation. *Nucleic Acids Res.* 44, 2873–2887
98. Li, J. *et al.* (2005) Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in *Arabidopsis*. *Curr. Biol.* 15, 1501–1507
99. Yu, B. *et al.* (2005) Methylation as a crucial step in plant microRNA biogenesis. *Science* 307, 932–935
100. Saito, K. *et al.* (2007) Pimet, the *Drosophila* homolog of HEN1, mediates 2′-O-methylation of Piwi-interacting RNAs at their 3′ ends. *Genes Dev.* 21, 1603–1608
101. Horwich, M.D. *et al.* (2007) The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* 17, 1265–1272
102. Reimão-Pinto, M.M. *et al.* (2015) Uridylation of RNA hairpins by tailor confines the emergence of microRNAs in *Drosophila*. *Mol. Cell* 59, 203–216
103. Liu, X. *et al.* (2014) A microRNA precursor surveillance system in quality control of MicroRNA synthesis. *Mol. Cell* 55, 868–879
104. Ha, M. and Kim, V.N. (2014) Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 15, 509–524
105. McManus, M.T. *et al.* (2000) *Trypanosoma brucei* guide RNA poly(U) tail formation is stabilized by cognate mRNA. *Mol. Cell. Biol.* 20, 883–891
106. Trippe, R. *et al.* (2006) Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA* 12, 1494–1504
107. Preti, M. *et al.* (2013) Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA. *Nucleic Acids Res.* 41, 4709–4723
108. Sikorski, P.J. *et al.* (2015) Distinct 18S rRNA precursors are targets of the exosome complex, the exoribonuclease RRP6L2 and the terminal nucleotidyltransferase TRL in *Arabidopsis thaliana*. *Plant J.* 83, 991–1004
109. Gazestani, V.H. *et al.* (2016) circTAIL-seq, a targeted method for deep analysis of RNA 3′ tails, reveals transcript-specific differences by multiple metrics. *RNA* 22, 477–486

PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Review

Cite this article: de Almeida C, Scheer H, Gobert A, Fileccia V, Martinelli F, Zuber H, Gagliardi D. 2018 RNA uridylation and decay in plants. *Phil. Trans. R. Soc. B* **373**: 20180163. http://dx.doi.org/10.1098/rstb.2018.0163

Accepted: 18 August 2018

One contribution of 11 to a theme issue '5′ and 3′ modifications controlling RNA degradation'.

Subject Areas:
molecular biology, plant science

Keywords:
RNA degradation, RNA decay, terminal nucleotidyltransferase, uridylation, mRNAs

Author for correspondence:
Dominique Gagliardi
e-mail: dominique.gagliardi@ibmp-cnrs.unistra.fr

†These authors contributed equally to this study.

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4244939.

THE ROYAL SOCIETY PUBLISHING

RNA uridylation and decay in plants

Caroline de Almeida[1,†], Hélène Scheer[1,†], Anthony Gobert[1], Veronica Fileccia[2], Federico Martinelli[2], Hélène Zuber[1] and Dominique Gagliardi[1]

[1]Institut de biologie moléculaire des plantes (IBMP), Centre national de la recherche scientifique (CNRS), Université de Strasbourg, 12 rue Zimmer, 67000 Strasbourg, France
[2]Dipartimento di Scienze Agrarie Alimentari Forestali, Università degli Studi di Palermo, viale delle scienze ed. 4, Palermo 90128, Italy

FM, 0000-0002-3981-0567; HZ, 0000-0002-2132-4536; DG, 0000-0002-5871-7544

RNA uridylation consists of the untemplated addition of uridines at the 3′ extremity of an RNA molecule. RNA uridylation is catalysed by terminal uridylyltransferases (TUTases), which form a subgroup of the terminal nucleotidyltransferase family, to which poly(A) polymerases also belong. The key role of RNA uridylation is to regulate RNA degradation in a variety of eukaryotes, including fission yeast, plants and animals. In plants, RNA uridylation has been mostly studied in two model species, the green algae *Chlamydomonas reinhardtii* and the flowering plant *Arabidopsis thaliana*. Plant TUTases target a variety of RNA substrates, differing in size and function. These RNA substrates include microRNAs (miRNAs), small interfering silencing RNAs (siRNAs), ribosomal RNAs (rRNAs), messenger RNAs (mRNAs) and mRNA fragments generated during post-transcriptional gene silencing. Viral RNAs can also get uridylated during plant infection. We describe here the evolutionary history of plant TUTases and we summarize the diverse molecular functions of uridylation during RNA degradation processes in plants. We also outline key points of future research.

This article is part of the theme issue '5′ and 3′ modifications controlling RNA degradation'.

1. RNA uridylation, a key post-transcriptional regulatory process

RNA uridylation is a post-transcriptional modification, which consists of the addition of uridines to the 3′ end of RNA. RNA uridylation plays a key role in the regulation of gene expression across eukaryotes, with the exception to date of *Saccharomyces cerevisiae*, which has lost the capacity to uridylate RNAs. U-tailing has been reported for a variety of RNA substrates: from mitochondrial transcripts in trypanosomes, to mRNAs and a plethora of non-coding RNAs in diverse organisms, including fission yeast, amphibians, insects, plants or mammals [1–9]. Uridylation targets transcripts produced by RNA polymerases I, II and III (see accompanying articles by Zigackova and Vanacova [10], Warkocki *et al*. [11] and [12–14]) and it emerges as a pervasive post-transcriptional process.

RNA uridylation plays diverse roles, which depend on the cellular compartment, the identity of the terminal uridylyltransferase (TUTase) or the RNA substrate itself [1,2,4–9]. The 3′ untemplated addition of uridines may facilitate processing of primary transcripts, stabilize RNA and possibly control translation of mRNAs. Yet, its chief role is to trigger degradation both by the 5′–3′ and 3′–5′ RNA degradation pathways [1,2,4–9].

Here, we summarize our current knowledge on RNA uridylation and decay in plants. We begin by describing the terminal nucleotidyltransferase (TNTase) family, to which TUTases belong. As an example, the organization of the small multigenic TNTase family is detailed for the model plant *Arabidopsis thaliana*. We then present a comprehensive evolutionary history of TUTases in Archaeplastida (i.e. all plants), which reveals that two TUTases have been maintained in the whole green lineage, suggesting specific and critical functions. We then review our current knowledge on how uridylation by these TUTases impacts the degradation of various classes of RNAs in plants.

© 2018 The Author(s) Published by the Royal Society. All rights reserved.

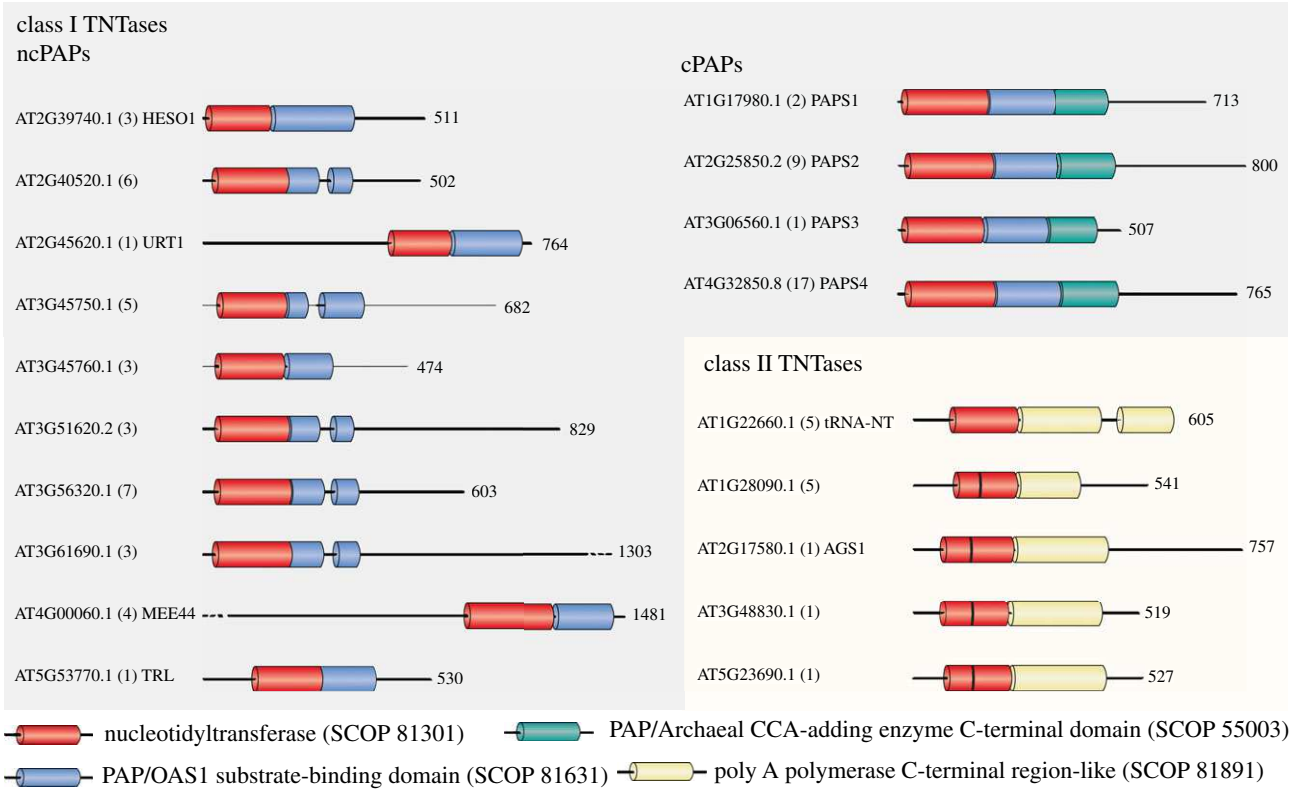


Figure 1. Domain organization of class I and class II TNTases of *A. thaliana*. The class I TNTase family is composed of 10 non-canonical poly(A) polymerases (ncPAPs) and four canonical poly(A) polymerases (cPAPs). The class II TNTase family contains the tRNA nucleotidyltransferase (tRNA-NT), also called the tRNA CCA-adding enzyme, and four bacterial PAP-like nucleotidyltransferases. Boxes represent conserved structural domains identified using the structural classification of proteins (SCOP) according to the superfamily database [25]. Non-conserved regions are drawn as lines. Each TNTase is identified by its AGI (*Arabidopsis* Genome Initiative) reference gene model. The numbers of all gene models are shown in parentheses. Finally, names are shown for the 10 TNTases that have been studied to date. The numbers on the right indicate the number of amino acids for each TNTase. The colour code for the Superfamily SCOP domains is indicated on the figure. The vertical black bar drawn in the nucleotidyltransferase domain of four of the five class II TNTases represents a motif discriminating bacterial PAP-like nucleotidyltransferases from bacterial tRNA-NT [15].

2. Characteristic features of TUTases in plants

RNA uridylation is catalysed by terminal uridylyltransferases (TUTases). TUTases belong to the superfamily of DNA polymerase beta-like nucleotidyltransferases [15]. This superfamily regroups enzymes that conjugate nucleotides to proteins, antibiotics or RNAs [15]. The nucleotidyltransferases that add untemplated nucleotides (adenosines, uridines, guanosines and cytidines) to the 3′ end of RNAs are called ribonucleotidyl transferases (rNTases) or terminal nucleotidyltransferases (TNTases).

(a) Classification of terminal nucleotidyltransferases

TNTases are split into two classes based on structural differences in their catalytic fold and in the domain responsible for nucleotide selection [15]. Class I includes the 'canonical' poly(A) polymerases (cPAPs), which are responsible for the co-transcriptional addition of stabilizing poly(A) tails to transcripts synthesized by RNA polymerase II (Pol II), the tRNA CCA-adding enzymes in Archaea, 2′-5′-oligo(A) synthetases (OAS) and a group of TNTases involved in the polyadenylation, uridylation, cytidylation and guanylation of diverse RNA substrates [3,15]. Class II TNTases correspond to bacterial poly(A) polymerases and tRNA CCA-adding enzymes found in eukaryotes and in certain bacteria.

(b) Eukaryotic TNTases are encoded by small multigenic families

TNTases are encoded by small multigenic families, whose complexity varies across eukaryotes. For instance, three canonical and 11 non-canonical PAPs (ncPAPs) are expressed in humans [16] (see also accompanying paper by Warkocki *et al.* [11]). TNTases are also encoded by small multigenic families in plants, as reported for the green algae *Chlamydomonas reinhardtii*, or for two land plants: *Zea mays* (maize) and *Arabidopsis thaliana* [17–24]. In *Arabidopsis*, 19 TNTase genes have been identified on the basis of sequence homology: 14 class I TNTases and 5 class II TNTases (figure 1). The class II of *Arabidopsis* TNTases contains a single tRNA CCA-adding enzyme also called tRNA-nucleotidyltransferase (tRNA-NT), which processes tRNAs encoded by the nuclear, plastidial and mitochondrial genomes [26], and four bacterial PAP-like nucleotidyltransferases, which are predicted to localize in mitochondria and plastids [18] (figure 1). The class I of *Arabidopsis* TUTases is composed of 10 non-canonical PAPs and four PAPS that contain the characteristic domains of canonical PAPs (PAPS1 to 4) [17–20]. Yet, PAPS3 is localized in the cytosol, is mostly expressed in pollen and does not contain the C-terminal extension found in PAPS1, S2 and S4 (figure 1) [19]. The role of PAPS3 remains to be characterized. By contrast, PAPS1, S2 and S4 correspond to the canonical PAPs involved in the co-transcriptional polyadenylation of RNA Pol II transcripts. Interestingly, PAPS1, S2

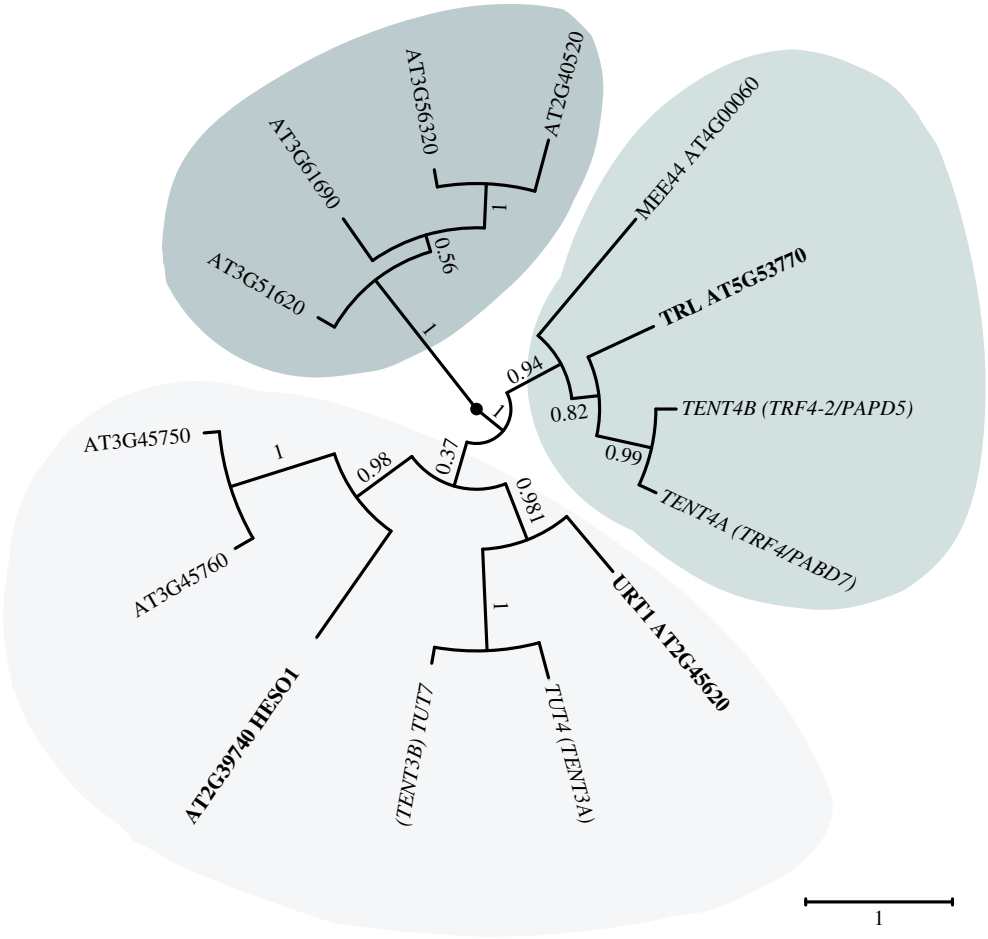


Figure 2. Phylogenetic relationship among *A. thaliana* class I ncPAPs and four human ncPAPs. The nucleotidyltransferase domains SCOP 81301 and the PAP/OAS1 substrate-binding domains SCOP 81631 of the 10 class I ncPAPs of *Arabidopsis* and four human ncPAPs were aligned with Muscle (v. 3.8.31) [37]. The maximum-likelihood tree was generated using PhyML (v. 3.1) on Phylogeny.fr [38] and edited using FigTree (v. 1.4.3, http://tree.bio.ed.ac.uk/software/figtree/). *Arabidopsis* and human ncPAPs are indicated in regular and italic characters, respectively. Support values (approximate likelihood-ratio statistical test, aLRT v 3.0) are shown on branches. The scale bar represents the number of substitutions per amino acid site. HESO1, URT1 and two other TNTases form a cluster with human TUT7 and TUT4. MEE44 and TRL form a cluster with human TENT4A and TENT4B. The remaining four TNTases form a separated cluster. HESO1, URT1 and TRL, the three class I ncPAPs that have been functionally characterized in *Arabidopsis*, are indicated in bold.

and S4 are functionally specialized and preferentially target subpopulations of transcripts [27–30].

Much remains to be discovered about the function of the 10 *Arabidopsis* class I ncPAPs. To date, functional data have been reported for three of them: TRF4/5-LIKE (TRL), HEN1 SUPPRESSOR 1 (HESO1) and UTP:RNA URIDYLYLTRANSFERASE 1 (URT1) [20,22,31,32]. TRL is a nuclear ncPAP, which adenylates rRNA maturation by-products and precursors to facilitate their degradation or processing by the RNA exosome [22]. TRL is an orthologue of TRF4 in *S. cerevisiae* or TENT4B (hTRF4–2, PAPD5) in humans [33–36]. Another *Arabidopsis* class I ncPAP, MEE44, is evolutionarily close to TRL and TENT4A/B. Indeed, a phylogenetic analysis using the nucleotidyltransferase domains of the 10 *Arabidopsis* ncPAPs aligned with that of four human ncPAPs (TENT4A, TENT4B and the 2 TUTases TUT4 and TUT7) reveals that TRL and MEE44 cluster with TENT4A/4B (figure 2). This analysis also indicates that four uncharacterized *Arabidopsis* ncPAPs form a distinct clade and finally that four proteins including HESO1 and URT1 cluster with the human TUTases TUT4/7 (figure 2). The nucleotide specificity of AT3G45750 and AT3G45760 has not been reported yet, but HESO1 and URT1 are indeed *bona fide* TUTases. Altogether, HESO1 and URT1 uridylate small RNAs, mRNAs and the 5′ fragments resulting from cleavage by the RNA-induced silencing complex (RISC) [20,31,32,39,40]. Of note, HESO1 is the homologue of MUT68, which was discovered in *C. reinhardtii*, as a ncPAP involved in the degradation of RISC-cleaved transcripts and small RNAs [41,42].

(c) Evolutionary analysis of TUTases in plants

A recent phylogenetic analysis of HESO1 and URT1 homologues, mostly from bryophytes, lycophytes and ferns, revealed that each TUTase forms a monophyletic group [43]. To extend our knowledge on the evolutionary history of TUTases in Archaeplastida (i.e. all plants), a comprehensive phylogenetic analysis was performed using URT1 and HESO1 homologues from 79 species representing all major groups of Archaeplastida, including glaucophytes, rhodophytes (red algae), chlorophytes and streptophyte algae, bryophytes (liverworts, hornworts and mosses), lycophytes and pteridophytes (e.g. ferns), gymnosperms (e.g. conifers and Gingko), and angiosperms (flowering plants; figure 3). To retrieve URT1 and HESO1 homologue sequences, we screened phytozome, NCBI TSA and NCBI EST databases by using either BLASTP or TBLASTN algorithms with the amino acid sequence of *Arabidopsis* URT1 (AT2G45620) and

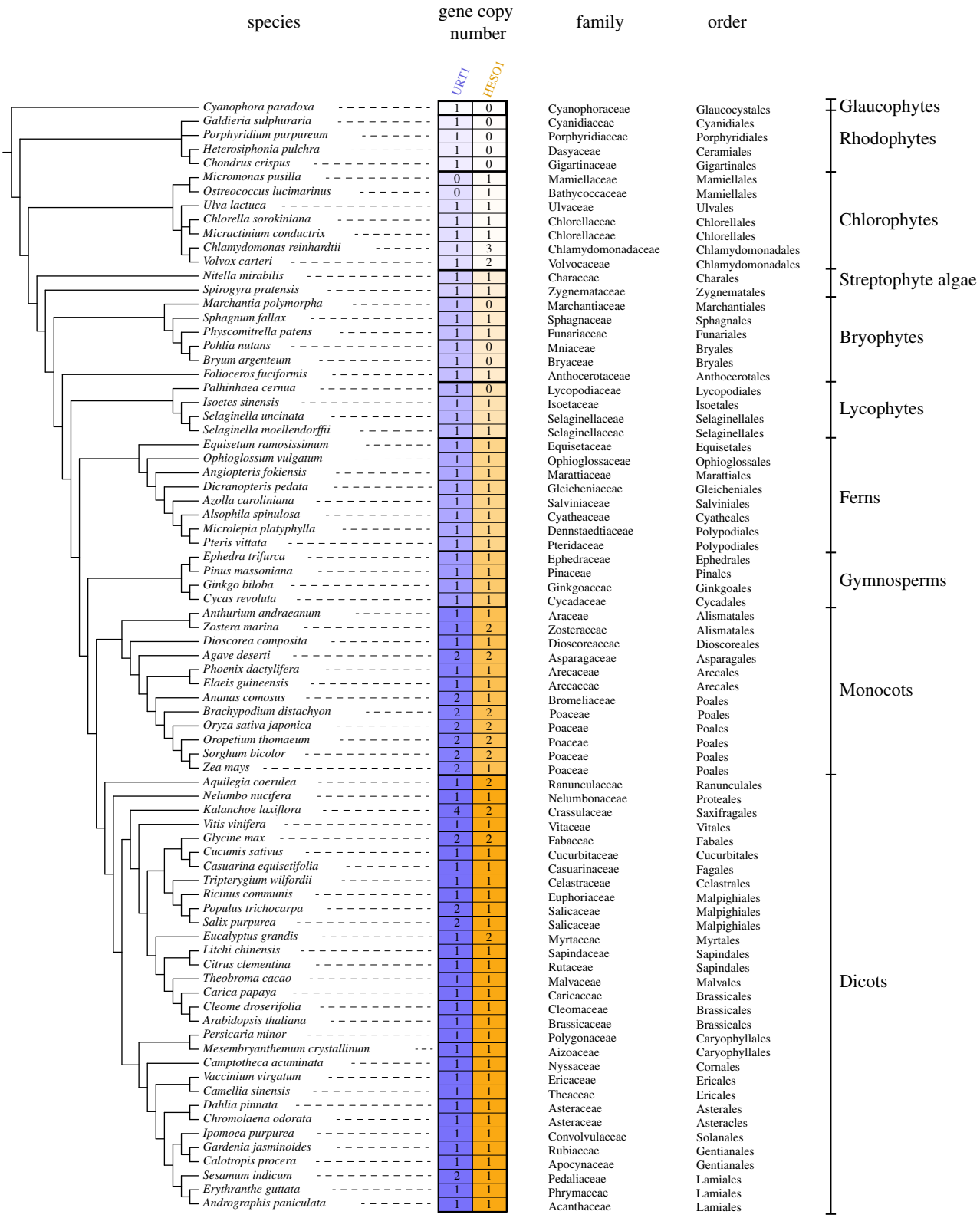


Figure 3. Copy number of HESO1 and URT1 in 79 representative species of Archaeplastida. The phylogenic relationship between the 79 species analysed in this study was visualized using Phylostatic [44]. Full species names are grouped by taxonomic clades indicated on the right. The colour code for URT1 and HESO1 in the different clades is conserved in figures 4 and 5. The numbers of copies detected for HESO1 and URT1 are indicated for each species. Sequences in FASTA format are given in electronic supplementary material, Datasets S1 and S2 for HESO1 and URT1, respectively.

HESO1 (AT2G39740). We also included in our analysis previously published URT1 and HESO1 sequences from bryophytes, lycophytes and ferns [43] (see electronic supplementary material, Dataset S1 and S2 for a compilation of HESO1 and URT1 sequences, respectively). URT1 and HESO1 homologous sequences for the representative species shown in figure 3 were separately aligned with MUSCLE [37]. Amino acids that did not align to the Catalytic Core Domain (CCD) (COG5260) identified in URT1 and HESO1 were trimmed. Finally, URT1 and HESO1 trimmed sequences were realigned altogether and the phylogenetic tree shown in figure 4 was generated using the maximum-likelihood method (v. 3.1/3.0 aLRT) and WAG substitution model implemented in PhyML [45,46] (see electronic supplementary

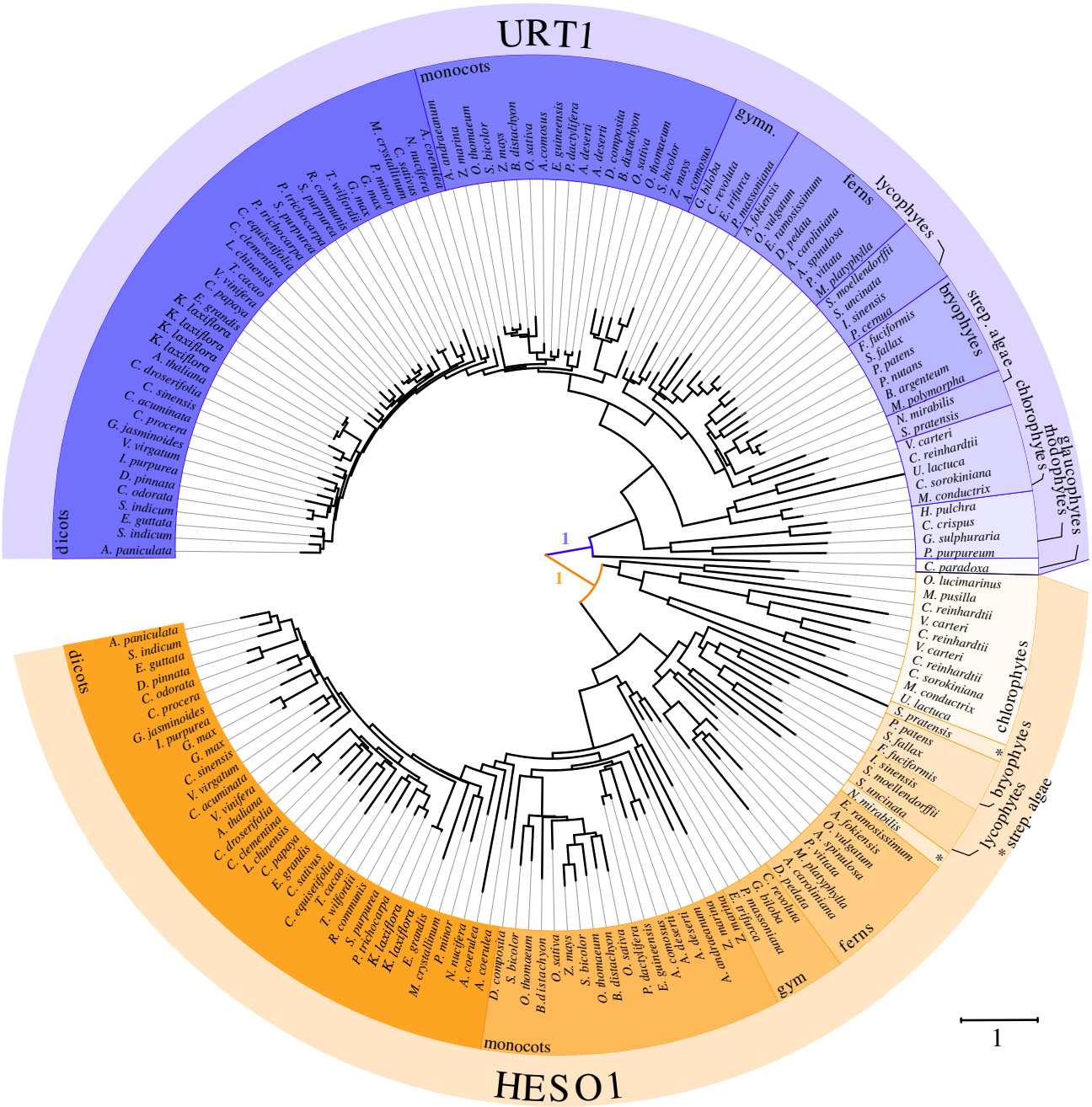


Figure 4. Phylogenetic relationship between URT1 and HESO1 sequences among 79 representative species of Archaeplastida. The phylogenetic tree was generated using the maximum-likelihood method and WAG substitution model implemented in PhyML (v. 3.1) [45,46]; see electronic supplementary material, Dataset S3 for sequence alignment). The tree was edited using iTOL (v. 4.2.1) [47]. Colour code for taxonomic clades is defined in figure 3. Statistical values for the first branches (approximate likelihood-ratio test, aLRT v. 3.0) support that URT1 and HESO1 proteins form two distinct clades. The scale bar represents the number of substitutions per amino acid site.

material, Dataset S3 for the final alignment of all sequences). The main outcome of this analysis is that a monophyletic group is observed for each TUTase, indicating an early divergence of HESO1 and URT1 that have been maintained in the green lineage (figure 4). In most species, homologues of HESO1 and URT1 are present each as a single copy (figures 3 and 4), raising the possibility that they are orthologues, as previously proposed for bryophyte, lycophyte and fern species [43]. Yet, several species either have multiple URT1 and/or HESO1 homologues, or have only one TUTase: either HESO1 or URT1. Interestingly, species representing the phyla Glaucophyta and Rhodophyta lack a HESO1 homologue (figures 3 and 4). Because glaucophytes and rhodophytes are early diverging in the Archaeplastida lineage, the absence of HESO1 homologues suggests that either URT1 homologues predate the apparition of HESO1, or glaucophytes and rhodophytes have lost HESO1. In addition, HESO1 was not detected in representative species of Marchantiales, Bryales and Lycopodiales (figures 3 and 4). Conversely, two species of the order Mamiellales, *Micromonas pusilla* and *Ostreococcus lucimarinus*, have no URT1 homologues. We further checked that URT1 is not detected in *Ostreococcus tauri*, another species of the order Mamiellales. Those three Chlorophyta green algae of the order Mamiellales are unicellular organisms with a reduced nuclear genome that seem to have lost URT1 during speciation. Of note, Mamiellales is the only order of all Viridiplantae lacking a URT1 homologue. Thus, the absence of either HESO1 or URT1 homologues appears extremely rare in plant species, and in such species, a single TUTase could be

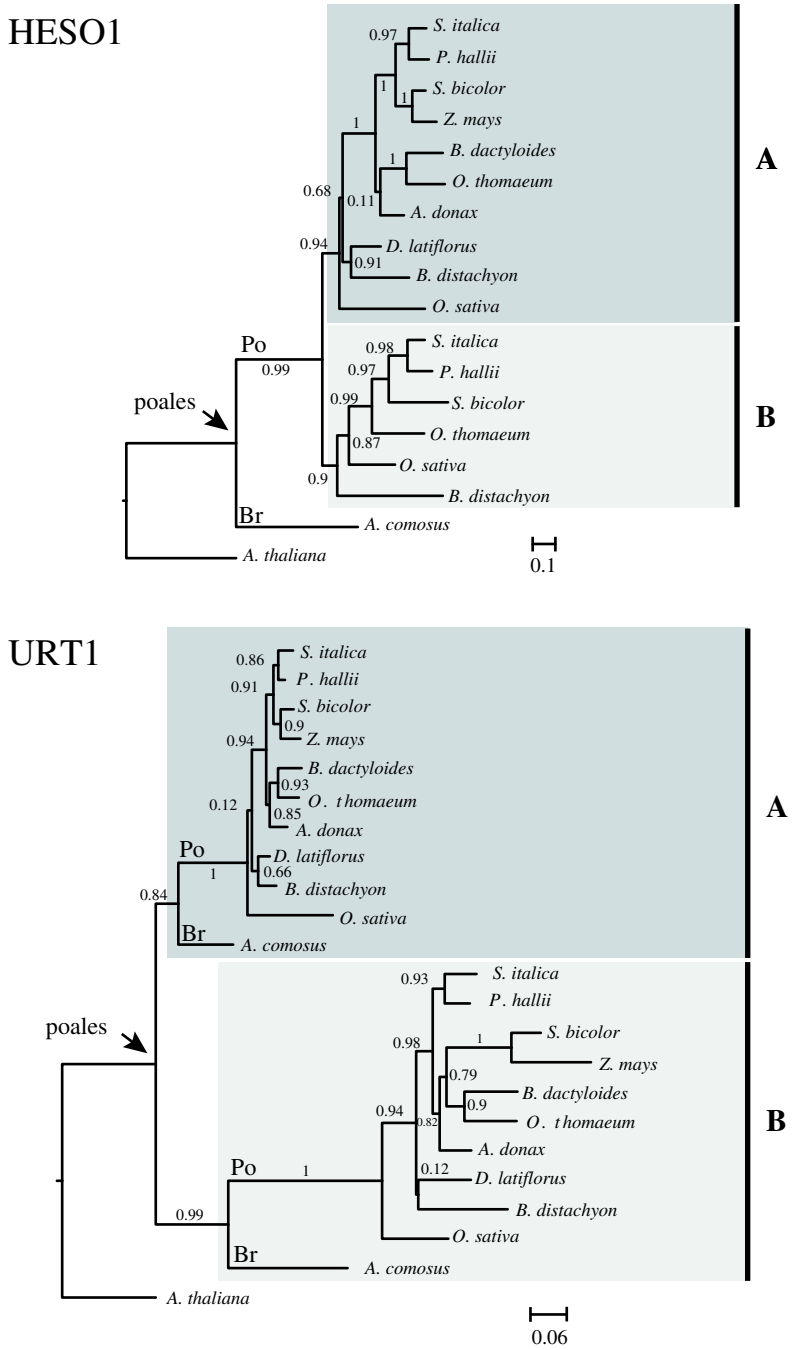


Figure 5. Phylogenetic relationship of URT1 and HESO1 isoforms among 11 species of Poales. Sequences of HESO1 and URT1 isoforms from 10 Poaceae (Po) and 1 Bromeliaceae (Br) were aligned separately. Full names of species are given in figure 3. The analysis was performed as described for figure 4 except that the trees were edited with FigTree (v. 1.4.3, http://tree.bio.ed.ac.uk/software/figtree/). Support values (approximate likelihood-ratio statistical test, aLRT v. 3.0) are shown on branches. The sequence alignments for HESO1 and URT1 used to build the trees are given in electronic supplementary material, Datasets S4 and S5, respectively.

responsible for the RNA uridylation catalysed by both HESO1 and URT1 in most plant species. Conversely, certain plant species have dual copies of URT1 and/or HESO1 because of either local or whole genome duplication (WGD) events, as for *Glycine max* (figures 3 and 4) [48]. Four out of five representative species of Poaceae chosen for our initial analysis shown in figures 3 and 4 also have at least two copies of URT1 and HESO1. To obtain a better view of the evolutionary history of TUTases in this family that regroups important crops such as maize (*Zea mays*), sorghum (*Sorghum bicolor*) or rice (*Oryza sativa*), a more focused phylogenetic analysis of URT1 and HESO1 homologues was performed from 11 species of Poales (10 Poaceae and the Bromeliaceae *Ananas comosus*) (figure 5; see electronic supplementary material, Datasets S4 and S5 for alignments of HESO1 and URT1 sequences, respectively). In fact, two copies of HESO1 were not systematically found (figure 5). For instance, maize and pineapple (*A. comosus*) are among the Poales species that lack the second HESO1 isoform, noted HESO1B. Usually, HESO1A, which is the isoform detected in all Poales, is constitutively expressed, and at a higher level than HESO1B (figure 6). Altogether, these data suggest that HESO1A could be the orthologue of the eudicot HESO1, while HESO1B is either dispensable in some Poales or may have acquired a specialized function or expression pattern in certain Poales species.

In contrast to HESO1 isoforms, two copies of URT1 are present in all 11 representative species chosen here and URT1 sequences form two well-defined clades, defining A

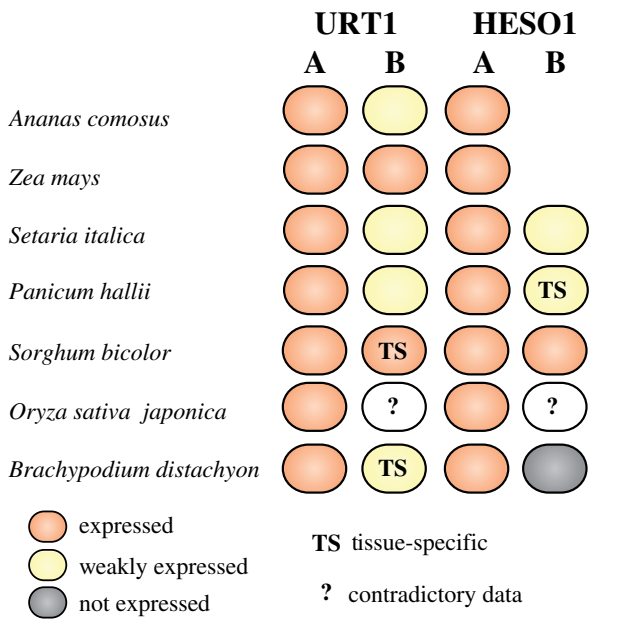


Figure 6. Expression of URT1 and HESO1 isoforms from selected species of the Poaceae family. The diagram was drawn based on the transcriptomic data deposited in Phytozome v. 12.1 (https://phytozome.jgi.doe.gov) [49], in Next-Gen Sequence DBs (https://mpss.danforthcenter.org/dbs) [50], in eFP browser (http://bar.utoronto.ca) [51] and from the RNA-seq data from [52]. Expression data were analysed for Poales species from the BOP clade (*B. distachyon*, *O. sativa*) and the PACMAD clade (*P. hallii*, *S. italica*, *S. bicolor*, *Z. mays*) of Poaceae and for a Bromeliaceae (*A. comosus*).

and B isoforms (figure 5). Clearly, URT1A homologues are more closely related to eudicot URT1 than the B homologues (figure 5). In addition, transcriptomic data in *Ananas comosus*, *Zea mays*, *Setaria italica*, *Panicum hallii*, *Sorghum bicolor*, *Oryza sativa japonica* or *Brachypodium distachyon* indicate that URT1A isoforms are constitutively expressed in all tissues, whereas the expression of URT1B isoforms is generally lower and sometimes restricted to specific tissues (figure 6). Both the phylogenetic and transcriptomic studies support the hypothesis that URT1A isoforms are orthologues of URT1 from eudicots, whereas B isoforms may be neo-functionalized. We cannot exclude that certain URT1B isoforms are in the process of pseudogenization. For instance, we failed to amplify fully spliced *URT1B* mRNAs by RT-PCR analysis in *Brachypodium* seedlings, in contrast to *URT1A* mRNAs. Yet, it is possible that *URT1B* mRNAs are effectively spliced only in response to environmental stimuli or at particular developmental stages. Future experimental work is needed to determine whether URT1B homologues are indeed neo-functionalized.

Overall, URT1 homologues seem present in almost all species of Archaeplastida. Except for the early branching Glaucophyta and Rhodophyta species that contain only URT1 homologues, the genomes of the vast majority of green algae and land plants encode both URT1 and HESO1 homologues. URT1 and HESO1 homologues form a monophyletic group for each TUTase. Several species have multiple isoforms of URT1 and HESO1 that sometimes have tissue-specific patterns of expression. It remains to be experimentally investigated which of these isoforms of TUTases have evolved specialized functions, and which are in the process of pseudogenization following gene duplication.

3. Uridylation accelerates the decay of small RNAs

Small RNAs can be tailed by untemplated nucleotides, mostly uridines and adenosines. 3′ adenylation of miRNAs seemed to slow down degradation *in vitro* using *Populus trichocarpa* cell extracts [53], but the role of small RNA adenylation in regulating stability remains to be fully elucidated. By contrast, uridylation triggers the degradation of small RNAs. This process, which was discovered in *Arabidopsis* [54], constitutes the best-documented role of uridylation in plants. Plant small RNAs are mostly 20–24 nt in length and are divided in two main families: microRNAs (miRNAs) and short interfering RNAs (siRNAs). miRNAs are processed from primary transcripts that contain a hairpin with an imperfectly paired stem, while siRNAs are processed from near-perfect double-stranded RNAs (dsRNAs) or fully paired dsRNAs. The fully paired dsRNAs are produced by RNA-dependent RNA polymerase (RDR), which uses the sense strand as a template to generate the dsRNA precursor. miRNA and siRNA precursors are processed by DICER-like (DCLs) enzymes into small RNA duplexes. The 3′ end of each strand of the duplexes is 2′-O-methylated by the methyltransferase HUA ENHANCER1 (HEN1) [55,56]. Only one strand of the duplexes is finally retained in complex with an ARGONAUTE (AGO) protein, while the passenger strand is discarded (reviewed in [57]). Except for rare exceptions mentioned below, virtually all mature small RNAs are thus methylated on their 3′ terminal ribose in plants.

Mutations in HEN1 result in pleiotropic developmental abnormalities in *Arabidopsis* because miRNA levels are drastically reduced [55,56]. Methylation by HEN1 is indeed necessary to protect the 3′ end of small RNAs (both miRNAs and siRNAs) from uridylation, which triggers their decay [54,58]. Besides *Arabidopsis*, mutation of HEN1 orthologues also induces uridylation-mediated destabilization of small RNAs in maize and rice [59,60]. Of note, small RNAs can be uridylated and adenylated in a wild-type context as reported in many species including *Arabidopsis*, tomato, *Medicago truncatula*, rice, maize and in the moss *Physcomitrella patens* [61]. However, untemplated tailing is mostly detected for 'off-size' small RNAs [61]. In addition, 'off-size' 23 nt heterochromatic siRNAs (hc-siRNAs, also called het- or he-siRNAs), as compared to canonical 24 nt hc-siRNAs, are mostly adenylated rather than uridylated, and this tailing is not increased by *hen1* mutation in *Arabidopsis* [61]. A possible explanation is that the untemplated nucleotides are added to hc-siRNAs precursors [61].

During the biogenesis of the vast majority of small RNAs, once small RNA duplexes have been generated by DCLs, the dsRNA binding domains of HEN1 position the 3′ ends of small RNA duplexes in the catalytic site to deposit the methyl group that prevents uridylation. It is worth noting that some miRNAs, like miR158 or miR319a in *Arabidopsis* and miR1510 in Phaseoleae species (e.g. soya bean), are substantially truncated and tailed even in a HEN1 wild-type context, suggesting that some small RNA duplexes could be poor substrates of HEN1 [32,60,62]. However, for the vast majority of small RNAs, the absence of HEN1 results in extensive nibbling and tailing, and the added nucleotides are mostly uridines [20,32,54–56,60,61]. Yet, the patterns of trimming and uridylation are different across miRNA families and specific patterns are conserved for the same miRNA family between maize, rice and *Arabidopsis hen1*

mutants [60]. Therefore, intrinsic features of miRNAs that are conserved between monocotyledons and dicotyledons could determine the extent of nibbling and tailing [60].

The TUTase HESO1 is the major TUTase uridylating both siRNAs and miRNAs to facilitate their decay in *Arabidopsis* [20,32]. Its orthologue in *C. reinhardtii*, MUT68, was previously shown to uridylate small RNAs to trigger their degradation, revealing the conservation of this process from algae to land plants [42]. HEN1 SUPPRESSOR 1 (HESO1) was identified in *Arabidopsis* by a forward genetic screen aiming at identifying suppressors of the *hen1* phenotype, but also by systematically testing which of the T-DNA mutants for the 10 class I ncPAPs of *Arabidopsis* was able to partially rescue the *hen1* phenotype [20,32]. In both studies, a *heso1* mutation in a *hen1* background partially restores miRNA levels and markedly reduces small RNA uridylation [20,32]. Altogether, these data show that HESO1 is the predominant TUTase uridylating small RNAs in plants. Yet, the residual uridylation of small RNAs in the double *heso1 hen1* mutant indicates that TUTases other than HESO1 are able to target small RNAs. Both forward and reverse genetic strategies identified URT1 as a secondary TUTase able to uridylate miRNAs in the *heso1 hen1* background [63,64]. URT1 was previously identified as the major TUTase uridylating deadenylated mRNAs in *Arabidopsis* (see §5) [31]. URT1 is localized in the cytosol, P-bodies and stress granules [31]. Its cytosolic localization likely explains why URT1 does not uridylate nuclear hc-siRNAs in a *heso1 hen1* background, but is restricted to miRNAs [63,64]. Importantly, the residual uridylation of hc-siRNAs in a *hen1 heso1 urt1* background reveals the existence of another TUTase yet to be characterized. In *Arabidopsis*, the two genes that cluster with HESO1 and URT1 in the phylogenetic analysis shown in figure 2 are possible candidates for encoding this additional TUTase activity.

HESO1 and URT1 uridylate miRNAs, but both TUTases act distinctively and cooperatively [63,64]. For instance, HESO1 uridylates full-length miR158 (which is poorly methylated by HEN1) while 1-nt truncated miR158 is mostly uridylated by URT1 [63]. Also, HESO1 has a broad role in uridylating miRNAs, while URT1 action seems restricted to fewer targets, likely explaining why the *urt1* mutation does not rescue the *hen1* phenotype, while the *heso1* mutation does [32,63].

Also, because mono-uridylated miRNAs accumulate in *hen1 heso1* background, it was proposed that URT1 could mono-uridylate unmethylated miRNAs to provide a U-terminating substrate, which is favoured by HESO1. HESO1 would then further tail the small RNA [20,32,57,63,64].

The role of HESO1 in uridylating small RNAs was identified in a *hen1* mutant, and that of URT1 in a *hen1 heso1* background, because both HESO1's and URT1's activities are inhibited by the methyl group deposited by HEN1 on the 3′ terminal ribose of small RNAs. In a wild-type context, tailing occurs mostly on nibbled small RNAs. In *Arabidopsis*, four 3′–5′ exoribonucleases, SMALL RNA DEGRADING NUCLEASES (SDN1 to 4) are responsible for nibbling small RNAs [65,66]. SDNs are only partially inhibited by 2′-O-methylation of small RNAs and they are able to remove the 3′ terminal methylated nucleotide of small RNAs, thereby generating truncated, unprotected substrates for HESO1 and URT1 [66]. In contrast to *heso1* and *urt1* single mutants, or a null double mutant (data not shown), which have no obvious phenotype, combining mutations in three out the four SDNs results in higher miRNA levels and pleiotropic developmental defects [65], indicating that nibbling by

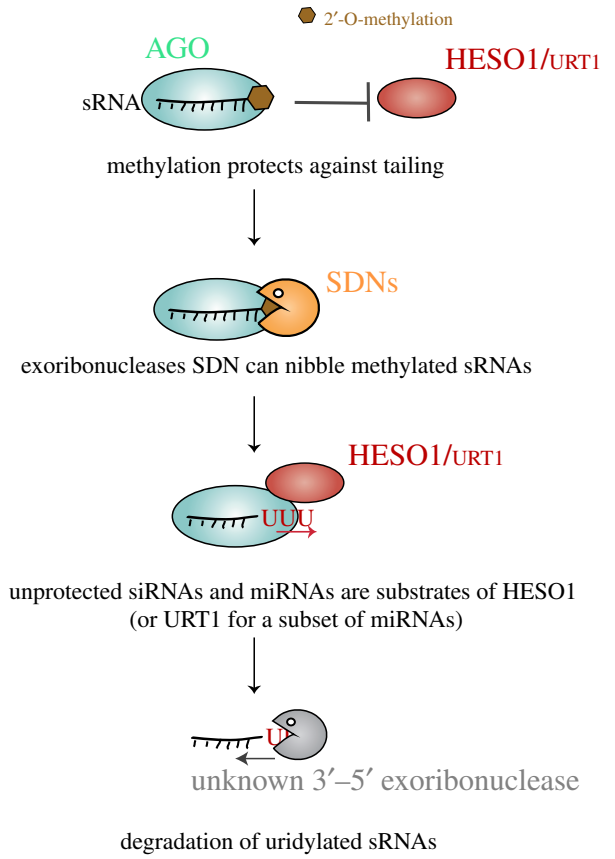


Figure 7. Small RNA uridylation and decay. 2′-O-methylation deposited by the methyltransferase HEN1 protects against uridylation by HESO1 or by URT1. The exoribonucleases' SDNs can nibble methylated sRNAs that are loaded in AGO, thereby generating nibbled, unprotected sRNAs. These unprotected siRNAs or miRNAs are targeted by HESO1 (or URT1 for a subset of miRNAs). The uridylated small RNAs are subsequently degraded by a 3′ to 5′ exoribonucleolytic activity(ies) that is unknown in *Arabidopsis* and was proposed to be RRP6 in *C. reinhardtii* [42].

SDNs is a limiting step in controlling miRNA decay as compared to uridylation.

Both SDNs and TUTases interact with AGO proteins [39,63,66], explaining why nibbling and tailing of miRNAs are AGO1-dependent [60]. In addition, uridylation antagonizes nibbling of small RNAs, likely revealing a competition between TUTases and SDNs to access the 3′ end of small RNAs [64]. Importantly, SDNs are unlikely to degrade uridylated small RNAs [65] and therefore a yet unidentified activity is responsible for the degradation of uridylated small RNAs in *Arabidopsis*. In *C. reinhardtii*, RRP6, a cofactor of the RNA exosome, was proposed to degrade uridylated small RNAs [42]. Our current view of small RNA degradation based mostly on the work in *Arabidopsis* is summarized in the model presented in figure 7.

Besides facilitating small RNA degradation, uridylation has additional roles in small RNA metabolism. The slicer activity of AGO1 is inhibited *in vitro* by the uridylation of miR165/6 in complex with AGO1 [63]. In addition, uridylation of miR158 by URT1 seems to impair its repression activity in a *hen1* mutant [63]. These observations, made so far either *in vitro* or in a *hen1* background, indicate that TUTases have the potential to control miRNA activity in plants. Such a regulatory role was reported in mice for miR-26, whose uridylation abrogates function without affecting stability [67]. A second alternative role of miRNA uridylation

in plants is to control the biogenesis of secondary siRNAs, which are triggered by cleavage of a target by certain 22 nt miRNAs. Indeed, mono-uridylation of miR170/1 to 22 nt isoforms triggers the production of phased, secondary siRNAs (phasiRNAs) in a *hen1* background [60]. Why other 22 nt miRNA isoforms also accumulating in *hen1* do not trigger phasiRNA production is unknown. Interestingly, the control of phasiRNA biogenesis by uridylation was recently identified for miR1510 in Phaseoleae species, which include common bean and soya bean [62]. miR1510 regulates several *nucleotide-binding and leucine-rich repeat protein* (*NB-LRR*) genes by triggering the production of phasiRNAs. In soya bean, and most other Phaseoleae species, the miR1510 duplex has a terminal mispairing that inhibits HEN1 activity [62]. As a result, unmethylated miR1510 is mono-uridylated into a 22 nt species able to trigger phasiRNA production [62]. By recapitulating miR1510 biogenesis in *Arabidopsis*, HESO1 was identified as the TUTase that mono-uridylates miR1510 [62]. This example illustrates that uridylation of miRNAs might evolve functions distinct from merely promoting small RNA degradation.

4. Uridylation of 5′ fragments of mRNAs cleaved by RISC

The repression of gene expression by post-transcriptional gene silencing (PTGS) is achieved either by repressing translation or by inducing mRNA degradation (reviewed in [68]). In plant PTGS, mRNA degradation is initiated by AGO1-mediated cleavage that is guided by sequence complementarity between the small RNA loaded in RISC and its target mRNA (reviewed in [57]).

RISC generates a 5′-cleavage fragment (5′CF) and a 3′-cleavage fragment (3′CF). The 3′CF is degraded by XRN4, the cytosolic 5′–3′ exoribonuclease in plants [69]. The 5′CF is eliminated both by the 5′–3′ and the 3′–5′ RNA degradation pathways. Interestingly, uridylation participates in the clearance of this fragment by stimulating degradation from both its 5′ and its 3′ ends. The addition of uridines to the 3′ end of the 5′ fragment is an evolutionarily conserved mechanism detected in *Arabidopsis*, mice or humans [70,71].

In *Arabidopsis*, HESO1, which acts on siRNAs and miRNAs, was identified as the major TUTase uridylating 5′CF resulting from RISC cleavage [39]. The immunoprecipitation of AGO1 by recombinant HESO1 suggests that uridylation of 5′CF (and small RNAs) may occur in the AGO complex [39]. Importantly, 5′CF for *MYB33* mRNAs accumulate in *heso1* mutants, showing that uridylation by HESO1 promotes the degradation of this fragment [39]. Of note, URT1 is responsible for the residual uridylation of *MYB33* 5′CF although its activity is not required to stimulate the degradation of this 5′CF [40]

Several observations indicate that uridylated 5′CF are degraded by the 5′–3′ pathway. The simultaneous analysis of 5′ and 3′ ends of 5′CF in *Arabidopsis* identified the presence of oligo(U) stretches at the 3′ end and showed a diversity of 5′ end positions for the analysed targets [70]. The authors suggested that uridylated 5′CF can be degraded from their 5′ end, implying that uridylation enhances 5′–3′ decay of the 5′ mRNA fragment produced by RISC. Indeed, uridylated *MYB33* 5′CF are preferentially uncapped in *Arabidopsis* [39]. These uncapped RNAs could be produced either by endoribonucleolytic cleavages or by decapping. This latter possibility is coherent with *in vitro* decapping assays in mammalian cell extracts that revealed uridylation as promoting decapping [72]. Finally, the accumulation of 5′CF of *MYB33* mRNAs in *Arabidopsis xrn4* mutants shows that the 5′–3′ RNA degradation pathway indeed participates to the elimination of the 5′ fragments generated by RISC cleavage [39].

5′CF are also incontestably cleared by the 3′–5′ RNA degradation pathway in plants [73]. Indeed, several 5′CF accumulate in *Arabidopsis ski2*, *ski3* and *ski8* mutants, SKI2/3/8 forming the Ski complex, the activator of the RNA exosome in the cytosol [73]. This observation strongly suggests the involvement of the RNA exosome in degrading 5′CF in *Arabidopsis*, as shown in *Drosophila* [74]. However, this involvement remains to be formally demonstrated using mutants affected in the function of core subunits of the *Arabidopsis* RNA exosome. Interestingly, secondary siRNAs are produced in the absence of the SKI complex for a number of miRNA targets and it was proposed that SKI2 could promote the rapid dissociation of RISC from the target mRNA, thereby restricting the production of transitive siRNAs [73]. In addition to this study in *Arabidopsis*, the RNA exosome was also suggested to participate in the degradation of 5′CF in *C. reinhardtii* [40]. MUT68, the TUTase that uridylates small RNAs in *C. reinhardtii*, was reported to facilitate the degradation of a mRNA targeted by artificial siRNAs [40]. No uridylation was detected at the sites of cleavage by RISC but at that time, a low-depth analysis was performed. In addition, to our knowledge, no endogenous miRNA-mediated cleavage was investigated. It is therefore unknown at present whether MUT68 uridylates 5′CF. Interestingly, oligo(A)-tailing was detected upstream of the siRNA-induced RISC cleavage sites in the *mut68* strain, suggesting that non-canonical polyadenylation tags 5′CF [40]. These oligo(A) tails were proposed to facilitate 3′–5′ degradation by the RNA exosome [40].

Of note, the degradation of RISC cleavage fragments likely promotes the dissociation of RISC from its target, which appears of prime importance for recycling RISC. This recycling involves recently identified 3′–5′ exoribonucleases that interact with AGO1 and AGO10: RICE1 and RICE2 [75]. The inactivation of catalytic residues of homohexameric RICE proteins leads to low levels of miRNAs and accumulation of 5′CF [75]. RICE1/2 likely initiates the degradation of 5′CF, thereby facilitating RISC dissociation and recycling [75].

The current model for the degradation of 5′CF is shown in figure 8. HESO1 and RICEs are recruited to RISC through the interaction with AGO (figure 8). HESO1 catalyses the uridylation of 5′CF, possibly promoting decapping and subsequent degradation by XRN4. Concomitantly or alternatively, RICE1/2 ensures the initiation of 3′–5′ degradation of RISC-associated 5′ fragments by starting to nibble uridylated 3′ ends of 5′CF. RICEs′action promotes RISC dissociation and therefore its recycling, but RICEs unlikely fully degrade the 5′CF. This is ensured by the RNA exosome assisted by the SKI complex (figure 8).

5. Intricate function for uridylation in the decay of plant mRNAs

Uridylation of deadenylated mRNAs is widely conserved among eukaryotes, including plants but excluding *S. cerevisiae* that has lost the capacity to uridylate all RNAs. Studies over the past years in *S. pombe*, *X. laevis*, *Aspergillus nidulans*, *A. thaliana*, *M. musculus*, *Patiria pectinifera* (starfish), *D. melanogaster* or *H. sapiens* have revealed that uridylation must be considered as an integral step in the degradation

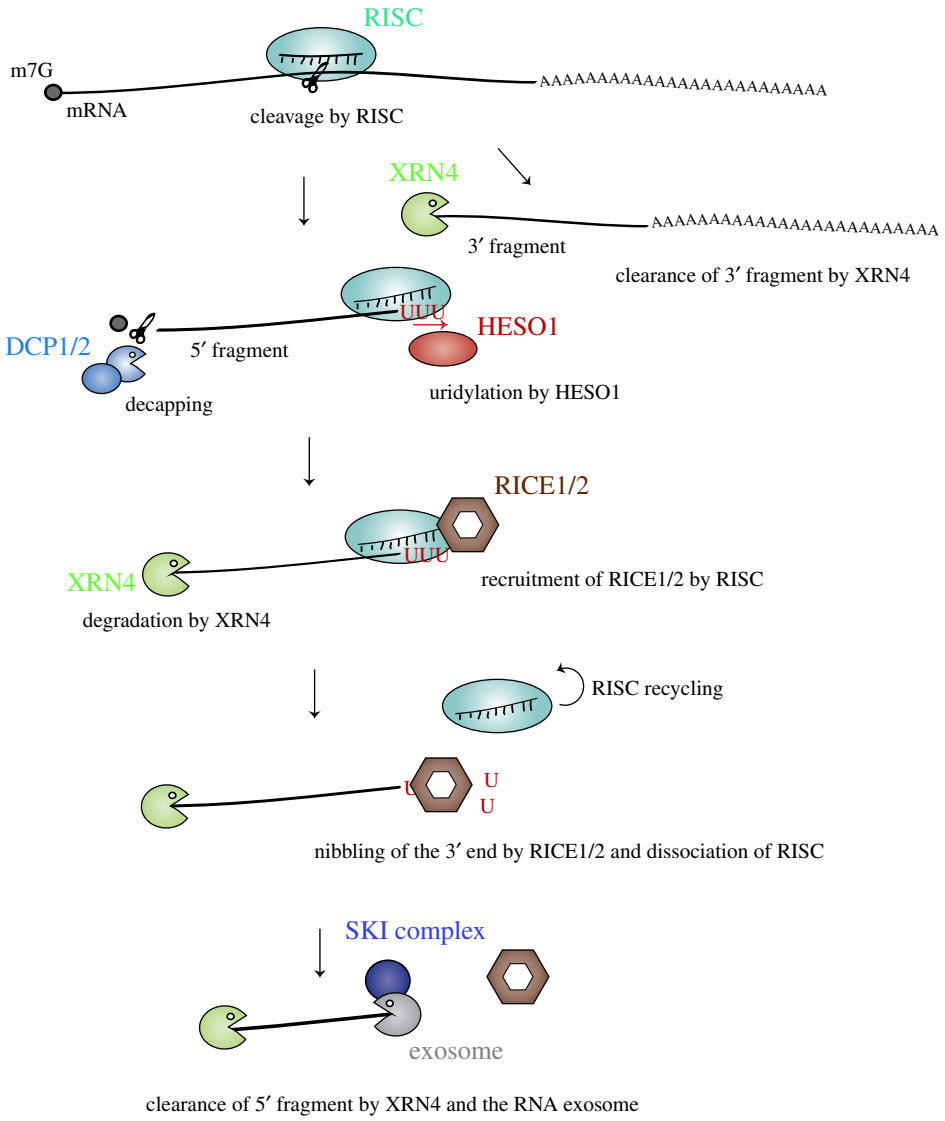


Figure 8. Uridylation by HESO1 promotes degradation of 5′ fragments of RISC-cleaved mRNAs. RISC cleavage of mRNAs generates a 3′ cleavage fragment that is degraded by XRN4, and a 5′ cleavage fragment. The 5′ cleavage fragment is uridylated by HESO1, which binds RISC, but can also be decapped by the DCP1/2 complex. The exoribonucleases RICE1/2, which are recruited by RISC, nibble the uridylated 5′ cleavage fragment. This nibbling helps RISC dissociation and RISC recycling. Finally, the 5′ cleavage fragment is degraded by XRN4 and the RNA exosome.

of mRNAs [3,13,31,76–83]. Tailing oligo(A) tails with a few uridines likely facilitates the binding of LSm1-7 complex, which binds preferentially to short oligo(A) tails and oligo(U) tails [72,84–86]. Binding of the LSm1-7 complex leads to the recruitment of the decapping complex and subsequent degradation by the cytosolic 5′ to 3′ exoribonuclease Xrn1. A similar process could occur in plants: the binding of the LSm1-7 complex could promote the recruitment of the decapping machinery triggering degradation of the uncapped RNA by XRN4. Moreover, U-tails can be directly recognized by Dis3L2 or the RNA exosome to promote 3′ to 5′ decay. Therefore, uridylation influences both the 5′–3′ and 3′–5′ degradation of mRNAs (see accompanying articles by Zigackova and Vanacova [10], Warkocki *et al*. [11] and recent reviews [2,4,5,8]).

In *Arabidopsis*, uridylation of oligo-adenylated mRNAs is mainly performed by URT1 [31,83]. Uridylation in *Arabidopsis* can be detected on uncapped mRNAs, as shown for *CCR2* and *LOM1* mRNAs [31,80] and originally described in *S. pombe* [82]. Yet, no experiment has demonstrated so far an influence of URT1 on global mRNA half-lives, possibly because its direct targets correspond to deadenylated mRNAs, which represent a very minor population among all mRNAs. Also, deadenylation is likely a rate-limiting step in the bulk decay of mRNAs, thereby masking the potential impact of URT1 on mRNA degradation, which is restricted to its targets, i.e. the minor sub-population of deadenylated mRNAs. Although an impact of uridylation in accelerating mRNA decay remains to be shown formally in plants, uridylation definitely has a role in the mRNA degradation process. URT1 prevents excessive deadenylation, as mRNAs with shorter oligo(A) tails accumulate in *urt1* mutants, whereas the complementation and overexpression of URT1 in *Arabidopsis* increase oligo(A) tail sizes [31,83]. Importantly, a global analysis of mRNA uridylation by TAIL-seq revealed that URT1 repairs oligo(A) tails to an average extension length of 16 nucleotides (nt) [83]. A similar sub-population of mRNAs with a oligo(A) tail size distribution centered at 16 nt exists for non-uridylated mRNAs in wild-type plants [83]. Both these uridylated and non-uridylated 16 nt extensions are recognized and bound by a Poly(A) Binding Protein (PABP) *in vitro* and *in vivo* [83]. It is at present unknown in plants whether translation can be initiated on mRNAs with uridylated oligo(A) tails bound by PABP, or whether uridylation would inhibit translation as proposed for reporter mRNAs co-expressed with TUTases in *Xenopus* oocytes [87]

or for Nonsense-Mediated Decay (NMD) targets in *A. nidulans* [80]. Besides a link between uridylation and translation, which remains to be explored in *Arabidopsis*, the recognition of uridylated oligo(A) tails by PABPs could modulate deadenylation [83]. Moreover, terminal uridines *per se* are likely to impede deadenylase activities, thereby contributing to slowing down deadenylation. By impeding deadenylation at the 3′ end and possibly stimulating decapping at the 5′ end as in other eukaryotes, URT1 could favour the 5′–3′ directionality of mRNA degradation. Such 5′–3′ directionality is crucial during co-translational decay and in line with this, uridylated mRNAs were detected on polysomes [31].

mRNA uridylation drops by 70–80% in null *urt1* mutants [31,83]. This shows that URT1 is the main TUTase uridylating mRNAs, but the residual uridylation observed in null *urt1* mutants also reveals the involvement of at least another TUTase [31,83]. HESO1 is a likely candidate as the secondary TUTase able to uridylate mRNAs. This possibility remains to be experimentally demonstrated. The TAIL-seq analysis of *urt1 xrn4* double mutant suggested that this URT1-independent uridylation has a different function in mRNA metabolism. URT1-independent uridylation targets mostly very short oligo(A) tails and, unlike URT1, does not seem able to restore a nucleotide extension of sufficient length allowing PABP binding. Interestingly, in *xrn4* mutant, only these short uridylated oligo(A) tails accumulate compared to WT, suggesting that this population could be targeted by XRN4 and rapidly degraded [83]. Hence, a complex role of uridylation in the metabolism of mRNAs is emerging in *Arabidopsis*. Uridylation by distinct TUTases or of different oligo(A) sizes could favour decapping, impede deadenylation or restore a PABP binding site. Whether mRNA uridylation impacts translation or mRNA storage has to be explored in plants.

6. Conclusion and future key points in plant RNA uridylation

The primary function of RNA uridylation in controlling RNA stability is conserved across eukaryotes, including plants. HESO1 and URT1 homologues are present in most plant species, and their roles in the metabolism of small RNAs and mRNAs could be conserved. Yet, we are just beginning to apprehend the diversity of roles played by RNA uridylation in plants. In addition, our current knowledge of RNA uridylation in plants has been gathered using mostly two model species, the flowering plant Arabidopsis *thaliana* and the green algae *Chlamydomonas reinhardtii*. Therefore, the diversity of specialized biological functions involving RNA uridylation remains to be explored in various plant species. The discovery that mono-uridylation of a miRNA triggers the biogenesis of phased secondary siRNAs in Phaseoleae species to regulate disease resistance genes illustrates the potential of exploring uridylation in diverse plant species [62].

The just-emerging picture drawn from our knowledge in *A. thaliana*, and to a lesser extent in *C. reinhardtii*, is that uridylation targets short and long non-coding RNAs, as well as mRNAs. It is certain that the RNA substrates identified to date represent just a fraction of what is left to discover. For instance, the uridylation of ribosomal or viral RNAs has been reported in plants, with no clues about the impact of U-tailing on these RNAs [22,88]. The systematic identification of RNA substrates of uridylation is a mandatory step towards discovering all functions of RNA uridylation in plants. In addition to the substrates, the identification of all the factors 'reading' the uridylation status of RNA and translating this information into biological outputs is required to decipher the molecular basis for the multiple roles played by uridylation in plant RNA metabolism.

Data accessibility. The datasets supporting this article have been uploaded as part of the electronic supplementary material.

Authors' contributions. D.G. and C.d.A. wrote the paper; H.S.,. C.d.A. and A.G. performed the evolutionary analyses; V.F., F.M. and H.Z. analysed TUTase expression patterns; C.d.A., H.S. and H.Z. prepared illustrations; D.G. acquired funding.

Competing interests. We declare we have no competing interests.

Funding. Work in D.G.'s laboratory is currently supported by the Centre National de la Recherche Scientifique (CNRS, France) and the Agence Nationale de Recherche (ANR, France) as part of the program d'Investissements d'Avenir in the frame of the LabEx NetRNA (ANR-2010-LABX-36) and ANR-15-CE12-0008-01 to D.G.

References

1. Aphasizhev R, Suematsu T, Zhang L, Aphasizheva I. 2016 Constructive edge of uridylation-induced RNA degradation. *RNA Biol.* **13**, 1078–1083. (doi:10.1080/15476286.2016.1229736)
2. De Almeida C, Scheer H, Zuber H, Gagliardi D. 2018 RNA uridylation: a key post-transcriptional modification shaping the coding and non-coding transcriptome. *WIREs RNA* **9**, e1440. (doi:10.1002/wrna.1440)
3. Kwak JE, Wickens M. 2007 A family of poly(U) polymerases. *RNA* **13**, 860–867. (doi:10.1261/rna.514007)
4. Łabno A, Tomecki R, Dziembowski A. 2016 Cytoplasmic RNA decay pathways - enzymes and mechanisms. *Biochim. Biophys. Acta* **1863**, 3125–3147. (doi:10.1016/j.bbamcr.2016.09.023)
5. Munoz-Tello P, Rajappa L, Coquille S, Thore S. 2015 Polyuridylation in eukaryotes: a 3′-end modification regulating RNA life. *BioMed Res. Int.* **2015**, 968127. (doi:10.1155/2015/968127)
6. Norbury CJ. 2010 3′ Uridylation and the regulation of RNA function in the cytoplasm. *Biochem. Soc. Trans.* **38**, 1150–1153. (doi:10.1042/BST0381150)
7. Norbury CJ. 2013 Cytoplasmic RNA: a case of the tail wagging the dog. *Nat. Rev. Mol. Cell Biol.* **14**, 643–653. (doi:10.1038/nrm3645)
8. Scheer H, Zuber H, De Almeida C, Gagliardi D. 2016 Uridylation earmarks mRNAs for degradation... and more. *Trends Genet.* **32**, 607–619. (doi:10.1016/j.tig.2016.08.003)
9. Scott DD, Norbury CJ. 2013 RNA decay via 3′ uridylation. *Biochim. Biophys. Acta* **1829**, 654–665. (doi:10.1016/j.bbagrm.2013.01.009)
10. Zigáčková D, Vaňáčová S. 2018 The role of 3′ end uridylation in RNA metabolism and cellular physiology. *Phil. Trans. R. Soc. B* **373**, 20180171. (doi:10.1098/rstb.2018.0171)
11. Warkocki Z, Liudkovska V, Gewartowska O, Mroczek S, Dziembowski A. 2018 Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism. *Phil. Trans. R. Soc. B* **373**, 20180162. (doi:10.1098/rstb.2018.0162)
12. Łabno A, Warkocki Z, Kuliński T, Krawczyk PS, Bijata K, Tomecki R, Dziembowski A. 2016 Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. *Nucleic Acids Res.* **44**, 10 437–10 453. (doi:10.1093/nar/gkw649)
13. Reimão-Pinto MM, Manzenreither RA, Burkard TR, Sledz P, Jinek M, Mechtler K, Ameres SL. 2016 Molecular basis for cytoplasmic RNA surveillance by

uridylation-triggered decay in *Drosophila*. *EMBO J.* **35**, 2417–2434. (doi:10.15252/embj.201695164)

14. Ustianenko D, Pasulka J, Feketova Z, Bednarik L, Zigackova D, Fortova A, Zavolan M, Vanacova S. 2016 TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. *EMBO J.* **35**, 2179–2191. (doi:10.15252/embj.201694857)
15. Martin G, Keller W. 2007 RNA-specific ribonucleotidyl transferases. *RNA* **13**, 1834–1849. (doi:10.1261/rna.652807)
16. Mroczek S *et al.* 2017 The non-canonical poly(A) polymerase FAM46C acts as an onco-suppressor in multiple myeloma. *Nat. Commun.* **8**, 619. (doi:10.1038/s41467-017-00578-5)
17. Hunt AG *et al.* 2008 *Arabidopsis* mRNA polyadenylation machinery: comprehensive analysis of protein–protein interactions and gene expression profiling. *BMC Genomics* **9**, 220. (doi:10.1186/1471-2164-9-220)
18. Lange H, Sement FM, Canaday J, Gagliardi D. 2009 Polyadenylation-assisted RNA degradation processes in plants. *Trends Plant Sci.* **14**, 497–504. (doi:10.1016/j.tplants.2009.06.007)
19. Meeks LR, Addepalli B, Hunt AG. 2009 Characterization of genes encoding Poly(A) polymerases in plants: evidence for duplication and functional specialization. *PLoS ONE* **4**, e8082. (doi:10.1371/journal.pone.0008082)
20. Ren G, Chen X, Yu B. 2012 Uridylation of miRNAs by hen1 suppressor1 in *Arabidopsis*. *Curr. Biol.* **22**, 695–700. (doi:10.1016/j.cub.2012.02.052)
21. Salinas-Giegé T, Cavaiuolo M, Cognat V, Ubrig E, Remacle C, Duchêne A-M, Vallon O, Maréchal-Drouard L. 2017 Polycytidylation of mitochondrial mRNAs in *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **45**, 12 963–12 973. (doi:10.1093/nar/gkx903)
22. Sikorski PJ, Zuber H, Philippe L, Sement FM, Canaday J, Kufel J, Gagliardi D, Lange H. 2015 Distinct 18S rRNA precursors are targets of the exosome complex, the exoribonuclease RRP6L2 and the terminal nucleotidyltransferase TRL in *Arabidopsis thaliana*. *Plant J.* **83**, 991–1004. (doi:10.1111/tpj.12943)
23. Yang H, Song J, Yue L, Mo X, Song J, Mo B. 2017 Identification and expression profiling of *Oryza sativa* nucleotidyl transferase protein (NTP) genes under various stress conditions. *Gene* **628**, 93–102. (doi:10.1016/j.gene.2017.06.038)
24. Zimmer SL, Schein A, Zipor G, Stern DB, Schuster G. 2009 Polyadenylation in *Arabidopsis* and *Chlamydomonas* organelles: the input of nucleotidyltransferases, poly(A) polymerases and polynucleotide phosphorylase. *Plant J.* **59**, 88–99. (doi:10.1111/j.1365-313X.2009.03853.x)
25. Wilson D, Pethica R, Zhou Y, Talbot C, Vogel C, Madera M, Chothia C, Gough J. 2009 SUPERFAMILY—sophisticated comparative genomics, data mining, visualization and phylogeny. *Nucleic Acids Res.* **37**, D380–D386. (doi:10.1093/nar/gkn762)
26. Schmidt von Braun S, Sabetti A, Hanic-Joyce PJ, Gu J, Schleiff E, Joyce PBM. 2007 Dual targeting of the tRNA nucleotidyltransferase in plants: not just the signal. *J. Exp. Bot.* **58**, 4083–4093. (doi:10.1093/jxb/erm267)
27. Czesnick H, Lenhard M. 2016 Antagonistic control of flowering time by functionally specialized poly(A) polymerases in *Arabidopsis thaliana*. *Plant J.* **88**, 570–583. (doi:10.1111/tpj.13280)
28. Kappel C *et al.* 2015 Genome-wide analysis of PAPS1-dependent polyadenylation identifies novel roles for functionally specialized poly(A) polymerases in *Arabidopsis thaliana*. *PLoS Genet.* **11**, e1005474. (doi:10.1371/journal.pgen.1005474)
29. Trost G, Vi SL, Czesnick H, Lange P, Holton N, Giavalisco P, Zipfel C, Kappel C, Lenhard M. 2014 *Arabidopsis* poly(A) polymerase PAPS1 limits founder-cell recruitment to organ primordia and suppresses the salicylic acid-independent immune response downstream of EDS1/PAD4. *Plant J.* **77**, 688–699. (doi:10.1111/tpj.12421)
30. Vi SL *et al. et al.* 2013 Target specificity among canonical nuclear poly(A) polymerases in plants modulates organ growth and pathogen response. *Proc. Natl Acad. Sci. USA* **110**, 13 994–13 999. (doi:10.1073/pnas.1303967110)
31. Sement FM, Ferrier E, Zuber H, Merret R, Alioua M, Deragon J-M, Bousquet-Antonelli C, Lange H, Gagliardi D. 2013 Uridylation prevents 3′ trimming of oligoadenylated mRNAs. *Nucleic Acids Res.* **41**, 7115–7127. (doi:10.1093/nar/gkt465)
32. Zhao Y, Yu Y, Zhai J, Ramachandran V, Dinh TT, Meyers BC, Mo B, Chen X. 2012 The *Arabidopsis* nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. *Curr. Biol.* **22**, 689–694. (doi:10.1016/j.cub.2012.02.051)
33. LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. 2005 RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell.* **121**, 713–724. (doi:10.1016/j.cell.2005.04.029)
34. Lubas M, Christensen MS, Kristiansen MS, Domanski M, Falkenby LG, Lykke-Andersen S, Andersen JS, Dziembowski A, Jensen TH. 2011 Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell.* **43**, 624–637. (doi:10.1016/j.molcel.2011.06.028)
35. Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W. 2005 A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol.* **3**, e189. (doi:10.1371/journal.pbio.0030189)
36. Wyers F *et al.* 2005 Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725–737. (doi:10.1016/j.cell.2005.04.030)
37. Edgar RC. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797. (doi:10.1093/nar/gkh340)
38. Dereeper A *et al.* 2008 Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465–W469. (doi:10.1093/nar/gkn180)
39. Ren G, Xie M, Zhang S, Vinovskis C, Chen X, Yu B. 2014 Methylation protects microRNAs from an AGO1-associated activity that uridylates 5′ RNA fragments generated by AGO1 cleavage. *Proc. Natl Acad. Sci. USA* **111**, 6365–6370. (doi:10.1073/pnas.1405083111)
40. Zuber H, Scheer H, Joly A-C, Gagliardi D. 2018 Respective contributions of URT1 and HESO1 to the uridylation of 5′ fragments produced from RISC-cleaved mRNAs. *Front Plant Sci.* (doi:10.3389/fpls.2018.01438)
41. Ibrahim F, Rohr J, Jeong W-J, Hesson J, Cerutti H. 2006 Untemplated oligoadenylation promotes degradation of RISC-cleaved transcripts. *Science* **314**, 1893. (doi:10.1126/science.1135268)
42. Ibrahim F, Rymarquis LA, Kim E-J, Becker J, Balassa E, Green PJ, Cerutti H. 2010 Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in *Chlamydomonas*. *Proc. Natl Acad. Sci. USA* **107**, 3906–3911. (doi:10.1073/pnas.0912632107)
43. You C, Cui J, Wang H, Qi X, Kuo L-Y, Ma H, Gao L, Mo B, Chen X. 2017 Conservation and divergence of small RNA pathways and microRNAs in land plants. *Genome Biol.* **18**, 158. (doi:10.1186/s13059-017-1291-2)
44. Stoltzfus A *et al.* 2013 Phylotastic! Making tree-of-life knowledge accessible, reusable and convenient. *BMC Bioinf.* **14**, 158. (doi:10.1186/1471-2105-14-158)
45. Anisimova M, Gascuel O. 2006 Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* **55**, 539–552. (doi:10.1080/10635150600755453)
46. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321. (doi:10.1093/sysbio/syq010)
47. Letunic I, Bork P. 2016 Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245. (doi:10.1093/nar/gkw290)
48. Panchy N, Lehti-Shiu M, Shiu S-H. 2016 Evolution of gene duplication in plants. *Plant Physiol.* **171**, 2294–2316. (doi:10.1104/pp.16.00523)
49. Goodstein DM *et al.* 2012 Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186. (doi:10.1093/nar/gkr944)
50. Nakano M, Nobuta K, Vemaraju K, Tej SS, Skogen JW, Meyers BC. 2006 Plant MPSS databases: signature-based transcriptional resources for analyses of mRNA and small RNA. *Nucleic Acids Res.* **34**, D731–D735. (doi:10.1093/nar/gkj077)
51. Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007 An 'Electronic Fluorescent Pictograph' browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* **2**, e718. (doi:10.1371/journal.pone.0000718)
52. Davidson RM, Gowda M, Moghe G, Lin H, Vaillancourt B, Shiu S-H, Jiang N, Buell CR. 2012 Comparative transcriptomics of three Poaceae

species reveals patterns of gene expression evolution. *Plant J.* **71**, 492–502. (doi:10.1111/j.1365-313X.2012.05005.x)

53. Lu S, Sun Y-H, Chiang VL. 2009 Adenylation of plant miRNAs. *Nucleic Acids Res.* **37**, 1878–1885. (doi:10.1093/nar/gkp031)
54. Li J, Yang Z, Yu B, Liu J, Chen X. 2005 Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in *Arabidopsis*. *Curr. Biol.* **15**, 1501–1507. (doi:10.1016/j.cub.2005.07.029)
55. Yang Z, Ebright YW, Yu B, Chen X. 2006 HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2′ OH of the 3′ terminal nucleotide. *Nucleic Acids Res.* **34**, 667–675. (doi:10.1093/nar/gkj474)
56. Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X. 2005 Methylation as a crucial step in plant microRNA biogenesis. *Science* **307**, 932–935. (doi:10.1126/science.1107130)
57. Yu Y, Jia T, Chen X. 2017 The 'how' and 'where' of plant microRNAs. *New Phytol.* **216**, 1002–1017. (doi:10.1111/nph.14834)
58. Yu B *et al.* 2010 siRNAs compete with miRNAs for methylation by HEN1 in *Arabidopsis*. *Nucleic Acids Res.* **38**, 5844–5850. (doi:10.1093/nar/gkq348)
59. Abe M, Yoshikawa T, Nosaka M, Sakakibara H, Sato Y, Nagato Y, Itoh J. 2010 WAVY LEAF1, an ortholog of *Arabidopsis* HEN1, regulates shoot development by maintaining MicroRNA and trans-acting small interfering RNA accumulation in rice. *Plant Physiol.* **154**, 1335–1346. (doi:10.1104/pp.110.160234)
60. Zhai J *et al. et al*. 2013 Plant microRNAs display differential 3′ truncation and tailing modifications that are ARGONAUTE1 dependent and conserved across species. *Plant Cell* **25**, 2417–2428. (doi:10.1105/tpc.113.114603)
61. Wang F, Johnson NR, Coruh C, Axtell MJ. 2016 Genome-wide analysis of single non-templated nucleotides in plant endogenous siRNAs and miRNAs. *Nucleic Acids Res.* **44**, 7395–7405. (doi:10.1093/nar/gkw457)
62. Fei Q, Yu Y, Liu L, Zhang Y, Baldrich P, Dai Q, Chen X, Meyers B. 2018 Biogenesis of a 22-nt microRNA in Phaseoleae species by precursor-programmed uridylation. *Proc. Natl Acad. Sci. USA* **115**, 8037–8042. (doi:10.1073/pnas.1807403115)
63. Tu B *et al. et al*. 2015 Distinct and cooperative activities of HESO1 and URT1 nucleotidyl transferases in microRNA turnover in *Arabidopsis*. *PLoS Genet.* **11**, e1005119. (doi:10.1371/journal.pgen.1005119)
64. Wang X, Zhang S, Dou Y, Zhang C, Chen X, Yu B, Ren G. 2015 Synergistic and independent actions of multiple terminal nucleotidyl transferases in the 3′ tailing of small RNAs in *Arabidopsis*. *PLoS Genet.* **11**, e1005091. (doi:10.1371/journal.pgen.1005091)
65. Ramachandran V, Chen X. 2008 Degradation of microRNAs by a family of exoribonucleases in *Arabidopsis*. *Science.* **321**, 1490–1492. (doi:10.1126/science.1163728)
66. Yu Y *et al*. 2017 ARGONAUTE10 promotes the degradation of miR165/6 through the SDN1 and SDN2 exonucleases in *Arabidopsis*. *PLoS Biol.* **15**, e2001272. (doi:10.1371/journal.pbio.2001272)
67. Jones MR, Quinton LJ, Blahna MT, Neilson JR, Fu S, Ivanov AR, Wolf DA, Mizgerd JP. 2009 Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat. Cell Biol.* **11**, 1157–1163. (doi:10.1038/ncb1931)
68. Iwakawa H, Tomari Y. 2015 The functions of microRNAs: mRNA decay and translational repression. *Trends Cell Biol.* **25**, 651–665. (doi:10.1016/j.tcb.2015.07.011)
69. Souret FF, Kastenmayer JP, Green PJ. 2004 AtXRN4 degrades mRNA in *Arabidopsis* and its substrates include selected miRNA targets. *Mol. Cell.* **15**, 173–183. (doi:10.1016/j.molcel.2004.06.006)
70. Shen B, Goodman HM. 2004 Uridine addition after microRNA-directed cleavage. *Science.* **306**, 997. (doi:10.1126/science.1103521)
71. Xu K, Lin J, Zandi R, Roth JA, Ji L. 2016 MicroRNA-mediated target mRNA cleavage and 3′-uridylation in human cells. *Sci. Rep.* **6**, 30242. (doi:10.1038/srep30242)
72. Song M-G, Kiledjian M. 2007 3′ Terminal oligo U-tract-mediated stimulation of decapping. *RNA.* **13**, 2356–2365. (doi:10.1261/rna.765807)
73. 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 Res.* **43**, 10 975–10 988. (doi:10.1093/nar/gkv1014)
74. Orban TI, Izaurralde E. 2005 Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA.* **11**, 459–469. (doi:10.1261/rna.7231505)
75. Zhang Z *et al.* 2017 RISC-interacting clearing 3′–5′ exoribonucleases (RICEs) degrade uridylated cleavage fragments to maintain functional RISC in *Arabidopsis thaliana*. *Elife.* **6**, e24466. (doi:10.7554/eLife.24466)
76. Chang H, Lim J, Ha M, Kim VN. 2014 TAIL-seq: genome-wide determination of poly(A) tail length and 3′ end modifications. *Mol. Cell* **53**, 1044–1052. (doi:10.1016/j.molcel.2014.02.007)
77. Lim J, Ha M, Chang H, Kwon SC, Simanshu DK, Patel DJ, Kim VN. 2014 Uridylation by TUT4 and TUT7 marks mRNA for degradation. *Cell.* **159**, 1365–1376. (doi:10.1016/j.cell.2014.10.055)
78. Malecki M, Viegas SC, Carneiro T, Golik P, Dressaire C, Ferreira MG, Arraiano CM. 2013 The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* **32**, 1842–1854. (doi:10.1038/emboj.2013.63)
79. Morozov IY, Jones MG, Razak AA, Rigden DJ, Caddick MX. 2010 CUCU modification of mRNA promotes decapping and transcript degradation in *Aspergillus nidulans*. *Mol. Cell. Biol.* **30**, 460–469. (doi:10.1128/MCB.00997-09)
80. Morozov IY, Jones MG, Gould PD, Crome V, Wilson JB, Hall AJW, Rigden DJ, Caddick MX. 2012 mRNA 3′ tagging is induced by nonsense-mediated decay and promotes ribosome dissociation. *Mol. Cell. Biol.* **32**, 2585–2595. (doi:10.1128/MCB.00316-12)
81. Ochi H, Chiba K. 2016 Hormonal stimulation of starfish oocytes induces partial degradation of the 3′ termini of cyclin B mRNAs with oligo(U) tails, followed by poly(A) elongation. *RNA.* **22**, 822–829. (doi:10.1261/rna.054882.115)
82. Rissland OS, Norbury CJ. 2009 Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. *Nat. Struct. Mol. Biol.* **16**, 616–623. (doi:10.1038/nsmb.1601)
83. Zuber H, Scheer H, Ferrier E, Sement FM, Mercier P, Stupfler B, Gagliardi D. 2016 Uridylation and PABP cooperate to repair mRNA deadenylated ends in *Arabidopsis*. *Cell Rep.* **14**, 2707–2717. (doi:10.1016/j.celrep.2016.02.060)
84. Chowdhury A, Mukhopadhyay J, Tharun S. 2007 The decapping activator Lsm1p-7p–Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA* **13**, 998–1016. (doi:10.1261/rna.502507)
85. Zhou L, Zhou Y, Hang J, Wan R, Lu G, Yan C, Shi Y. 2014 Crystal structure and biochemical analysis of the heptameric Lsm1–7 complex. *Cell Res.* **24**, 497–500. (doi:10.1038/cr.2014.18)
86. Sharif H, Conti E. 2013 Architecture of the Lsm1-7-Pat1 complex: a conserved assembly in eukaryotic mRNA turnover. *Cell Rep.* **5**, 283–291. (doi:10.1016/j.celrep.2013.10.004)
87. Lapointe CP, Wickens M. 2013 The nucleic acid-binding domain and translational repression activity of a *Xenopus* terminal uridylyl transferase. *J. Biol. Chem.* **288**, 20 723–20 733. (doi:10.1074/jbc.M113.455451)
88. Huo Y, Shen J, Wu H, Zhang C, Guo L, Yang J, Li W. 2016 Widespread 3′-end uridylation in eukaryotic RNA viruses. *Sci. Rep.* **6**, e25454. (doi:10.1038/srep25454)

Since the publication of these reviews, several papers about uridylation of coding and non-coding RNAs have been published:

- TUT4/7-mediated uridylation shapes the mouse maternal transcriptome by tagging mRNAs for elimination during oocyte growth (Morgan et al., 2017). Uridylation by TUT4/7 also regulates mRNA levels in pachytene spermatocytes and contributes to the meiotic progression in mammals (Morgan et al., 2019). Therefore, TUT4/7 have a key function in the determination of male and female germlines and are essential for neonatal survival. The latter study also highlights that a common feature of both spermatocyte and maternal TUT4/7-regulated transcripts are AU-rich elements in the 3' UTRs.
- In *C. elegans*, PUP1 (CDE-1), PUP2 and PUP3 uridylation activities are crucial for proper germline development and fertilization (Li and Maine, 2018).
- A new mechanism called Tail-U-Mediated Repression of miRNA basepairing (TUMR) has been proposed by Yang et al., 2019. The paper identifies mRNAs with non-canonical target sites that are regulated by TUT4/7-uridylated miRNAs. Using miR27a as a model, they showed that uridylation enables the correct base-pairing of the miRNA seed sequence to its targets.
- Using biochemical experiments and structure-guided mutagenesis, Kroupova et al., showed that an arginine residue in the active site of Drosophila Tailor is important for its selectivity for RNAs with terminal guanosine and uridines (Kroupova et al., 2019). This clarified how Tailor recognises and uridylates mono-uridylated substrates including mirtron RNAs to promote their degradation by Dis3L2.
- TUT4/7 cooperates with the helicase MOV10 to restrict the retrotransposition of LINE-1 elements in humans (Warkocki et al., 2018b). This study shows that uridylation is a pervasive modification of the LINE-1S mRNA but does not impact its stability. They suggested that uridylation of LINE-1S mRNAs prevents the initiation of retrotranscription.
- Uridylation has a major role in antiviral defense in both *C. elegans* and mammals (Le Pen et al., 2018). This paper showed that CDE-1 mediated uridylation promotes the degradation of the genomic RNA of Orsay virus, a natural pathogen of *C. elegans*. In mammalian cells, TUT4(7) uridylates the influenza A virus mRNA which also contributes to the regulation of its stability.
- Finally, RNA uridylation is the topic of 4 recent reviews. The first review describes the main factors involved in RNA uridylation and summarizes the impact of uridylation on RNA

metabolism on the cellular and organism level (Zigáčková and Vaňáčová, 2018). A second publication focuses on the function of oligouridylation in the metabolism of histone mRNAs in mammals (Meaux et al., 2018). The third publication concentrates on the mechanisms by which human TUTases regulate the metabolism of specific RNAs (Yashiro and Tomita, 2018). The fourth review focuses on human terminal nucleotidyl transferases (TENTs) (Warkocki et al., 2018a).

2. Essential protective elements of eukaryotic mRNAs.

The prime function of mRNA uridylation is to trigger degradation by both 5’-3’ and 3’-5’ decay pathways. Recent discoveries now suggest that uridylation also influences mRNA deadenylation, translatability and storage. To better comprehend the different mechanisms by which uridylation affects mRNA metabolism, I will summarise in the next chapters the key steps of an mRNA’s life cycle.

2.1. Co-transcriptional addition of the 5' cap and 3' polyA-tails.

During transcription, essential protective elements are added to eukaryotic messenger transcripts. These elements are recognised and bound by various protein factors that orchestrate mRNA processing, export into the cytoplasm and translation.

2.1.1. The 5’ cap structure.

The first major modification that occurs on nascent transcripts during transcription by RNA polymerase II (Pol II) is the **addition** of a **cap structure** to the **5’ end of the RNA**. The capping enzyme is specifically recruited to Pol II transcripts through the C-terminal domain CTD of the polymerase (Komarnitsky et al., 2000). The conventional capping reaction consists of 3 subsequent enzymatic reactions. First, the 5’ triphosphate of the nascent mRNA is converted into a 5’ diphosphate by the RNA triphosphatase (TPase). Second, the RNA guanylyltransferase (GTase) uses GTP (guanosine triphosphate) as cofactor to add a guanosine monophosphate group to the 5’ diphosphate of the mRNA via a 5’-5’ triphosphate linkage. Third, the N7 methyltransferase (MTase) adds a methyl group to the N7 amine position of the guanine (Ghosh and Lima, 2010; Ramanathan et al., 2016). The 7-methylguanosine cap (^{m7}GpppRNA) is termed cap0 and is a characteristic feature of all Pol II transcripts including mRNAs. In metazoans and eukaryotic viruses, the cap0 is further methylated at the ribose 2'-*O* position of the first transcribed nucleotide (^{m7}GpppN$_m$pRNA, cap1) and occasionally also at the second nucleotide (^{m7}GpppN$_m$pN$_m$pRNA, cap2) (Furuichi et al., 1975; Wei et al., 1975; Werner et al., 2011). These additional 2'-*O* methylations occur in the nucleus (cap1) and in the cytosol (cap2) and are crucial for the distinction between self- and non-self RNAs, in other words, endogenous and exogenous RNAs (Daffis et al., 2010; Devarkar et al., 2016; Lässig and Hopfner, 2017; Züst et al., 2011). The recognition of foreign RNAs involves cytoplasmic sensor proteins such as MDA5 and RIG-I and leads to an innate immune response that can induce antiviral responses (Daffis et al., 2010;

Decroly et al., 2012; Züst et al., 2011). To avoid that the innate immune response is triggered by endogenous but incompletely capped mRNAs, an efficient quality control mechanism involving decapping proteins and the 5'-3' exoribonuclease DXO eliminates mRNAs with m7G caps that are not 2'-O methylated (Jiao et al., 2013; Picard-Jean et al., 2018).

The m7G cap structure at the 5' extremity is important for several subsequent processes in the mRNA life cycle that are described below in Chapter 2.2.

2.1.2. The 3' polyA tail.

Another crucial modification of mRNAs is the **polyadenylation** of the **3' end** which is catalysed by canonical polyA polymerase (cPAP) following co-transcriptional cleavage of the primary transcript. Recent data suggest that, under some circumstances, polyA tails synthesized by cPAP can promote the degradation of mRNAs retained in the nucleus, similar to non-canonical tails that are added to non-coding RNAs (Tudek et al., 2018). However, canonical poly-A tails are best known for their important roles in the stabilisation, maturation, nuclear export and translation of mRNAs (Dreyfus and Régnier, 2004; Edmonds, 2002).

The maturation of the 3' end of mRNAs has been intensively studied in humans and yeast which led to the identification of multiple factors and elements that are indispensable for the correct formation of mature mRNAs. The 3' ends of the transcribed mRNAs are recognised and cleaved by a conserved polyadenylation complex (Figure 2), named CPF in yeast and CPSF in human. The CPF/CPFS complex is composed of multiple factors that confer RNA binding or endonuclease, protein phosphatase and polyadenylase activities (Tudek et al., 2018). Plant genomes encode for orthologs of all main subunits of the polyadenylation complex (Hunt et al., 2012; Millevoi and Vagner, 2009). Higher plants possess expanded gene families encoding these factors, and tandem affinity purification (TAP) coupled to mass spectrometry has identified additional subunits of the polyadenylation complex in plants (Zhao et al., 2009).

The first step of mRNA 3' maturation is the recognition of a cis-element in the 3'UTR by the polyadenylation complex. In higher animal species, a conserved AAUAAA hexamer sequence is essential for the identification of the cleavage site (Edmonds, 2002). It recruits the polyadenylation complex and dictates the cleavage site thanks to well-conserved distances to U/GU rich downstream sequence elements (DSE) and upstream G-rich elements (Tudek et al., 2018). In yeast and plants, the consensus polyadenylation motifs have a different organization (Hunt, 2008). In these organisms, the AAUAAA element is less dominant as polyadenylation signal. As compared

to human, the enhancing elements (named FUE for far upstream elements) are located further upstream of the cleavage site (Loke, 2005). In *Arabidopsis thaliana*, a uridine-rich FUE is located up to 125 nucleotides upstream of the cleavage site. A "near upstream element" (NUE) that is rich of adenosines and resembles the AAUAAA elements of human mRNA, is located up to 10 nucleotides upstream of the cleavage site. It was suggested that conserved secondary structures present in the region surrounding the cleavage site are essential for correct 3' end maturation, mainly because mutations in the NUE that appear to impact these structures (as judged from RNA secondary structure prediction) have a negative effect on the cleavage of the mRNA (Loke, 2005).

A particularity of higher eukaryotes and especially of plants is the high ratio of alternative polyadenylation (APA). The proportion of mRNA transcripts with APA is 70% in Arabidopsis, 80% in rice and 50% in Chlamydomonas (Shen et al., 2008a, 2008b; Wu et al., 2011). APA often results in the production of mRNA isoforms with different 3'UTRs that contain or lack specific motifs for the binding of proteins or regulatory RNAs. mRNA isoforms produced by APA can differ in their posttranscriptional regulations, their stability or their translation efficiency. Thus, APA is an important layer of control accounting for much of the complexity in the expression of eukaryotic genomes. This notion is further emphasized by the observation that APA occurs specifically in selected tissues and at defined developmental stages (Lutz and Moreira, 2011). Of note, recent studies in Arabidopsis have shown that APA sites are also present in introns, in the coding sequence and even in the 5'UTR (Guo et al., 2016; Zhu and Vaughn, 2018).

After cleavage at the polyadenylation site, the canonical polyA polymerase cPAP adds a long stretch of adenosines to the 3' extremity of the freshly transcribed mRNA (Fig. I2). The newly synthetized polyA tail is instantly bound by polyA binding proteins (PABP) that have several functions in the biogenesis of mature mRNA and that actively regulate the extent of the polyadenylation by cPAP (detailed below in Chapter 2.2.2) (Kühn et al., 2009).

2.2. The cap binding complex and polyA binding proteins.

Throughout its life cycle, mRNA transcripts interact with a plethora of protein factors, making it in fact a messenger ribonucleoprotein mRNP. The formation, composition and dynamics of mRNPs are crucial for mRNA maturation, transport, translation, and stability. The m7G cap and polyA tail elements play an essential role in the formation of mRNPs because they represent landing platforms for factors implicated in all kind of mRNA processes. In the following, I will discuss the

two main types of co-factors that are bound to Pol II transcripts throughout their lifetime, the cap binding complexes (CBC) and PolyA binding proteins (PABP).

2.2.1. The nuclear cap binding complex CBC.

As detailed above, mRNA 5' capping occurs co-transcriptionally and results in the addition of a **m7G cap** to the nascent mRNA extremity. In yeast, animals and plants, the m7G cap is recognised by the **nuclear cap binding complex nCBC**, a heterodimer composed of the cap binding proteins CBP20 and CBP80 (Daszkowska-Golec, 2018). CBP20 directly binds the m7G cap, whereas CBP80 stabilises this interaction (Fortes et al., 2015; Kierzkowski et al., 2009).

The binding of the nCBC protects the 5' end against decapping enzymes, e.g. DXO, and positively stimulates the cleavage at the polyadenylation site and mRNA 3' end maturation (Flaherty et al., 1997). Moreover, nCBC is important for transcription regulation and several downstream events in mRNA maturation. For example, the binding of yeast nCBC to certain nascent mRNAs promotes the formation of the pre-initiation complex (PIC) at the respective promoter regions (Lahudkar et al., 2011). nCBC has been shown to promote pre-mRNA splicing by interacting with spliceosomal small nucleolar RNPs in mammals (Izaurralde et al., 2007; Pabis et al., 2013). In Arabidopsis and mammals, nCBC can also bind the multifunctional "RNA effector protein" SERRATE (SE), or ARS2 in human. nCBC-SE/nCBC-ARS2 determines transcript fate by interacting with splicing, RNA export or RNA degradation machineries. For example, nCBC and SE interacts with LUC7 to control alternative splicing (Amorim et al., 2018; Laubinger et al., 2008; Raczynska et al., 2014). nCBC and SERRATE have also important roles in both splicing and processing of capped primary miRNAs in plants (Kim et al., 2008; Laubinger et al., 2008). nCBC-SE /ARS is also crucial for the correct nuclear export of mRNA into the cytoplasm. nCBC interacts directly with ALY (or REF, Yra1 in yeast), a component of the TREX complex, a mRNA export machinery conserved in yeast, Arabidopsis, Drosophila and humans. TREX is composed of the THO core complex that interacts with DEAD-box RNA helicases and mRNA export adaptors (Cheng et al., 2006; Chi et al., 2013; Nojima et al., 2007; Rodrigues et al., 2012; Sen et al., 2019). Interestingly, the binding of the export factor ALY to human nCBC-ARS2 is mutually exclusive to the binding of nCBC-ARS2 to the RNA helicase hMTR4 which promotes RNA degradation by nuclear RNA surveillance (Schulze et al., 2018). The interaction between nCBC and SE with both RNA export and RNA degradation factors is also conserved in plants (Lange et al., 2014; Raczynska et al., 2014).

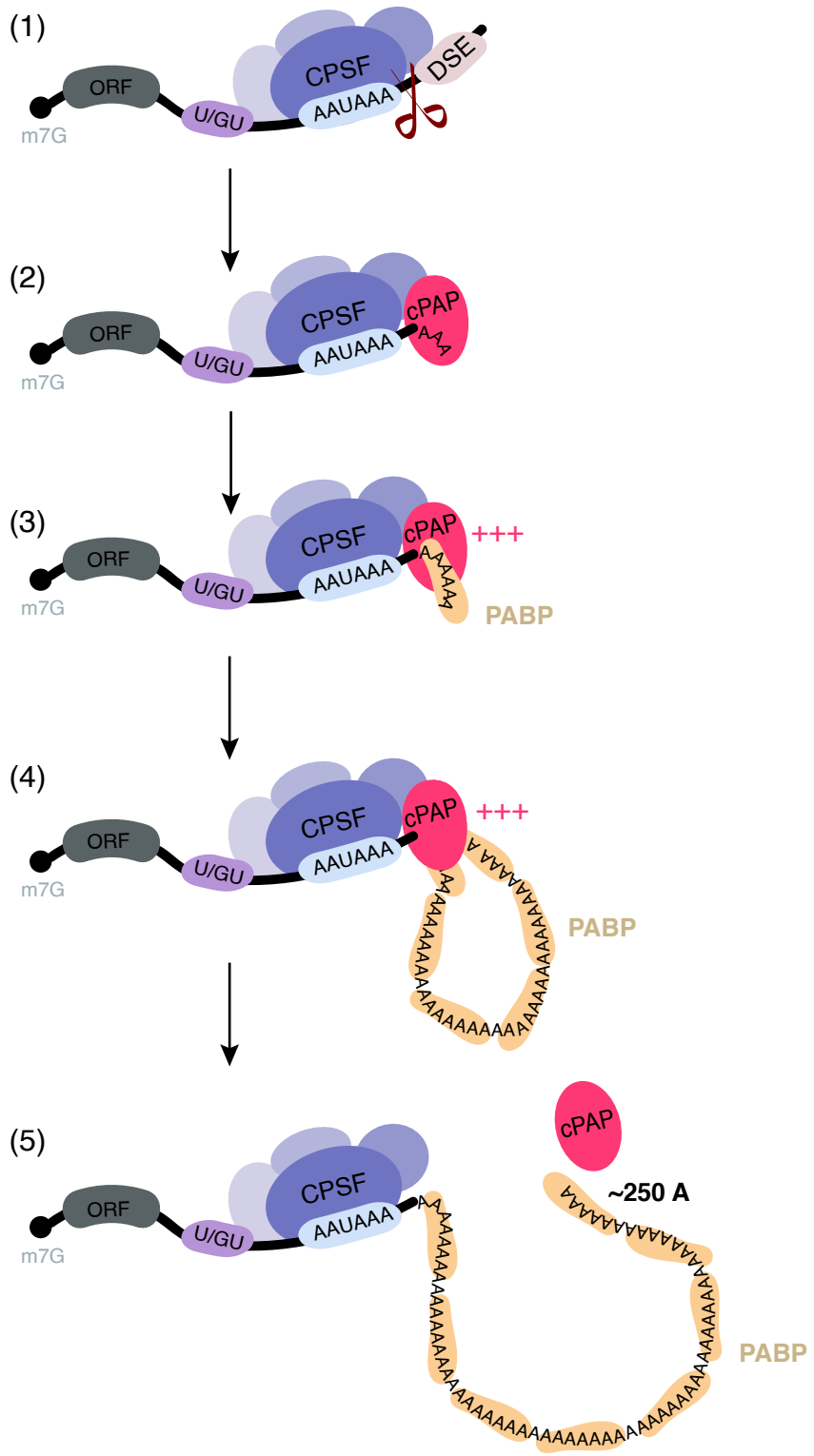


Figure 2: Polyadenylation of mRNAs in humans.
In humans, the polyadenylation site comprises an upstream U or GU rich element, the conserved AAUAAA hexamer sequence, and a downstream sequence element DSE. **(1)** The cleavage and polyadenylation specificity factor (CPSF) binds the polyadenylation site and cleaves the pre-mRNA at the cleavage site that is dictated by well conserved distances between the sequence elements. **(2)** Canonical PolyA polymerase cPAP is recruited by CPSF to the cleavage site and oligoadenylates the 3' ends, initially with low affinity. **(3)** The nascent oligoA tail is bound by nuclear polyA binding protein PABPN1, which also directly interacts with cPAP. **(4)** The stable interactions between CPSF-cPAP-PABPN1 positively stimulates the activity of cPAP, leading to the production of long polyA tails which bind more PABPs.
(5) The polyA tail bound to PABP is suggested to form a spherical structure, that changes when the polyA tail size exceeds *circa* 250 nucleotides. PABP can no longer support the interaction between CPSF and cPAP, leading to the dissociation of the cPAP from the mRNA.
m7G: 7-methylguanosine cap; ORF: open reading frame. Adapted from Kühn et al., 2009.

Once exported to the cytoplasm, nCBC is replaced by the cytoplasmic m7G cap binding factor eIF4F. Similar to nCBC, eIF4F protects mRNA 5' ends against decapping and interacts with various binding partners to promote translation initiation (more information in chapter 2.2.5).

2.2.2. Nuclear polyA binding proteins PABPN.

Similar to the m7G cap, the polyA tail is coated by proteins that have several roles in mRNA maturation, export, translation and stability. All eukaryotes possess both nuclear and cytoplasmic **polyA binding proteins PABP**. In human and yeast, two non-orthologous nuclear PABP (PABPN) bind and regulate the synthesis of the polyA tail (Hector et al., 2002; Kaufmann et al., 2004; Kühn and Wahle, 2004; Kühn et al., 2009; Wahle, 1991). During polyA tail synthesis, PABPN binds nascent polyA tails of a minimal length of 10-12 nucleotides (Wahle, 1991). Additional PABPNs are bound as the tail gets longer. In plants, nuclear PABPs remained largely uncharacterized and only nuclear PABPN1 of *Citrus sinensis* has been definitely shown to bind mRNA polyA tails (Domingues et al., 2015).

Human PABPN1 binds oligo-adenosines through an RRM domain in the C-terminal part of the protein (Kerwitz et al., 2003). PABPN1 binding has several functions: It prevents the binding of the polyadenylation complex to the polyA tail thereby restricting it to the AAUAAA sequence, it facilitates the interaction between cPAP and the polyadenylation complex, and it increases the affinity of cPAP for RNA (Kühn et al., 2009). When cPAP is first recruited by the polyadenylation complex CPSF it has only a low affinity for RNA 3' end (Figure 2). The interaction of cPAP with both CPSF and PABPNs increases the affinity of the cPAP for RNA (Kerwitz et al., 2003) and permits the synthesis of long tails in a processive mode. *In vitro* synthesized polyA tails reach a length of *circa* 250 nucleotides (*or* 60 nucleotides in yeast). The current hypothesis is that the PABP-polyA tail adopts a spherical structure that can accommodate a certain number of PABPNs. This structure is disrupted when tail length exceeds 250 adenosines and PABPN1 can no longer support the interaction between CPSF and cPAP, and results in the release of the cPAP from the polyA tail (Cheng and Gallie, 2013; Kühn et al., 2009). Therefore, PABPN1 both measures and defines polyA tail length.

The primary functions of PABPNs is the protection of the newly synthetized polyA tails from 3'-5' exoribonucleases. Yet, the yeast PABPN Nab2p has been shown to interact with the exoribonuclease Rrp6p (Browning and Bailey-Serres, 2015). Rrp6 is one of the enzymatically active subunits of the nuclear exosome (Jackson et al., 2010). The exosome is a major 3'-5' RNA

degradation machine present in both nucleus and cytoplasm of all eukaryotic cells (see Chapter 4.1). The observation that PABPN can recruit RRP6 and the nuclear exosome to polyA tails suggests that PABPN is, similar to nCBC, also important for nuclear mRNA surveillance (Schmid et al., 2012).

Upon export to the cytoplasm, polyA tails are bound by cytoplasmic PABPs (PABPC). PABPCs play a major role in translation initiation (see Chapter 2.2.4). However, a prime function of PABPCs is to prevent the premature shortening of the polyA tail. Yet, PABPCs also regulate the recruitment of polyA specific nucleases. The role of PABPs in controlling RNA stability is discussed in Chapter 3 of the introduction.

2.2.3. The cytoplasmic cap binding complex eIF4F.

As mentioned above, both the cytoplasmic cap-binding complex eIF4F and PABPCs bound to 5' and 3' ends of mRNAs, respectively, have crucial roles for the translation initiation.

Typically, the nuclear cap binding complex is replaced by the cytoplasmic cap binding complex once the mRNA is exported to the cytosol. However, in mammals, spliced mRNAs still bound to nuclear CBC can undergo a pioneer round of translation (Chiu et al., 2004; Lejeune et al., 2004; Oh et al., 2007). This first round of translation is linked to mRNA export and allows rapid responses in case of environmental or physiological changes. The dissociation of nCBC involves the binding of importin β, and eventually also importin α. Importins recognise proteins with nuclear localisation signals and promote their import into the nucleus (Cautain et al., 2015). The binding of importin β changes the affinity of nCBC for its mRNA substrate and leads to its dissociation from the m7G cap (Dias et al., 2009). The m7G cap is subsequently bound by eIF4F (eukaryotic initiation factor 4F), a heterotrimeric protein complex composed of eIF4A, eIF4E and eIF4G.

The eIF4F complex serves as a platform for the assembly of other translation initiation factors and the recruitment of the pre-initiation complex (PIC). The subunit that directly binds the m7G cap is eIF4E, whereas the large multidomain protein eIF4G is responsible for PIC recruitment. Some plants possess a second eIFiso4F complex with analogous eIFiso4E and eIFiso4G subunits. eIFiso4G lacks the N-terminal region of eIF4G and is conserved in all Viridiplantae. By contrast, eIFiso4E is restricted to angiosperms (Mayberry et al., 2011; Patrick and Browning, 2012). Both eIF4F and eIFiso4F cap-binding complexes initiate translation *in vitro*, but show different affinities for specific groups of mRNAs (Gallie and Browning, 2001). Moreover, the isoform of the scaffold protein eIFiso4G seems to be essential for plant growth in Arabidopsis (Lellis et al., 2010). However, the function of these isoforms has not been fully clarified yet.

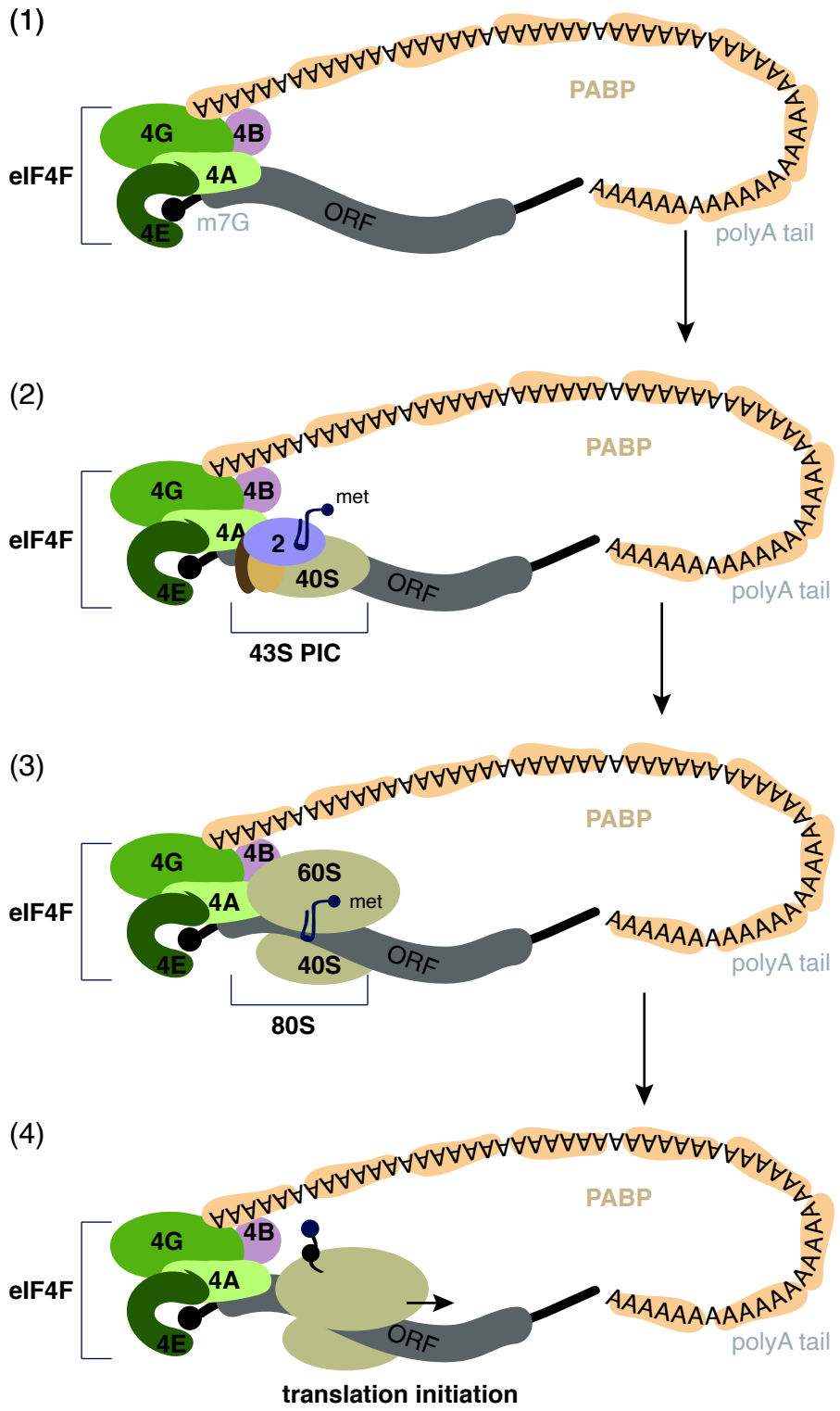


Figure 3: Simplified scheme of translation initiation in eukaryotes.
(1) The m7G cap at the 5' extremity of the mRNA is bound by the eIF4F complex composed of eIF4E, eIF4A and eIF4G. eIF4F and eIF4B interact with the PABPs bound to the 3' polyA tail, leading to the formation of a ternary complex polyA tail-PABP-eIF4F that results in the circularisation of the mRNA. **(2)** The 43S pre-initiation complex (40S small ribosome subunit, eIF2 initiation factor and methionyl initiator MET-tRNA$_i^{MET}$) is recruited to the ternary complex. **(3)** 43S PIC scans the 5' UTR until reaching the AUG initiation codon. Dissociation of eIF2 allows the binding of the 60S ribosome subunit and formation of the 80S initiation complex. **(4)** eIF factors are released and mRNA translation is initiated. eIF4: eucaryotic initiation factors 4; met: methionine; ORF: open reading frame.
Adapted from Jackson et al., 2010

2.2.4. Cytoplasmic polyA binding proteins (PABPC).

Cytoplasmic PABPs (PABPC) are conserved among eukaryotes. As compared to yeast, the gene family expanded in metazoans and plants (Goss and Kleiman, 2013; Mangus et al., 2003). *Arabidopsis thaliana* has 8 PABPCs that can be sorted into 3 classes based on their sequence and expression patterns (Belostotsky, 2003; Gallie, 2017; Gallie and Liu, 2014). Class I is formed by PABP1, PABP3 and PABP5 that are expressed in reproductive tissues. Class II PABP2, PABP4 and PABP8 are highly expressed and are suggested to have general and mostly redundant functions in plant growth and development (Belostotsky, 2003; Gallie, 2017). Class III comprises the weakly expressed proteins PABP6 and PABP7. PABPCs possess 4 highly conserved tandemly repeated RRM motifs. With its four RRMs, a single PABPC binds typically 23-27 adenosines (Baer and Kornberg, 1980; Livak and Schmittgen, 2001; Smith et al., 1997). However, a PABPC can already efficiently bind polyA tails of only 12 nucleotides with its RRM1 and RRM2 domains (Baer and Kornberg, 1980; Deo et al., 1999; Kühn and Pieler, 1996; Sachs et al., 1987). In addition to the RRMs, PABPC contain a C-terminal region with a proline-rich unstructured sequence linker and a α-helical peptide-binding domain. The C-terminal linker sequence is required for the cooperative binding of multiple PABPC to the polyA tail. The cooperative binding is prompted by the RRM2 motifs that contact the proline rich region of the linker sequence of the neighbouring PABPC (Melo et al., 2003; Sawazaki et al., 2018). The PABC at the N-terminus contains the conserved amino acids MLLE that form the recognition site for proteins with a PAM2 (polyA-binding protein interacting motif 2) (Sawazaki et al., 2018; Xie et al., 2014). Interestingly, PABPC can also bind to 3'UTRs notably on AU rich sequences such as polyadenylation signals, mostly through RRM3 and RRM4 that have less specificity for homopolymeric polyA (Kini et al., 2016; Sladic et al., 2004; Webster et al., 2018a). A recent study using *in vitro* deadenylation assays has shown that each RRM of yeast Pab1 protects 8 nucleotides of the polyA tail and that, as the polyA tail is shortened, Pab1 can bind the 3'UTR (Webster et al., 2018b).

2.2.5. eIF4F and PABPC promote translation initiation.

Both eIF4F and PABPC are necessary for translation initiation (Figure 3). Once bound to the m7G cap, eIF4F recruits other initiation factors to the 5' end of mRNAs. For instance, eIF4B directly binds the eIF4G scaffold protein and assists the ATP-dependent helicase eIF4A of the complex to unwind the secondary structures at the 5' end (Gallie, 2014). Furthermore, PABPC interact with eIF4B and eIF4G of the cap binding complex (Cheng and Gallie, 2013). Previous studies have shown that the binding of multiple PABPC to the polyA tail and the interaction of PABPC and eIFG4

promote the formation of a stable ternary complex and leads to the circularisation of the mRNA. They suggest that the formation of this ternary complex is necessary for the formation of the 43S pre-initiation complex (PIC) (Tarun and Sachs, 1995; Wells et al., 1998). Of note, recent studies revisit this model and show that pre-translational mRNPs have a linear organisation and rarely show interacting 5'-3' ends (Adivarahan et al., 2018; Pierron and Weil, 2018; Rajyaguru and Parker, 2012). The PIC is formed by the 40S small ribosome subunit and initiation factors including eIF2, which interacts with the methionyl initiator tRNA MET-$tRNA_i^{MET}$ (reviewed in (Browning and Bailey-Serres, 2015)). The 43S PIC scans the 5'UTR from 5' to 3' until it reaches the AUG initiation codon. The hydrolysis of GTP-eIF2 leads to a change in the conformation and the formation of the 48S complex that is recognized by the 60S subunit of the ribosome (Jackson et al., 2010). The eIF factors are released from the PIC and the so-formed 80S initiation complex proceeds to the translation of the mRNA into proteins.

Hence, the collaboration of eIF4F and PABPC is essential to assure the translation initiation of mRNAs. In addition, both factors are important to control the balance between mRNA translation and mRNA decay, which is discussed in the next Chapter.

3. Deadenylation, the rate-limiting step of bulk mRNA decay.

During their lifetime, mRNAs are bound by many proteins that regulate their maturation and function. Following maturation and export, mRNAs are classically bound by the translation machinery to produce proteins. The regulation of translation participates to the fine tuning of gene expression and is done at various steps including translation initiation and elongation. mRNA degradation is, *inter alia*, an efficient means to control mRNA steady-state level and regulate protein production in response to internal and external stimuli. I will describe the steps and key factors of the main cytoplasmic mRNA degradation pathways and discuss the relation between decay and translation efficiency.

Deadenylation has been described throughout the literature as the initiation step of mRNA degradation for most transcripts. Typically, deadenylation proceeds in two distinct phases, the initial trimming phase that removes the first adenosines of the polyA tail, and the mRNA turnover phase, in which further shortening of the polyA tails leads to the gradual dissociation of the PABPs from the polyA tail and to the detachment of the PABPs from the eIF4F complex bound to the 5' m7G cap. These processes expose the mRNA extremities to the decapping machinery and to 5' to 3' and 3' to 5' exoribonucleolytic pathways. The progressive removal of the terminal adenosines of the polyA tail is performed by several enzymes that differ in their catalytic properties and their subcellular localisations, and are each recruited and stimulated by specific factors (Yan, 2014). Four deadenylation enzymes have been described in eukaryotes: polyA ribonuclease (PARN), Nocturnin (NOC), the polyA-specific ribonuclease complex PAN2/3 and the CCR4 (carbon catabolite repressor 4)/NOT (Negative on TATA) complex (CCR4/NOT).

3.1. The polyA ribonuclease PARN.

PARN is a 3'-5' exoribonuclease of the DEDD superfamily which is characterised by the presence of three aspartic acids (D) and one glutamic acid (E) as conserved catalytic residues in the activity domains (Goldstrohm and Wickens, 2008; Zuo, 2001). PARN was initially described in mammalian cells and is conserved in many eukaryotes, with the notable exceptions of *Saccharomyces cerevisiae* and *Drosophila melanogaster* (Parker and Song, 2004). PARN mostly forms a homodimer and recognises polyA tails through its RRM domain (Nilsson et al., 2007; Wu et al., 2005). A particularity of PARN is that it can directly interact with the m7G cap structure of the

mRNA. This interaction stimulates the processivity of PARN, whereas the presence of cap binding complexes or PABPs inhibit its activity (Balatsos et al., 2006; Dehlin et al., 2000; Gao et al., 2000; Körner and Wahle, 1997; Martînez et al., 2001). The current model is that PARN targets capped mRNA with accessible m7G extremities not protected by any cap binding complex and with long polyA tails that have a low PABP occupancy.

PARN has diverse biological functions (Godwin et al., 2013; Virtanen et al., 2013; Wu et al., 2005; Yan, 2014). Early studies indicated a key role for PARN during early development of Xenopus oocytes (Copeland and WORMINGTON, 2001; Körner et al., 1998). PARN was also suggested to deadenylate transcripts undergoing non-sense mediated decay (NMD) (Lejeune et al., 2003) or mRNA with AU-rich elements (ARE) (Lai et al., 2003). Furthermore, PARN has been shown to deadenylate mRNAs involved in the DNA damage response (DDR) and mRNAs targeted by the miRNA-induced silencing pathway (Cevher et al., 2010; Zhang et al., 2015). Recent studies show that PARN is highly expressed in cancer cells and is activated by tumour suppressors (Maragozidis et al., 2015; Shukla et al., 2019; Zhang et al., 2013). Of note, PARN contributes also to the processing and stability of several non-coding RNAs in mammalian cells, including small Cajal body RNAs (scaRNAs) and small nucleolar RNAs (snoRNA), human telomerase RNA, Piwi-interacting RNAs (piRNAs) and Y RNAs (Berndt et al., 2012; Izumi et al., 2016; Moon et al., 2015; Nguyen et al., 2015; Shukla and Parker, 2017; Shukla et al., 2016; Son et al., 2018; Tang et al., 2016; Tseng et al., 2018). Additionally, PARN regulates the stability of diverse miRNAs in human cells by trimming their oligoadenylated 3' ends, thereby preventing the recognition and rapid degradation by DIS3L2 (Shukla et al., 2019).

In *A. thaliana*, PARN is essential for early development and has been proposed to target a subset of embryo-specific mRNAs (Chiba et al., 2004; Reverdatto et al., 2004). The expression of Arabidopsis PARN is upregulated upon abscisic and salicylic acid treatment suggesting that PARN is implicated in the response to abiotic stresses such as drought and high salinity (Nishimura et al., 2005).

In humans, Xenopus and plants, PARN shuttles between the nucleus and the cytoplasm (Godwin et al., 2013). In Arabidopsis, the subcellular localisation of PARN has first been shown to be in the cytoplasm and nucleus (Moreno et al., 2013). However, Arabidopsis PARN also has a well conserved mitochondria-targeting sequence (MTS) at its N-terminus (Hirayama et al., 2013). Indeed, when fused to a C-terminal GFP tag, PARN is located in mitochondria and regulates in

cooperation with the polyA polymerase AGS1 the polyadenylation status of mitochondrial mRNAs (Hirayama, 2014).

3.2. CCR4-like Nocturnin (NOC).

The deadenylase nocturnin belongs to the EEP superfamily and its nuclease domain shares similarities with the nuclease domain of CCR4 (Godwin et al., 2013). The deadenylase activity of Xenopus and mouse NOC has been confirmed *in vitro*. However, the leucine rich repeats (LRR) that are present at the N-terminal part of CCR4 and mediate the interaction of CCR4 and CAF in the CCR4-NOT complex are absent in NOC. Yet, a study suggested that NOC associates to the CCR4-NOT complex in Drosophila (Temme et al., 2010).

The particularity of NOC, or CCRN4L in humans and mouse, is that its expression is controlled by the circadian clock in metazoans. Because maximal expression levels are reached in the night, the protein was named nocturnin (Green and Besharse, 2002; Li et al., 2008). NOC is expressed in many different tissues throughout various organisms (Baggs and Green, 2003; Garbarino-Pico et al., 2007; Godwin et al., 2013) and is a key regulator of many metabolic and developmental processes (Godwin et al., 2013; Hughes et al., 2018; Müller et al., 2013; Stubblefield et al., 2012, 2018). The plant orthologue of NOC, HESPERIN (HESP), has been described only recently and regulates the steady-state levels of mRNAs encoding circadian clock proteins. In addition, HESP is implicated in the regulation of plant growth and development in response to oxidative stress (Delis et al., 2015).

3.3. PolyA specific ribonuclease complex PAN2/3.

PAB-dependent polyA ribonucleases (PAN) have been first described in yeast (Boeck et al., 1996; Brown and Sachs, 1998; Sachs and Deardorff, 1992). PAN proteins are recruited to mRNAs by interacting with PABPs which stimulates PAN nuclease activity (Boeck et al., 1996; Mangus et al., 2004). The active form of PAN is a heterotrimeric complex of the DEDD-family exoribonuclease PAN2 and two PAN3 subunits (Schäfer et al., 2019; Uchida et al., 2004; Wolf et al., 2014). PAN3 possesses a short sequence motif, PAM2, that is recognised by the C-terminal MLLE of PABP (Siddiqui et al., 2007). In addition, a recent study showed that PAN2/3 recognises the interfaces formed by PABP oligomerisation and that its deadenylation efficiency depends on the number of such interfaces that are present on a polyA tail (Schäfer et al., 2019).

PAN2/3 is not essential and is probably at least partially redundant with the CCR4/NOT complex (Parker and Song, 2004; Wahle and Winkler, 2013; Wolf and Passmore, 2014). In most eukaryotes, the predominant role of PAN2/3 is to initiate the shortening of the polyA tail (Beilharz and Preiss, 2007; Brown and Sachs, 1998; Mangus et al., 2004; Schäfer et al., 2019). A recent study has now shown that the inactivation of PAN2/3 in human cell lines leads to a slight but reproducible increase of polyA tails longer than 150 nucleotides, but has globally no impact on the mean polyA tail length and no considerable impact on the mRNA steady-state level (Yi et al., 2018). By contrast, depletion of subunits of the CCR4/NOT complex leads in general to a drastic change in mRNA tail length and an increase of the steady-state level (Parker and Song, 2004; Wahle and Winkler, 2013; Wolf and Passmore, 2014). These data support the current hypothesis that CCR4/NOT is the main deadenylase in eukaryotic cells and impacts the vast majority of the mRNA population. Of note, PAN2 is conserved in Chlorophytes and Bryophytes, but has no homologue in higher plants (Pavlopoulou et al., 2013)

3.4. The main deadenylase complex CCR4/NOT.

The main deadenylation activity is provided by the multi-subunit protein complex CCR4/NOT. CCR4/NOT consists of a well conserved core complex with two catalytic subunits, CCR4 and CAF1, which belong to the EEP and DEDD superfamilies, respectively. CCR4 (Carbon Catabolite Repressor 4) was originally identified in a genetic screen for factors involved in the glucose repression of yeast alcohol dehydrogenase 2 (ADH2) (Denis, 1984). The NOT (Negative On TATA-less) genes were initially identified in a reporter screen based on promoters that lack canonical TATA boxes (Collart and Struhl, 1994). The CCR4/NOT complex acts at various steps of the mRNA life cycle and is a central regulator of gene expression (Chapat and Corbo, 2014; Collart, 2016; Collart and Panasenko, 2012; Maillet et al., 2000; Miller and Reese, 2012; Shirai et al., 2014). CCR4/NOT's biological functions are important for embryogenesis and development, germline maintenance, cell proliferation and immune responses in yeast, mammals and Xenopus (Chapat and Corbo, 2014). In yeast, NOT1 is an essential protein and the CCR4/NOT complex is required for viability (Maillet et al., 2000).

3.4.1. Architecture of the CCR4/NOT complex.

The architecture of CCR4/NOT complexes have been studied in Drosophila, humans and yeast (Bhaskar et al., 2015; Maillet et al., 2000; Villanyi and Collart, 2016). An overall structure has been proposed from low-resolution electron microscopy images (Nasertorabi et al., 2011). The human

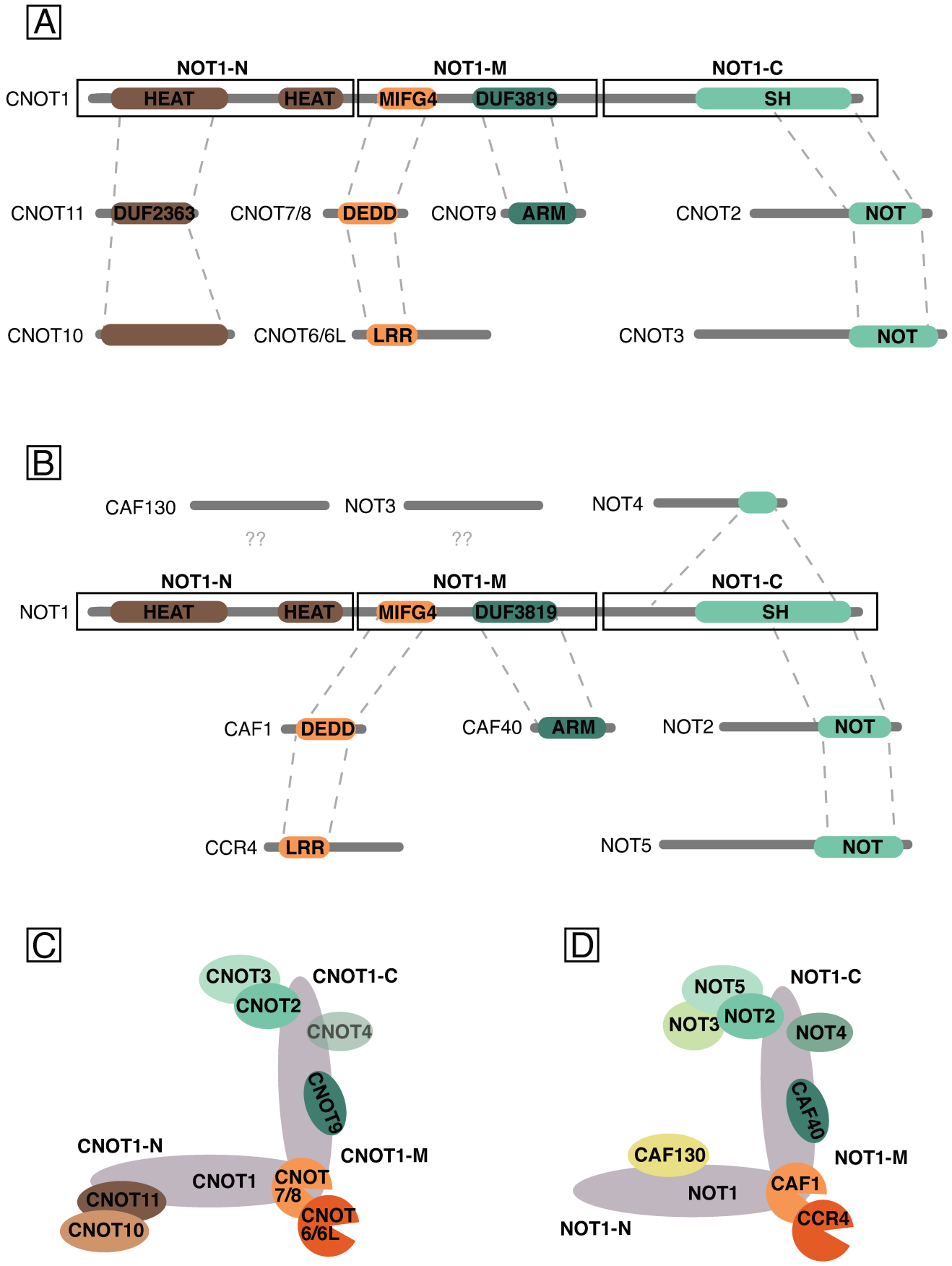


Figure 4: The CCR4/NOT core subunits in human and yeast.
A) and B) Linear representations of the interaction map of CCR4/NOT core subunits in humans **(A)** and yeast **(B)**. **C) and D)** Cartoon representations of the predicted L-shaped CCR4/NOT complex based on the low resolution electron microscopy images by Nasertorabi et al., 2011 in humans **(C)** and yeast **(D)**.
Adapted from Collart, 2016 and Xu et al., 2014.

CCR4/NOT core complex is composed of 8 CNOT proteins (Figure 4A and C). The central scaffold protein CNOT1 is a large protein with multiple domains that mediate its interactions with the CAF1 deadenylases (CNOT7/8 in human) and with other NOT proteins (Xu et al., 2014). CNOT10 and CNOT11 are recruited through an interaction between CNOT11s domain of unknown function DUF2363 and a segment in the N-terminal part of CNOT1 (NOT1-C), probably the conserved HEAT domain (Bawankar et al., 2013; Mauxion et al., 2013). CNOT9 has a conserved armadillo repeat domain that bind the DUF3819 of CNOT1. CNOT3 and CNOT2 are recruited to CNOT1 by a direct interaction between the NOT box domain of CNOT2 with NOT1s HEAT-like repeat domain SH (Bhaskar et al., 2015; Boland et al., 2013; Chen et al., 2014; Mathys et al., 2014). Finally, the CAF1 homologues CNOT7/8 are recruited via the MIF4G domains present in NOT1. CCR4 proteins (CNOT6/CNOT6L in human) are recruited via their interaction with CAF1, mediated by a conserved leucine rich repeat LRR at the N-terminal part of CCR4 and the DEDD domain of CAF1 proteins (Figure 4A and C) (Bai et al., 1999; Draper et al., 2015). All core NOT proteins have homologues in Drosophila whose CCR4/NOT complex architecture is similar to the human complex. By contrast, *Saccharomyces cerevisiae* lacks homologs of CNOT10 and CNOT11 (see Table S1 in supplementary data for an overview of the core CCR4/NOT subunits and Figure 4B and D). Instead, the yeast CCR4/NOT complex contains an additional CAF130 protein that has no homologue in humans, and a NOT4 protein that interacts with the C-terminal part of NOT1 (NOT1-C) (Bhaskar et al., 2015; Chen et al., 2001). Of note, a NOT4 homologue is conserved in humans, but is apparently not stably associated with the CCR4/NOT complex. Finally, yeast has 2 orthologues of CNOT3, NOT3 and NOT5, the latter of which is recruited by interacting with NOT2. How NOT3 is recruited to yeast CCR/NOT remains unclear to date (Figure 4B and D). Finally, yeast CCR4 is recruited by CAF1 similar to the situation in human (Basquin et al., 2012; Petit et al., 2012).

3.4.2. Functions of the CCR4/NOT complex.

The functions of yeast and mammal CCR4/NOT complexes have been intensively studied in the last decade. CCR4/NOT is present in both the nucleus and the cytoplasm, and subunits of the core complex have been detected associated to polysomes and in P-bodies (Halter et al., 2014; Standart and Weil, 2018; Teixeira and Parker, 2007). CCR4/NOT acts as regulation platform of gene expression and is implicated in many different pathways (Chapat and Corbo, 2014; Collart, 2016; Collart and Panasenko, 2012; Maillet et al., 2000; Miller and Reese, 2012; Shirai et al., 2014). I will shortly summarise the main roles of the CCR4/NOT complex of the nucleus and cytoplasm before explaining in more detail how CCR4/NOT acts on mRNA polyA tails.

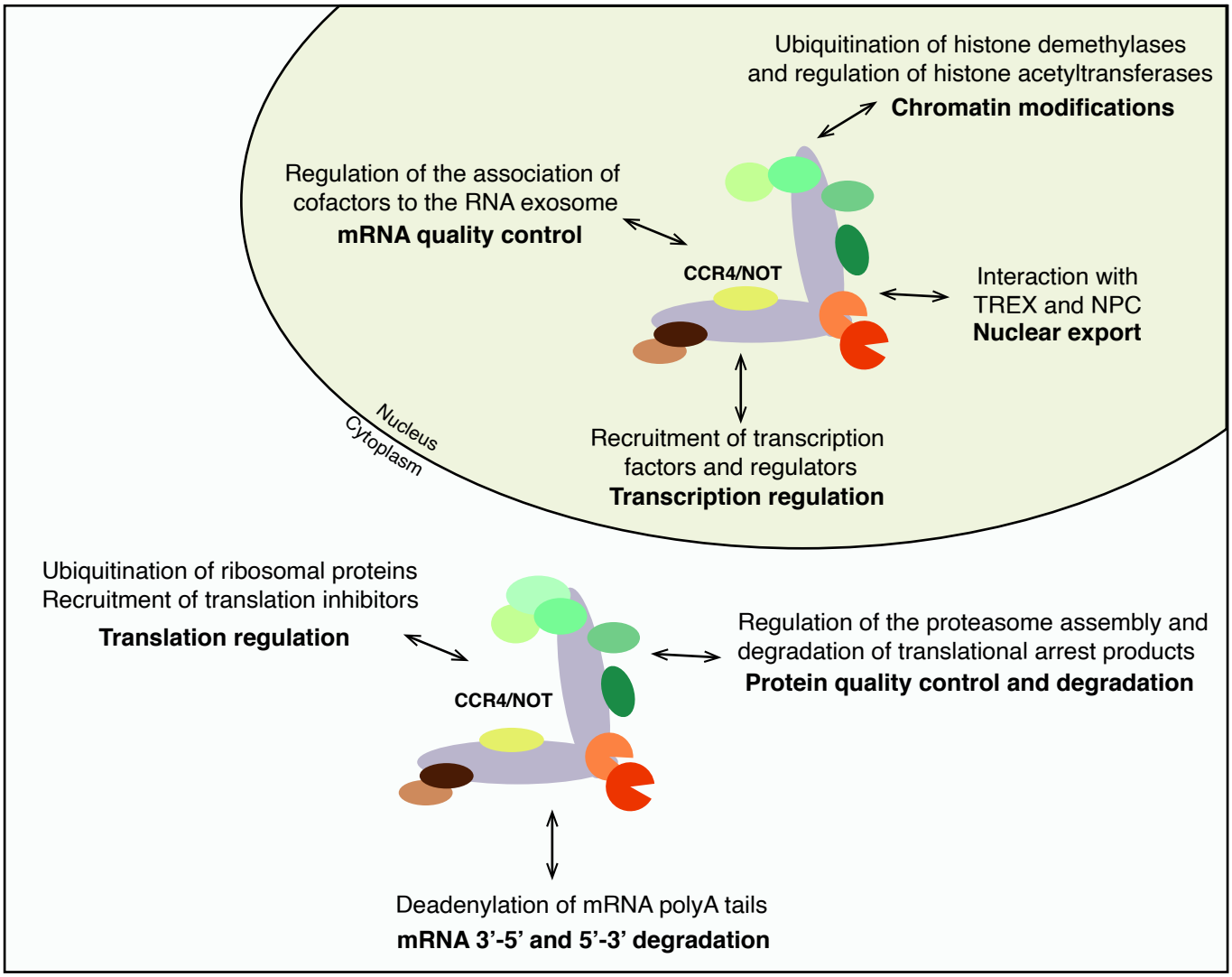


Figure 5: Functions associated to the CCR4/NOT complexes in nucleus and cytoplasm.
Reviewed in Maillet et al., 2000; Collart and Panasenko, 2012; Miller and Reese, 2012; Chapat and Corbo, 2014; Shirai et al., 2014; Martine A Collart, 2016.

In the nucleus, CCR4/NOT regulates chromatin modifications though the stimulation of histone acetyltransferases. In addition, the NOT4 subunit is a RING E3 ligase implicated in the ubiquitination of histone demethylases (Figure 5). Nuclear CCR4/NOT fine-tunes transcription initiation and elongation by recruiting transcription factors and regulators to the mRNA promoter regions. CCR4/NOT also interacts with subunits of the mRNA export machinery TREX and the nuclear pore complex (NPC) suggesting an involvement in the export of mature mRNAs to the cytoplasm (Figure 5). Interestingly, nuclear CCR4/NOT has been shown to influence the association of the RNA helicase Mtr4 and exoribonuclease Rrp6 to the exosome core complex, the main 3'-5' RNA degradation machinery in the nucleus. These data indicate that CCR4/NOT also participates to the elimination of polyadenylated transcripts by nuclear RNA surveillance.

In the cytoplasm, CCR4/NOT regulates the assembly of the proteasome protein degradation machinery by recruiting chaperones to the proteasome. The RING E3 ligase domain NOT4 was suggested to ubiquitinylate truncated proteins produced by translational arrest and mark ribosomal proteins for degradation by the proteasome.

CCR4/NOT can also interact with the translation machinery. The translation termination factors eRF1/3 triggers the deadenylation of mRNAs by recruiting CCR4/NOT (Funakoshi et al., 2007), and CCR4/NOT can associate with translation inhibitors such as DDX6/Dhh1 or the GYF protein GIGYF2 (Amaya Ramirez et al., 2018). Dhh1 directly binds to ribosomes and marks mRNAs with low codon optimality and slow ribosome elongation for translation inhibition and decay (Radhakrishnan et al., 2016). Finally, the catalytic subunits CCR4 and CAF1 are deadenylases that shorten mRNA polyA tails, which ultimately leads to the degradation of the mRNA (Figure 5). The main biological functions of deadenylation by CCR4/NOT include bulk mRNA degradation, miRNA-mediated repression and translational repression during translational initiation. Moreover, mRNA deadenylation is tightly connected to translation.

3.4.3. Recruitment of CCR4/NOT to mRNA.

Besides a putative RRM domain in NOT4, the CCR4/NOT complex has no defined RNA binding motifs in their subunits. The recruitment of the CCR4/NOT complex to specific mRNAs involves RNA binding proteins RBP such as Pumilio (PUF) and Tristetraprolin (TTP) proteins that recognise specific sequence elements (Webster et al., 2018a). PUF proteins bind to Pumilio-response elements (PRE) (Figure 6), while TTP proteins recognise ARE elements in the 3’ UTR. TTP proteins directly interact with NOT1 and NOT9. Recruitment of CCR4/NOT is also promoted by covalent

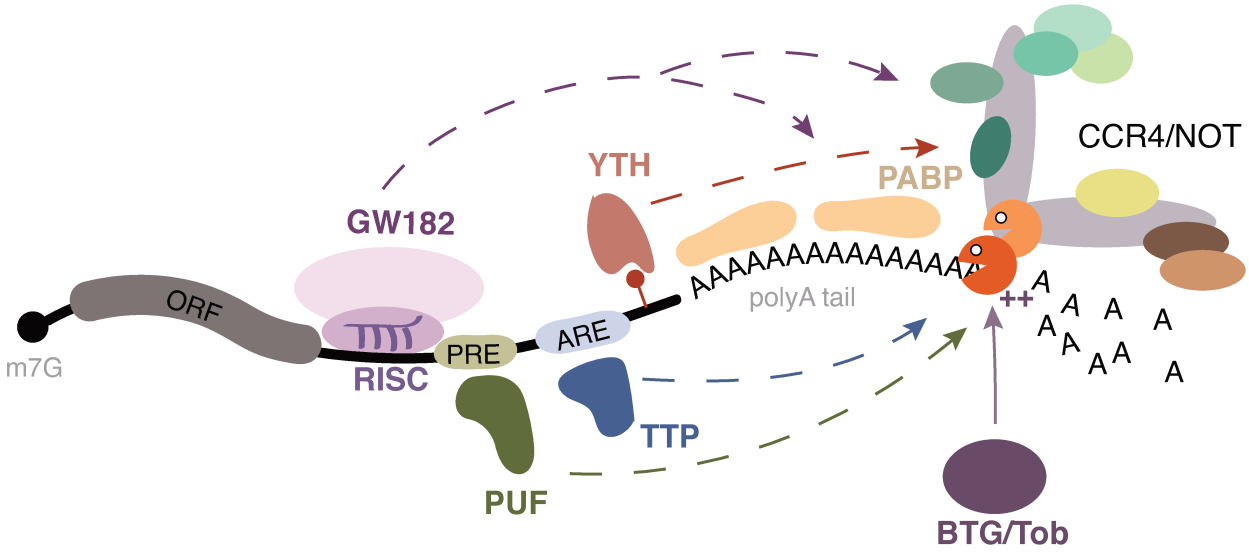


Figure 6: Recruitment of CCR4/NOT by specific RNA binding proteins to the mRNA 3' ends.
CCR4/NOT can unspecifically interact with PABPs bound to polyA tails. In addition, CCR4/NOT can be recruited by m6A readers YTH proteins, and by sequence-specific RNA binding proteins such as tristetraprolin (TTP) proteins or Pumilio (PUF) proteins. Following assembly of a RISC complex guided by a miRNA or siRNA, CCR4/NOT can be recruited by GW182 (glycine-tryptophan repeats). CCR4/NOTs deadenylase activity is stimulated by members of the BTG/Tob protein family that interact with both PABP proteins and CAF1 deadenylases.
PRE: Pumilio-response elements; ARE: AU-rich elements; m7G: 7-methylguanosine cap
Adapted from Shirai et al., 2014

RNA modifications. For instance, YTHDF2 proteins that read m6A modifications can directly bind to NOT1 which triggers the degradation of the mRNA (Du et al., 2016). Both CCR4/NOT and PAN2/3 deadenylases can be recruited by GW182 proteins incorporated in RNA silencing complexes RISC (Behm-Ansmant et al., 2006; Braun et al., 2013; Fabian et al., 2011). Finally, members of the mammalian BTG/Tob protein family that associate to both PABPs and cytoplasmic polyadenylation element-binding CPEB proteins recruit CCR4/NOT by direct interaction with the CAF1 deadenylases (Mauxion et al., 2008; Ogami et al., 2014; Stupfler et al., 2016).

In addition to their role in protecting polyA tails from precocious degradation, PABPs have also key functions in monitoring mRNA deadenylation and degradation. The role of PABPs in controlling deadenylation was revealed by showing that mRNA polyA tails in cells depleted for PABP are longer as compared to wild type cells (Caponigro and Parker, 1995). Subsequent work showed that PABPs can directly interact with PAN2/3 (Boeck et al., 1996; Schäfer et al., 2019; Siddiqui et al., 2007; Uchida et al., 2004). Pull down assays indicated that PABPs can also interact with the CCR4 subunit (Webster et al., 2018b). In addition, the presence of PABPs on the polyA tails positively stimulates deadenylation by CCR4/NOT in yeast and in human cell lines (Webster et al., 2018b; Yi et al., 2018). These results implicate that PABPs can recruit deadenylases to polyA tails, which triggers their deadenylation independently of specific sequence element within the body of the mRNA.

3.4.4. Deadenylation by the CCR4/NOT complex.

Using a biochemical system with the fully reconstituted CCR4/NOT complex of *S. pombe*, Lori Passmore's research group has elucidated how CCR4 and CAF1 deadenylases act on polyA tails (Webster et al., 2018b). Their data indicate that deadenylation proceeds in a stepwise manner. Moreover, the deadenylation rate drops at defined regularly spaced intervals of 8 nucleotides. This periodicity is dictated by PABPs that are bound to the polyA tail stretches. The PABP footprints are defined by the individual RRMs, each of which binds *circa* 8 nucleotides of the polyA tail (Burd et al., 1991; Kühn and Pieler, 1996; Nietfeld et al., 1990). RRM1, RRM2 and RRM3 bind specifically to homopolymeric polyA stretches, whereas RRM4 is less specific and can also bind to 3'UTRs to stabilise the PABP binding (Webster et al., 2018b). Notably, CCR4 can deadenylate polyA tails with or without PABPs at similar rates. Thus, CCR4 can initiate the shortening of long polyA tails bound by PABP, which successively displaces the PABPs from the polyA tail (Figure 7). Webster et al. propose that the 8 nucleotide footprints in their deadenylation assays are caused by ducking of CCR4 into each RRM domain of the PABPs, which slows down deadenylation until the RRM is fully

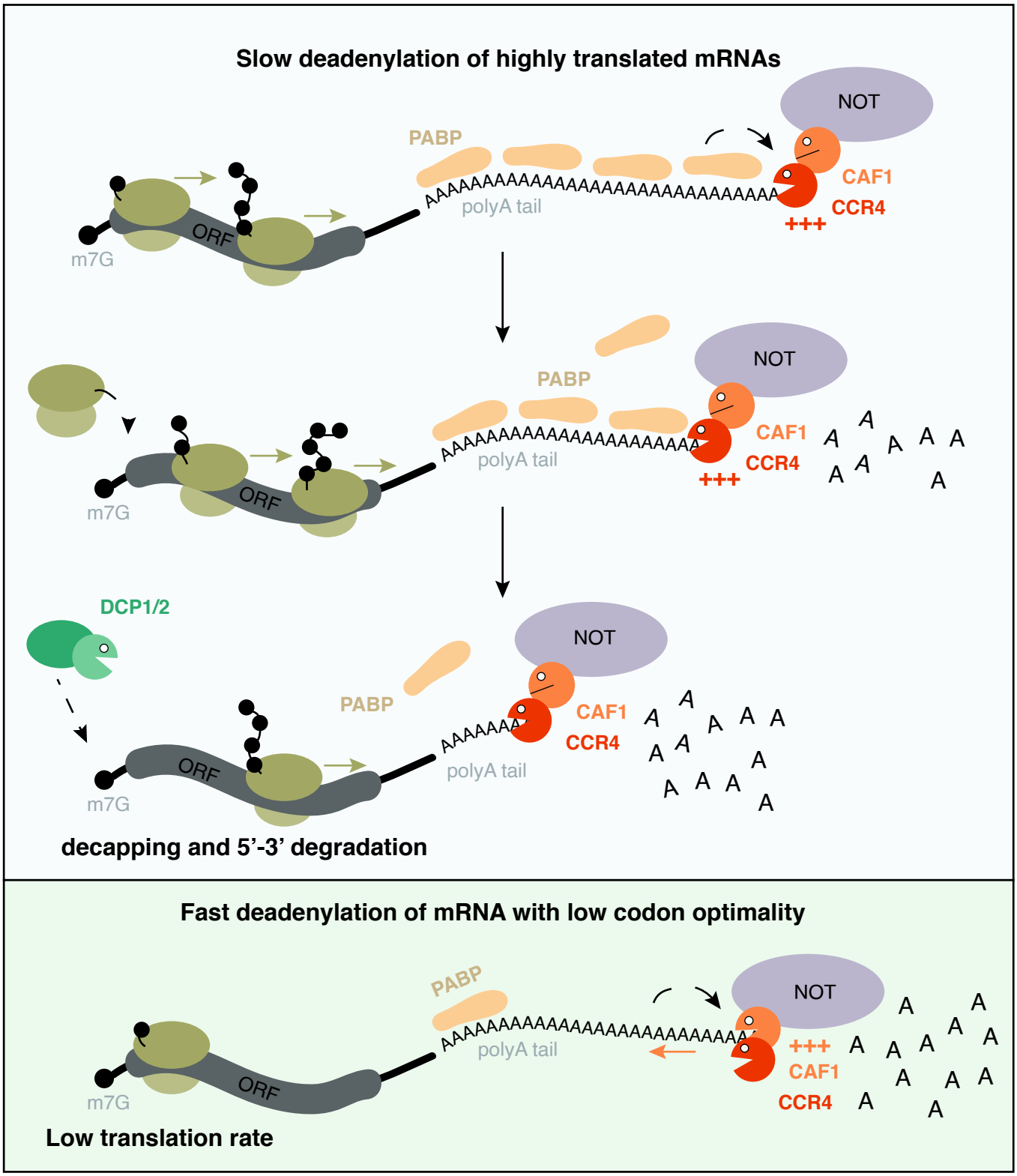


Figure 7: Model of mRNA deadenylation by CCR4/NOT.
Upper panel: Codon-optimised mRNAs tend to have polyA tails that are highly covered by PABPs. CAF1 is inhibited by the presence of PABPs, but CCR4 can deadenylate polyA tails bound by PABPs. It slowly deadenylates polyA tails and displaces the PABPs by stacking into each of the 4 RRM domains (not illustrated). As polyA tails are shortened, translation initiation rates drop and the 5' ends become accessible for the decapping factors DCP1/2.
Lower panel: mRNAs with non-optimal codon composition are poorly translated and have a polyA tail that is less bound by PABPs. CAF1 deadenylase preferentially deadenylates polyA tails that are unbound by PABPs. The concerted actions of CAF1 and CCR4 leads to the fast shortening of the polyA tail of mRNAs that have a low translation rate, leading eventually to the rapid degradation of the mRNA.
DCP1/2: Decapping complex; m7G: 7-methylguanosine cap; ORF: open reading frame.
Adapter from Webster et al., 2018

released from the polyA tail. They further suggested that deadenylation pushes the last remaining PABPs into the 3'UTR which delays the complete dissociation of the last PABP. Finally, Webster et al. showed that the last PABP remains bound until the last RRM domain is displaced, suggesting that 8 adenosines are sufficient for PABP binding (Webster et al., 2018b).

In contrast to CCR4, the activity of CAF1 is inhibited by PABPs bound to the polyA tail. An interesting observation is that the PABP occupancy correlates with the codon optimality of the upstream open reading frame (ORF) (Webster et al., 2018b). Poor codon optimality leads to slow translation rates and to poor PABP coverage (Hanson and Coller, 2018). As CAF1 degrades polyA tails poorly covered by PABPs faster than tails with high PABPs occupancy, CAF1 selectively accelerates the deadenylation of non-optimal mRNAs that are poorly translated, which contributes to the globally shorter half-lives of such mRNAs (Webster et al., 2017) (Figure 7). Highly translated mRNAs with high PABP occupancy on the contrary are degraded essentially by CCR4, that deadenylates its substrates slowly.

Importantly, CCR4/NOT recruitment and mRNA deadenylation does not inevitably induce immediate degradation. New high-throughput sequencing methods without a polyA-based purification step enabled the sequencing of mRNA populations of all polyA tail sizes and polyA tail size determination. An important outcome is that stable mRNAs can have short polyA tails (<30 nucleotides), and that most mRNAs have shorter polyA tails than previously expected (Choi and Hagedorn, 2003; Meijer et al., 2007). The median tail length of polyA tails is of 50-100 adenosines in mammals and *C. elegans*, and around 30 adenosines in yeast (Castellano and Bazzini, 2017; Chang et al., 2014; Lima et al., 2017; Nicholson and Pasquinelli, 2019; Subtelny et al., 2014). Considering that the initial size of newly synthesized polyA tails is estimated to 250 nucleotides in mammals and 90 nucleotides in yeast, the new data suggest that most mRNAs undergo deadenylation in the cytoplasm, regardless of their stability. Moreover, a recent study showed that, across eukaryotes, mRNAs encoding highly expressed and well-translated genes have shorter polyA tails (< 70As) than less translated mRNAs (Lima et al., 2017). A negative effect on translation is only seen for mRNAs with excessively shortened polyA tails of less than 20 nucleotides, at least in somatic HeLa cells (Park et al., 2016). Yet, this relation between polyA tail length and translation could indeed depend of the cellular context. For instance, mRNAs with longer tails are more efficiently translated in frog and zebrafish embryos, while this correlation is lost at later developmental stages (Subtelny et al., 2014). Moreover, cytoplasmic polyadenylation of particular mRNAs has been shown to reactivate translation (Lim et al., 2016; Subtelny et al., 2014).

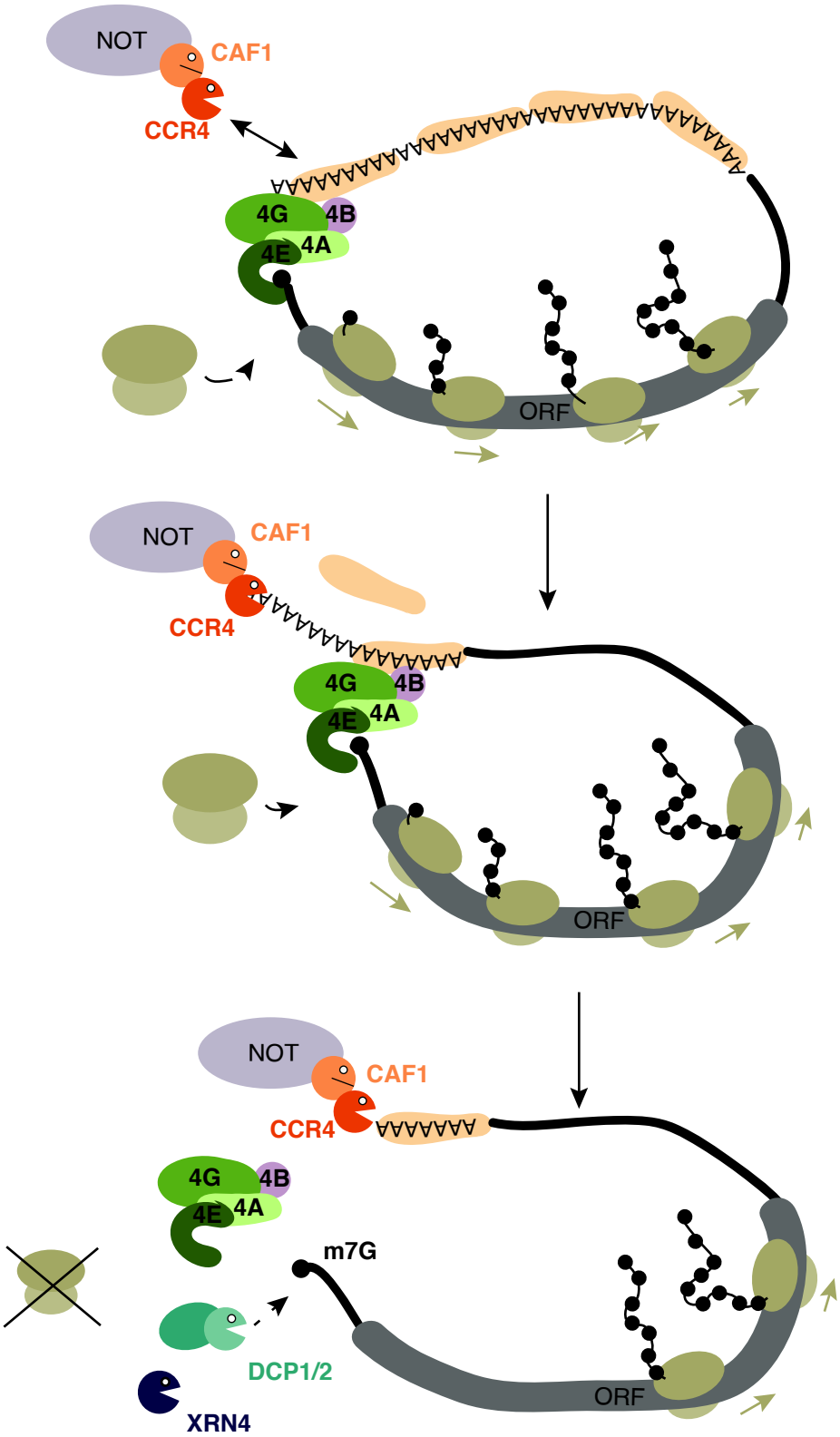


Figure 8: mRNA de-circularisation by deadenylation.
The polyA tail size distribution of highly translated mRNAs show a phased pattern that is likely caused by the CCR4-mediated deadenylation and dislocation of PABPs. A recent study proposed that one single PABP bound to the mRNA oligoA tail is sufficient to maintain the circularised structure of mRNAs that has been shown to promote the formation of the translation initiation complex PIC. The dislocation of the last PABP by CCR4 leads to the dissociation of the closed-loop structure and triggers translation inhibition and decapping and degradation of the mRNA.
Adapted from Lima et al., 2017.

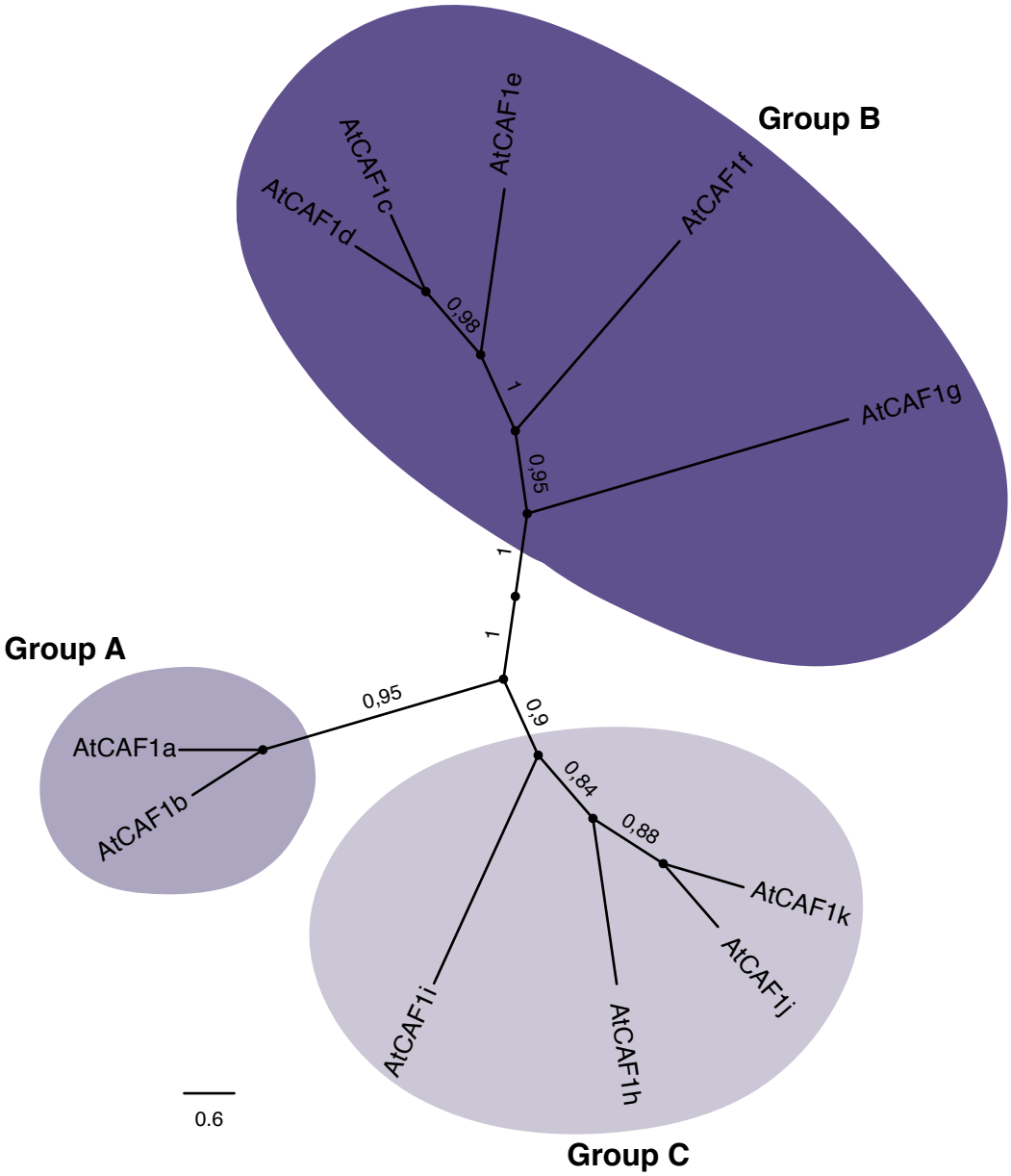


Figure 9: Phylogenetic relationship between *Arabidopsis thaliana* CAF1 proteins. Cladogram of the 11 CAF1 homolog encoded in the genome of *Arabidopsis thaliana*. The multiple alignment was done with Muscle (v. 3.8.31) and the maximum- likelihood tree was generated using PhyML (v. 3.1) on Phylogeny.fr. The cladogram was edited using FigTree (v. 1.4.3, http://tree.bio.ed.ac.uk/software/figtree/).

Surprisingly, highly expressed genes show a phased pattern of *circa* 30 adenosines in their polyA tail size distribution profiles, resembling footprints of PABPs (Lima et al., 2017). As CCR4 has been shown to deadenylate genes with high codon optimality and polyA tails with a dense PAPB coverage, this phasing pattern results likely from CCR4-mediated deadenylation and PABP dislocation (Webster et al., 2018b). Highly expressed mRNAs with polyA tails of less than 29 nucleotides are hardly detected, likely because they undergo fast degradation as they are too short to stably accommodate a single PABP. These results suggest that one single PABP is enough to limit excessive deadenylation and assure mRNA translation. Lima *et al.* propose a model where one PABP is enough to assure mRNA circularisation, recruit PIC and promote mRNA translation (Figure 8).

Less translated genes tend to have longer polyA tails that lack phasing patterns (Castellano and Bazzini, 2017; Lima et al., 2017; Nicholson and Pasquinelli, 2019). This fits to the observation that low expressed genes with little codon optimality degraded by CAF1 have polyA tails poorly bound by PABPs (Webster et al., 2018b). Still, as low expressed transcripts tend to have short half-lives, it is surprising to observe longer polyA tails. A possible explanation is that the deadenylation of such transcripts induces their fast elimination, that would explain why deadenylated tails are less detectable at steady-state level.

Taken together, the emerging picture is that deadenylation is a key step that determines how, when, and at which rate the target mRNA is degraded.

3.4.5. CCR4/NOT complexes in plants.

All core subunits of the human CCR4/NOT complex have homologues in *Arabidopsis thaliana*, rice, and other plant species. First co-immunoprecipitation (IP) experiments using FLAG-CCR4 and yeast two-hybrid assays confirmed the binding of CAF1 and CCR4 to Arabidopsis CCR4/NOT (Arae et al., 2019). Interestingly, the CAF1 gene family is significantly expanded in angiosperms. Plant CAF1 proteins can be classified into 3 distinct clades in Arabidopsis (Arae et al., 2019; Pavlopoulou et al., 2013). Group A comprises CAF1a and CAF1b, group B includes CAF1c, CAF1d, CAF1e, CAF1f and CAF1g, and group C is composed of CAF1h, CAF1i, CAF1j and CAF1k (Figure 9). Yeast two-hybrid assays revealed that Group A CAF1s interact with the MIF4G domain of NOT1, but not with CCR4 (Arae et al., 2019). The CAF1 proteins of Group B do not interact with NOT1, and Group C proteins interact with both CCR4 and NOT1. Similar results have been obtained for the CAF1 homologues in rice: OsCAF1b, the orthologue that clusters near Arabidopsis group A, did not bind to

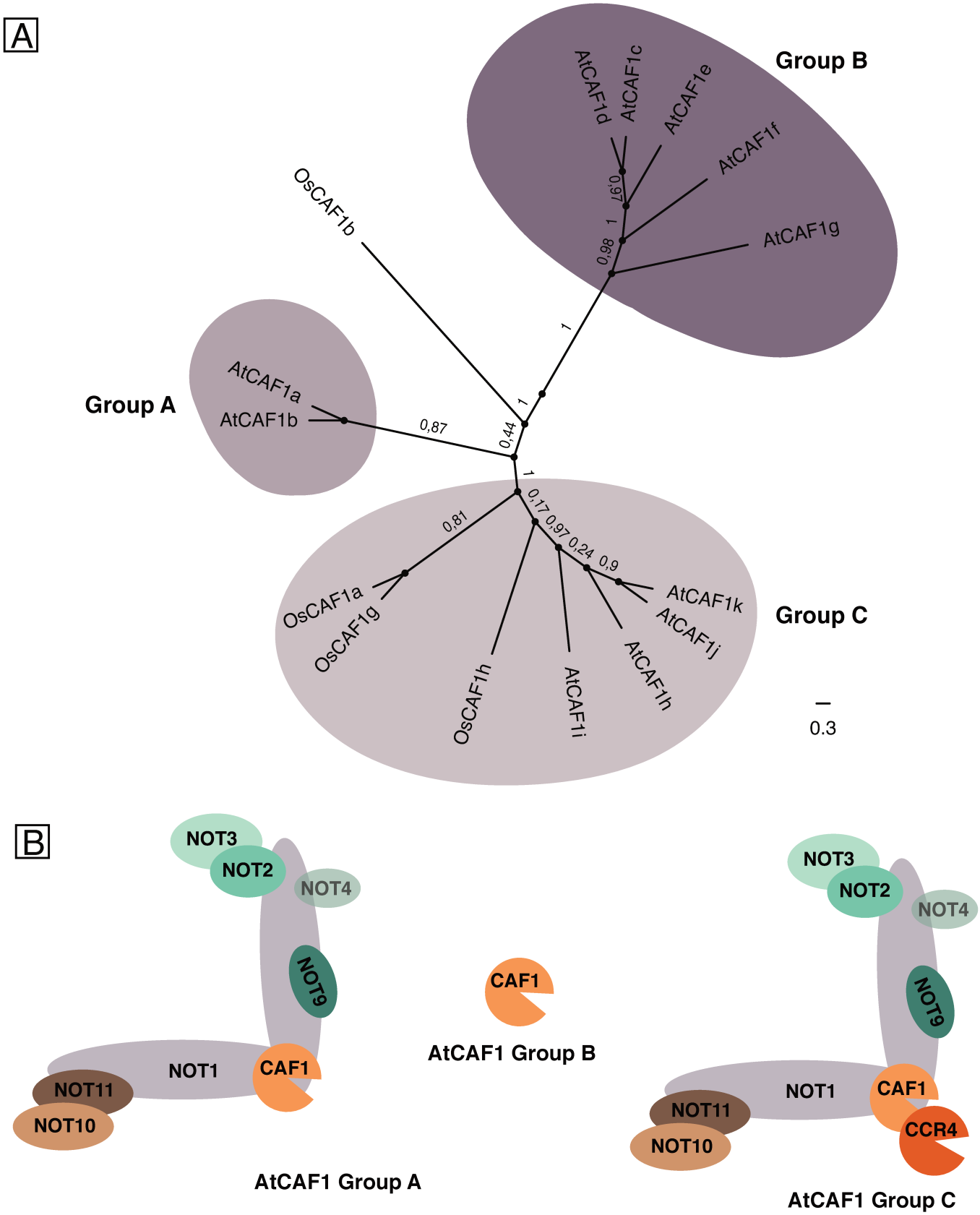


Figure 10: Plant CCR4/NOT complex.
A) Phylogenetic relationship between *Arabidopsis thaliana* and *Oryza sativa* CAF1 homologues. The multiple alignment was done with Muscle (v. 3.8.31) and the maximum- likelihood tree was generated using PhyML (v. 3.1) on Phylogeny.fr. The cladogram was edited using FigTree (v. 1.4.3, http://tree.bio.ed.ac.uk/software/figtree/).
B) Cartoon illustrating the possible compositions of plant CCR4/NOT complexes based on the results of Arae et al., 2019 and Chou et al., 2016.

OsCCR4a/b, whereas OsCAF1a, OsCAF1h and OsCAF1g, that cluster with Arabidopsis CAF1 proteins of group C, interacted with both OsNOT1 and OsCCR4 proteins (Figure 10A) (Chou et al., 2016). Taken together these results suggest that plants possess two versions of CCR4/NOT complexes, one that comprises only CAF1a/b while the other one contains both CAF1h/i/j/k and CCR4. The CAF1 proteins of group B possibly act independent of CCR4/NOT as standalone enzymes (Figure 10A and B). These results gave first insights into the molecular architecture of the CCR4/NOT complex in plants and need to be confirmed.

Arabidopsis has also two homologs of CCR4, CCR4a and CCR4b (Dupressoir et al., 2001). Interestingly, the N-terminal parts of CCR4a and CCR4b do not contain the LRR motifs that are essential for the interaction with CAF1 and the recruitment to the CCR4/NOT complex in yeast and humans. Thus, the interaction between AtCCR4 and CAF1 must involve other motifs not identified yet. In rice, a MYND-like domain at the N-terminal part of CCR4 and a PxLxP motif in CAF1 have been implicated in the interaction of OsCCR4 and OsCAF1 (Chou et al., 2016). However, only the MYND domain is conserved in AtCCR4, while no PxLxP motif is present in AtCAF1.

Because PAN2/3 are not systematically conserved in plants, CCR4/NOT is thought to be the main enzyme that both initiates and completes deadenylation of plant RNAs (Pavlopoulou et al., 2013). However, the physiological functions of the plant CCR4/NOT complex remain largely unexplored. Plant deadenylases display different expression profiles to specific signals that imply overlapping and specific functions in mRNA metabolism (Chou et al., 2014; Walley et al., 2010). Plant CAF1 and CCR4 deadenylases have been shown to be stress-responsive deadenylases whose expression is increased upon stress or developmental stimuli (Chen et al., 2016; Liang et al., 2009; Sarowar et al., 2007; Suzuki et al., 2015; Walley et al., 2007; Zhang et al., 2013). Deregulation of CCR4 and CAF1 impacts mRNA steady-state levels of genes implicated in starch metabolism, plant development, abiotic stress responses, and pathogenesis-related genes. Interestingly, both AtCCR4a/b in Arabidopsis and OsCAF1B in rice are localised in P-bodies (Suzuki et al., 2015).

These studies provided a first insight into the architecture of plant CCR4/NOT complexes, but further studies are necessary to characterise composition, activity and physiological functions of the main deadenylation enzyme in plants.

4. The main factors involved in bulk mRNA decay.

Shortening of the polyA tail is an important and probably the main mechanism to initiate bulk mRNA decay. Other mechanisms to initiate mRNA degradation include endonucleolytic cleavages induced by translation-coupled mRNA quality control pathways such as non-sense mediated decay (NMD), no-stop decay (DSD) or no-go decay (NGD). mRNAs can also be cleaved by the miRNA-guided RNA-induced silencing complex. Each of these possible mechanisms involves proteins that recruit decay factors and trigger the elimination of the mRNA by exoribonucleolytic pathways that degrade RNA from either 3'-5' or from 5'-3'.

4.1. The 3’ to 5’ degradation pathway.

3’-5’ degradation is carried out by two distinct enzymes, the RNA exosome and the exoribonuclease DIS3L2. The RNA exosome is an essential multisubunit complex composed of a conserved core complex of 9 subunits (Exo9) which associates to exoribonuclease, RNA helicases and other co-factors in a compartment specific manner (łabno et al., 2016; Zinder and Lima, 2017). Exo9 is composed of six RNase PH-like proteins that form a barrel-like structure resembling bacterial polynucleotide phosphorylases (PNPases), and three RNA-binding proteins forming a cap-like structure (Evguenieva-Hackenberg et al., 2014; Januszyk and Lima, 2014; Schneider and Tollervey, 2013). In contrast to archaeal exosomes that possess three phosphorolytic sites inside the central channel (Evguenieva-Hackenberg et al., 2014), Exo9 of yeast and human exosome is catalytic inactive and the enzymatic activity is solely provided by the associated hydrolytic exoribonucleases RRP6, DIS3 (RRP44) and DIS3L (Dziembowski et al., 2007; Januszyk and Lima, 2014; Siwaszek et al., 2014; Staals et al., 2010; Tomecki et al., 2010; Zinder and Lima, 2017). Plant exosome complexes exhibit an additional phosphorolytic activity provided by the RRP41 subunit of Exo9 (Sikorska et al., 2017).

Exosome complexes are present in both nuclear and cytoplasmic compartments of all eukaryotic cells. Associated with dedicated exosome targeting complexes that aid in RNA recognition and exosome binding to other protein factors, the nuclear exosome contributes to rRNA processing and is a key player of RNA surveillance by degrading pre-mRNAs, mis-spliced or 3'- extended transcripts, snoRNA precursors and a variety of presumably non-functional RNA generated from intergenic regions or by premature transcription termination (Callahan and Butler, 2010; Falk et al., 2014; Hématy et al., 2016; Lange et al., 2008, 2014; Lubas et al., 2013; Meola et al., 2016;

Sikorski et al., 2015; Vanacova et al., 2007; Wyers et al., 2005). The cytoplasmic exosome participates in the turnover of mRNAs and has specific roles in degrading the substrates of NMD, NGD (No-Go Decay) and NSD (No-Stop Decay) (Shoemaker and Green, 2012) and the 5' fragments produced by RISC cleavage (Branscheid et al., 2015; Orban and Izaurralde, 2005). All functions of the cytoplasmic exosome require its association with SKI7 and the SKI complex (Zinder and Lima, 2017). The SKI complex can directly bind to ribosomes and recruits the exosome complex to mRNAs stalled in translation (Schmidt et al., 2016). In yeast, a SKI-Ska1 complex promotes the degradation of ribosome free RNA regions including 3' UTR's of mRNAs (Zhang et al., 2019). In plants, the cytoplasmic exosome is associated with SKI7 and RST1, a large ARM-repeat protein, and RIPR (for RST1 INTERACTING PROTEIN) which links Exo9-SKI7-RST1 to the Ski complex (Lange et al., 2019).

Degradation by the exosome requires accessible 3' extremities generated by endoribonucleolytic cleavages or by removing protecting proteins 3' ends such as PABPs or LSm complexes. Furthermore, the degradation of non-coding nuclear exosome substrates and pre-mRNAs is triggered by the non-templated addition of oligoA tails by non-canonical polyA polymerases (Kuai et al., 2004; LaCava et al., 2005; Vaňáčová et al., 2005). Recent studies show that cytoplasmic exosome targets can also get marked by both oligoadenylation and uridylation, including mRNAs (Lim et al., 2016), 5' RISC-cleaved fragments (Ibrahim et al., 2006), Ago-bound small RNA (Pisacane and Halic, 2017) and pre-miRNAs (Liu et al., 2014).

DIS3L2 is an exoribonuclase of the RNase D family that is absent in *S. cerevisiae* but conserved in *S. pombe*, metazoa and plants. DIS3L2 is a paralogue of the exosome-associated exoribonuclase RRP44, but lacks the PIN domain that mediates the interaction of RRP44 with the exosome. Indeed, DIS3L2 degrades its RNA targets independently of the exosome (Lubas et al., 2013; Malecki et al., 2013; Zhang et al., 2010). Mutations in human DIS3L2 are linked with the Perlman syndrome of foetal overgrowth and a predisposition for various tumour development (Astuti et al., 2012; Hunter et al., 2018; Morris et al., 2013; Towler et al., 2016). DIS3L2 is a cytoplasmic protein which plays a crucial role in the quality control of non-coding RNAs including unprocessed and highly-structured RNAs (Labno et al., 2016; H.-M. Chang, Triboulet, Thornton, & Gregory, 2013; Faehnle, Walleshauser, & Joshua-Tor, 2014; B. Kim et al., 2015; Lubas et al., 2013; Malecki et al., 2013; Pirouz, Du, Munafò, & Gregory, 2016; Reimão-Pinto et al., 2016; Ustianenko et al., 2013, 2016). The plant homologue SOV (SUPPRESSOR OF VARICOSE) was initially identified as a suppressor of the developmental phenotype of Arabidopsis plants of the Col-0 accession mutated in the

decapping factor VARICOSE (VCS) (Zhang et al., 2010). This observation and a recent study analysing decay rates indicated that SOV contributes to the degradation of cytoplasmic mRNAs, most of which are also substrates of the 5'-3' pathway (Sorenson et al., 2018). In fact, due to the high level of redundancy between 3’-5’ and 5’-3’ degradation pathways, the contribution of individual pathways is generally difficult to assess in all model organisms. However, due to a point mutation in the SOV gene, the Col-0 accession that is used in most Arabidopsis studies including my thesis, does not possess a fully functional SOV exoribonuclease: in contrast to the SOV version of most Arabidopsis accession, the Col-0 version of SOV fails to rescue the severe phenotype of the varicose decapping mutant (Zhang et al., 2010). Therefore, the Col-0 accession is probably one of the best suited Arabidopsis accessions to study processes upstream of exoribonucleolytic decay.

Interestingly, DIS3L2 has a high affinity for 3’ uridylated tails. Drosophila DIS3L2 even associates with the TNTase Tailor in the terminal RNA uridylation-mediated processing (TRUMP) complex (Faehnle et al., 2014; Lin et al., 2017). Whether SOV in Arabidopsis is also stimulated by 3’ uridylation of its targets has not been demonstrated yet.

4.2. The 5’ to 3’ degradation pathway.

5’-3’ decay is thought to be the main pathway for the turnover of mRNA in the cytoplasm. The key steps of the 5’-3’ degradation pathway are the removal of the m7G cap, followed by 5'-3' exoribonucleolytic degradation. The substrates of 5'-3' exoribonucleolytic decay are predominantly mRNAs with short poly A tails produced by deadenylation, indicating that for most mRNAs, deadenylation precedes decapping. However, specialised pathways that trigger decapping without prior deadenylation do also exist.

4.2.1. Deadenylation-independent decapping of mRNAs.

In yeast, where polyA tails are shorter than in mammals, Edc3 can directly bind to mRNAs to trigger decapping and degradation without a previous deadenylation step (Badis et al., 2004; Dong et al., 2010; He et al., 2014; Kolesnikova et al., 2013). However, the best studied example of deadenylation-independent decapping is the elimination of mRNAs harbouring specific features such as premature stop codons or long 3' UTRs by the nonsense-mediated decay pathway NMD (Conti and Izaurralde, 2005; Hu et al., 2010; Isken and Maquat, 2007; Rebbapragada and Lykke-Andersen, 2009). Such RNAs are bound by the RNA helicase and NMD factor UPF1, with recruits the decapping activators Edc3, Pat1 and Dcp2 (Dehecq et al., 2018; Swisher and Parker, 2011). Recent data obtained in our lab showed that Arabidopsis UPF1 associates also with the decapping

factors and translational repressors DCP5 and Dhh1 (Chicois et al., 2018). Interestingly, ribosome dissociation and deadenylation-independent decapping and degradation of NMD targets in *Aspergillus nidulans* is promoted by uridylation (Morozov et al., 2012).

4.2.2. Deadenylation-dependent mRNA decapping.

As mentioned, most RNAs undergo deadenylation before they are decapped and degraded from the 5' end. Deadenylated mRNAs are recognised by Lsm1-7, a doughnut-shaped complex composed of seven distinct Sm-like proteins (Sharif and Conti, 2013; Zhou et al., 2014). While Lsm1 is a cytoplasmic protein, the proteins Lsm2-7 shuttle between cytoplasm and nucleus and are also components of the nuclear Lsm2-8 complex (Bouveret et al., 2000; Tharun et al., 2000). The Lsm2-8 complex protects the 3' extremities of the U6 snRNAs and is required for spliceosome assembly. Lsm1-7 is a key factor of mRNA decapping and 5’-3’ degradation in the cytoplasm (Achsel et al., 1999; Bouveret et al., 2000; Licht et al., 2008; Tharun and Parker, 2001; Tharun et al., 2000; Vindry et al., 2017).

Interestingly, Lsm1-7 has a high affinity for deadenylated and uridylated oligoA tails (Chowdhury et al., 2007; Song and Kiledjian, 2007). In Drosophila, Lsm1-7 associates with the CCR4/NOT complex, which further facilitates its recruitment to deadenylated targets RNAs (Haas et al., 2010). Lsm1-7 binds to the 3’ end of deadenylated mRNAs and recruits the decapping activator Pat1 and the decapping complex Dcp1/2 (Beelman et al., 1996; Bouveret et al., 2000; Chowdhury et al., 2007) The binding of Pat1-Lsm1-7 to the deadenylated 3’ extremities protects them from excessive deadenylation (He and Parker, 2001). Mutations in Lsm1-7 or Pat1 lead to the accumulation of capped and deadenylated mRNAs showing that they are also required for decapping and 5’-3’ degradation (Bouveret et al., 2000; Tharun et al., 2000). In addition to Lsm1-7, Pat1 interacts with many other decay factors including the decapping activators and translational repressors Dhh1 and Edc3 and the 5’-3’ exoribonuclease Xrn1 (Braun et al., 2010; Haas et al., 2010; Ozgur et al., 2010). Taken together, Pat1-Lsm1-7 connects deadenylation with translation inhibition, decapping and 5'-3' degradation.

The decapping holoenzyme Dcp1/2 is weakly active on its own. Its decapping activity is strongly enhanced by the binding of the enhancer proteins Edc1, Edc2 and Edc3 (Charenton et al., 2016; Nissan et al., 2010; Steiger et al., 2003). Edc3 interacts with Dcp2, the catalytic subunit of the decapping complex. Dcp1 supports Dcp2 activity by recruiting Edc1 and Edc2 (Borja et al., 2011; Mugridge et al., 2018). Hydrolysis of the cap by Dcp2 liberates m7GDP. The remaining

monophosphate at the mRNAs 5’ extremity is recognised by 5'-3' exoribonucleases (Li and Kiledjian, 2010). Other activators that interact directly with Dcp1-2 in yeast and in plants are Scd6 and DCP5, respectively. They belong to a protein family that has a conserved LSm domain at the N-terminal part and an FDF (phenylalanine/aspartic acid) domain flanked by two RG/RGG (arginine/glycine) domains at their C-terminus (Roy and Rajyaguru, 2018). Scd6 and DCP5 have first been suggested as decapping activators because of their direct interaction with the decapping complex (Barbee et al., 2006; Xu and Chua, 2009). However, Scd6 and its human homolog LSM14 have also been shown to be implicated in translation inhibition and interact directly with eIF4G of the CBC, hindering the formation of PIC at the 5’ end of mRNAs (Rajyaguru et al., 2012). A recent study suggests that Scd6 recruits Dhh1 to repress the translation initiation and activate Dcp2-mediated mRNA decay *in vivo* (Zeidan et al., 2018). The direct interaction between LSM14/DCP5/Scd6 and DDX6/Dhh1 has been shown in a wide range of eukaryotes, and is critical for the formation of P-bodies in humans (Ayache et al., 2015).

The conserved main cytoplasmic 5'-3' exoribonuclease is Xrn1 (Pacman in Drosophila), a highly processive enzyme responsible for the bulk turnover of mRNA and a wide range of non-coding RNAs (Nagarajan et al., 2013). In metazoa, Xrn1 degrades also the 5' fragments produced by RISC cleavage. Its paralog Xrn2 is located in the nucleus and degrades rRNA precursors and various other types of nuclear RNA species with accessible 5' extremities (Miki and Großhans, 2013; Nagarajan et al., 2013). Arabidopsis has no homolog of XRN1, but 3 orthologs of XRN2, names XRN2, XRN3 and XRN4 (Kastenmayer and Green, 2002). XRN2 and XRN3 are nuclear proteins involved in ribosomal rRNA processing and degradation of non-coding RNAs, respectively (Kurihara, 2017; Kurihara et al., 2012; Zakrzewska-Placzek et al., 2010). XRN4 is the main 5'-3' exoribonuclease in the cytoplasm and degrades mRNAs as well as 3’ fragments and possibly 5’ fragments produced by RISC (Gregory et al., 2008; Gy et al., 2007; Olmedo et al., 2006; Potuschak et al., 2006; Rymarquis et al., 2011; Souret et al., 2004; Zhang et al., 2017a). Recent evidence indicates that degradation by Xrn1 or XRN4 can occur co-translationally on polysomes (Hu et al., 2009; Merret et al., 2015; Pelechano et al., 2015; Yu et al., 2016).

The predominant 5'-3' polarity of co-translational RNA decay is an important mechanism that prevents the production of truncated proteins potentially produced by the translation of 3' trimmed mRNAs. The CCR4/NOT complex plays an important role in initiating the deadenylation of translated mRNAs and the recruitment of decapping enzymes and 5’-3’ degradation factors to

polysomes. Understanding whether and how mRNA uridylation contributes to maintain the 5'-3' polarity of mRNA degradation is a main interest of our team's research.

Thesis objectives

Thesis objectives.

Uridylation of deadenylated mRNAs emerges as an integral step of the general degradation pathways of mRNAs in eukaryotes (with the notable exception of *S. cerevisiae*). In *H. sapiens*, *S. pombe* and *A. nidulans*, uridylation has been proposed to promote the recruitment of the LSm1-7 complex to deadenylated polyA tails, which activates decapping and triggers 5'-3' degradation of the mRNA by Xrn1. Alternatively, uridylation can also lead to rapid degradation by the exosome complex or by Dis3L2, the latter of which was indeed demonstrated to prefer uridylated transcripts (Aphasizhev et al., 2016; Łabno et al., 2016; Lim et al., 2014; Munoz-tello et al., 2015; Read et al., 2011; Scott and Norbury, 2013; Zigáčková and Vaňáčová, 2018).

In Arabidopsis, URT1 is the main uridylyltransferase that targets deadenylated mRNAs (Sement et al., 2013; Zuber et al., 2016). In *urt1* null mutants, the mRNA uridylation levels are reduced by 70-80%. Interestingly, *urt1* mutants accumulate excessively deadenylated mRNAs with tails of less than 10 nucleotides, and the expression of URT1 restored this molecular phenotype (Sement et al., 2013; Zuber et al., 2016). These results suggested that URT1-mediated uridylation prevents the excessive deadenylation of mRNAs. Moreover, immunoprecipitation experiments realised by a former PhD student Hélène Scheer revealed that URT1 co-purifies with the decapping and translation regulator DCP5. Other proteins detected in URT1 IPs were the translation repressor Dhh1 and the CCR4-NOT complex. These data suggest a molecular interaction network that links uridylation to deadenylation, translation regulation and mRNA degradation by decapping and 5'-3' exoribonucleolytic decay.

Although recombinant URT1 can synthesize long poly-uridine tails *in vitro*, it has a distributive activity for the first added nucleotides (Sement et al., 2013). This result fits well to the observation that *in vivo*, most uridylated mRNAs possess only 1-2 U at their 3' ends. Our group has characterised the 3' ends of mRNAs with short polyA tails by high throughput sequencing techniques (Zuber et al., 2016). This revealed that non-uridylated polyA tails of 0-30 nucleotides have a tail size centred around 16 nucleotides in wild-type plants. Importantly, the length of uridylated tails (i.e. the oligo-A stretch plus the added uridines) was also about 16 nucleotides. By contrast, loss of URT1 led to shorter polyA tails of less than 10 nucleotides. These results led us to propose that uridylation by URT1 repairs deadenylated mRNAs to a polyA tail size of 16 nucleotides. Interestingly, this particular tail size extension is determined by the binding of one

single PABP to the oligo-adenylated and uridylated polyA tails *in vitro* and *in vivo* (Zuber et al., 2016).

How the binding of one single PABP to the 3' end of Arabidopsis mRNAs impacts mRNA stability, translation or storage is unknown. However, the obtained results suggested that the uridylation-dependent repair of deadenylated polyA tails and the binding of PABP can slow down deadenylation and delay the formation of 3' truncated mRNAs. Thus, the polyA tail length of oligo-adenylated mRNAs might be controlled by the combined actions of URT1-mediated uridylation and deadenylase activities. The impact of URT1-mediated uridylation on longer polyA tails could not be investigated in this pioneer work, largely due to technical limitations in the correct estimation of long homopolymeric stretches by HTS sequencing methods. The depth of these initial experiments was also not sufficient to investigate the role of uridylation for the regulation of mRNA stability in the context of developmental or environmental adaptations.

Another open question is how URT1 is recruited to the deadenylated mRNAs. For instance, URT1 could gain access to the mRNA extremities when the deadenylase activity switches from progressive to distributive as the polyA tail gets shorter, as seen for yeast CCR4 *in vitro* (Viswanathan et al., 2003). URT1 could also intrinsically prefer shorter oligoA tails, as shown for human TUT4/7 (Lim et al., 2014). Finally, we like to understand the mechanism(s) by which uridylation impairs deadenylation. Uridylation could favour the binding of a protective element to the 3' end, for instance, PABP. Another possibility is that the binding of URT1 itself blocks the access for deadenylases to the 3' extremity. Finally, it could be the presence of terminal uridines *per se* that inhibits the activity of deadenylases.

The main aim of my thesis was to better understand impact of uridylation on deadenylation.

In particular,

- I contributed to the optimisation of the TAIL-seq and 3'RACE-seq methods that enable the in-depth characterisation of mRNA 3' ends in Arabidopsis.
- I used a transient expression system in leaves of *Nicotiana benthamiana* to investigate the effects of URT1 expression on size distribution and uridylation status of a reporter mRNA.
- I used *in vitro* deadenylation assays to show that the addition of 1-2 uridines inhibits the catalytic activity of the deadenylase CAF1b.

- I employed 3'RACE-seq to study the role of URT1-mediated uridylation for the control of polyA tails length in Arabidopsis.
- I started a series of experiments designed to understand the different functions of URT1 and HESO, the second Arabidopsis TUTase that is able to uridylate mRNA under certain conditions.

Together, the experiments I am describing here substantially enhance our understanding of the dynamic equilibrium between uridylation and deadenylation that controls the length of polyA tails in plant and present an important step towards a better comprehension of how polyA tails size impacts mRNA storage, translation and stability.

Results

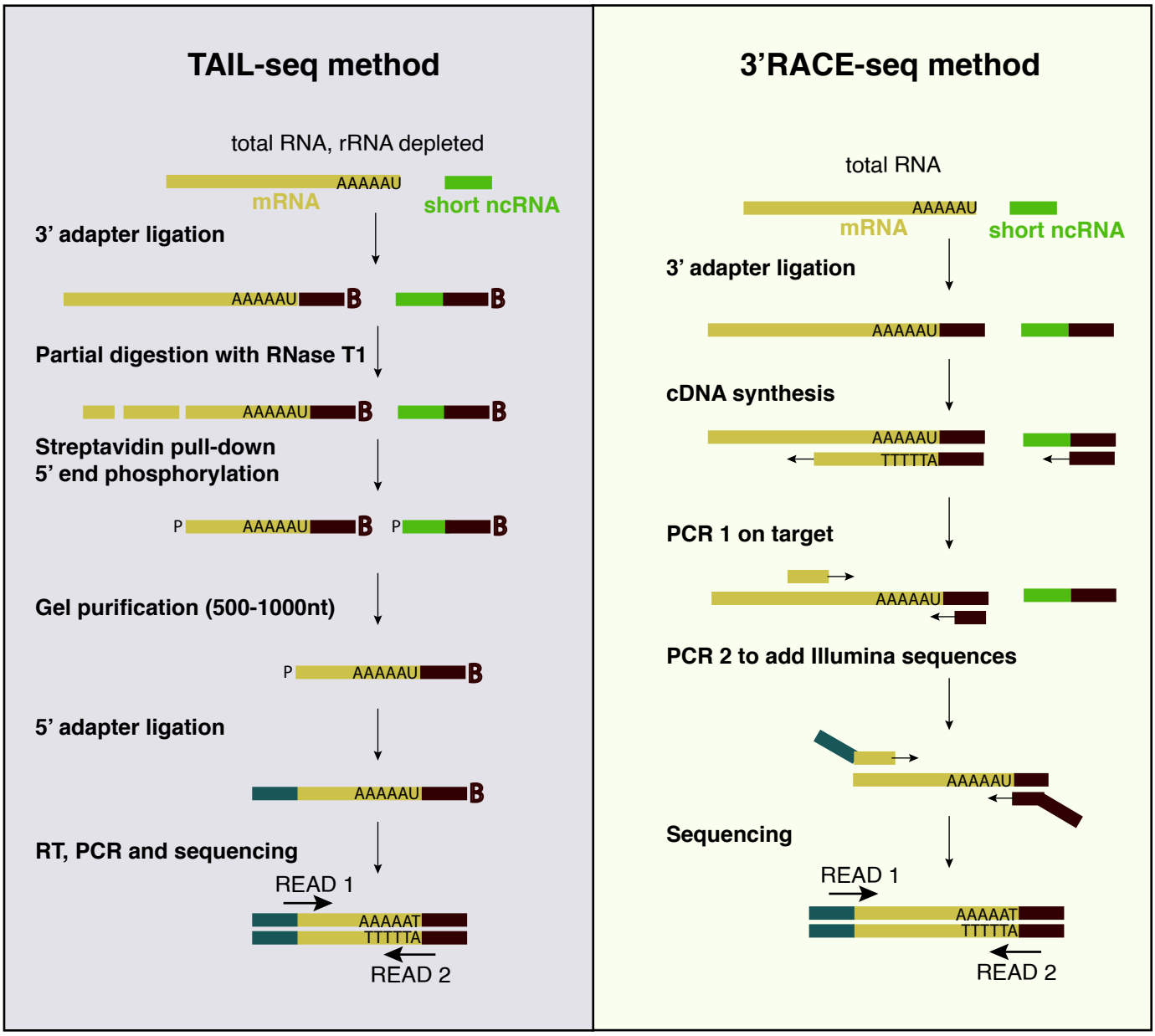


Figure 11: Schematic description of the experimental procedures of the TAIL-seq and 3'RACE-seq methods.
Left panel: Flow chart of the TAIL-seq method. Total RNA is depleted from ribosomal RNAs and ligated to a biotinylated 3' adapter, followed by a partial digestion with RNase T1. Ligated 3' end fragments are then purified using streptavidin beads and phosphorylated at the 5' ends. mRNA fragments are size-selected by separation and elution from a 6% UREA-PAGE. After ligation of a 5' adapter, reverse-transcription is performed using a primer complementary to the 3' adapter. The TAIL-seq libraries are amplified by PCR and paired-end sequenced using Illumina sequencing technology.
Right panel: Flow chart of the 3'RACE-seq method. Total RNA (no ribodepletion required) is ligated with a 3' adapter. A primer complementary to this adapter is used for reverse-transcription and cDNA synthesis. The 3' ends of the mRNA of interest are amplified by nested PCRs using primers that include Illumina sequences needed for the hybridization of the libraries on the flow cell surface.

1. Optimisation of TAIL-seq and 3'RACE-seq methods to analyse mRNA polyA tail length and 3' modifications.

PAL-seq and TAIL-seq are two recently developed high-throughput sequencing methods designed to analyse mRNA polyA tail sizes at a transcriptome-wide level. The PAL-seq technique (or polyA tail length profiling by sequencing) was developed by the lab of David Bartel and is based on the incorporation of biotin-conjugated desoxy-UTPs with fluorescent tags during the sequencing (Subtelny et al., 2014). PAL-seq allows accurate polyA size measurement directly on cDNAs, which avoids a possible bias towards shorter polyA tails by PCR amplification. As the polyA tails are not directly sequenced, the difficulties in sequencing homopolymers are dodged. The inconvenience of this method is that it cannot detect terminal uridines added to the 3' ends, because a splint oligonucleotide with 8 thymidines is used as anchor for the 3' adaptor ligation. Thus, PAL-seq is not suited to investigate the link between uridylation and deadenylation.

By contrast, TAIL-seq can detect both, polyA tails lengths and 3' modifications. TAIL-seq was developed by the lab of Narry Kim (Chang et al., 2014) and then adapted for Arabidopsis samples by our lab (Zuber et al., 2016). The TAIL-seq library preparation starts with the removal of ribosomal RNAs (Figure 11, left panel) and the ligation of a biotinylated adapter to the 3' ends of the remaining RNAs. The adapter comprises a delimiter sequence which ensures that only true 3' ends are analysed, and a random sequence of 15 nucleotide that allows the elimination of PCR duplicons from the analysis. The ligated RNAs are partially fragmented using RNase T1 that cleaves selectively after guanosine residues. The 3' end fragments are purified using streptavidin beads and size-selected by separation and elution from a 6% UREA-PAGE. An adapter is ligated to the 5' end of the purified RNA fragments before cDNAs are synthesised using an oligonucleotide complementary to the 3' adaptor, amplified by PCR and sequenced using Illumina sequencing technology. The 5' and 3' extremities of TAIL-seq libraries consist of the Illumina sequences that assure the hybridization of the libraries on the flow cell surface and the sequencing. Libraries are paired-end sequenced: Read 1 identifies the mRNA, while read 2 comprises the mRNA 3' end with the polyA tail. To add sequence diversity to the sample, 25-30% of a Phix control v3 library is sequenced in parallel. Spike-in sequences with known polyA tail lengths are added to evaluate the quality of the polyA tails size estimation during the analysis.

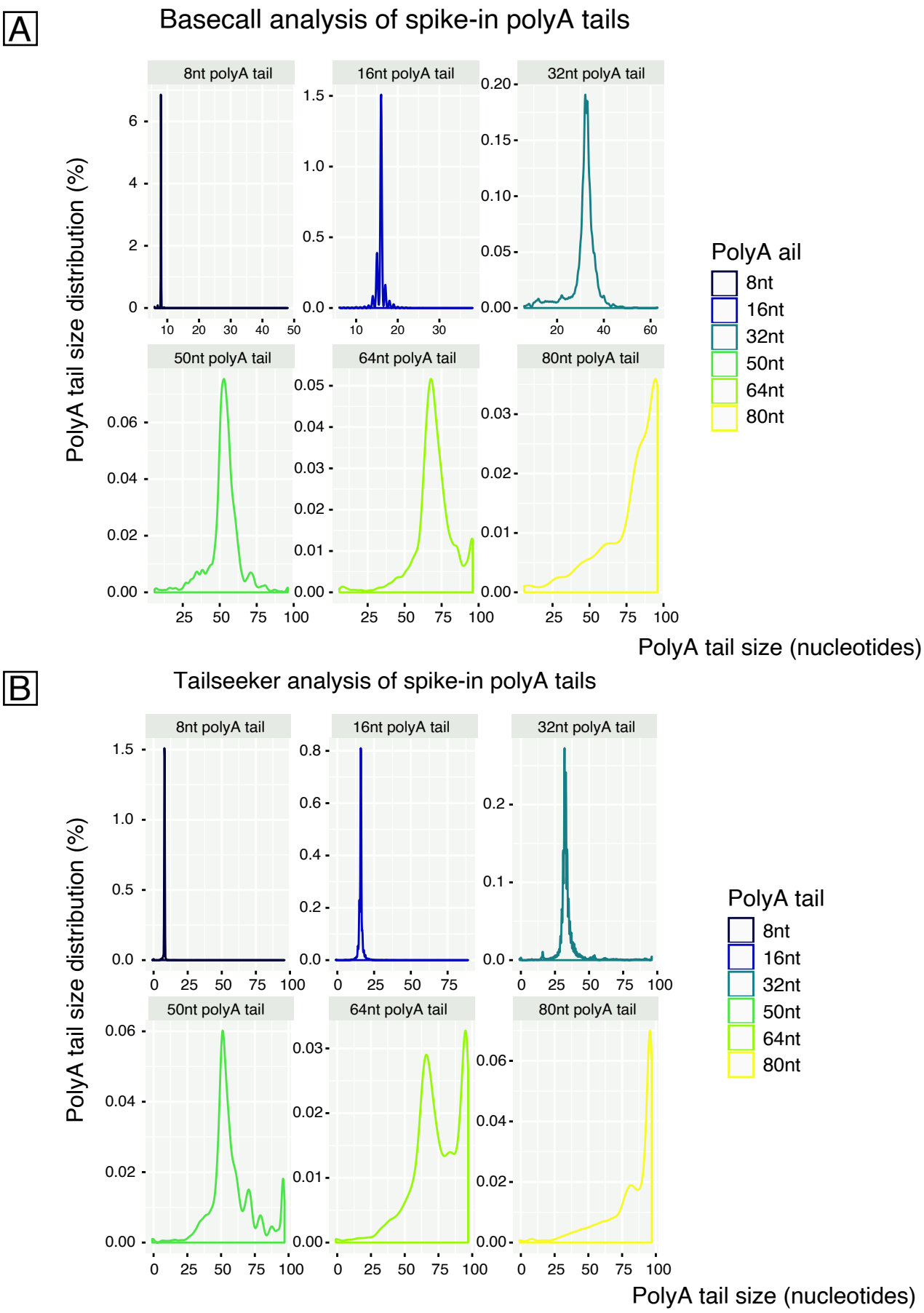


Figure 12: Estimation of polyA tail sizes using the Tailseeker or basecalling pipelines on the TAIL-seq dataset.

A) The TAIL-seq sequencing data was analysed using the classical basecall analysis. The density of the polyA tails of spike-in sequences was analysed to evaluate the accuracy of the polyA size estimation.

B) The TAIL-seq sequencing data was analysed using the tailseeker analysis adapted by Chang et al., 2017. The density of the polyA tails of spike-in sequences was analysed to evaluate the accuracy of the polyA size estimation.

TAIL-seq libraries of good quality and depths are crucial to examine modulations of polyA tail sizes and 3' tails in different mutant backgrounds, tissues, or in response to environmental stimuli. For example, our laboratory wants to use TAIL-seq to compare mRNA 3' ends in WT, *urt1* and *heso1* mutant plants to determine how URT1- and HESO1- mediated uridylation sculpts the global polyA tail size distribution and participates to mRNA 3' end processing. The first TAIL-seq analysis that we performed provided many interesting information (Zuber et al., 2016). However, the depth of this first analysis was not sufficient to allow a gene to gene analysis. During my PhD, I participated in optimising the TAIL-seq method, in particular by improving the library preparation. I managed to produce the first TAIL-seq dataset for Arabidopsis with a depth sufficient for a gene to gene analysis. This library was produced from wild-type *A. thaliana* plants (flower material) and was sequenced on an Illumina MiSeq system. The preparation of additional libraries from independent biological replicates is currently in progress.

1.1. PolyA tails size estimation by Tailseeker and basecalling analysis pipelines.

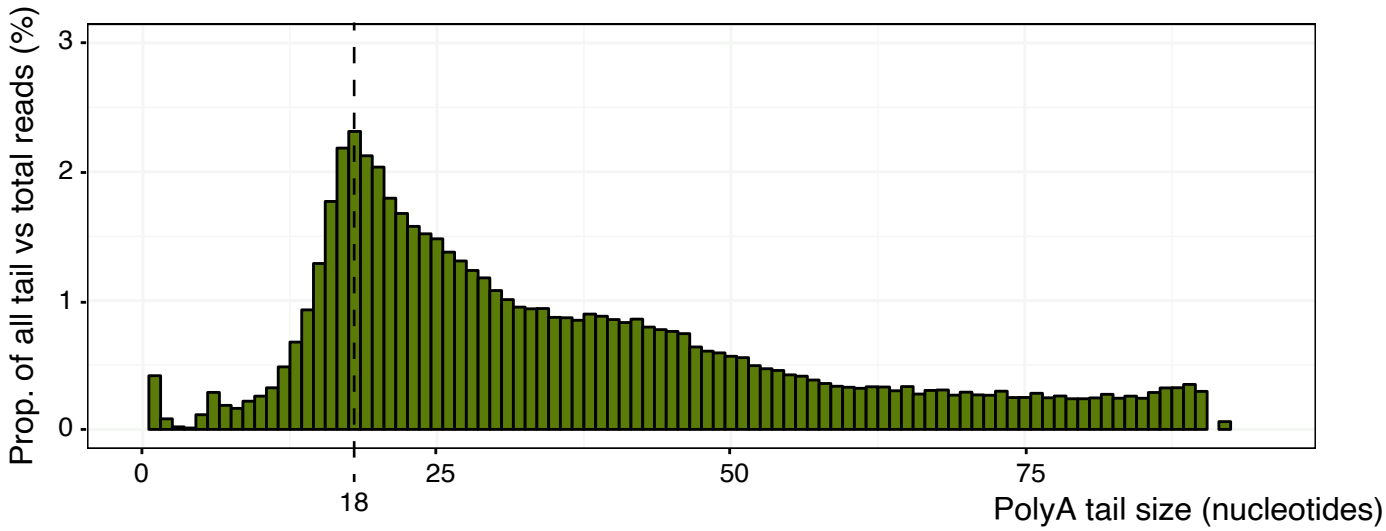
To analyse TAIL-seq libraries, Hélène Zuber developed a bioinformatic pipeline based on basecalling analysis (Zuber et al., 2016) that she continuously optimised. Yet, when Chang *et. al.* developed the TAIL-seq method, they observed important biases in the length estimation for long polyA tails (>30As) when using an analysis pipeline based on basecalling (Chang et al., 2014). To counteract these biases, they developed a software called Tailseeker that uses quantitative fluorescence data to detect the border between the mRNA body and the polyA tail and estimate the polyA tail length. The analysis pipeline based on Tailseeker provided a more accurate estimation of the length of long polyA tails.

Version 3 of the Tailseeker software (https://github.com/hyeshik/tailseeker) was recently implemented on the IBMP's computing cluster. We therefore compared our basecalling based pipeline and the Tailseeker pipeline by analysing the polyA tail sizes of spike-in molecules that were sequenced alongside our TAIL-seq library. These spike-in sequences had polyA tails of fixed sizes, *i.e.* 8, 16, 32, 50, 64 and 80 nucleotides. For spike-ins with 8 to 32 As, polyA sizes were correctly measured by both the basecalling based pipeline (Figure 12A) and the Tailseeker pipeline (Figure 12B). The length of longer polyA tails was overestimated by both methods as previously observed. Yet, our basecalling based pipeline estimated the size of long polyA tails more precisely than the Tailseeker pipeline, particularly for the spike-in with 64 adenosines. Thus, we analysed

A Number of mRNAs

all reads	14 612
min. 20 reads	2 044
min. 100 reads	387
min. 200 reads	149
min. 500 reads	36

B Genome-wide polyA tail size distribution in wild-type *A. thaliana* plants (flowers)



C Global distribution of uridylation level

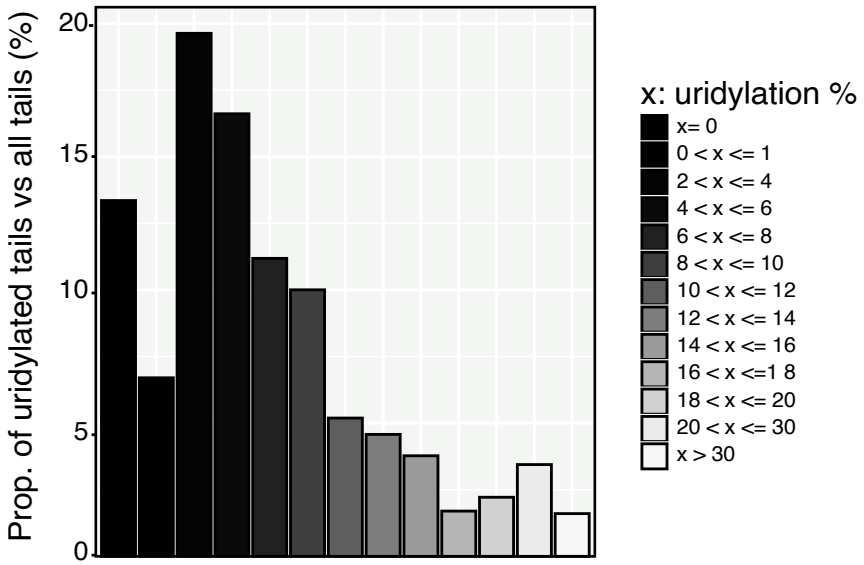


Figure 13: Analysis of the TAIL-seq wild-type dataset.
A) Number of mRNAs detected in the sequencing data.
B) PolyA tail size distribution of all pooled mRNAs.
C) Global distribution of the uridylation level of the mRNAs that have at least 20 reads.

all our TAIL-seq and 3'RACE-seq libraries using our basecalling based pipeline to minimise the biases associated with long polyA size estimations.

1.2. First analysis of the new Arabidopsis TAIL-seq dataset.

3' ends of more than 14 612 mRNAs were detected in the new TAIL-seq library that I prepared from Arabidopsis flowers (Figure 13A). Among them, 2 044 were supported by at least 20 reads, 387 by at least 100 reads, and 36 by at least 500 reads. The depth was considerably increased compared to the first TAIL-seq libraries described in Zuber et al., 2016.

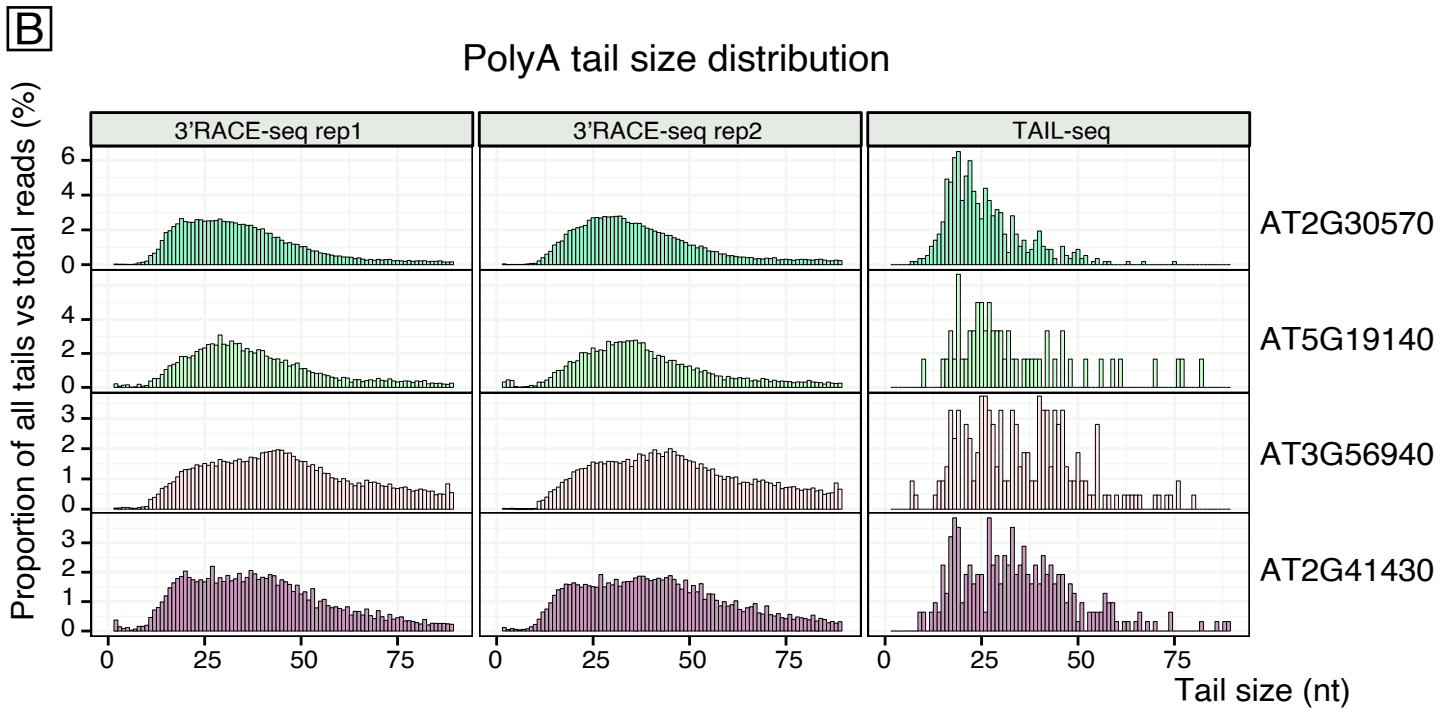
To obtain a global overview of the average length of polyA tails, we combined the data for all mRNAs and plotted the size distribution of all polyA tails. We observed polyA tails from 6–90 nucleotides with the majority of tails between 13 and 50 nt. The size distribution had a clear peak at 18 nucleotides (Figure 13B). This is very similar to the peak at 16 nucleotides that we determined previously by analysing polyA tails of <30 nucleotides by TAIL-seq (Zuber et al., 2016). Only a small proportion of polyA tails were longer than 50 nucleotides. For the interpretation of this observation, it is important to keep in mind that the TAIL-seq protocol comprises several steps that potentially introduce biases against long polyA tails, *ie.* PCR and flow cell clustering (Morgan et al., 2017). Therefore, the TAIL-seq method likely underestimates the proportion of long polyA tails, despite the optimised analysis pipeline. However, the TAIL-seq method is well suited to compare changes in the polyA tail size distributions among different samples. To obtain accurate polyA size distributions for individual mRNAs, we want to further improve the depth of the TAIL-seq libraries. Several steps in the library preparation can be optimized. The main complication we encountered is the poor efficiency and reproducibility of the ribosomal RNA depletion. Partial removal of the rRNAs severely impacts the depth, most likely because the rRNAs outcompete the poor clustering of sequences with long polyA tails in the Illumina flow cell.

The sequencing depth obtained by TAIL-seq was sufficient to determine the proportions of uridylated transcripts for each of the mRNAs that were detected with at least 20 reads (over 2 000 mRNAs, Figure 13A). Uridylated transcripts were detected for 87% of these mRNAs (Figure 13C), which confirmed that mRNA uridylation is widespread in Arabidopsis (Zuber et al., 2016). Interestingly, the uridylation levels varied remarkably between different mRNAs. 20% of the mRNAs had uridylation levels of 2-4%, and 16% of 4-6%, but only 5% of the mRNAs had uridylation levels higher than 20% (Figure 13C). Although these preliminary results must be confirmed in the

A

Number of reads

	total			uri		
Samples	rep1	rep2	TAIL	rep1	rep2	TAIL
AT2G30570	55771	42845	569	585	329	20
AT2G41430	30461	38346	312	1333	1446	31
AT5G19140	14771	9738	60	284	198	3
AT3G56940	80732	70207	214	1930	1935	12



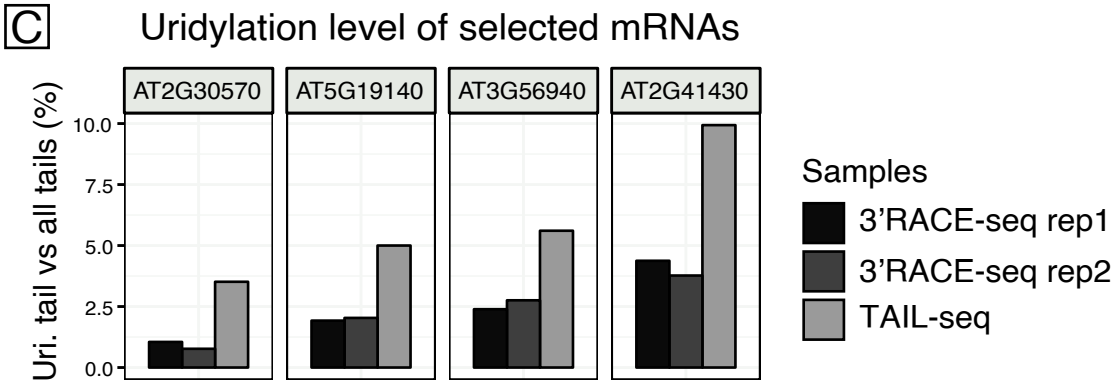


Figure 14: Comparison between the 3'RACE-seq datasets and the TAIL-seq dataset.
A) Number of reads for each of the mRNAs in both 3'RACE-seq (rep1 and 2) and the TAIL-seq libraries.
B) PolyA size distribution of 4 selected mRNAs.
C) Uridylation levels of the 4 selected mRNAs.

biological replicates that are currently in preparation, these data suggest that the levels of uridylated transcripts are specific to each mRNA and/or regulated.

This library allowed us to have a first subset of highly uridylated mRNAs which would be interesting to investigate in different mutant backgrounds of the main TUTases of *A. thaliana*.

1.3. 3'RACE-seq enables the deep analysis of target mRNA 3' ends.

To determine both polyA tail size and 3' modification of selected mRNAs, our group as well as many others have developed the 3'RACE-seq method. 3'RACE-seq combines the classical method to determine mRNA 3' ends by PCR, termed rapid amplification of cDNA 3' ends (3'RACE), with the high throughput sequencing of the amplicons using an Illumina sequencing system (see supplementary data, Scheer et al., 2020, published article in Methods in Molecular Biology). The development of the 3'RACE-seq method has been initiated by Hélène Zuber and Hélène Scheer. I helped to improve the library preparation in order to increase sequencing depths. As compared to the sequencing of cloned RACE fragments, 3'RACE-seq has considerably improved the analysis of 3' ends of different types of target mRNA molecules including rRNA precursors (Sikorska et al., 2017), RISC-cleaved transcripts (Zuber et al., 2018), and viral RNAs (in preparation). The enormous depth and the high reproducibility of 3'RACE-seq also enabled the accurate analysis of mRNA 3' ends. 3'RACE-seq starts with the ligation of a 3' adapter on total RNAs, followed by a reverse transcription using a primer complementary to the adapter (Figure 11, right panel). The mRNA 3' ends are amplified by two rounds of PCR using nested primers. The second PCR increases the specificity and adds the Illumina adapter sequences on both 5' and 3' extremities. Individual barcodes (also called Index regions) added with the reverse primers allow multiplexing and the simultaneous sequencing of different samples. The amplicons are then paired-end sequenced on the IBMP's Illumina MiSeq system.

As compared to TAIL-seq, 3'RACE-seq has a much better sequencing depth. For less expressed mRNAs such as the *AT5G19140* mRNA, only 60 reads were detected by TAIL-seq, whereas 10 000 reads were detected by 3'RACE-seq (Figure 14A). Furthermore, 3'RACE-seq has less biases against long polyA tails and is highly reproducible (Figure 14B). For each of the four examples shown in the figure, the profiles of the polyA tails size distributions obtained by 3'RACE-seq vary only slightly between replicate 1 and 2. Thus, the 3'RACE-seq method allows to obtain highly reproductible polyA size distributions. This is crucial for the correct size estimation of the polyA tail and

indispensable for the analysis of small variations in the polyA tail length in different genetic backgrounds.

Interestingly, the uridylation levels of the four mRNAs were lower in the 3'RACE-seq libraries than in the TAIL-seq library (Figure 14C). Due to the bias of TAIL-seq against longer polyA tails, the levels of short polyA tails are overestimated. Because uridylation occurs predominantly on short tails, TAIL-seq also overestimates the uridylation levels. Altogether, these data show that both methods are highly complementary. TAIL-seq is a very powerful method to identify mRNAs whose uridylation levels vary according to mutant background or conditions. By contrast, 3'RACE-seq is the method of choice to precisely analyse the polyA tail size distributions and reliably quantify 3' end modification of selected mRNAs and to compare those parameters in different genetic backgrounds and conditions.

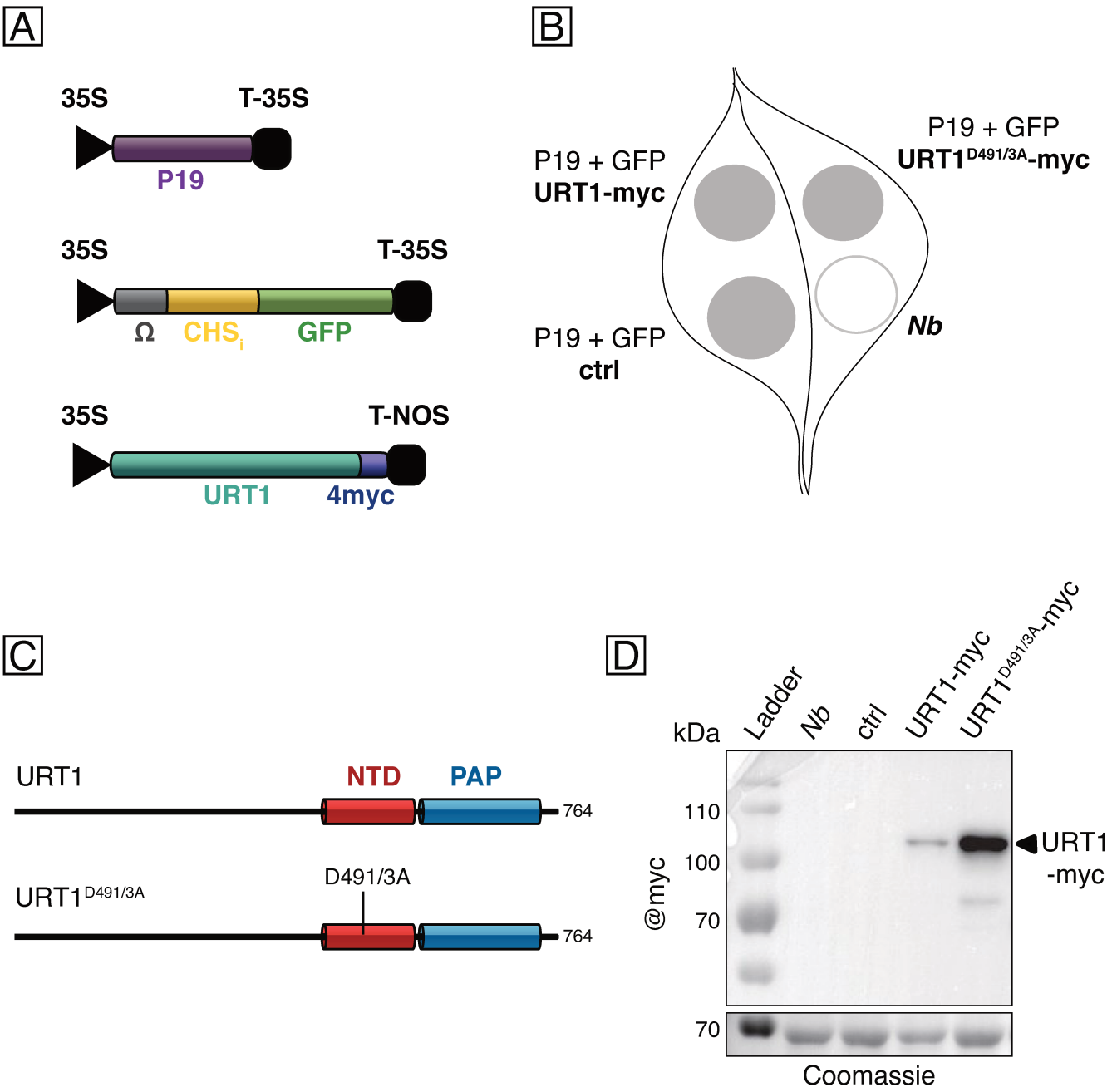


Figure 15: Transient expression of URT1-myc and URT1$^{D491/3A}$-myc in *N. benthamiana.*
A) Diagrams of the constructs used to express the silencing suppressor P19, the reporter *GFP* mRNA and the URT1-myc proteins in *N. benthamiana.*
35S: CaMV 35S promoter, Ω: omega 5' leader sequence of TuMV, CHS$_i$: intron of the petunia chalcone synthase gene, T-35S: CaMV 35S terminator, T-NOS: nopaline synthase gene terminator.
B) Illustration of the infiltrated patches on *N. benthamiana* leaves. ctrl: control patches infiltrated only with P19 and GFP; *Nb*: *Nicotiana benthamiana* uninfiltrated leaf patch.
C) Diagrams illustrating the domain organisations and mutations of the different URT1-myc constructs. NTD: polymerase β-like nucleotidyltransferase domain; PAP: PolyA polymerase-associated domain.
D) Western bot analysis of the protein extracts obtained from the different leave patches harvested four days after infiltration (4 DAI). Monoclonal @myc antibodies were used for the detection of the URT1-myc constructs. The Coomassie stain of the membrane is shown as loading control.

2. Assessing the role of URT1 by transient expression assays in *Nicotiana benthamiana* leaves.

Previous results of our group have shown that URT1 mostly acts on deadenylated, short polyA tails (Sement et al., 2013). Uridylation by URT1 leads to the repair of the 3' ends of excessive deadenylated mRNAs to allow the binding of a single PABP (Zuber et al., 2016). Because URT1 competes with deadenylases for the same RNA substrates, *i.e.* mRNAs that are subject to deadenylation, my main objective was to assess whether URT1-mediated uridylation can directly impact mRNA deadenylation. To this end, our strategy was to disrupt the equilibrium between uridylation and deadenylation by transiently overexpressing URT1 in *Nicotiana benthamiana* leaves. I used 3'RACE-seq to determine the polyA tail sizes and the uridylation status of a *GFP* reporter mRNA which I co-expressed with URT1 in *N. benthamiana*. I also determined polyA tail sizes and the uridylation status of an endogenous *N. benthamiana* mRNA. This experimental system was also used to express mutated versions of URT1 such as catalytic mutants and mutants lacking the interaction motifs present in the N-terminal IDR of URT1.

2.1. Overexpression of AtURT1 in *Nicotiana benthamiana* changes the polyA tail distribution of reporter *GFP* mRNAs.

To analyse how URT1 influences the deadenylation step of a reporter mRNA, C-terminal tagged versions of URT1-myc were co-expressed in *N. benthamiana* leaves with a *GFP* reporter construct. The reporter includes the sequence of the first intron of the Petunia chalcone synthase (CHS) gene inserted 5' of the *GFP* coding sequence (Figure 15A). This CHS intron allows us to differentiate the unspliced pre-messenger RNA from mature *GFP* mRNA using RNA blots. Thus, the signal of the mature form of the *GFP* mRNA can be normalised against the pre-mRNA. Furthermore, the reporter contains the 5' leader sequence (omega sequence) of the tobacco mosaic virus TuMV as 5' UTR sequence (Figure 15A). This sequence has been shown to promote the translation of reporter genes by enhancing the binding of Hsp101 (Gallie, 2002; Gallie et al., 1987). In addition to URT1-myc and the *GFP* reporter, we systematically co-expressed the P19 silencing suppressor to prevent silencing. All constructs were expressed under the control of the cauliflower mosaic plant virus (CaMV) 35S promoter and the Agrobacterium "Nopaline Synthase" terminator (Figure 15A).

Each of the three constructs was transformed into *Agrobacterium tumefaciens*. Transient co-expression was achieved by infiltrating *N. benthamiana* leaves with the pre-mixed Agrobacterium cultures. One leaf patch was infiltrated with the *GFP* reporter and P19, the second patch co-expressed the *GFP* reporter, P19 and wild-type URT1-myc, and the third patch received the *GFP* reporter, P19, and the inactive form URT1$^{D491/3A}$-myc (Figure 15B and C) in which two aspartic acids within the catalytic nucleotidyltransferase domain are mutated into alanines (D491A and D493A) (Sement et al., 2013). Four days after the infiltration, both proteins and total RNAs were extracted from infiltrated leaf patches.

The analysis of protein extracts by western blots using myc-antibodies revealed large differences between the expression levels of the different URT1-myc constructs (Figure 15D) that were observed in all replicates. The inactive version URT1$^{D491/3A}$-myc was systematically more expressed than the active form, URT1-myc. The differential accumulation of active *versus* inactive URT1-myc versions has important implications for the interpretation of the 3'RACE-seq results as discussed below. Of note, Hélène Scheer demonstrated that the expression of URT1 leads to the repression of the translation of reporter *GFP* in *N. benthamiana* (Scheer, 2018). Thus, the different protein levels of URT1-myc and inactive URT1$^{D491/3A}$-myc may be linked to the translation repression of its own mRNA by active version of URT1.

After adapter ligation and cDNA synthesis, the 3' ends of the *GFP* reporter mRNAs were sequenced by 3'RACE-seq using a paired-end sequencing approach. The 5' ends were sequenced up to position 41 to identify the amplified target. The 3' ends were sequenced with 111 cycles to investigate the length and 3' end modifications of the polyA tails.

2.1.1. Overexpression of active URT1 increases the uridylation level of the reporter GFP mRNA.

To demonstrate the robustness of the 3'RACE-seq analysis, I first present the results obtained for 6 different biological replicates. Replicates 1-3 and 4-6 were sequenced separately. For each sample, we obtained 50 000 to 100 000 total polyA tail reads for the *GFP* transcript (Suppl. Figure S1). PolyA tails with terminal uridines at the 3' ends were classified as uridylated tails.

The first observation to mention is that in the control sample (*GFP* and P19), *circa* 3% the *GFP* reporter mRNAs were uridylated (Figure 16A). This basal uridylation level is likely due to the endogenous activity of the *N. benthamiana* URT1 homolog, NbURT1. In samples expressing active URT1-myc, *circa* 15% of the *GFP* reporter mRNAs had uridylated polyA tails. Analysing the number

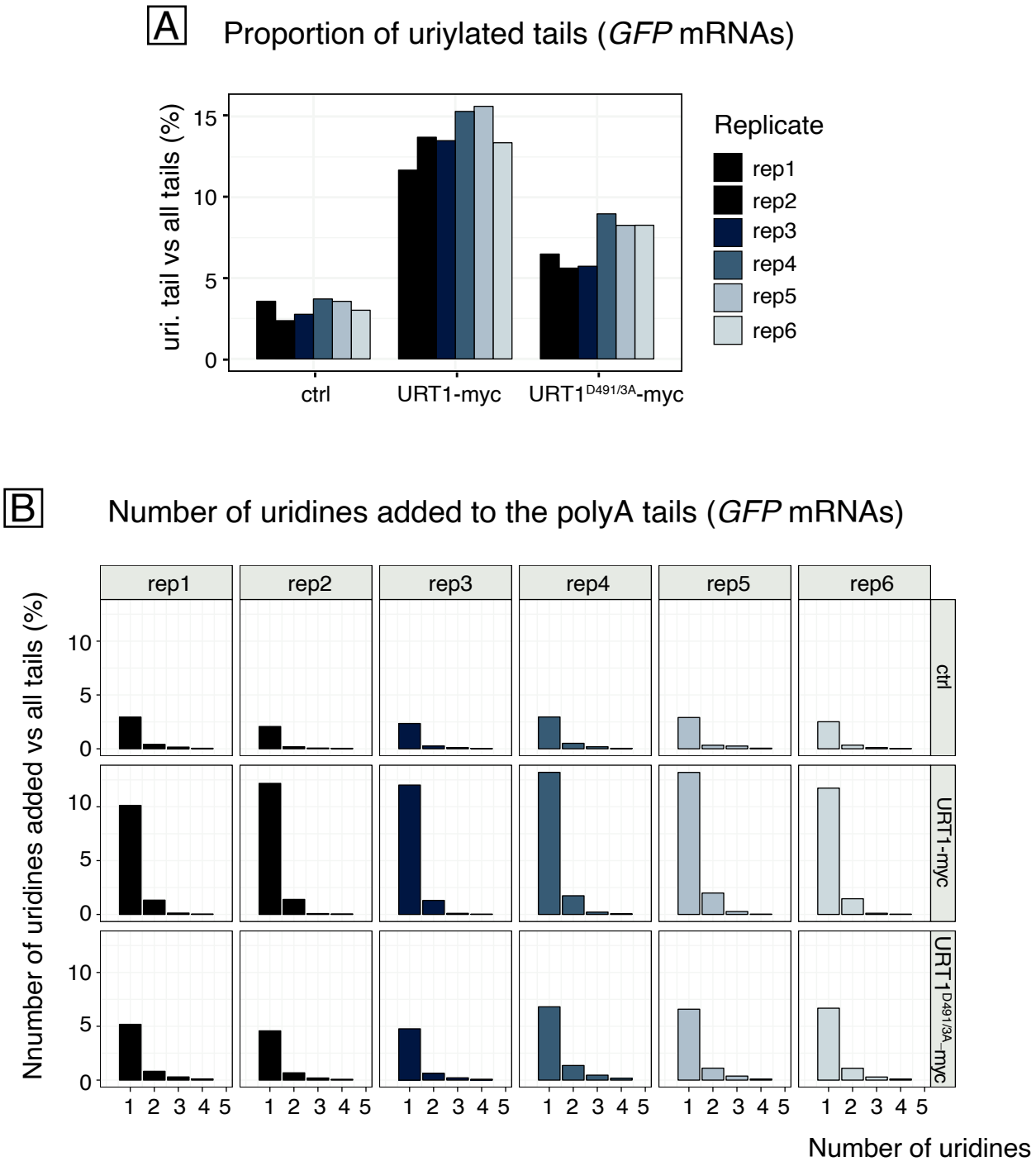


Figure 16: Uridylation level of reporter *GFP* mRNAs.
The 3' ends of the GFP reporter mRNA were analysed by 3' RACE -seq on libraries prepared from the infiltrated leave patches
A) Uridylation percentage of the reporter *GFP* transcript in the control patches, URT1-myc or URT1$^{D491/3A}$-myc patches for 6 biological replicates.
B) Number of uridines added to the 3' end of uridylated tails.

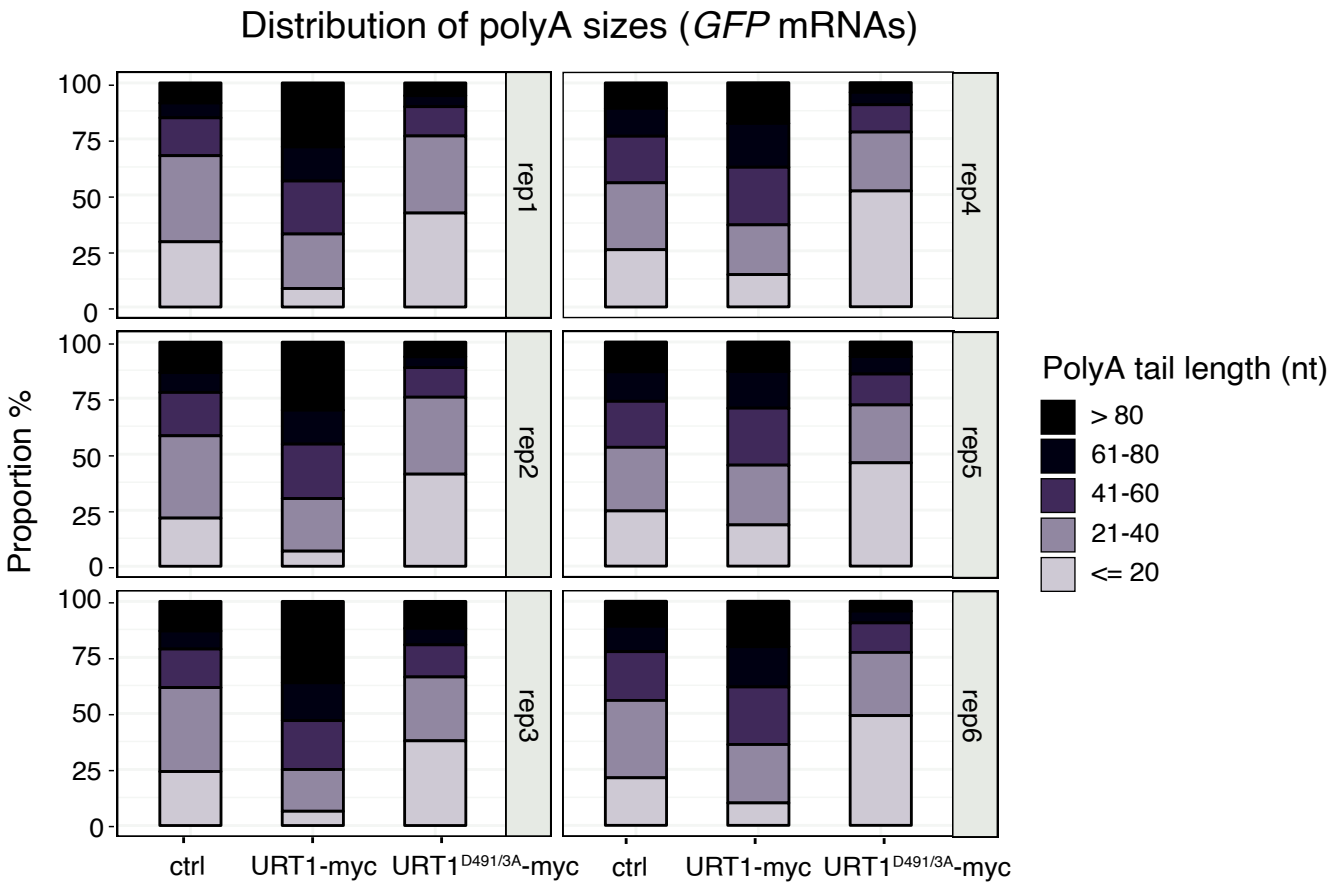


Figure 17: Size repartition of the polyA tails of reporter *GFP* mRNAs.
Proportion of polyA tail length below 20 nucleotides, between 21 and 40 nucleotides, 41 and 60 nucleotides, 61 and 80 nucleotides, and over 80 nucleotides.

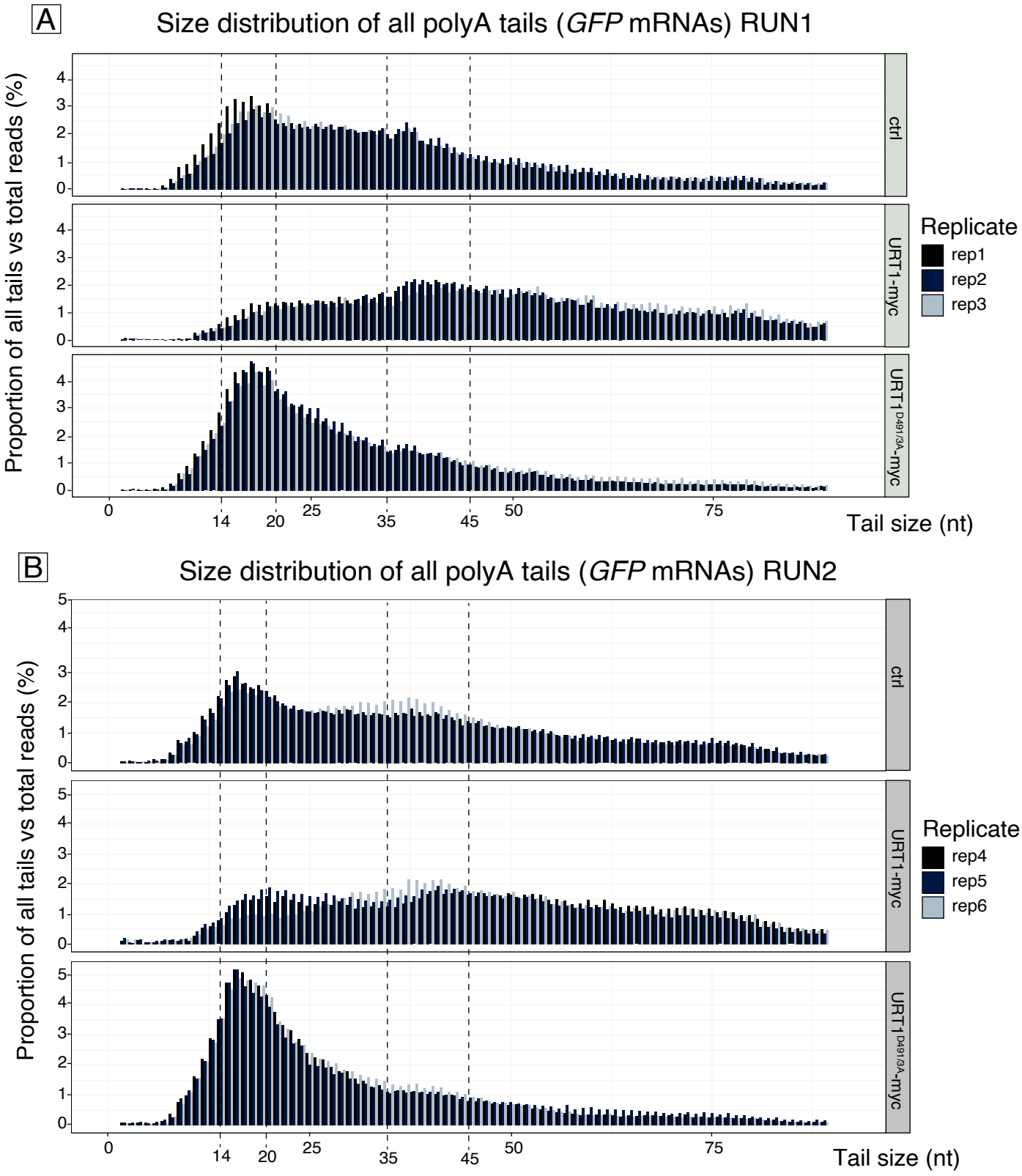


Figure 18: Size distribution of all polyA tails of reporter *GFP* mRNAs for replicates 1-3 **(A)** and 4-6 **(B)** as histograms. The 3' end amplicons are sequenced with 111 cycles. The beginning of the amplicon corresponds to the random region and the delimiter sequence of 20 nucleotides. Hence, the MiSeq system can precisely determine the polyA-tail length of tails up to 91 nucleotides while all larger polyA tails can only be classified as > 91 nt. To have an accurate size distribution, only polyA tails from 1 - 90 nucleotides were plotted.

of uridines that were added to the *GFP* reporter polyA tails revealed that in all samples, most transcripts contained a single uridine. Hence, the overexpression of URT1-myc does not lead to an artificial elongation of the U-tract at the mRNA 3' ends (Figure 16B). These observations demonstrate that overexpressed Arabidopsis URT1 protein is active and uridylates the polyA tails of the *GFP* reporter mRNA in *N. benthamiana*.

Surprisingly, the percentage of uridylated tails in the URT1$^{D491/3A}$-myc sample was not similar to the control sample but was reduced to only approximately 7%. The observation that samples expressing URT1$^{D491/3A}$-myc had higher levels of uridylated *GFP* mRNAs than the control samples was unexpected and suggested that the massive overexpression of the URT1$^{D491/3A}$-myc protein stabilises transcripts that have been uridylated by an endogenous TUTase activity. This possibility is addressed below (see Results 2.2.2. and 2.4.).

2.1.2. Expression of URT1-myc changes the polyA tail size distribution of reporter *GFP* mRNA.

To understand whether the overexpression of URT1-myc proteins influences the polyA tail length of reporter *GFP*, I compared the size distribution of the polyA tails in the different samples (presented as bar charts in Figure 17 and as histograms in Figure 18). In the controls, about 30% of all tails were 21 to 40 nucleotides long. In samples expressing active URT1-myc, 40-50% of all polyA tails were longer than 80 nucleotides, and only a few tails were shorter than 20 nucleotides. The proportion of polyA tails longer than 80 nucleotides was drastically reduced in patches expressing inactive URT1$^{D491/3A}$-myc. Moreover, the pattern of polyA tails in URT1$^{D491/3A}$-myc differed from the control: 40% of the tails were shorter than 20 nucleotides (Figure 17) and thus represent deadenylated transcripts.

The alternative representation of the polyA size distribution as histograms (Figure 18) highlights the reproducibility of the experiments. The distribution profiles were almost identical among both replicates and the two separate MiSeq runs (Figure 18A and B). In control samples, a large proportion of transcripts had short tails of 14 to 20 nucleotides. The size of 14-20 As is remarkably similar to the sizes of short polyA tails observed in the TAIL-seq datasets in Arabidopsis, which was about 17-19 nucleotides (see Result 1.2.). The population of transcripts with short tails of 14-20 over-accumulated upon expression of URT1$^{D491/3A}$-myc. By contrast, samples expressing URT1-myc had a spread distribution of polyA tail sizes with a high proportion of polyA tails of 35-45

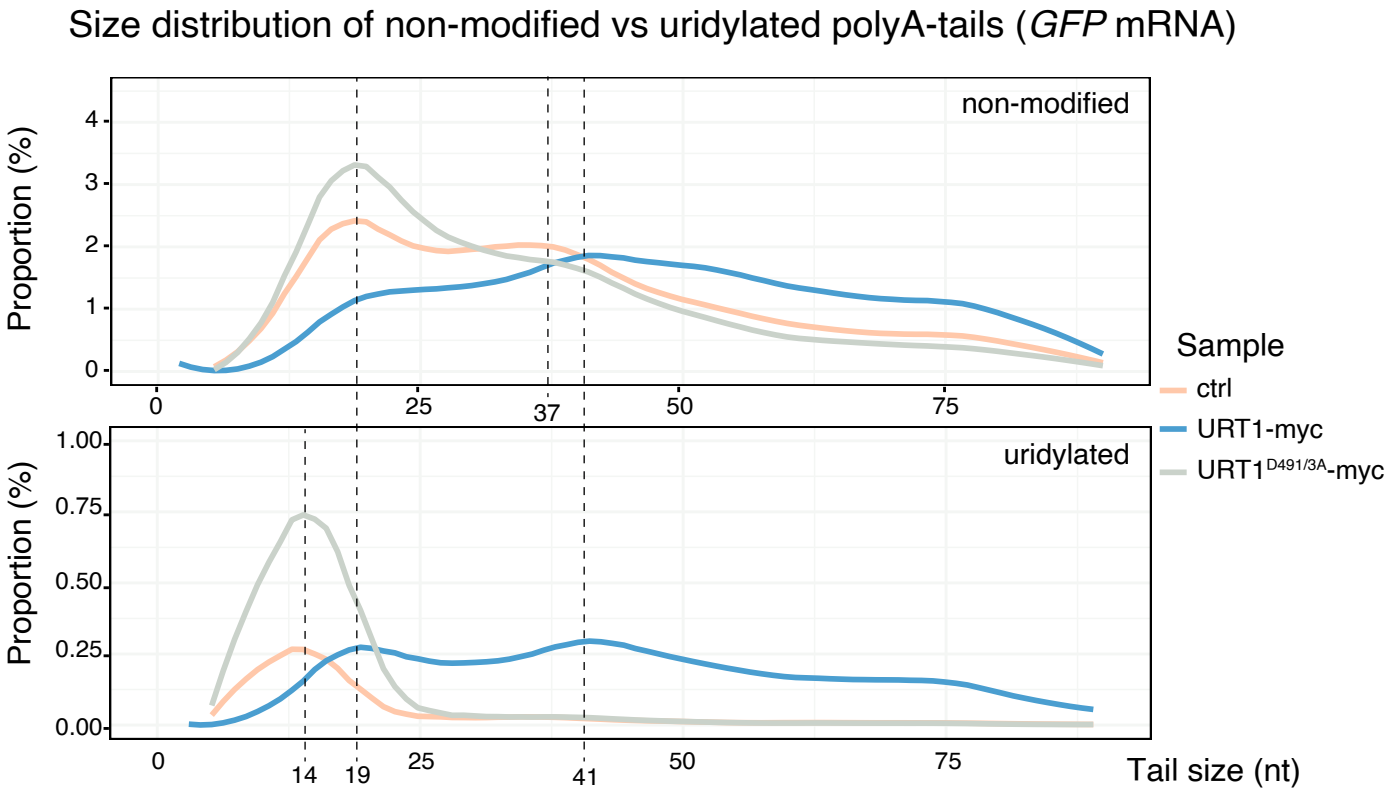


Figure 19: Size distribution of unmodified polyA tails and uridylated polyA tails.
The 3'RACE seq data obtained from six biological replicates for each of the control, URT1-myc and URT1$^{D491/3A}$-myc experiments in 2 independent Miseq runs were merged and are represented as line charts. The lines were smoothed using a Loess regression method. The top panel shows the size distribution of non-modified polyA tails. The bottom panel shows the size distribution of uridylated tails comprising both the homopolymeric A-stretch and the added uridines.

nucleotides, while the proportion of transcripts with short polyA tails was clearly decreased (Figure 18).

Taken together, the analysis of the polyA tails size distribution indicated that the accumulation of reporter transcripts with long polyA tails is linked to the catalytic activity of URT1. By contrast, the overexpression of URT1$^{D491/3A}$-myc does not only affect uridylated transcripts, but has a more general effect on transcripts with short polyA tails. Both the accumulation of uridylated transcripts and the accumulation of deadenylated transcripts are signs of compromised mRNA degradation, which I discuss later in this chapter.

2.1.3. The polyA size distribution is modulated by URT1-mediated uridylation.

In order to understand whether uridylation *per se* influences the size of mRNA polyA tails, we next compared the size distribution of uridylated tails *versus* non-modified tails. As the profiles obtained from 6 biological replicates were highly similar, the data were pooled and plotted as line charts (Figure 19). In control samples, the peak for the size distribution of unmodified polyA tails was at 19 nucleotides. A second, but smaller peak was detected at about 37 nucleotides and longer polyA tails were readily detected. In URT1-myc samples, the population of unmodified polyA tails of about 19 nucleotides was clearly decreased. Instead, the main peak of the size distribution was detected at 41 nucleotides, and the population of transcripts with even longer polyA tails was clearly increased. The profiles of unmodified polyA tails in samples expressing URT1$^{D491/3A}$-myc confirmed the accumulation of transcripts with polyA tails of about 19 nucleotides.

In both control and URT1$^{D491/3A}$-myc samples, almost all uridylated tails were shorter than 25 nucleotides and the majority of the uridylated tails were only about 14 nucleotides (*i.e*. mostly 12-13 adenosines with 1-2 uridines). This observation is in line with previous results obtained in Arabidopsis, which showed that in a wild-type situation, uridylation occurs predominantly on short polyA tails (Sement et al., 2013; Zuber et al., 2016). The profiles of uridylated polyA tails in samples expressing URT1$^{D491/3A}$-myc revealed a pronounced increase of transcripts with short uridylated polyA tails of *circa* 14 nucleotides. Thus, the higher frequency of uridylated *GFP* mRNA in the URT1$^{D491/3A}$-myc samples is due to the accumulation of this uridylated polyA tail population. The main peak of the population with short uridylated tails was shifted from 14 nucleotides in the controls and URT1$^{D491/3A}$-myc to 19 nucleotides in the URT1-myc samples. Furthermore, a second peak appeared at about 41 nucleotides, and longer uridylated polyA tails were abundantly detected.

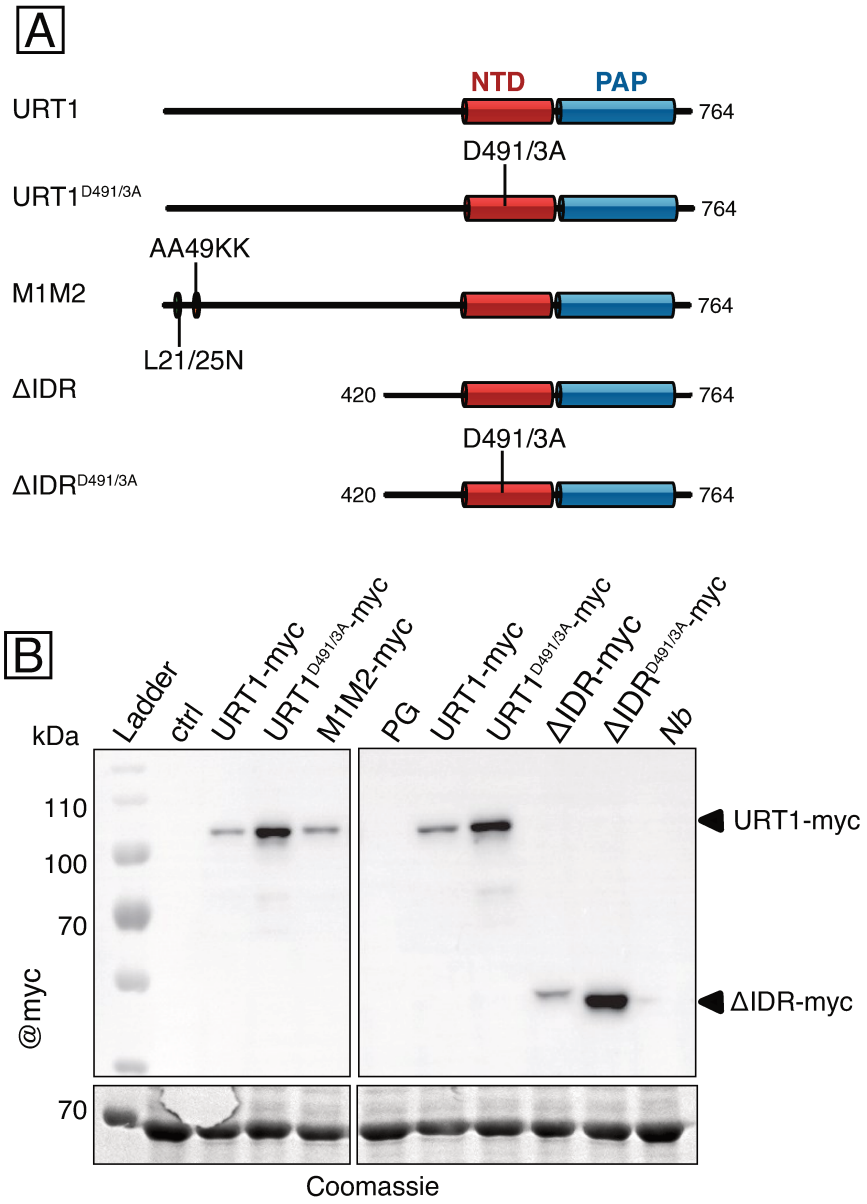


Figure 20: Expression of additional M1M2-myc, ΔIDR-myc and ΔIDR$^{D491/3A}$-myc constructs in *N. benthamiana*.

A) Diagrams illustrating the domain organisations and mutations of the different URT1-myc constructs. NTD: polymerase β-like nucleotidyltransferase domain; PAP: PolyA polymerase-associated domain; IDR, intrinsically disordered region. M1M2 refers to two short conserved sequence motifs in the IDR.

B) Western bot analysis of the protein extracts (rep2 and 3) obtained from the different leave patches harvested four days after infiltration (4 DAI). Monoclonal @myc antibodies were used for the detection of the URT1-myc constructs. The coomassie stain of the membrane is shown as loading control.

ctrl: control patches infiltrated only with P19 and GFP; *Nb*: *Nicotiana benthamiana* uninfiltrated leaves

Taken together, these results confirmed that the expression of URT1-myc results in the accumulation of uridylated and non-uridylated transcripts with longer polyA tails, suggesting that uridylation by URT1 can influence the overall polyA tail length. Two hypotheses which are not necessarily mutually exclusive can explain the shift towards longer tails that we observe for both unmodified and uridylated polyadenylated *GFP* transcripts:

A) Uridylation by URT1-myc protects polyA tails from deadenylation and degradation. The population of transcripts with long uridylated and non-uridylated tails increases because uridylation slows down deadenylation.

B) Uridylation by URT1-myc induces the rapid elimination of transcripts with short polyA tails, for example by recruiting degradation factors to this mRNA polyA tail population.

Both possibilities were experimentally investigated. The results are presented in Results part 2.4. and 3.)

2.2. Functional implication of the intrinsically disordered region of URT1 in the turnover of short polyA tails.

As detailed in the introduction, URT1 has a long intrinsically disordered region (IDR) interspersed with short conserved motifs in its N-terminus. Previous results obtained in the group indicated that the M1 and M2 motifs of the IDR are conserved in all URT1 orthologs of the green lineage. (Ferrier, 2013; Scheer, 2018). Additionally, Hélène Scheer showed during her PhD that the M1 motif is required for the interaction of recombinant URT1 with the decapping factor DCP5. The function of the M2 motif is still unknown.

To assess whether and how the N-terminal IDR and the M1 and M2 motifs influence the polyadenylation and uridylation profiles of reporter *GFP* mRNAs, we performed other series of infiltration experiments using additional mutant versions of URT1-myc (Figure 20A). The construct M1M2-myc carries mutations in both conserved short motifs. In the M1 motif, two leucines at position 21 and 25 were changed to asparagines (L21/25N). In the M2 motif, two alanines at position 49/50 were changed to lysines (AA49KK). ΔIDR-myc and $\Delta IDR^{D491/3A}$-myc are active and inactive versions of URT1-myc lacking the entire N-terminal IDR domain (Figure 20A). All constructs were co-expressed with the reporter *GFP* mRNA and P19 in *N. benthamiana* as described above. Three independent replicates were sequenced for the control, URT1-myc and

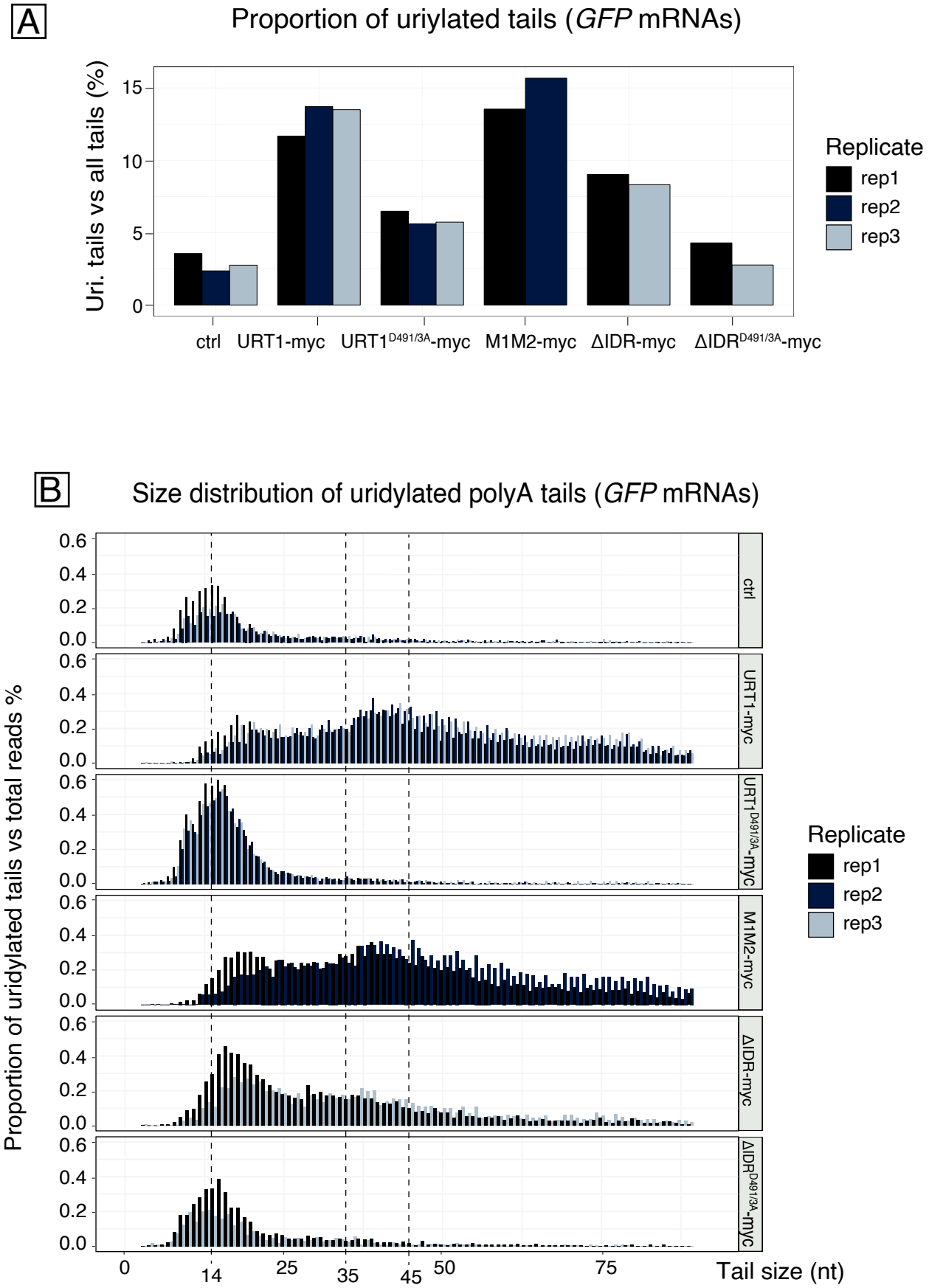


Figure 21: The N-terminal IDR is required for the accumulation of transcripts with short uridylated tails upon overexpression of URT1$^{D491/3A}$.

A) The bar graph shows the proportion of uridylated polyA tails of the *GFP* reporter mRNA determined by 3' RACE-seq in 3 biological replicates.

B) The histogram shows the size distribution of the uridylated polyA tails. The accumulation of transcripts with short tails of around 14 nucleotides induced by URT1$^{D491/3A}$ is not observed upon expression of ΔIDR$^{D491/3A}$.

URT1$^{D491/3A}$-myc samples, and the M1M2-myc sample has been sequences for replicates 1 and 2, whereas active and inactive ΔIDR-myc samples were sequences for replicates 1 and 3.

Similar to the first experiment, the catalytic inactive URT1$^{D491/3A}$-myc and ΔIDR$^{D491/3A}$-myc proteins were expressed at higher levels compared to the active versions URT1-myc, M1M2-myc and ΔIDR-myc (Figure 20B).

2.2.1. The removal of the IDR reduces uridylation levels of reporter *GFP* mRNA.

The 3'RACE-seq analysis revealed that the expression of URT1-myc and M1M2-myc resulted in similar proportions of uridylated *GFP* reporter mRNAs of around 15% (Figure 21A). Hence, the mutations in the M1 and M2 motifs do not reduce the catalytic activity of URT1-myc. By contrast, the proportion of uridylated polyA tails was reduced to 10% in the ΔIDR-myc sample. The mutation of the activity residues in ΔIDR$^{D491/3A}$-myc led to a further reduction of uridylated polyA tails to the residual levels observed in the control samples (Figure 21A). Previous enzyme assays had shown that the N-terminal part of URT1 is dispensable for its *in vitro* uridylation activity (Ferrier, 2013). However, as compared to full-length URT1-myc, the reduced proportion of uridylated tails detected upon expression of ΔIDR-myc strongly suggested that the IDR in the N-terminal part of URT1 has an important function *in vivo*. For instance, the IDR could be important for the recruitment of URT1 to the messenger RNA, or implicated in an interaction with a protein partner that positively stimulates the activity of URT1.

This hypothesis was further supported by the size distributions of the uridylated polyA tails (Figure 21B). The profiles detected in samples expressing URT1-myc or the catalytic mutant URT1$^{D491/3A}$-myc were identical to those observed in the previous experiment, *i.e.* a high proportion of tails with 35-45 nucleotides in URT1-myc and a strong accumulation of transcripts with short tails of about 14 nucleotides in URT1$^{D491/3A}$-myc (compare 21B to 19). The profiles for uridylated polyA tails in M1M2-myc samples resembled the profiles detected in URT1-myc samples, *i.e.* we observed the accumulation of transcripts with long uridylated polyA tails and a reduction of the population of transcripts with short uridylated polyA tails (Figure 21B). By contrast, the size distribution of uridylated polyA tails in ΔIDR-myc samples displayed an intermediate profile. As compared to control and URT1$^{D491/3A}$-myc samples, expression of ΔIDR-myc resulted in a mild accumulation of reporter transcripts with short uridylated polyA tails as well as a mild increase in the proportion of transcripts with long uridylated polyA tails. Finally, the size distribution profile in ΔIDR$^{D491/3A}$-myc samples did not show the pronounced accumulation of short uridylated tails

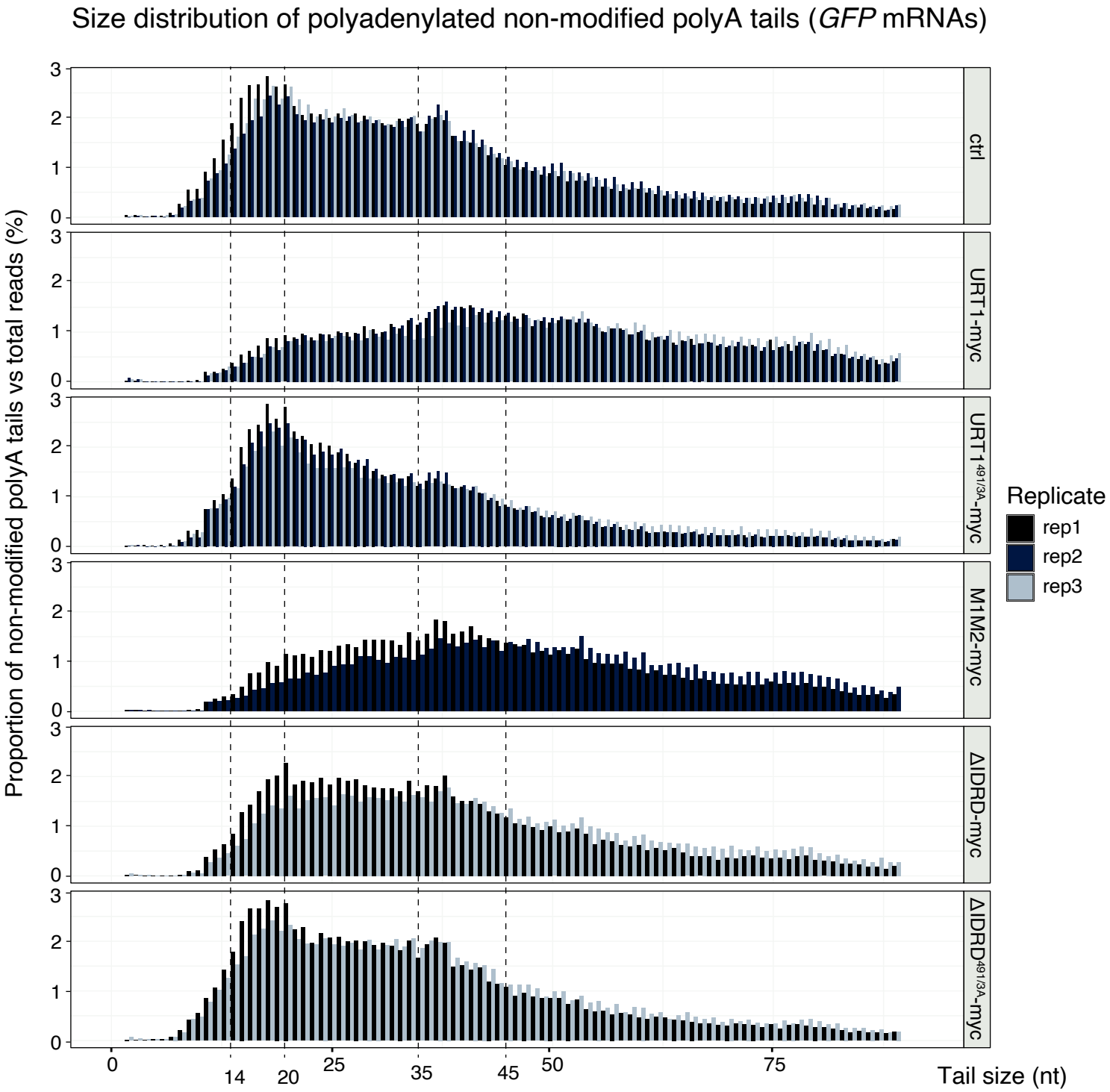


Figure 22: The N-terminal IDR is important for the regulation of polyA tail sizes by URT1. The bar graph shows the proportion of non-modified polyA tails of the *GFP* reporter mRNA determined by 3' RACE-seq in 3 biologial replicates. Samples expressing ΔIDR-myc had reduced levels of transcripts with long polyadenylated tails as compared to URT1-myc.

observed in URT1$^{D491/3A}$-myc samples, but was identical to the profiles observed in the control samples (Figure 21B). Because ΔIDR$^{D491/3A}$-myc and URT1$^{D491/3A}$-myc proteins were expressed at similar levels (Figure 20B), these results indicate that the intrinsically disordered region of URT1$^{D491/3A}$-myc is responsible for the accumulation of short uridylated polyA tails observed upon massive overexpression of catalytic inactive URT1.

2.2.2. The N-terminal IDR is important for the regulation of polyA tail sizes by URT1.

Next, I analysed the size distribution of homopolymeric, non-modified polyA tails (Figure 22). In control samples the largest proportion of tails were detected between 14-45 nucleotides. Expression of URT1-myc resulted in reduced levels of transcripts with short polyA tails of less than 35 nucleotides, and in higher levels of transcripts with long polyA tails of more than 45 nucleotides. By contrast, samples expressing URT1$^{D491/3A}$-myc accumulated transcripts with short polyA tails of 14-20 nucleotides and had low levels of transcripts with long polyA tails. As observed for uridylated tails, the size distribution of non-modified polyA tails was almost identical in the samples expressing the catalytic active versions URT1-myc or M1M2-myc. However, samples expressing ΔIDR-myc had reduced levels of transcripts with long polyadenylated tails (Figure 22). The size distribution of non-modified tails in samples expressing ΔIDR$^{D491/3A}$-myc was identical to the control (Figure 22).

Taken together, these data indicated that the IDR is important for the regulation of polyA tail sizes by URT1. Interestingly, the IDR is required for the accumulation of transcripts with short tails that is observed upon overexpression of catalytically inactive URT1, regardless of the uridylation status of the tails.

2.2.3. *GFP* mRNAs with A-rich tails and G/A or U/A extensions accumulate when inactive forms of URT1-myc are overexpressed.

When we analysed the first 3' RACE datasets, we noticed that a significant proportion of the tails were A-rich and contained a homopolymeric polyA tail stretch followed by mostly G/A and U/A rich extensions (Figure S2A). The proportion of mRNAs with such A-rich tails was typically 7% in control samples and 13% in samples expressing URT1-myc (Figure 23A). Interestingly, we observed a pronounced accumulation of *GFP* mRNAs with A-rich tails in samples expressing the inactive full-length version URT1$^{D491/3A}$-myc (Figure 23A). Up to 22% of all tails were A-rich in these samples.

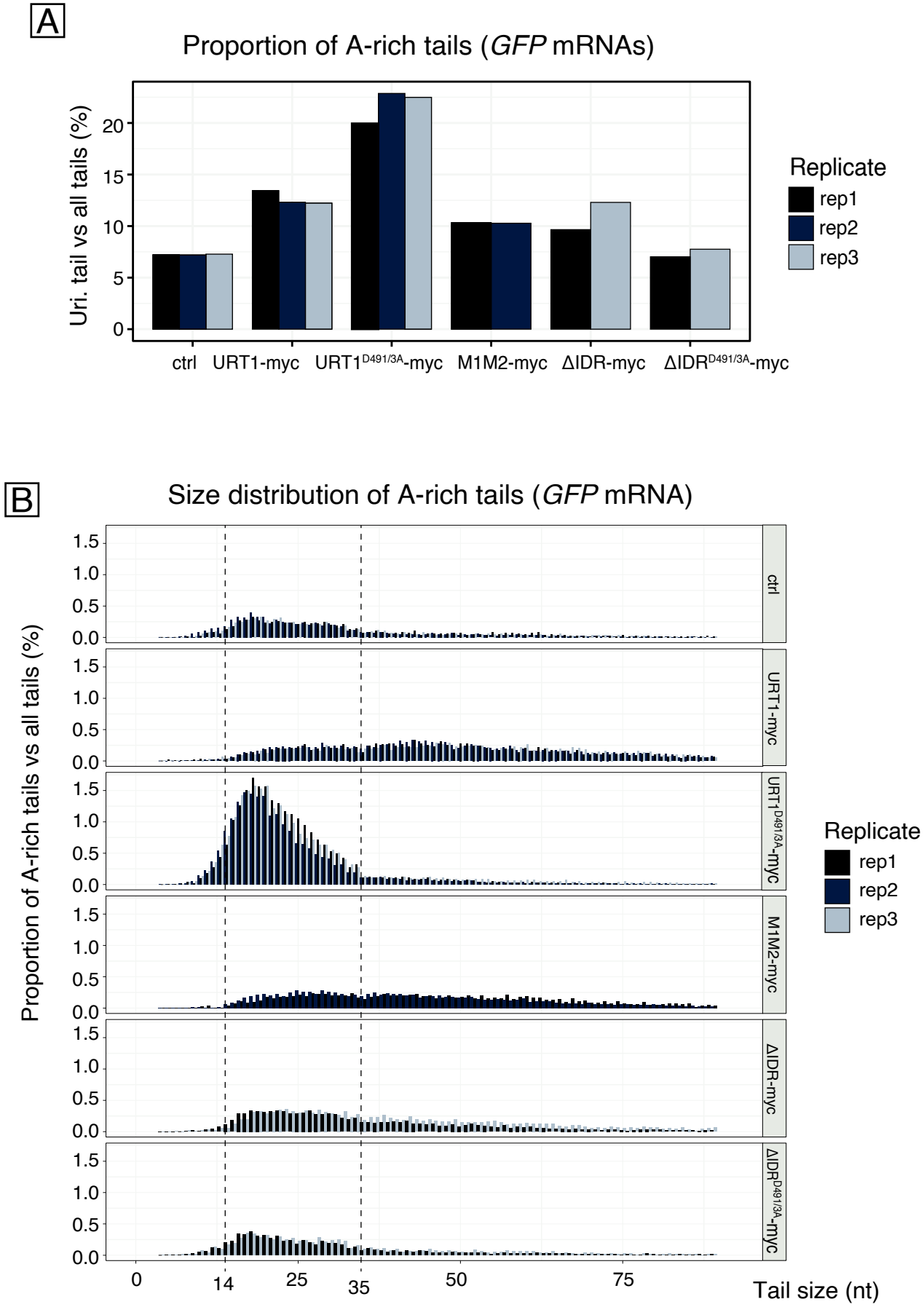


Figure 23: Expression of URT1$^{D491/3A}$-myc leads to accumulation of short A-rich tails. The bar graph shows the proportion of A-rich tails of the *GFP* reporter mRNA determined by 3'RACE-seq in 3 biological replicates. Samples expressing URT1$^{D491/3A}$-myc have an important accumulation of A-rich tails of 14-35 nucleotides.

By contrast, samples expressing ΔIDR$^{D491/3A}$-myc had 7% of A-rich tails, similar to controls (Figure 23A).

The overall size distribution of A-rich tails was similar to the size distribution of uridylated tails (Figure 23B compared to Figure 21B). In controls, the majority of the tails were 14-35 nucleotides, and expressing catalytically active URT1-myc and M1M2-myc resulted in increased proportions with longer tails of more than 35 nucleotides (Figure 23B). As seen for uridylated tails, the expression of ΔIDR-myc results in a slightly reduced level of long polyA tails under 35 nucleotides long. Interestingly, if we compared the two catalytic mutants URT1$^{D491/3A}$-myc and ΔIDR$^{D491/3A}$, the removal of the N-terminal region led to a drastic change in the distribution of these A-rich tails (Figure 23B). URT1$^{D491/3A}$-myc expression resulted in a vast accumulation of transcripts with A-rich tails of about 18 nucleotides. The expression of ΔIDR$^{D491/3A}$-myc however resulted in an important decrease in the proportion of these short A-rich tails (Figure 23B). These results are very similar to what we observed for both short uridylated tails (Figure 21) and short non-modified tails (Figure 22) and suggest that overexpression of the IDR affects the turnover of mRNAs with tails of *circa* 18 nucleotides.

A plausible explanation for the production of A-rich tails is that both deadenylated and uridylated tails can become substrates of terminal nucleotidyltransferases of *N. benthamiana* with other nucleotide specificities if they are not rapidly degraded. The absence of short A-rich tails in samples expressing catalytic active versions of URT1-myc is in line with the hypothesis that (1) uridylation by URT1 protects long polyA tails from degradation and thus prevents the production of shorter tails or (2) uridylation by URT1 favours the degradation of mRNAs with short tails. The pronounced accumulation of A-rich tails in URT1$^{D491/3A}$-myc suggests that the degradation of deadenylated transcripts is impeded in these samples, which renders such tails more accessible to other TNTases. The fact that they are accessible also suggests that URT1$^{D491/3A}$-myc is not protecting deadenylated transcripts by binding to their 3' ends. In light of the antagonistic effect of IDR$^{D491/3A}$-myc, a more likely explanation is that the degradation of transcripts with short polyA tails requires a factor that is sequestered by the binding to the IDR upon massive overexpression of URT1$^{D491/3A}$-myc.

2.2.4. Uridylation by URT1 does not destabilise the *GFP* reporter mRNA.

To examine the effect of URT1-myc, URT1$^{D491/3A}$-myc, M1M2-myc, ΔIDR-myc and ΔIDR$^{D491/3A}$ expression on the steady-state levels of the *GFP* reporter mRNA, I analysed total RNA samples by

Steady-state level of reporter *GFP* mRNAs

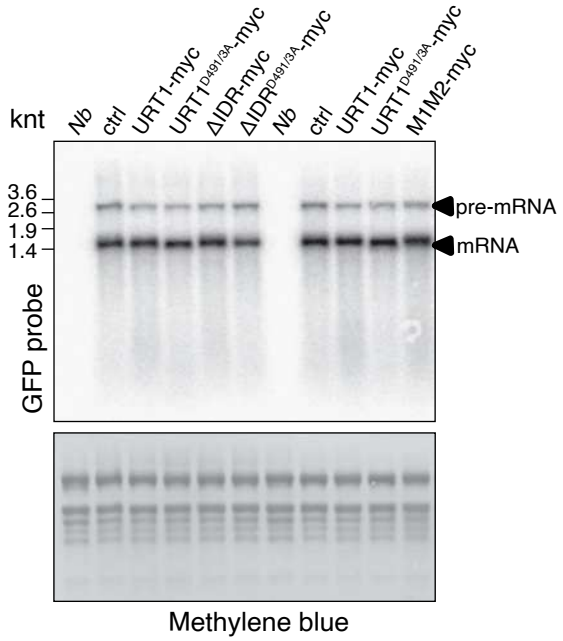


Figure 24: Northern blot analysis of reporter *GFP* mRNAs.
Northern blot analysis of total RNA (rep 2 and 3) hybridised with a probe specific to the GFP-encoding regions of the reporter *GFP* mRNAs. The arrows indicate unspliced and mature forms of the *GFP* mRNA. A size marker is indicated at the left. The methylene blue stain of the membrane is shown as loading control. The size shift that is observed between samples expressing active and inactive versions of URT1 (but not in Δ IDR$^{D491/3A}$-myc) reflects the virtual absence of *GFP* mRNAs with long polyA tails and the accumulation of deadenylated mRNAs with short polyA tails of 14-35 nt that we observed by 3' RACE-seq (compare to Figure 23).

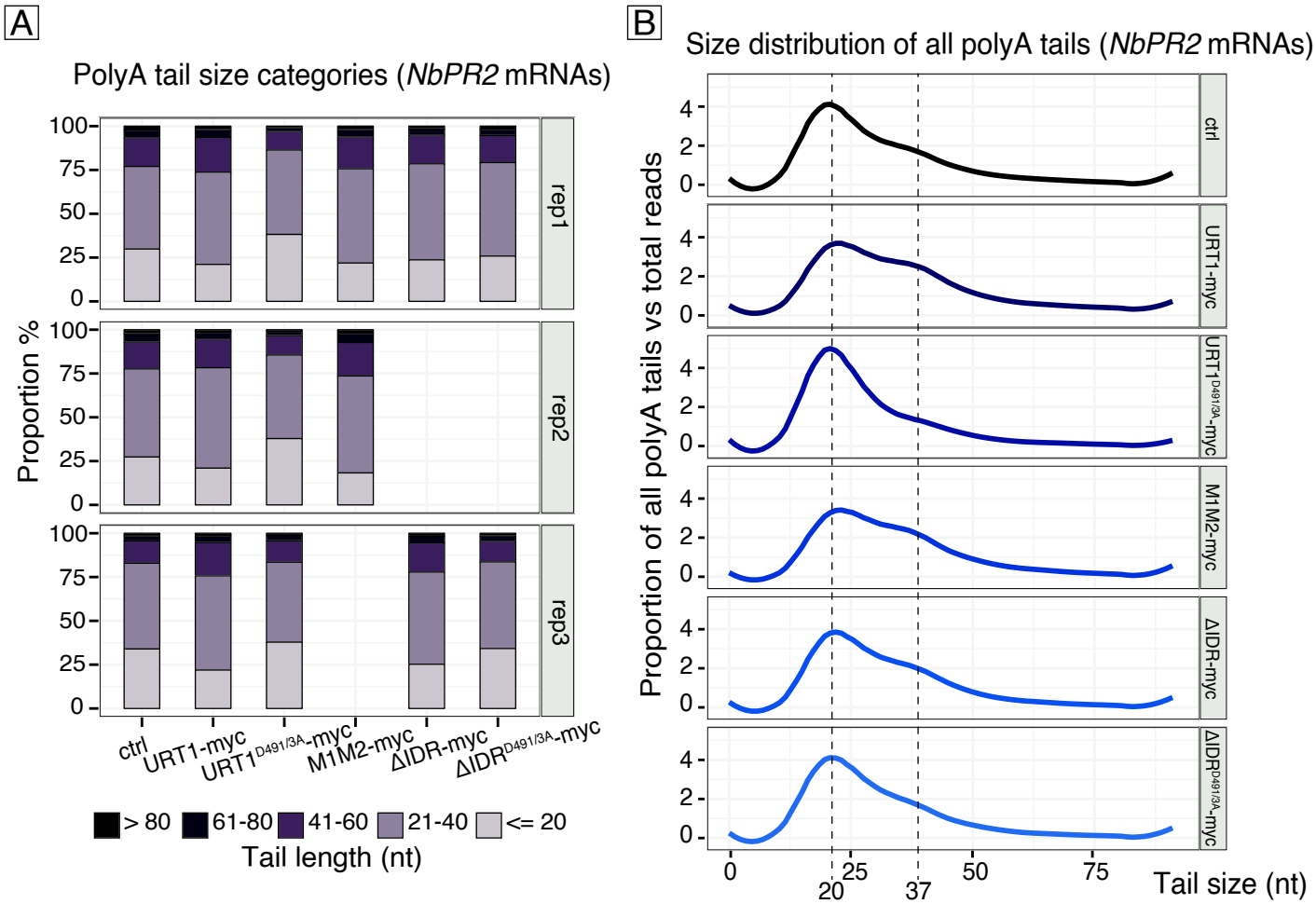


Figure 25: Size distribution of all tails on *NbPR2* mRNA.
A) Proportion of polyA tail length below 20 nucleotides, between 21 and 40 nucleotides, 41 and 60 nucleotides, 61 and 80 nucleotides, and over 80 nucleotides.
B) PolyA size distributions of all tails. The 3'RACE-seq data obtained from three biological replicates were merged and are represented as line charts. The lines were smoothed using a Loesss regression method.

northern blot (Figure 24). This revealed that the steady state levels of *GFP* mRNAs were similar in all samples. Hence, the drastic changes in the polyA tail size distribution or uridylation levels do not impact the steady-state levels of the mature *GFP* messenger RNAs. This suggests a fast and efficient homeostasis of *GFP* mRNAs production or that URT1 is not important to define the *GFP* mRNA's half-live in our transient expression system. Interestingly, *GFP* mRNAs have a slightly reduced apparent size in both URT1$^{D491/3A}$-myc samples, which likely reflects the noticeable accumulation of transcripts with short tails in these samples.

2.3. Overexpression of active and inactive versions of URT1 impacts tail size and uridylation levels of endogenous *PR2* mRNAs in *Nicotiana benthamiana*.

We also examined the effects of URT1-myc expression on an endogenous mRNA of *N. benthamiana*. We chose to analyse the mRNA of the *Niben101Scf04869g03002.1* gene encoding the pathogenesis related protein 2 (PR2). The PR2 protein is a β-1,3- glucanase required for the degradation of bacterial cell walls during bacterial infections of plants. Interestingly, PR2 expression is induced upon agro-infiltration. The 3' ends of *NbPR2* mRNAs of 3 independent replicates were sequenced in parallel with the 3' ends of the *GFP* reporter transcript in the MiSeq Run1. As previously, M1M2-myc samples have been sequenced for replicate 1 and 2 and the ΔIDR-myc samples for replicates 1 and 3. For each sample, we obtained more than 50 000 total reads for *NbPR2* mRNAs (Figure S3).

2.3.1. Expression of URT1 leads to accumulation of *NbPR2* transcripts with longer polyA tails.

An important observation is that overall, *NbPR2* mRNAs had shorter polyA tails than the reporter *GFP* transcript. In control samples, 75% of all *NbPR2* mRNAs had tails shorter than 40 nucleotides (Figure 25A). However, we observed similar effects on the polyA tail size distribution as we observed for the GFP-reporter mRNA, although the effects were less pronounced. The proportion of *PR2* transcripts with polyA tails longer than 40 nucleotides increased upon expression of URT1-myc, and decreased upon expression of URT1$^{D491/3A}$-myc. The representation of the same data in line charts (Figure 25B, for distribution of the individual replicates, see supplementary Figure S4) revealed a peak in the tail size distribution at 20 nucleotides in all samples. A second, but smaller peak at 37 nucleotides was observed, particular in URT1-myc and M1M2-myc samples. The

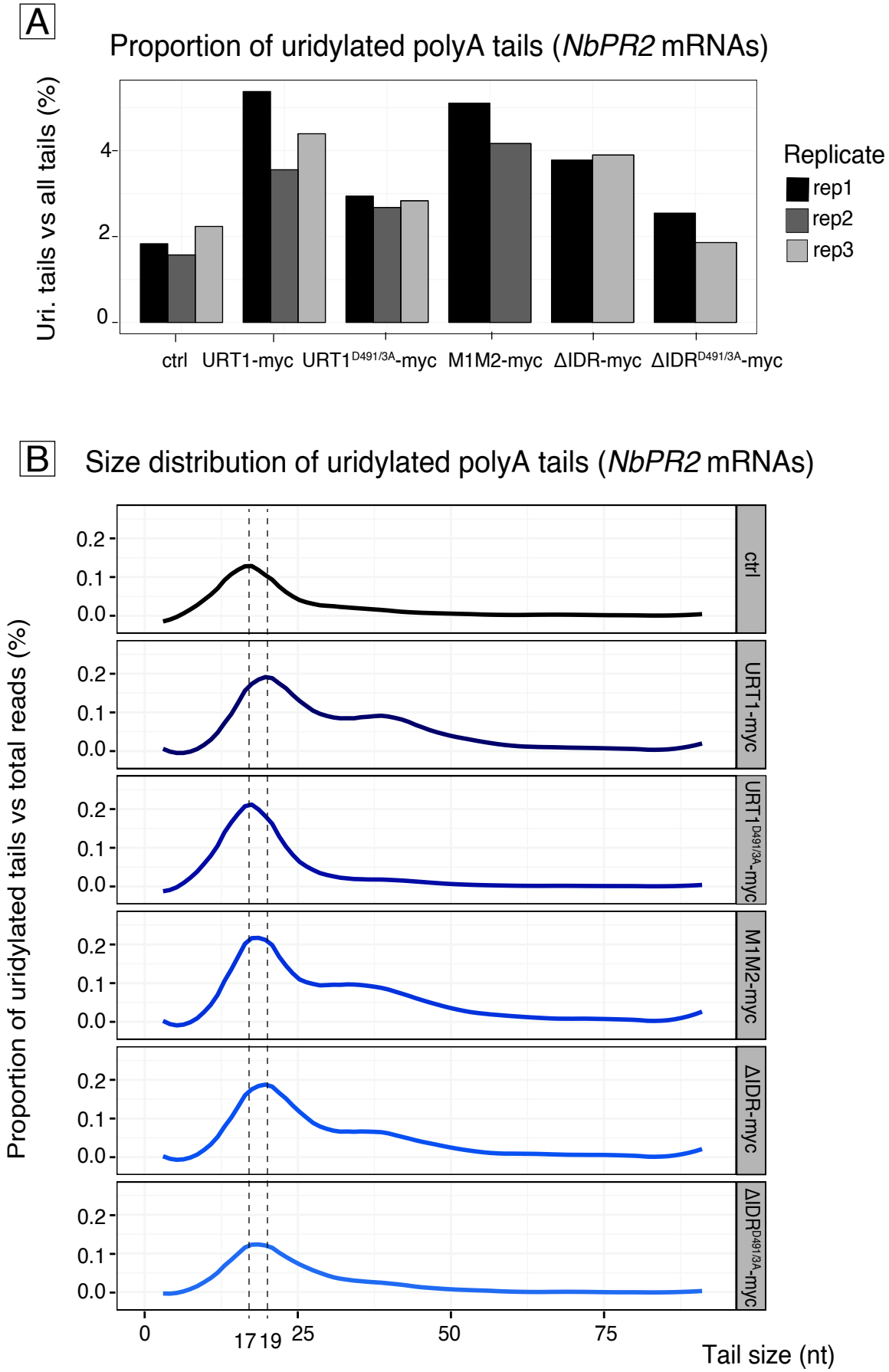


Figure 26: Level of uridylated tails and tail size distribution of uridylated tails of *NbPR2* transcripts.

A) Proportion of uridylated tails of *NbPR2* mRNAs.

B) Size distribution of uridylated polyA tails of *NbPR2*.

expression of URT1$^{D491/3A}$-myc reduced the number of *PR2* transcripts with longer polyA tails and lead to the accumulation of transcripts with tails of mainly 20 nucleotides. The expression of truncated ΔIDR$^{D491/3A}$-myc supressed this accumulation of transcripts with short polyA tails and the profiles resembled to the once observed for control samples (Figure 25B). Taken together, these data confirmed that the expression of active versions of URT1-myc increased the polyA size distribution of the *PR2* transcripts and that the overexpression of inactive URT1$^{D491/3A}$-myc stabilised the polyA tail population of 20 nucleotides.

2.3.2. Expression of URT1 increases the uridylation level of endogenous *PR2* transcript.

Next, we examined the proportion of uridylated *PR2* mRNAs. In control samples, 2% of *PR2* transcripts were uridylated, hence the basal level of uridylation was similar as for the *GFP* reporter transcript (Figure 26A). Expression of the catalytic active URT1-myc, M1M2-myc or ΔIDR-myc constructs increased the proportion of uridylated *PR2* mRNAs to approximately 4%. This indicates that the *PR2* transcript is less frequently uridylated by URT1 than the *GFP* reporter transcript. We cannot exclude that the uridylated *PR2* mRNAs may have a faster turnover. Similar to *GFP* mRNAs, the overexpression of URT1$^{D491/3A}$-myc reduced the uridylation level, without reaching the basal level that is detected in the control samples (Figure 26A). The uridylation levels in ΔIDR$^{D491/3A}$-myc however were similar to the control samples. The analysis of the size distribution profiles (Figure 26B, Figure S5) revealed a main peak of 17 nucleotides in the control samples. In URT1-myc samples, this peak was shifted to 19 nucleotides and the number of transcripts with longer uridylated tails (around 38 nucleotides) increased. Both, the shift of the main peak and the increase in the population of transcripts with long uridylated tails, were also visible in M1M2-myc or ΔIDR-myc samples. By contrast, URT1$^{D491/3A}$-myc expression induced the accumulation of *PR2* mRNAs with short uridylated tails of mostly 17 nucleotides (Figure 26B). This accumulation was not observed upon expression of ΔIDR$^{D491/3A}$-myc. These data confirmed that the N-terminal IDR is required for the accumulation of *PR2* mRNAs with short uridylated tails that is observed upon massive overexpression of URT1$^{D491/3A}$-myc.

2.3.3. Overexpression of URT1$^{D491/3A}$-myc results in increased levels of *PR2* mRNAs with A-rich tails.

As for the *GFP* reporter mRNA, many of the polyA tails of *PR2* mRNAs were classified as A-rich tails with GA and UA rich extensions (Figure S3). In control samples, 7% of all reads corresponded to

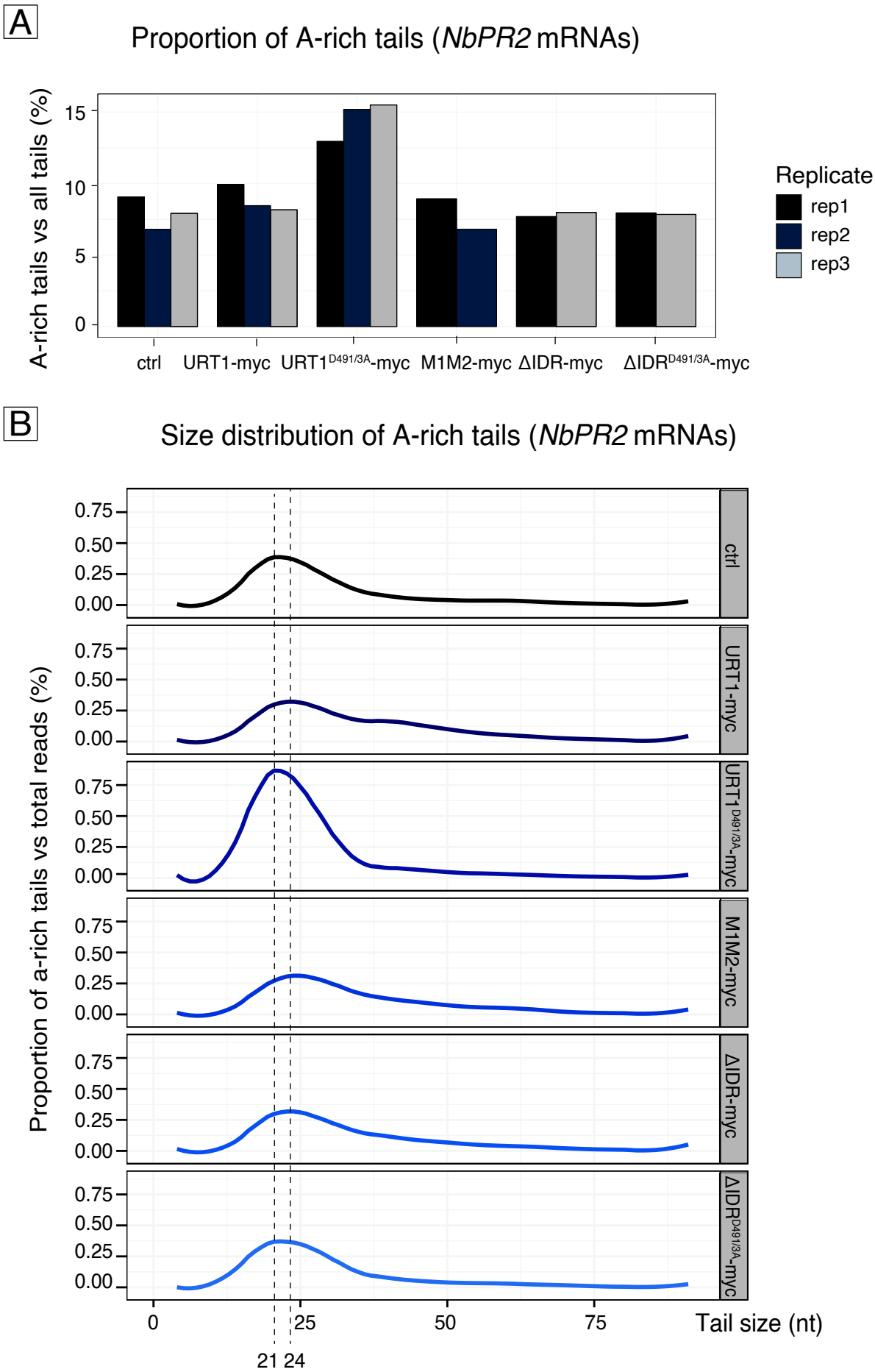


Figure 27: Size distribution and proportion of A-rich tails of *NbPR2* mRNA.
A) Proportion of A-rich tails of *NbPR2* mRNA.
B) Size distribution of A-rich tails. Short A-rich tails of *circa* 21 nucleotides accumulate when inactive URT1$^{D491/3A}$-myc is expressed.

such A-rich tails (Figure 27A). Similar proportions were detected in samples expressing URT1-myc, M1M2-myc, or ΔIDR-myc. By contrast, samples expressing URT1$^{D491/3A}$-myc had increased proportions of A-rich tails of up to 15% (Figure 27A). This accumulation was supressed by removal of the IDR (ΔIDR$^{D491/3A}$). The size distribution analysis revealed a main peak at about 21 nucleotides in the control and at 24 nucleotides in samples expressing active versions of URT1-myc (Figure 27B and Figure S6 for individual replicates). *PR2* transcripts with short A-rich tails of mostly 21 nucleotides accumulated in URT1$^{D491/3A}$-myc, but not in IDR$^{D491/3A}$-myc.

To conclude, the RACE-seq analysis of endogenous *PR2* mRNAs retrieved similar results as the analysis of the *GFP* reporter mRNAs. For both *GFP* and *PR2* mRNAs, uridylation is mainly detected on short polyA tails of 15-20 nucleotides in the controls and inactive samples. The observation that the main peak of the size distribution is at 14 nucleotides for *GFP* mRNAs and at 17 nucleotides for *PR2* mRNAs underlines that the peaks in the size distribution of deadenylated tails can be different for each transcript. However, longer uridylated tails are readily detected and their proportions largely increase upon overexpression of URT1-myc. Importantly the overexpression of active URT1-myc also affects the length of unmodified polyA tails. For both *GFP* and *PR2* transcripts, URT1-myc expression increases the proportion with long polyA tails and reduced the proportions with short tails. This observation is in line with the idea that uridylation protects or slows down the deadenylation of mRNAs.

The accumulation of mRNAs with uridylated and A-rich tails in the URT1$^{D491/3A}$-myc sample is most likely linked to the high expression levels of the URT1$^{D491/3A}$-myc protein (Figure 20B). One possible explanation for the accumulation of transcripts with short tails is that the inactive URT1 protein binds to *GFP* and *PR2* mRNAs and protects the 3' extremities from degradation. However, the pronounced accumulation of transcripts with A-rich tails argues against this scenario, as the heterogenous extensions are most likely added when the turnover of the mRNAs is compromised. Hence, the increased levels of transcripts with short uridylated or A-rich could be due to the depletion of a degradation factor that binds to the IDR of URT1.

2.3.4. NbPR2 mRNAs accumulate in URT1$^{D491/3A}$-myc samples.

To examine the effect of URT1-myc, URT1$^{D491/3A}$-myc, M1M2-myc, ΔIDR-myc and ΔIDR$^{D491/3A}$-myc expression on the steady-state levels of *NbPR2* mRNAs, I analysed total RNA samples by northern blot (Figure 28). This revealed that the steady state levels of *NbPR2* mRNAs were similar in all samples except in the URT1$^{D491/3A}$-myc samples. We can clearly see an important

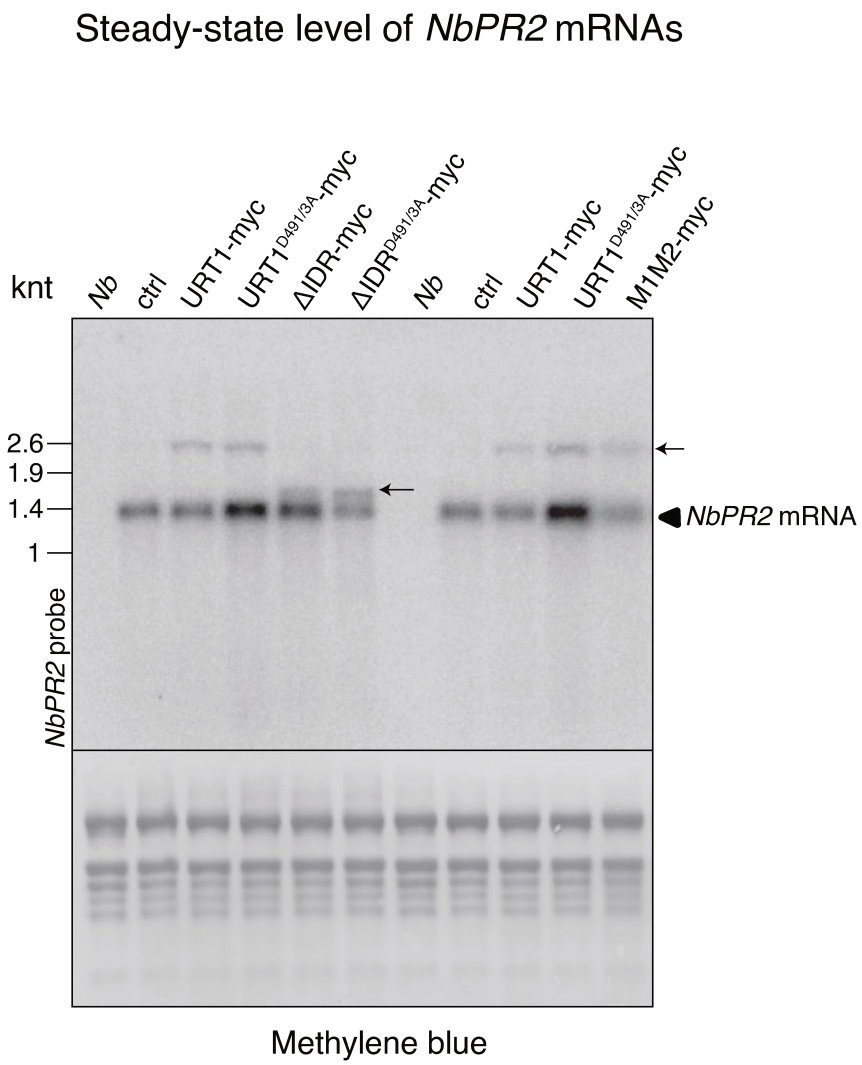


Figure 28: Northern blot analysis of reporter *NbPR2* mRNAs.
Northern blot analysis of total RNA (rep2 and rep3) hybridised with a probe specific to the *NbPR2* mRNAs. The arrows indicate mature forms of the *NbPR2* mRNA. A size marker is indicated at the left. The methylene blue stain of the membrane is shown as loading control. Of note, the membrane has been previously hybridised with a specific probe to the *URT1* mRNAs and stripped before hybridising with the *NbPR2* specific probe. The residual signals that are indicated by arrows correspond to the *URT1* mRNAs of the full-length and ΔIDR constructs.

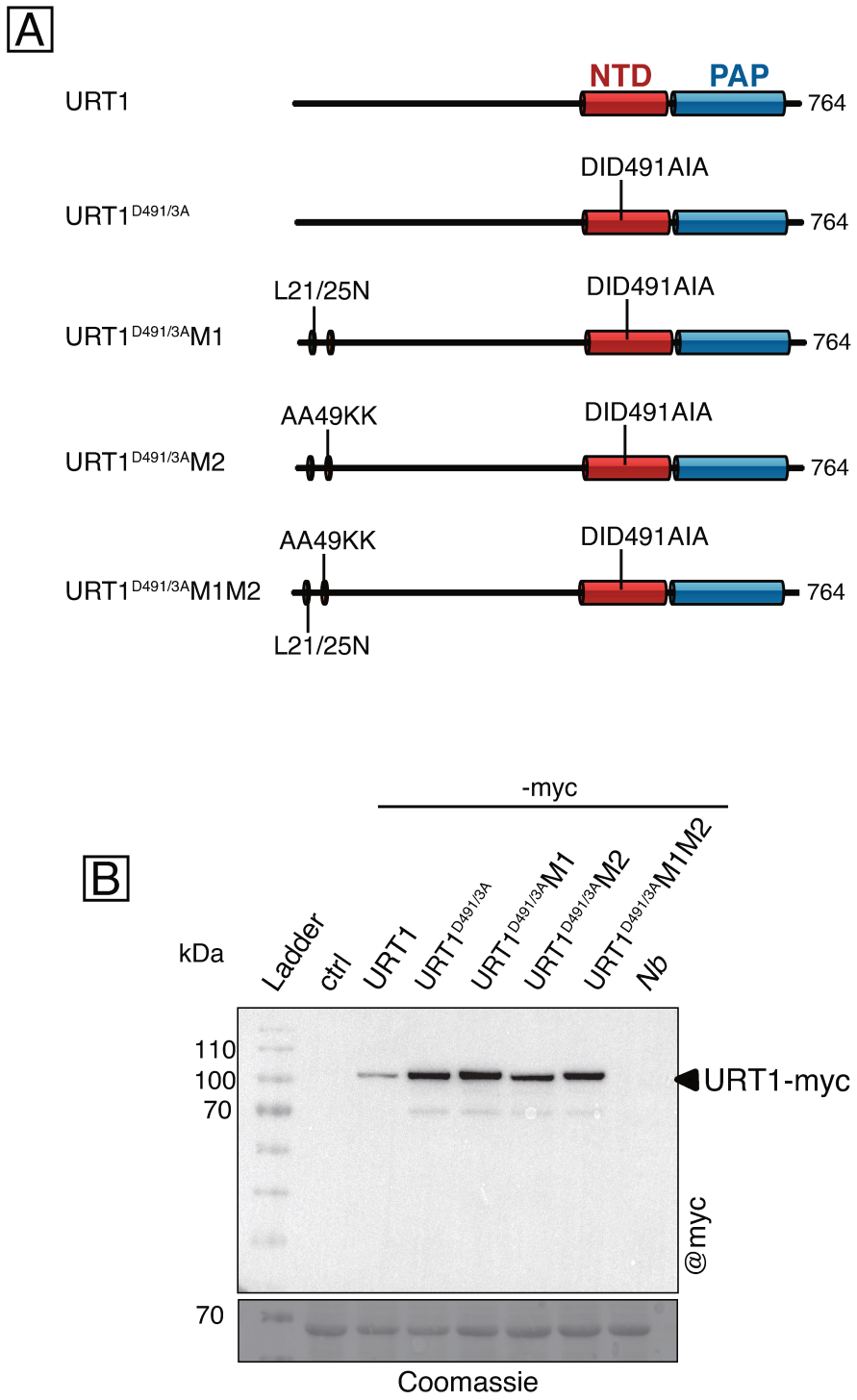


Figure 29: Transient expression of URT1$^{D491/3A}$-myc mutated in the M1 and/or M2 motifs in *N. benthamiana.*

A) Diagrams of the URT1 constructs designed to investigate the role of the M1 and M2 motifs for the accumulation of mRNAs with short polyA tails.

NTD: polymerase β-like nucleotidyltransferase domain; PAP: PolyA polymerase-associated domain.

B) Western bot analysis of the protein extracts obtained from the different patches harvested four days after infiltration (4 DAI). Wild-type and mutated versions of URT1-myc proteins were detected with monoclonal @myc antibodies. The Coomassie stain of the membrane is shown as loading control.

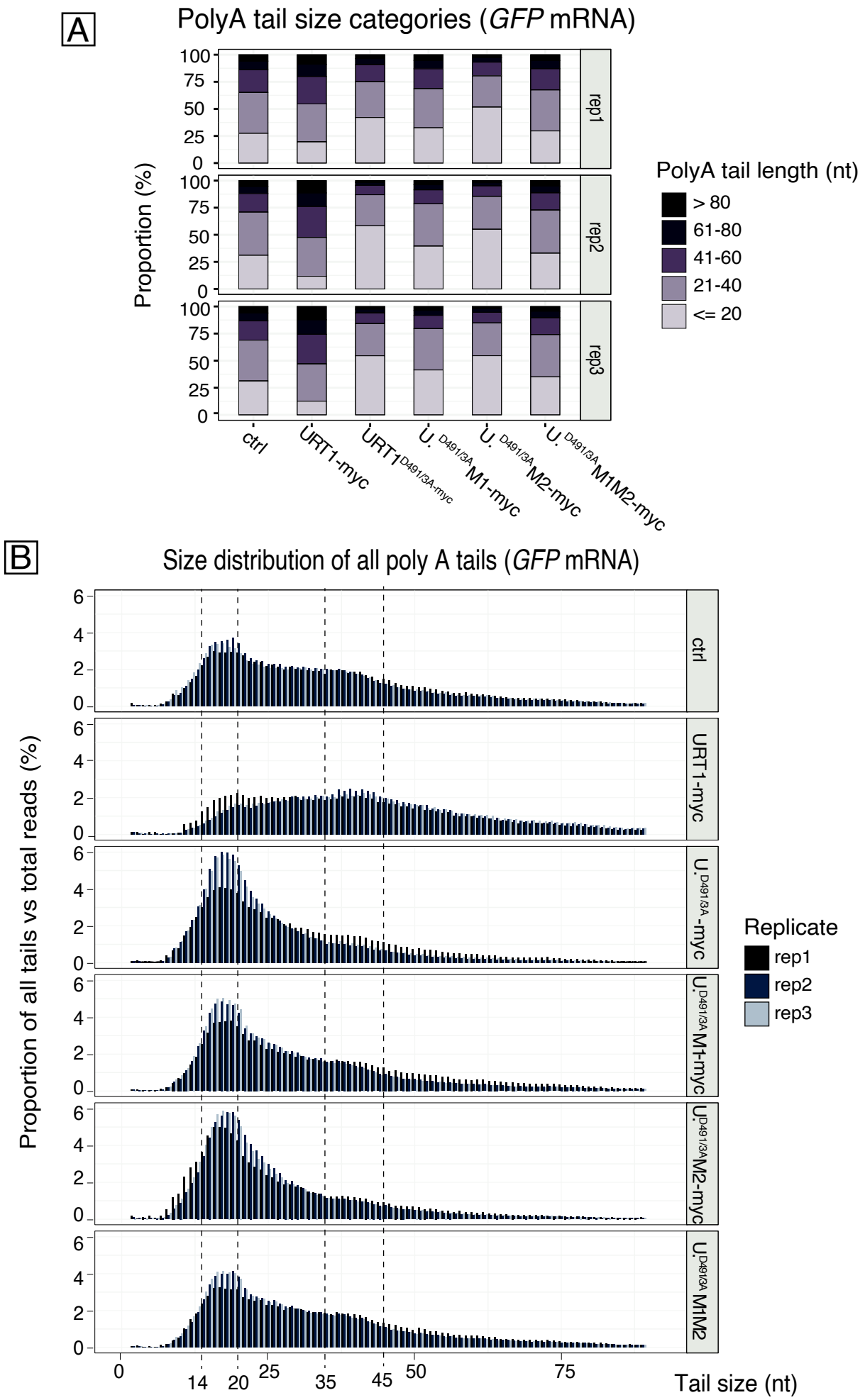


Figure 30: Size distribution of all tails of *GFP* mRNAs.

A) Proportion of polyA tails length below 20 nucleotides, between 21 and 40 nucleotides, 41 and 60 nucleotides, 61 and 80 nucleotides, and over 80 nucleotides.

B) Size distribution of all polyA tails. Expression of URT1$^{D491/3A}$M1 does not lead to the accumulation of mRNAs with short polyA tails (14-20 nt) that is observed upon expression of URT1$^{D491/3A}$.

overaccumulation of *PR2* mRNAs in the two analysed replicates. In contrast to the *GFP* reporter mRNA, the steady-state levels of the *NbPR2* messenger RNAs seems to be highly affected when inactive URT1$^{D491/3A}$-myc construct is overexpressed. I propose that *NbPR2* messengers are more subjected to degradation that the *GFP* reporter transcripts. Thus, the depletion of a degradation factor by the N-terminal IDR part of URT1$^{D491/3A}$-myc may lead to important accumulations of *PR2* transcripts.

2.4. The M1 motif is required for the accumulation of mRNAs with short polyA tails upon overexpression of URT1$^{D491/3A}$-myc.

As compared to URT1-myc, the removal of the IDR or the mutation of the M1 and M2 motifs in the active versions of URT1-myc had no major effects on the regulation of polyA tail sizes. By contrast, the removal of the IDR from URT1$^{D491/3A}$-myc (ΔIDR$^{D491/3A}$-myc) completely abolished the accumulation of mRNAs with short tails observed upon URT1$^{D491/3A}$-myc overexpression. This observation suggests an important function for the IDR in the turnover of mRNAs with this tail population. Interestingly, previous *in vitro* pulldown assays had shown that the M1 motif of the N-terminal IDR is crucial for the interaction between URT1 and the decapping activator DCP5.

Therefore, I introduced the L21/25N and/or AA49KK mutations in M1 and M2 motifs, respectively, in the URT1$^{D491/3A}$-myc construct (Figure 29A). Transient co-expression with the *GFP* reporter mRNA and P19 was achieved as described before. Importantly, URT1$^{D491/3A}$-myc, URT1$^{D491/3A}$M1-myc, URT1$^{D491/3A}$M2-myc and URT1$^{D491/3A}$M1M2-myc constructs were all expressed to similar levels and accumulated to higher levels as compared to URT1-myc (Figure 29B).

2.4.1. The M1 motif is implicated in the regulation of polyA tail length.

Three independent replicates were analysed by 3'RACE-seq. For each replicate, we obtained more than 100 000 total reads (Figure S7). In this run, we observed that the polyA tail size distribution was identical in replicate 2 and 3, but differed in replicate 1. These slight differences are probably linked to the library preparations that were performed simultaneously for both replicates 2 and 3, but separately for replicate 1. However, the results were overall comparable to the results obtained in runs 1 and 2. The co-expression of active URT1-myc notably increased the polyA tail sizes of *GFP* reporter mRNAs (Figure 30A and B). As observed before, the population of transcripts with short polyA tails of 14-20 nucleotides accumulated upon expression of inactive URT1$^{D491/3A}$-myc. Strikingly, this population was drastically reduced in the URT1$^{D491/3A}$M1-myc and

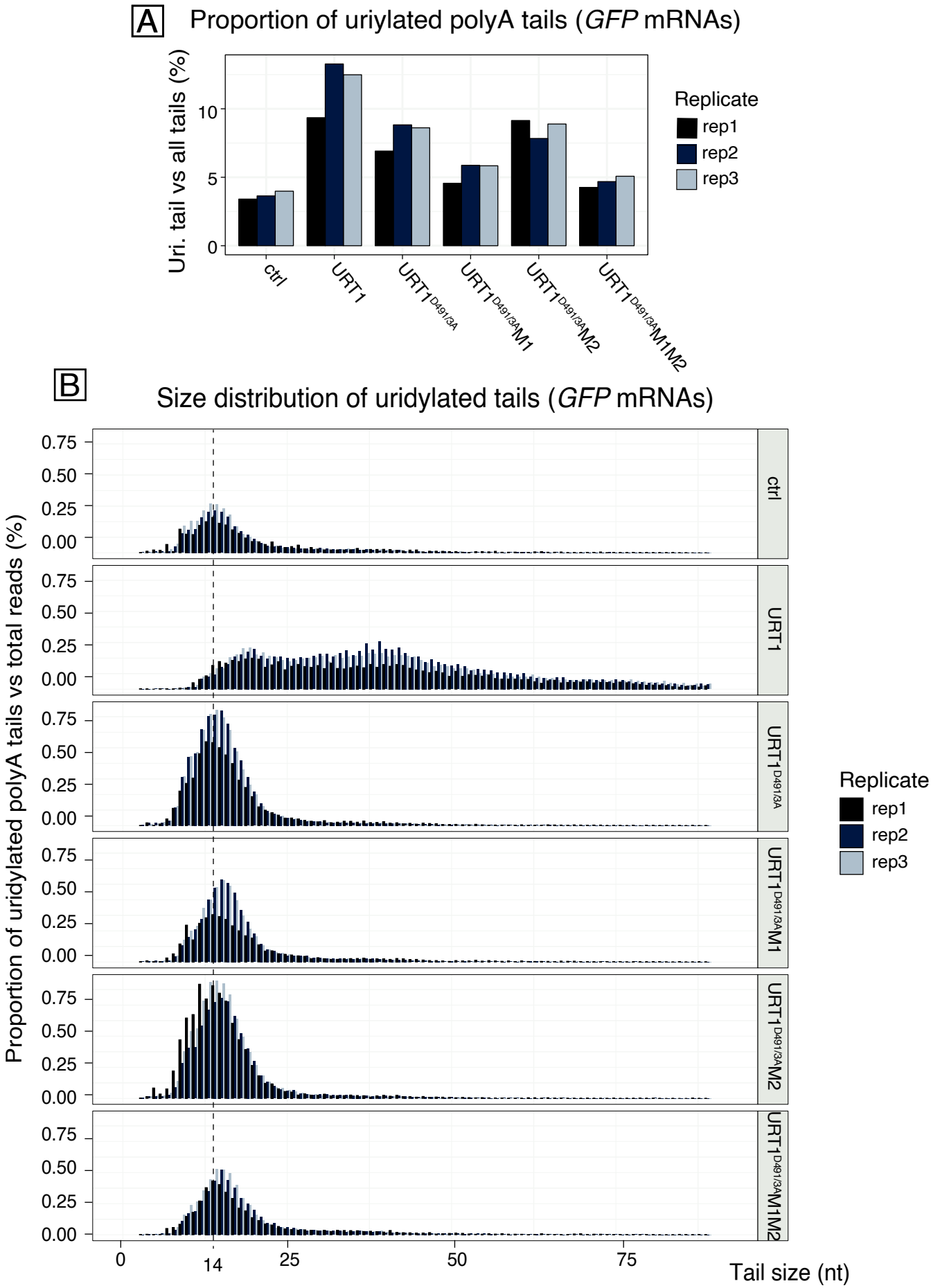


Figure 31: Size distribution of uridylated polyA tails of *GFP* mRNA.
A) Proportion of uridylated tails.
B) Size distribution of uridylated polyA tails. Mutations of the M1 motif in URT1$^{D491/3A}$-myc reduces the accumulation of short uridylated polyA tails compared to URT1$^{D491/3A}$-myc.

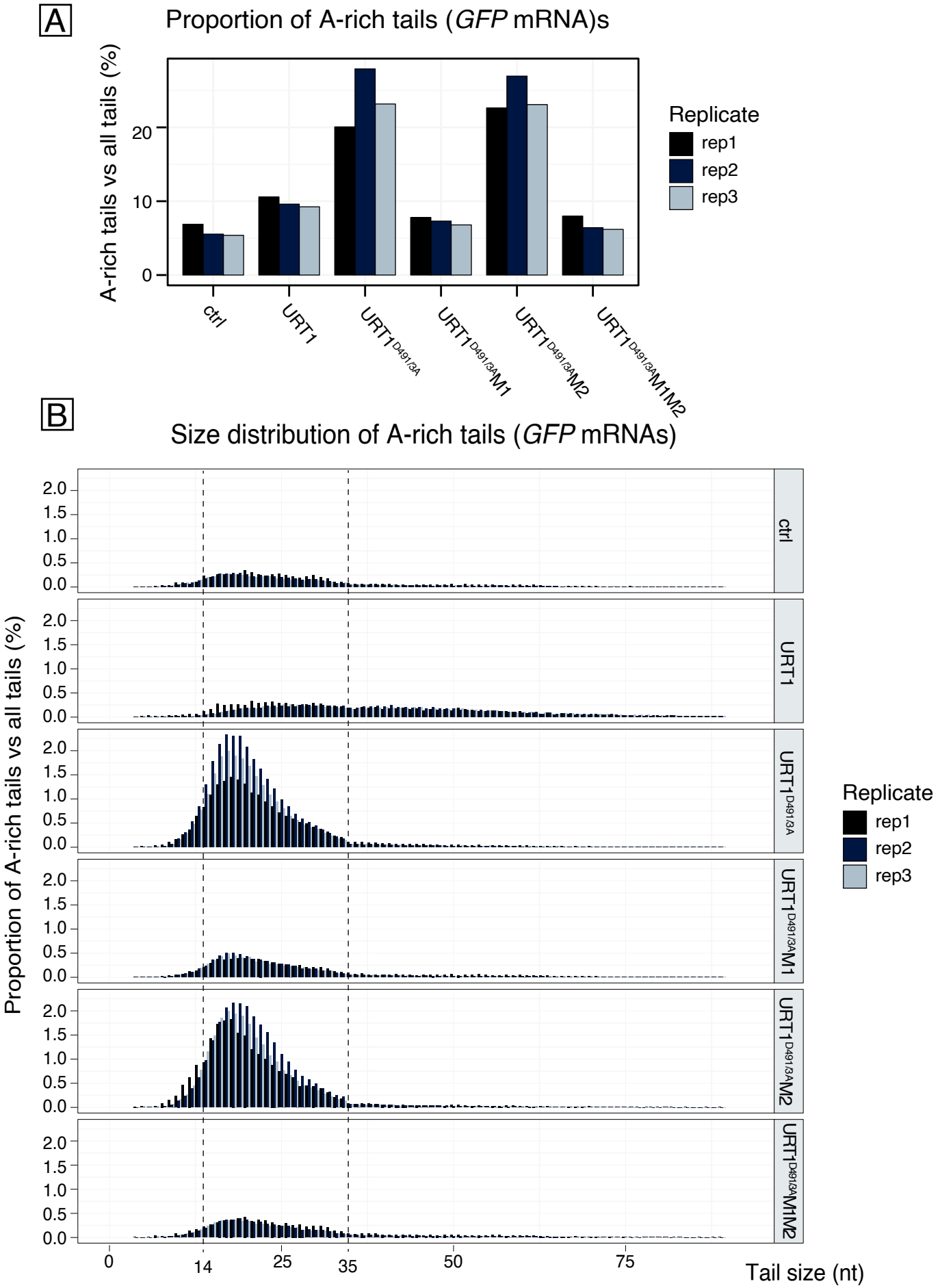


Figure 32: Size distribution of A-rich tails of *GFP* mRNAs.
A) Proportion of A-rich tails.
B) Size distribution of A-rich tails. Mutations of the M1 motif in URT1$^{D491/3A}$-myc leads to the reduction of short A-rich tails compared to URT1$^{D491/3A}$-myc.

URT1$^{D491/3A}$M1M2-myc samples (Figure 30A and B). By contrast, the mutation of M2 had no obvious impact on the polyA tail size distribution of the *GFP* mRNAs. Hence, these results clearly indicate that overexpression of the M1 motif affects the turnover of short deadenylated polyA tail of 20 nucleotides.

2.4.2. The M1 motif is required to suppress the overaccumulation of uridylated short polyA tails induced by overexpression of inactive URT1-myc.

The uridylation level of *GFP* mRNAs is *circa* 4% in control samples possessing only the endogenous NbURT1 activity (Figure 31A). Overexpression of URT1-myc led to an important increase of the uridylation percentage to approximately 10%. As seen previously, expression of inactive URT1$^{D491/3A}$-myc did not completely reverse the proportion of uridylated tails to endogenous levels, because transcripts with short uridylated polyA tails of about 14 nucleotides highly accumulate in this sample (Figure 31B). Strikingly, the additional mutation of the M1 motif led to a clear reduction of the population of transcripts with short uridylated polyA tails. Mutation of the M2 motif had no considerable impact on the proportion of short uridylated tails, and the accumulation of transcripts with short uridylated polyA tails upon simultaneous mutation of both M1 and M2 motifs resembled the profile observed in URT1$^{D491/3A}$M1-myc.

These results demonstrate that the M1 motif of the N-terminal IDR of URT1-myc is important for the accumulation of uridylated tails when URT1$^{D491/3A}$-myc is overexpressed.

2.4.3. Overexpression of URT1$^{D491/3A}$M1-myc leads to the accumulation of A-rich polyA tails.

Next, we assessed the accumulation of A-rich tails. As observed in the previous experiments, low levels of A-rich tails were detectable in the controls, and their size distribution shifted towards longer tails upon overexpression of URT1-myc (Figure 32A and B). Overexpression of URT1$^{D491/3A}$-myc induced the accumulation of A-rich tails of 14-35 nucleotides. This accumulation was reverted upon overexpression of URT1$^{D491/3A}$M1-myc and URT1$^{D491/3A}$M1M2-myc (Figure 32A and B). Compared to URT1$^{D491/3A}$-myc, the level of A-rich tails decreased by a factor of 4.

These results confirmed that the M1 motif is important for the effect that is observed on the accumulation of short polyA tails when inactive URT1$^{D491/3A}$-myc is overexpressed. Because the M1 motif has been shown to mediate the interaction between URT1 and DCP5, these results support the idea that the accumulation of short polyA tails seen in URT1$^{D491/3A}$-myc samples is linked to the

sequestration of the *N. benthamiana* DCP5 homolog and eventually DCP5-associated degradation factors.

Conclusions

The transient expression system in *N. benthamiana* leaves proved to be a fast and efficient system to assess the effects of wild-type and mutated versions of URT1 on the regulation of polyA tail length by 3'RACE-seq. One of the main results that we obtained with this system is that the overexpression of catalytically inactive URT1 results in the stabilisation of transcripts with short polyA tails and A-rich tails, and that the M1-motif in the N-terminal IDR is implicated for their correct turnover. Taking previous results of our team into account, the most straightforward explanation for this effect is that overexpression of the URT1 protein sequesters DCP5 and eventually other degradation factors that would be required for the turnover of transcripts with short polyA tails. Interestingly, the steady-state levels of *GFP* reporter mRNAs and *NbPR2* transcripts seem to be differently affected. Whereas the GFP transcript show a reduced apparent size that most likely mirrors the accumulation of short polyA tails in the URT1$^{D491/3A}$-myc samples (Figure 24), *NbPR2* mRNAs are clearly overaccumulating in these samples (Figure 28). This may be explained by the fact that *NbPR2* mRNAs have much shorter polyA tails than *GFP* reporter mRNAs. Thus, *NbPR2* transcripts could be more subject to decapping and the 5'-3' decay pathway than *GFP* mRNAs. The overexpression of URT1$^{D491/3A}$-myc and the depletion of the decapping machinery will consequently further affect the stability of *PR2* mRNAs.

The second main result of these experiments is that uridylation is required to control polyA tail length for endogenous *PR2* and reporter *GFP* mRNAs. Previous data of our group indicated that uridylation repairs the 3' ends of deadenylated mRNAs to enable the binding of at least one PABP. Our new results obtained in *N. benthamiana* demonstrate that uridylation by URT1 lead to the accumulation of mRNAs with long uridylated and non-uridylated polyA tails. This finding suggests that uridylation *per se* can slow down deadenylation, probably by inhibiting CCR4/NOTs deadenylase activity.

3. Terminal uridines impede the activity of CAF1b deadenylase.

The results obtained in *N. benthamiana* suggest that URT1-mediated uridylation can impede mRNA deadenylation. Therefore, I wanted to test if uridylation *per se* can impact the activity of deadenylases using *in vitro* assays. The goal was to determine if uridines present at the 3' end of an oligoadenylated RNA substrate could intrinsically hinder deadenylases from shortening polyA tails.

To test this, I needed to produce tagged versions of the main deadenylases of Arabidopsis. Arabidopsis has two CCR4 and eleven CAF1 homologs. Only very few biochemical assays have been performed to study plant deadenylases to date, and which of these proteins are indeed active deadenylases is still unclear. I chose to perform *in vitro* activity tests with CCR4a, CCR4b, and the 6 Arabidopsis CAF1 homologs that have been suggested to interact with NOT1 and/or CCR4 (see introduction): CAF1a, CAF1b, CAF1h, CAF1i, CAF1j and CAF1k. For each of these candidate deadenylases, I produced both wild-type and mutated proteins, the latter of which had substitutions of key residues in their catalytic sites. All proteins were purified from *E. coli* using N-terminal hisMBP or a hisGST tags, for CCR4 or CAF1 proteins, respectively. The deadenylation activity of each construct was tested at several protein concentrations and in different reaction buffers containing various concentrations of Mg^{2+}, glycerol or potassium chloride (Figure S8). Under the various conditions that I tested, only hisGST-CAF1b had a detectable deadenylation activity. The final reaction buffer used for the deadenylation tests was 20mM MOPS at pH7.2, 5mM of $MgCl_2$, 50mM of KCl, 7% of glycerol and 0.1% of Tween 20. Because CCR4 and CAF1 proteins are suggested to interact *in vivo*, I also tested combinations of CCR4a/b and CAF1a/b/h/i/j/k. None of the combinations except those containing CAF1b displayed deadenylation activity. Unfortunately, the removal of the N-terminal hisMBP and hisGST tag resulted in the destabilisation of the purified proteins. Therefore, I could not test the activity of the proteins devoid of their tags. Truncated proteins comprising only the EEP activity domains of CCR4a/b were also inactive in the tested conditions.

I used two independent replicates of hisGST-CAF1b for the activity tests. The first replicate was batch-purified using a glutathione sepharose resin (rep1). To reduce eventual contaminating proteins, a portion of replicate 1 was further batch-purified using NiNTA resin beads (rep1bis). The

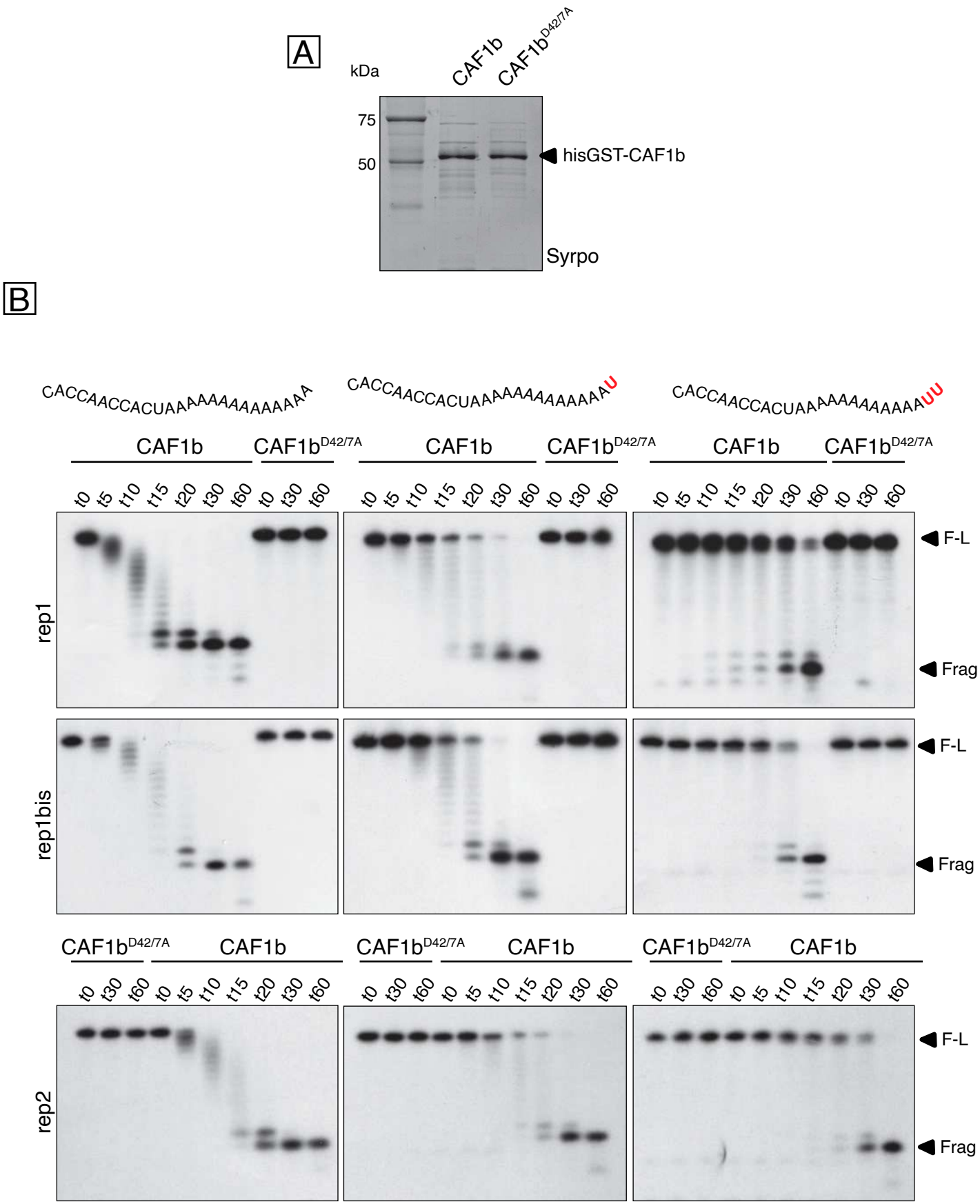


Figure 33: Deadenylation reaction by hisGST-CAF1b.
A) SDS-PAGE analysis of purified protein samples of active and inactive hisGST-CAF1b (rep1bis). Proteins were seperated on a 12% gel and colored with the SYPRO Ruby protein gel stain.
B) Deadenylation reaction for two independent replicates of CAF1b on CACCAACCACU-A_{14}, CAC-CAACCACU-$A_{13}U_1$ and CACCAACCACU-$A_{12}U_2$ substrates. 5' ^{32}P-labeled RNA substrates were incubated up to 1h with active and inactive hisGST-CAF1b proteins.
F-L: full-length, Frag: deadenylated CACCAACCAC fragment

second independent replicate was purified by affinity chromatography on glutathione sepharose coupled to size exclusion chromatography with an AKTA purification system (rep2). The purified protein samples showed some contaminants on SDS-PAGE analysis (Figure 33A), but no contaminated ribonuclease activity was detected in all deadenylation assays as shown below when incubating with inactive hisGST-CAF1b$^{D42/7A}$.

Three RNA substrates that differed in their 3' end were tested in the activity tests. Each of the three contained a CACCAACCACU sequence followed by either 14 adenosines, 13 adenosines and 1 single terminal uridine or 12 adenosines and 2 terminal uridines (Figure 33B). The RNA substrates were labelled at their 5' ends using radioactive γ^{32}P-ATP and T4 polynucleotide kinase PNK. Purified hisGST-CAF1b and inactive hisGST-CAF1b$^{D42/7A}$ were incubated with the different substrates, and aliquots of the reactions were collected at 0, 5, 10, 15, 20, 30 and 60 minutes. The samples were separated on a 6% acrylamide TBE-urea gel and analysed by autoradiography using X-ray films or a phosphorimager system.

Identical results were obtained for the three hisGST-CAF1b samples (Figure 33B). As expected, hisGST-CAF1b quickly degrades the polyA stretch of the substrate with 14 homopolymeric adenosine stretch. Initial shortening of the polyA tail was already observed after 5 minutes, and full-length CACCAACCACU-A_{14} substrates were no longer detectable even after this short incubation time. After 20 min, the majority of the RNA substrates were entirely deadenylated. Because no activity was ever detected for hisGST-CAF1b$^{D42/7A}$, we are confident that the activity observed in these experiments is due to hisGST-CAF1b and not to an RNase contamination.

A single uridine at the 3' end of the polyA tails reproducibly delayed the deadenylation activity of hisGST-CAF1b (Figure 33B). The first reduction of full-length CACCAACCACU-$A_{13}U_1$ and a ladder of bands representing degradation intermediates was observed after 10 minutes, and full deadenylation was achieved after 60 minutes. Importantly, full-length CACCAACCACU-$A_{13}U_1$ substrates were readily detected at all but the last time-points. Hence, once the 3' uridine is removed, deadenylation proceeds quickly. This result demonstrates that the removal of the 3' uridine is the rate-limiting step in this *in vitro* test system. The presence of 2 uridines severely inhibited the catalytic reaction in each of the 3 replicates. Less than half of the CACCAACCACU-$A_{12}U_2$ substrates was deadenylated after 20 minutes, and residual amounts of the full-length substrates were still detected after 60 minutes (Figure 33B).

Taken together, these results demonstrate that the presence of terminal uridines severely impacts the deadenylation activity of hisGST-CAF1b *in vitro*. The kinetics of the deadenylation activities of

Arabidopsis CCR4 or CAF1 are likely modulated when associated to the CCR4/NOT complex. Yet, yeast CAF1 and CCR4 display a very similar deadenylation activity as standalone proteins or as integrated subunits of the CCR4/NOT complex (Webster et al., 2018b). Hence, the intrinsic property of terminal uridines on their deadenylase activity is likely conserved for CAF1b associated to the whole complex. Moreover, these data are in good agreement with the hypothesis that uridylation by URT1 slows down the deadenylation of *GFP* reporter and *PR2* mRNAs in *N. benthamiana,* which results in the accumulations of transcripts with longer polyA tails. Uridylation *per se* is thus probably slowing down deadenylation *in vivo*.

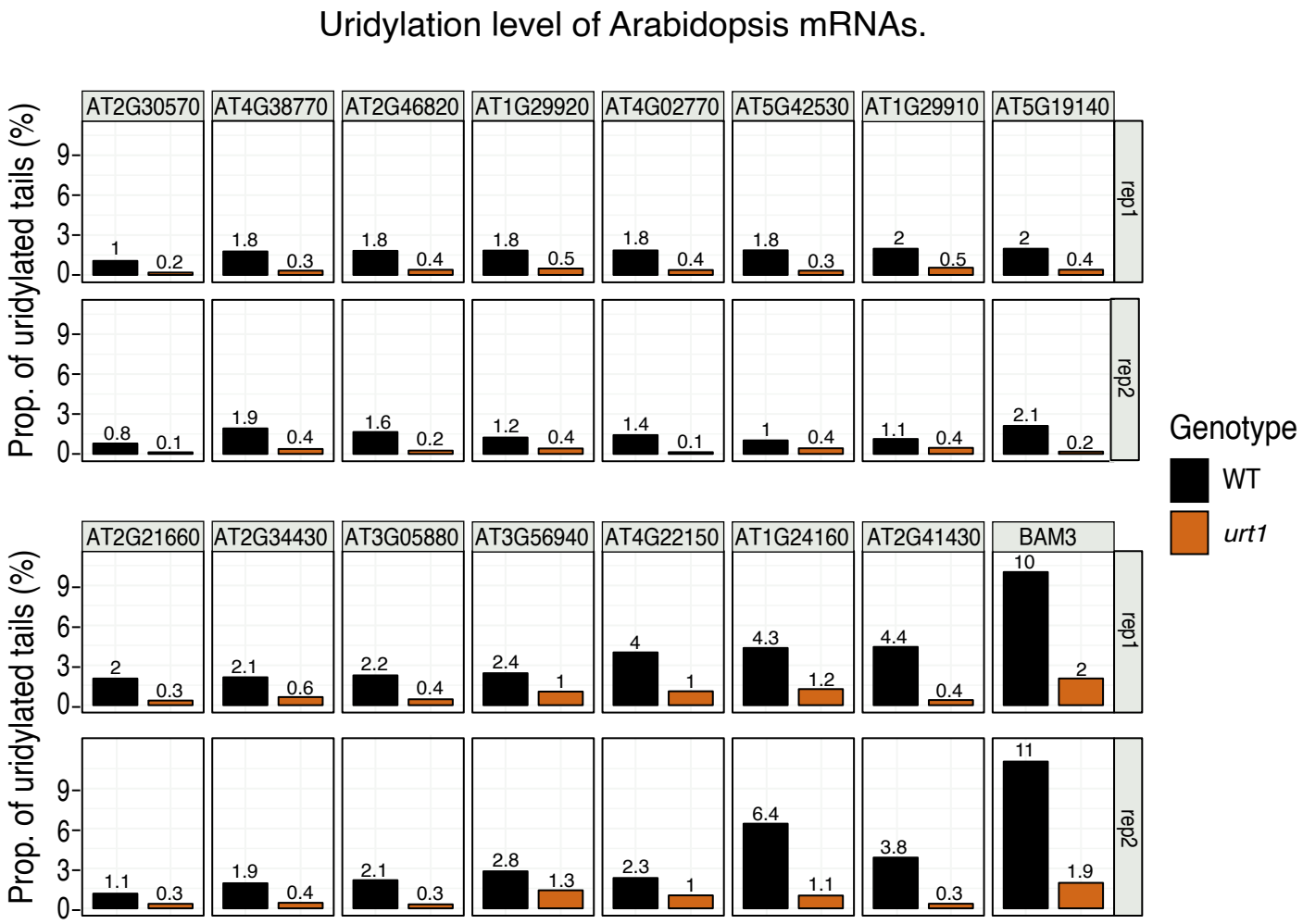


Figure 34: Uridylation level of analysed Arabidopsis mRNAs.
3'RACE-seq analysis of 16 Arabidopsis mRNAs. The bar graphs show the proportion of uridylated mRNAs in the two independent replicates. mRNAs with low uridylation frequencies are displayed at the top side, mRNAs with high uridylation levels are shown at the bottom right side.

4. Expression of URT1 leads to longer polyA tail profiles in *Arabidopsis thaliana.*

In the following experiments, I investigated the impact of URT1-mediated uridylation on polyA tail profiles in *Arabidopsis thaliana*. To this end, I compared the polyA tail sizes of 16 mRNAs in 2 biological replicates of WT and *urt1-1* mutants by 3'RACE-seq. Previous data suggested that uridylation is widespread and affects the majority of mRNAs. Yet, the proportion of uridylated transcripts versus non uridylated transcripts varies remarkably between individual mRNAs. At the time this experiment was performed, the TAIL-seq data (Results 1.2, Figure 13-14) were not available yet. Hence, we chose to analyse the 3' ends of *BAM3* and *AT1G24160* mRNAs because we knew from previous experiments that both mRNAs are frequently uridylated by URT1 (Sement et al., 2013; Zuber et al., 2016). The remaining 14 mRNAs were selected based on their expression levels (highly expressed mRNAs) and average polyA tail length (short and long polyA size distribution for comparison) (Subtelny et al., 2014).

4.1. Uridylation frequencies vary between individual mRNAs.

Of the 16 selected mRNAs, 12 mRNAs had uridylation frequencies below 3%. The remaining 4 mRNAs had uridylation frequencies of 4 – 10% (Figure 34). Importantly, these frequencies were largely similar between both replicates. Some variation was observed for *AT4G22150* mRNAs that had an uridylation level of 4% in rep1, while only 2% were uridylated in rep2. For *AT1G24160* mRNAs, we observed uridylation rates of 5% in rep1 and 6% in rep2. As expected, the overall uridylation frequency decreased in the *urt1* mutant for all analysed mRNAs. This confirms that URT1 is the main TUTase that uridylates mRNAs in Arabidopsis.

There is some discrepancy between these new results and older data obtained by 3'RACE data. This difference is likely explained by technical reasons. Using an experimental strategy that included the PCR amplification of 3' ends followed by cloning and Sanger sequencing, Zuber *et al.* found that *circa* 20% of *AT1G24160* and *BAM3* mRNAs were uridylated (Zuber et al., 2016). The new 3'RACE-seq strategy employed here firstly allows the deduplication of PCR amplicons, and secondly avoids the selection of short amplicons that may be introduced by the cloning step.

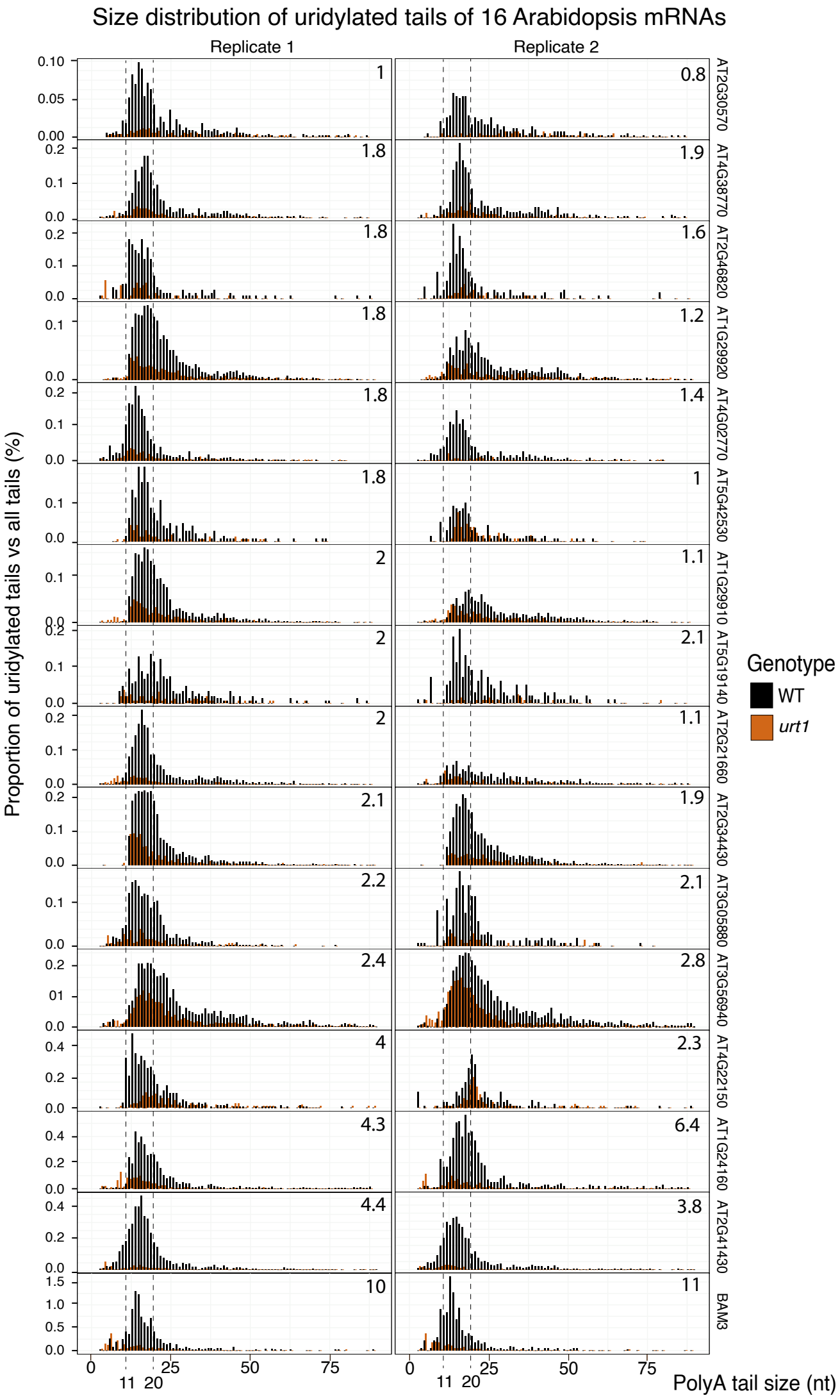

Size distribution of uridylated tails of 16 Arabidopsis mRNAs
Replicate 1
Replicate 2
Proportion of uridylated tails vs all tails (%)
PolyA tail size (nt)
AT2G30570
AT4G38770
AT2G46820
AT1G29920
AT4G02770
AT5G42530
AT1G29910
AT5G19140
AT2G21660
AT2G34430
AT3G05880
AT3G56940
AT4G22150
AT1G24160
AT2G41430
BAM3
Genotype
WT
urt1

Figure 35: Size distribution of uridylated tails of Arabidopsis mRNAs.

2 independent replicates were analysed by 3'RACE-seq for 16 mRNAs. The bar plots show the polyA size distribution of uridylated tails and are classified from the least to the most uridylated mRNA of replicate 1.

Hence, the uridylation levels of 5% and 9% that are now detected for *AT1G24160* and *BAM3* mRNAs, respectively, are likely a better estimation of the *in vivo* uridylation levels of these transcripts.

4.2. mRNAs with higher uridylation frequencies have longer polyA tails.

To evaluate an eventual link between uridylation and polyA tail sizes in *A. thaliana*, I analysed the polyA tail size distributions for the 16 candidate mRNAs in both WT and *urt1*, and sorted the polyA tail profiles of the mRNAs according to their uridylation frequency (Figures 35 and 36). The mRNAs shown at the top of the figures have low uridylation frequencies, while the mRNAs with high uridylation frequencies are displayed at the bottom. Figure 35 displays the size distribution of only uridylated tails, while Figure 36 displays the size distribution of homopolymeric polyA tails.

The vast majority of uridylated transcripts had short polyA tails, between 11 and 20 nucleotides, with slightly different mean sizes between the individual mRNAs (Figure 35). However, terminal uridines could also be observed on longer polyA tails (> 20 nucleotides), suggesting that URT1 can uridylate longer polyA tails *in vivo*. As previously observed, some uridylated transcripts are still detected in *urt1-1* mutants. Surprisingly, a considerable proportion of uridylated transcripts is detected for *AT3G56940* (Figure 35). Unpublished results from the group have shown by 3'RACE-seq that the residual level of uridylation is drastically decreased in *urt1 heso1* double mutants. As the *urt1-1* mutant is a null mutant, this residual level of uridylation is likely catalysed by HESO1, the second known TUTase of *A. thaliana*.

Loss of URT1 results in a shift towards shorter polyA tails, which is most obvious for highly uridylated transcripts. For example, *urt1* mutants accumulate *AT4G22150*, *AT1G24160*, *AT2G41430* and *BAM3* mRNAs with excessively shortened polyA tails of less than 10 nucleotides which are rarely detected in wild-type plants (Figure 35 and 36). Interestingly, the shift towards short polyA tails is still observed in *urt1-1* for the *AT3G56940* mRNA that has an important residual

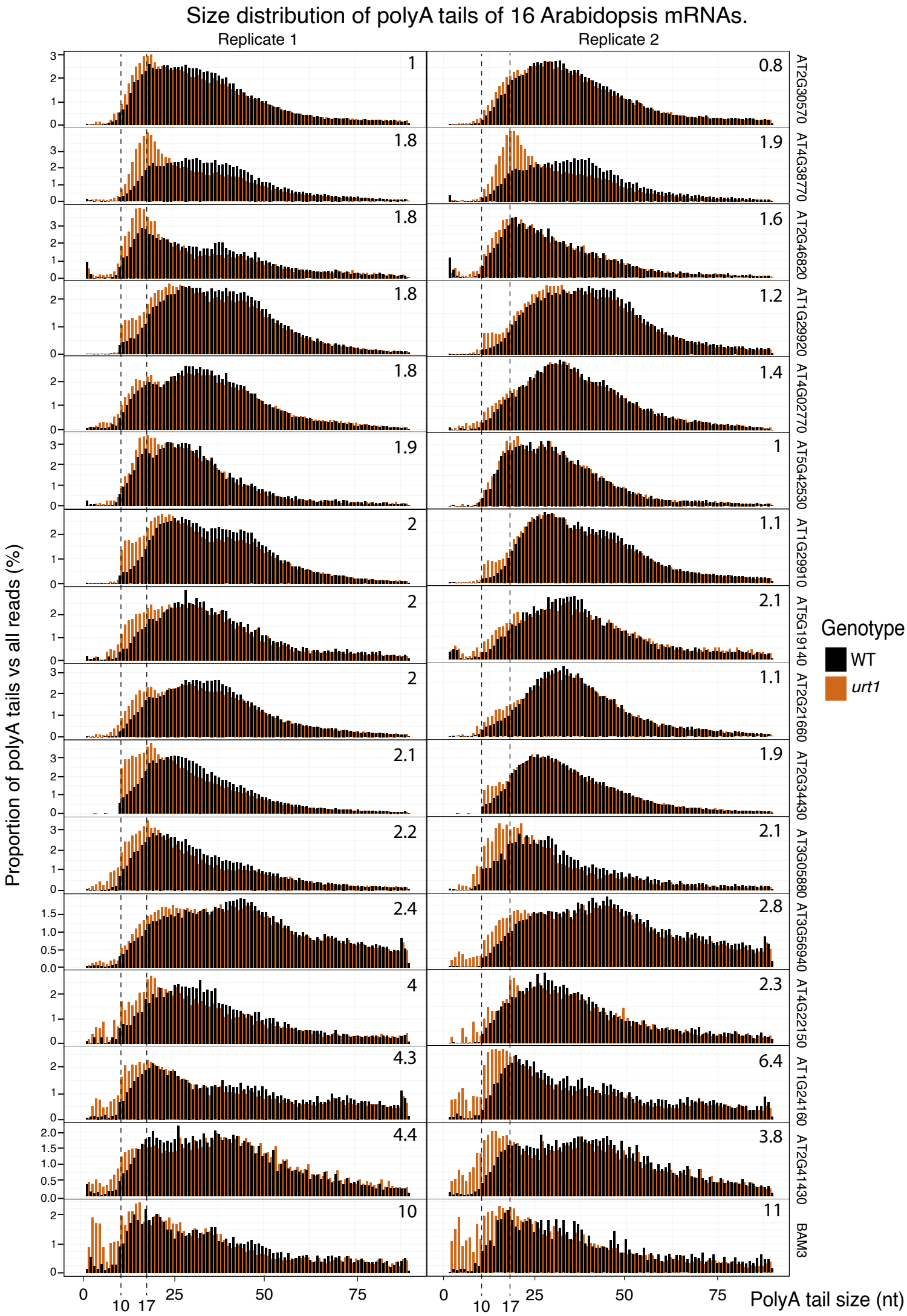

Size distribution of polyA tails of 16 Arabidopsis mRNAs.
Replicate 1
Replicate 2
Proportion of polyA tails vs all reads (%)
PolyA tail size (nt)
AT2G30570
AT4G38770
AT2G46820
AT1G29920
AT4G02770
AT5G42530
AT1G29910
AT5G19140
AT2G21660
AT2G34430
AT3G05880
AT3G56940
AT4G22150
AT1G24160
AT2G41430
BAM3
Genotype
WT
urt1

Figure 36: Size distribution of polyA tails of Arabidopsis mRNAs.

2 independent replicates were analysed by 3'RACE-seq for 16 mRNAs. The bar plots show the polyA size distribution of homopolymeric polyA tails and are classified from the least to the most uridylated mRNA of replicate 1.

uridylation level. This observation strongly suggests that the residual uridylation activity by HESO1 or another yet unknown TUTase cannot compensate for the lack of URT1. Altogether, these observations reinforce the hypothesis that the main role of URT1-mediated uridylation is to prevent the accumulation of excessively deadenylated mRNAs.

The most striking result from these analyses was observed when considering the distribution of homopolymeric polyA tails for the 16 mRNAs in wild-type plants (Figure 36). In fact, diverse profiles were observed for the different analysed mRNAs. Even though uridylation occurs mostly on short deadenylated tails below 25 nucleotides for all targets (Figure 35), the mRNAs with high proportions of shorter tails are not necessarily those that are the most uridylated (Figure 36). By contrast, mRNAs with high proportions of uridylated tails tend to have a polyA tail distribution that is spread across a wide size range (Figure 36). Such transcripts tend to have short polyA tails, but also very long polyA tails as compared to transcripts with low *in vivo* uridylation frequencies. Of note, the profiles of the homopolymeric tails show a phased polyA size distribution for some mRNAs. Surprisingly, this phased size distribution is best seen for highly uridylated mRNAs. As suggested in a few studies, this phasing can correspond to the footprints of PABPs that bind highly expressed mRNAs and control the deadenylation process by CCR4/NOT (Lima et al., 2017; Webster et al., 2018b).

Overall, our results revealed that URT1 has a major role in preventing the accumulation of excessive deadenylated mRNAs. The expression of URT1 leads to an overall accumulation of longer polyA tails for all tested mRNAs. This observation recalls the data obtained by transiently expressing URT1-myc in *N. benthamiana* and reinforces the hypothesis that uridylation by URT1 can protect mRNA polyA tails (> 10 nucleotides) from excessive deadenylation in Arabidopsis by slowing down deadenylation and favouring the degradation from 5'-3' of short polyA tails over 10 nucleotides. Interestingly, most uridylated tails are observed for mRNAs that have a disperse polyA size distribution and that shown signs of a phasing pattern in the size distribution. This profile is suggested to result from the degradation of polyA tail that are highly bound by PABPs. These aspects are considered in more details in the discussion.

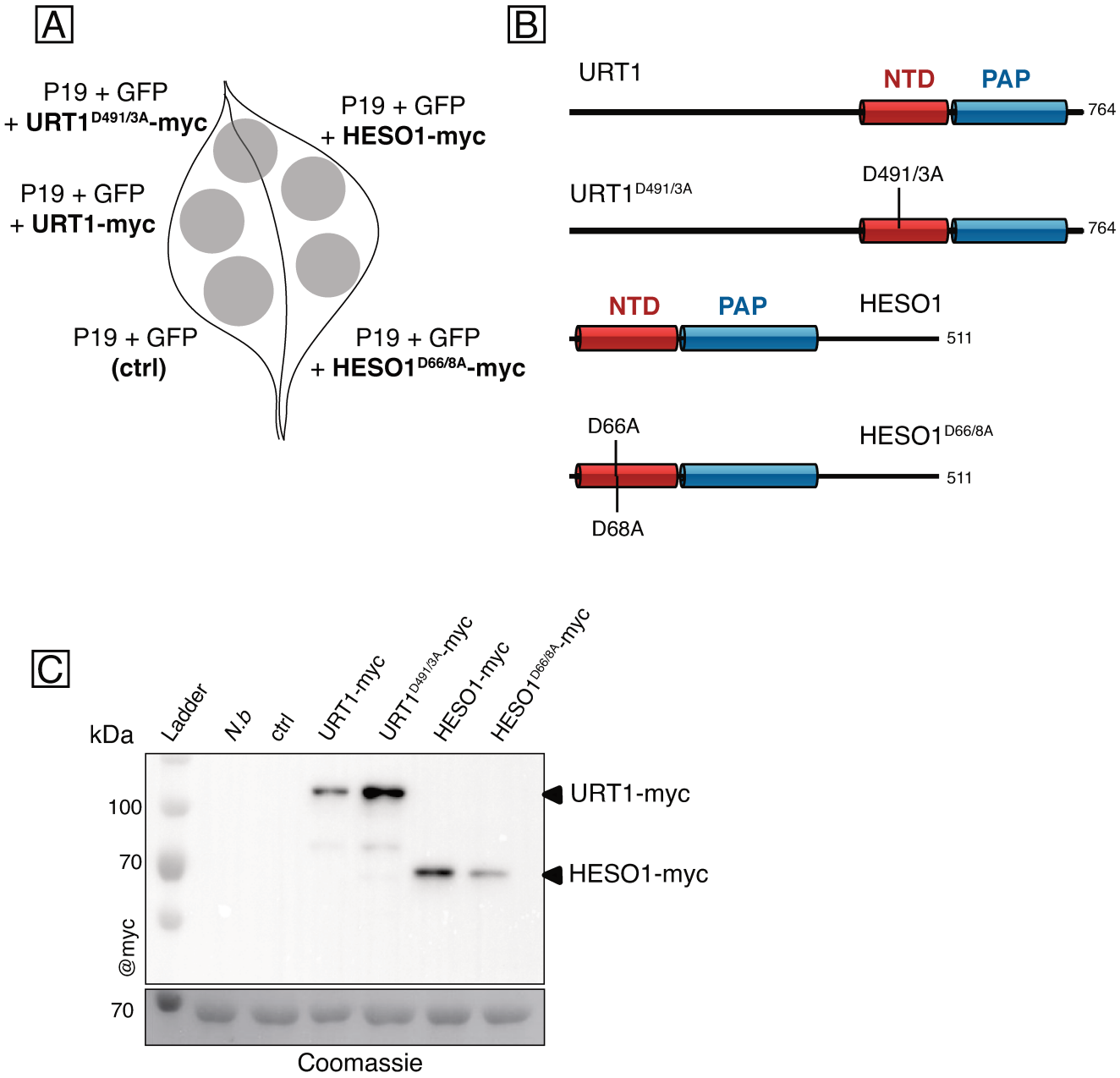


Figure 37: Transient expression of URT1-myc and HESO1-myc constructs in *N. benthamiana.*

A) Illustration of the infiltrated patches on *N. benthamiana* leaves. ctrl: control patches infiltrated only with P19 and GFP.

B) Diagrams illustrating the domain organisations and mutations of the different URT1-myc constructs. NTD: polymerase β-like nucleotidyltransferase domain; PAP: PolyA polymerase-associated domain.

C) Western blot analysis of the protein extracts obtained from the different leave patches harvested four days after infiltration (4 DAI). Monoclonal @myc antibodies were used for the detection of the URT1-myc constructs. The Coomassie stain of the membrane is shown as loading control. *Nb*: *Nicotiana benthamiana* uninfiltrated leaf patch.

5. The expression of HESO1-myc changes the polyA size distributions and decreases the steady-state level of reporter *GFP* mRNAs.

Our data clearly demonstrate that URT1 is responsible for most of the mRNA uridylation activity that is detected in wild-type plants of *A. thaliana* (Figure 34 and 35, Zuber et al., 2016). Yet, we observe residual uridylation of mRNAs in *urt1* null mutants. The TUTase responsible for this residual uridylation activity if HESO1. The prime function of HESO1 is to induce the rapid decay of unmethylated miRNAs and siRNAs (Ren et al., 2012; Tu et al., 2015; Wang et al., 2015; Zhao et al., 2012a). However, HESO1 and URT1 also share common substrates such as 5' RISC-cleaved mRNA fragments and some miRNAs (Branscheid et al., 2015; Orban and Izaurralde, 2005; Tu et al., 2015; Wang et al., 2015; Zhang et al., 2017b; Zuber et al., 2018). Here, I investigate the potential impact of HESO1-mediated uridylation on polyA tail length using the transient expression system in *N. benthamiana* that I also used to study the role of URT1.

To assess how HESO1 expression influences the polyadenylation and uridylation profiles of reporter *GFP* and *NbPR2* mRNAs, URT1-myc, URT1$^{D491/3A}$-myc, HESO1-myc and HESO1$^{D66/8A}$-myc were co-expressed with the *GFP* reporter mRNA and the silencing suppressor P19 in *N. benthamiana* leaves (Figure 37A). To abolish the catalytic activity of HESO1, two conserved aspartic acids (at position 66 and 68) in the nucleotidyltransferase domain were mutated to alanines (HESO1$^{D66/8A}$-myc) (Figure 37B). This mutation has been previously reported to abolish the uridylation capacity of HESO1 (Zhao et al., 2012a).

The analysis of protein extracts obtained from leaf patches by western blot confirmed that all proteins were expressed (Figure 37C). As observed before, the URT1$^{D491/3A}$-myc protein accumulated to higher levels than the active form URT1-myc. The levels of HESO1-myc were comparable or slightly higher than the levels of URT1-myc. By contrast, the inactive HESO1$^{D66/8A}$-myc protein was reproducibly less expressed than HESO1-myc. This observation must be taken into account for the interpretation of the results, but the reason for the poor expression of HESO1$^{D66/8A}$-myc is unknown.

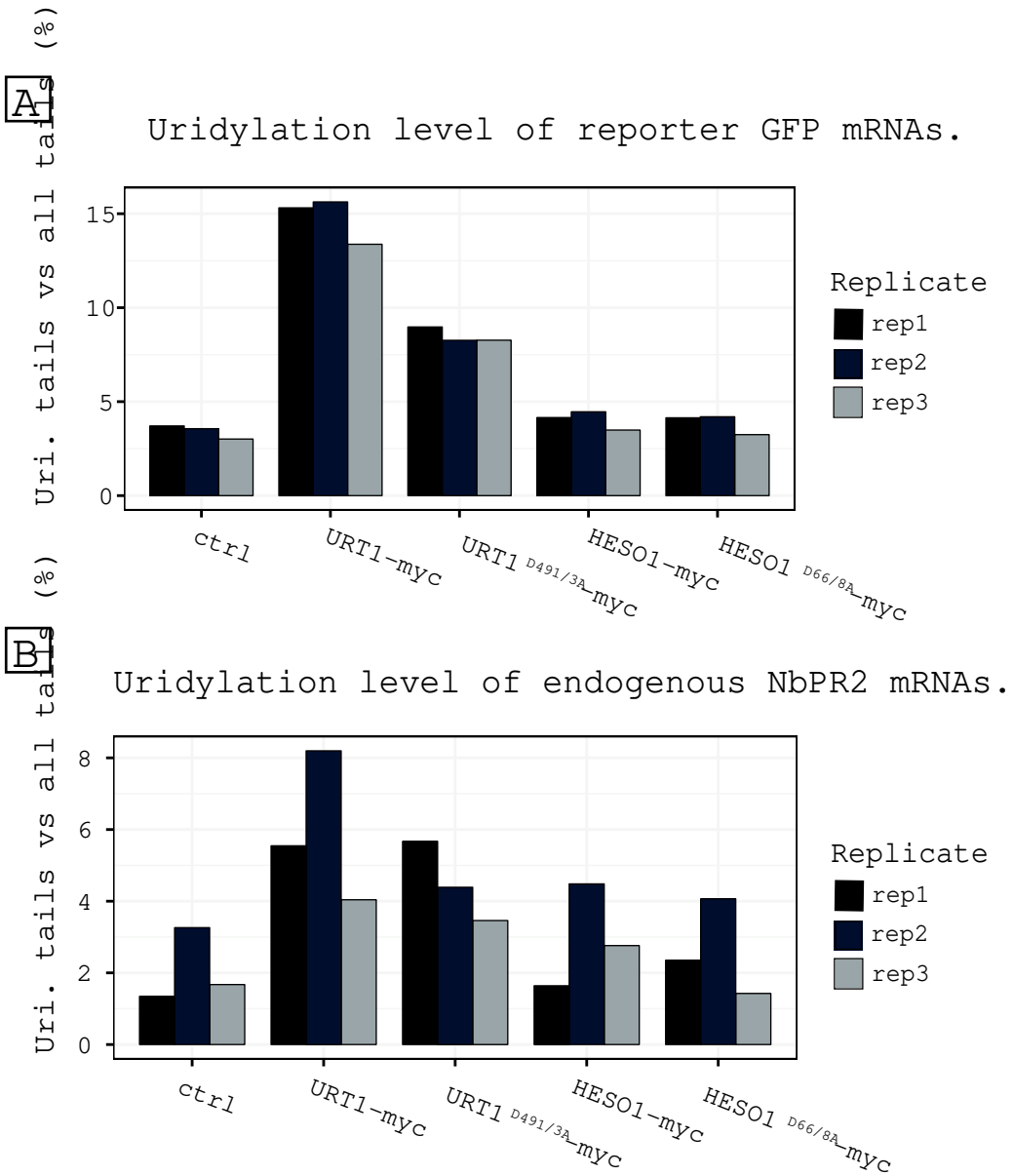


Figure 38: Uridylation level of GFP reporter and endogenous NbPR2 mRNA s.
A) Uridylation level of GFP reporter transcript.
B) Uridylation level of endogenous NbPR2 mRNA.

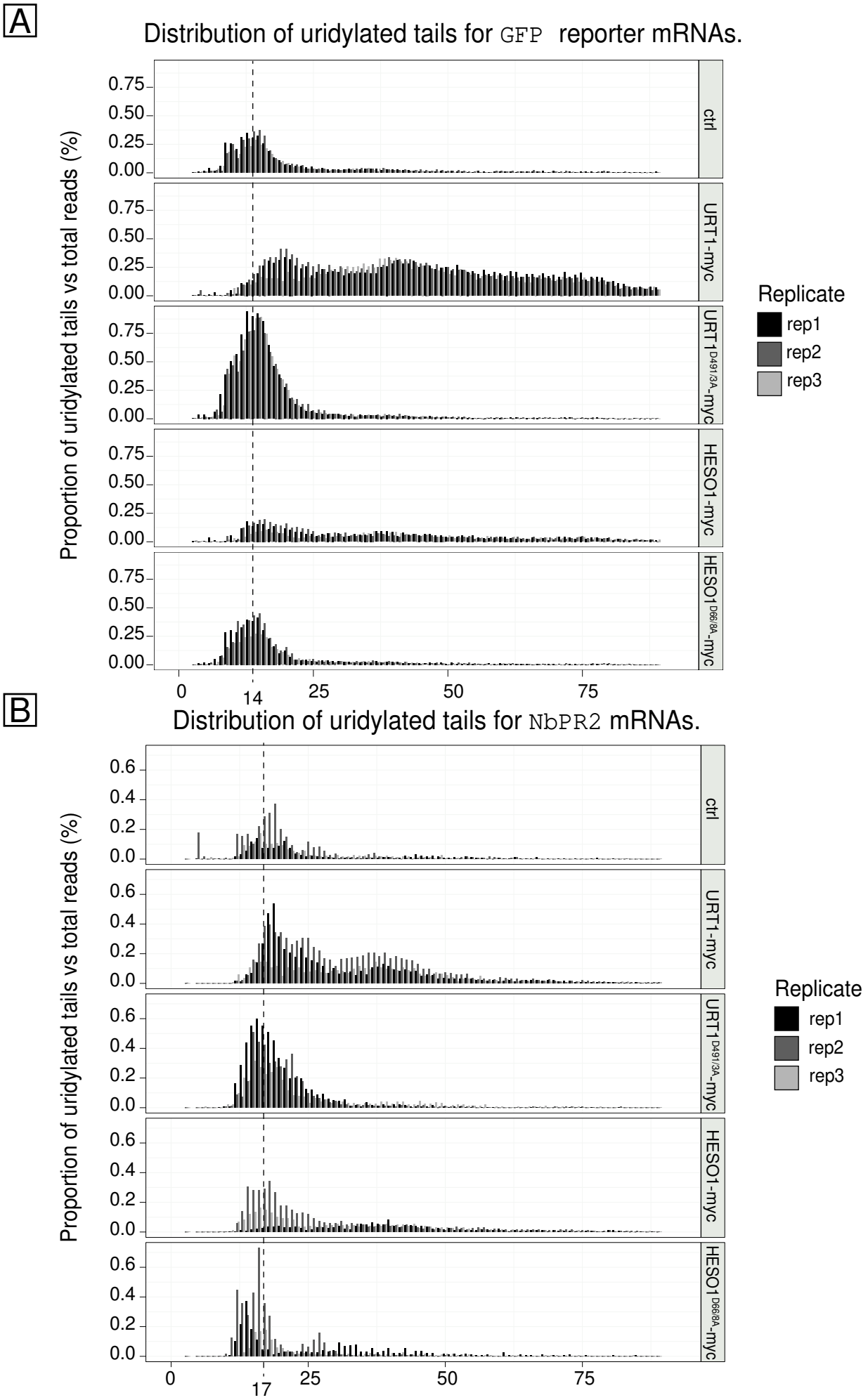


Figure 39: Distribution of uridylated polyA tails of reporter GFP and endogenous NbPR2 mRNAs.

A) Size distribution of uridylated polyA tails of reporter GFP transcript.

B) Size distribution of uridylated polyA tails of NbPR2 transcript.

5.1. Expression of HESO1-myc does not increase mRNA uridylation levels but impacts the size distribution of polyA tails.

I analysed the 3' ends of both the *GFP* reporter mRNA and the endogenous *NbPR2* mRNA by 3'RACE-seq in three biological replicates. As observed before, the proportion of uridylated *GFP* mRNA was drastically increased upon expression of URT1-myc. In URT1$^{D491/3A}$-myc samples, we observed the mild increase in the total proportion of uridylated transcripts due to the accumulation of short uridylated transcripts that we observed before. By contrast, the expression of HESO1-myc and HESO1$^{D66/8A}$-myc had no effect on the proportion of uridylated *GFP* reporter mRNAs (Figure 38A) and no systematic effect on the proportion of uridylated *PR2* mRNA (Figure 38B).

The analysis of the size distributions of uridylated polyA tails (Figure 39) confirmed the previously observed shift toward larger tails upon expression of URT1-myc and the pronounced overaccumulation of short uridylated tails in the URT1$^{D491/3A}$-myc samples. Interestingly, the size distribution profiles observed in HESO1-myc samples differed from the profiles observed in the control samples. Notably, we observed reduced levels of short uridylated polyA tails and a mild increase of transcripts with long uridylated polyA tails (Figure 39). The profiles observed upon expression of HESO1$^{D66/8A}$-myc resembled the profiles observed in the controls. Taken together, these observations indicated that the HESO1-myc protein is indeed catalytic active in the *N. benthamiana* system, as only the active HESO1-myc version changes the size distribution of uridylated polyA tails. The difference to URT1 may be that uridylation by HESO1 leads to the rapid degradation of its targets in *N. benthamiana*, which would explain why the uridylation level is not increased when expressing HESO1-myc (Figure 38). Of note, HESO1 has been shown to add longer uridine extensions to miRNAs and 5' RISC-cleaved fragments, whereas URT1 adds mostly 1 or 2 uridines (Tu et al., 2015; Zhao et al., 2012a; Zuber et al., 2018). In mammals, oligouridylation efficiently triggers the fast degradation of pre-miRNA let7, whereas monouridylation has been shown to repair the 3' end overhang of let7 for the formation of mature and functional miRNAs (Balzeau et al., 2017). HESO1 could thus oligouridylate *GFP* and *PR2* mRNAs to trigger their fast degradation. However, similar to URT1-myc, most uridylated tails were detected with 1 or 2 uridines at the 3' ends in the HESO1-myc samples (Figure S9). Either, HESO1 is adding only a few uridines to its target (similar to URT1), or the oligouridylated tails have a fast turnover rate and cannot be detected at steady-state level.

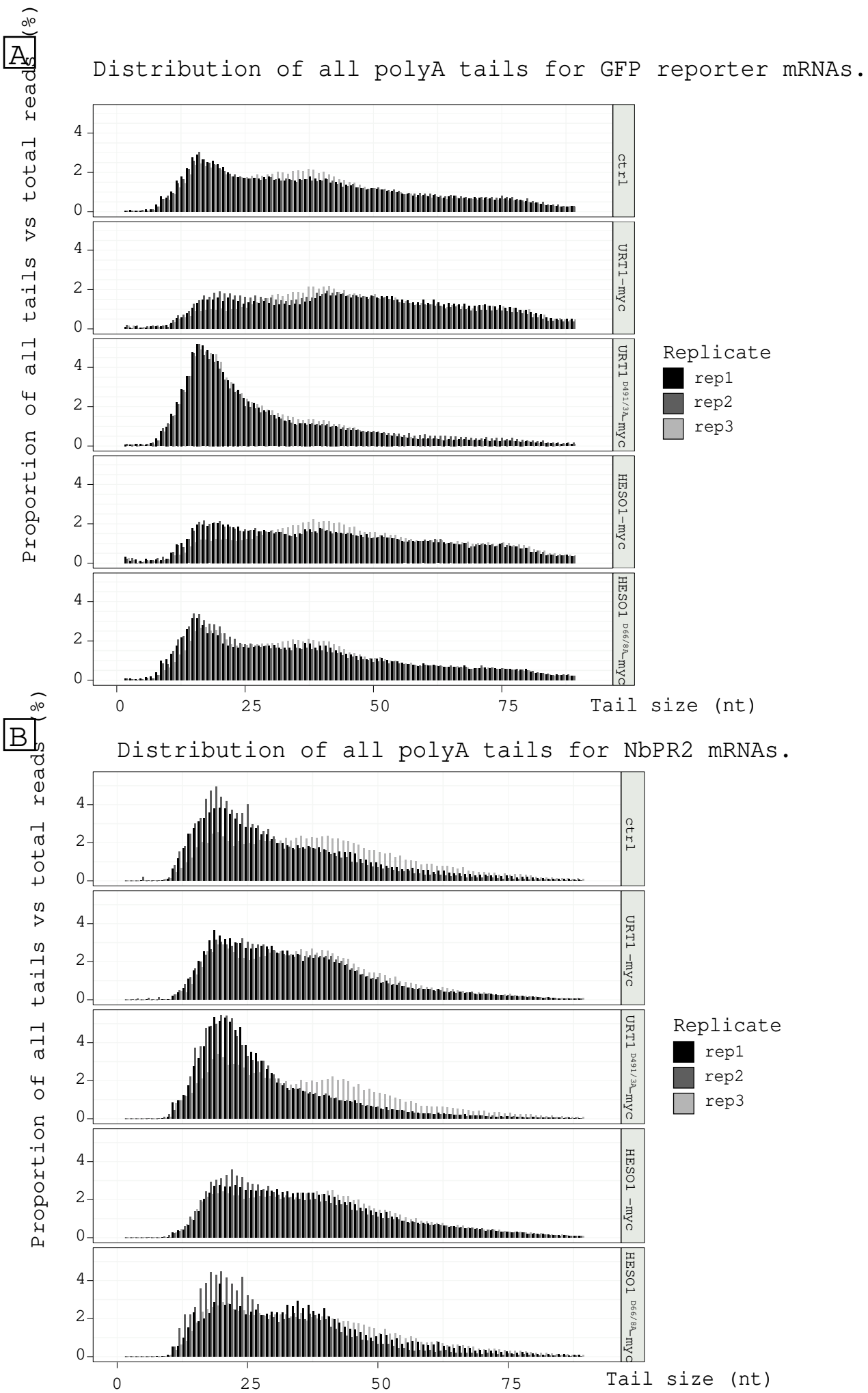


Figure 40: Distribution of all GFP and NbPR2 mRNAs.
A) Size distribution of uridylated polyA tails of reporter GFP transcript.
B) Size distribution of uridylated polyA tails of NbPR2 transcript.

Taken together, these data showed that HESO1 can uridylate both *GFP* reporter and *PR2* mRNAs, and that expression of HESO1 changes the size profile of uridylated tails without impacting the overall uridylation levels. These results reveal functional differences between URT1 and HESO1.

5.2. Expression of HESO1-myc affects polyA tail size distributions similar to the expression of URT1-myc.

Next, I analysed the sizes of all polyA tails regardless of their uridylation status (Figure 40). As previously observed, the expression of URT1-myc reduced the proportion of *GFP* and *PR2* mRNAs with short polyA tails and increased the proportion of transcripts with long polyA tails. Interestingly, the distribution of the polyA tail sizes in the HESO1-myc samples were strikingly similar to the pattern observed in URT1-myc samples (Figure 40), even though the uridylation level is not increased as in the URT1-myc sample (Figure 38). However, the polyA size distribution in inactive HESO1$^{D66/8A}$-myc samples was identical to the distribution detected in control samples. The reduction of transcripts with short polyA tails and the shift towards longer tails observed in HESO1-myc could thus be linked to the uridylation activity of HESO1-myc. Yet, as the inactive HESO1$^{D66/8A}$-myc construct is less expressed than the HESO1-myc sample, we cannot exclude that the change in the polyA size distribution between the samples expressing active and inactive construct is due to the differential accumulation of the HESO1-myc and HESO1$^{D66/8A}$-myc proteins.

The expression of HESO1$^{D66/8A}$-myc did not result in increased proportions of transcripts with short polyA tails as observed in URT1$^{D491/3A}$-myc (Figure 40). Moreover, the high accumulation of A-rich tails that was observed upon overexpression of URT1$^{D491/3A}$-myc and suppressed by expression of ΔIDR$^{D491/3A}$-myc was not detected in HESO1$^{D66/8A}$-myc samples (Figure S10). A possible explanation is that HESO1$^{D66/8A}$-myc does not deplete the decapping machinery as does URT1$^{D491/3A}$-myc. However, this negative result may also be explained by the lower expression levels of HESO1$^{D66/8A}$-myc as compared to all versions of catalytically inactive URT1$^{D491/3A}$-myc.

Taken together, these results indicate that URT1-myc and HESO1-myc expression have similar outcomes on the size distribution of all polyA tails, but different effects on the accumulation of uridylated tails. While the dominant effect of URT1-myc expression is the important increase of transcripts with long uridylated polyA tails, the most obvious effect of HESO1-myc expression is the decrease of the population of transcripts with short uridylated polyA tails. Together with the fact that HESO1-myc expression does not increase the total proportion of uridylated tails, this observation suggests that transcripts with uridylated polyA tails are less stable in HESO1-myc

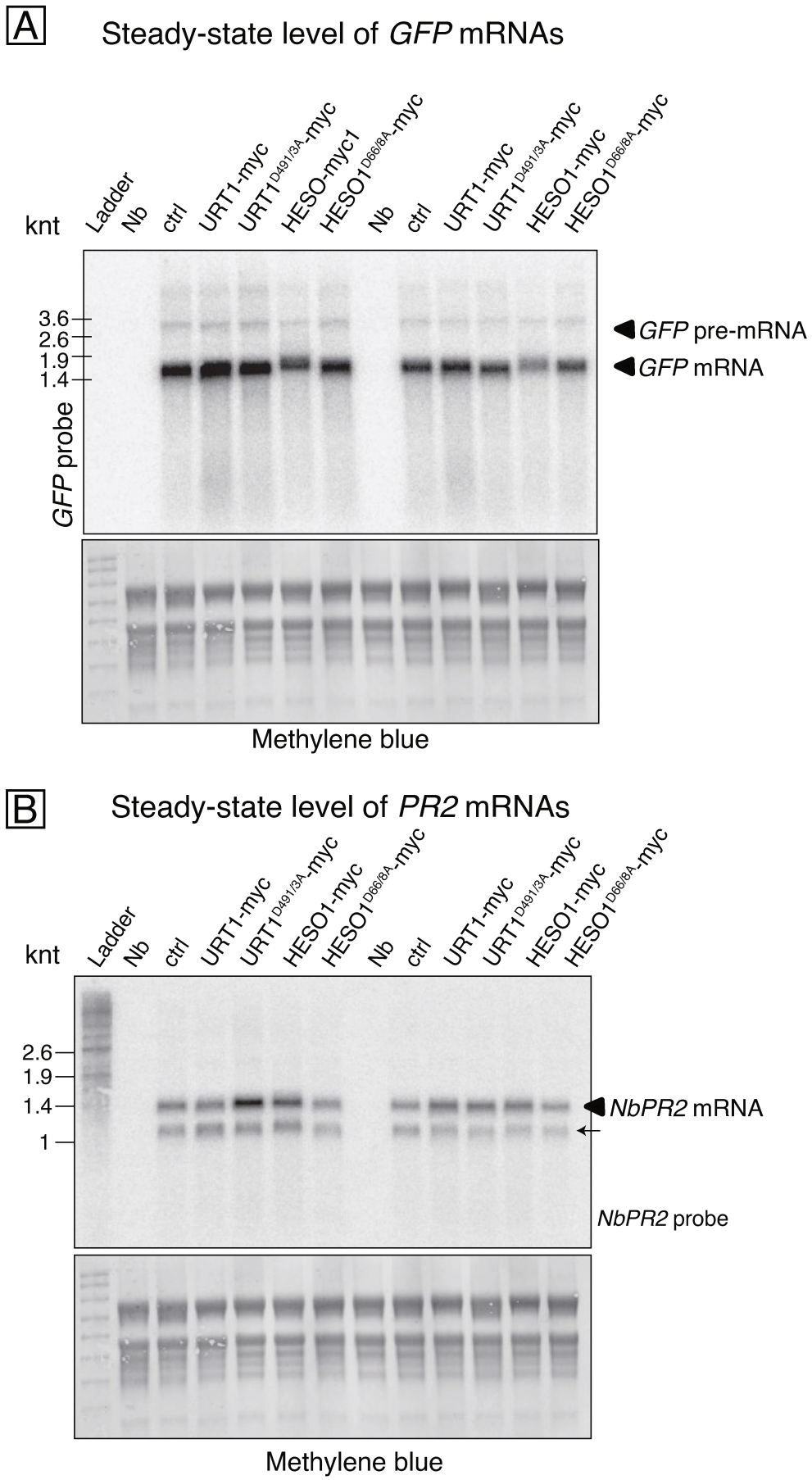


Figure 41: Northern blot analysis of reporter *GFP* and *NbPR2* mRNAs.
A) Northern blot analysis of total RNA (rep1 and 2) hybridised with a probe specific to the *GFP* mRNAs. The arrows indicate mature forms of the *GFP* mRNA. A size marker is indicated at the left. The methylene blue stain of the membrane is shown as loading control.
B) Northern blot analysis of total RNA (rep1 and 2) hybridised with a probe specific to the *NbPR2* mRNAs. Of note, the membrane has been previously hybridised with a specific probe to the *P19* mRNAs and stripped before hybridising with the *NbPR2* specific probe. The residual signals that are indicated by arrows correspond to the *P19* mRNAs.

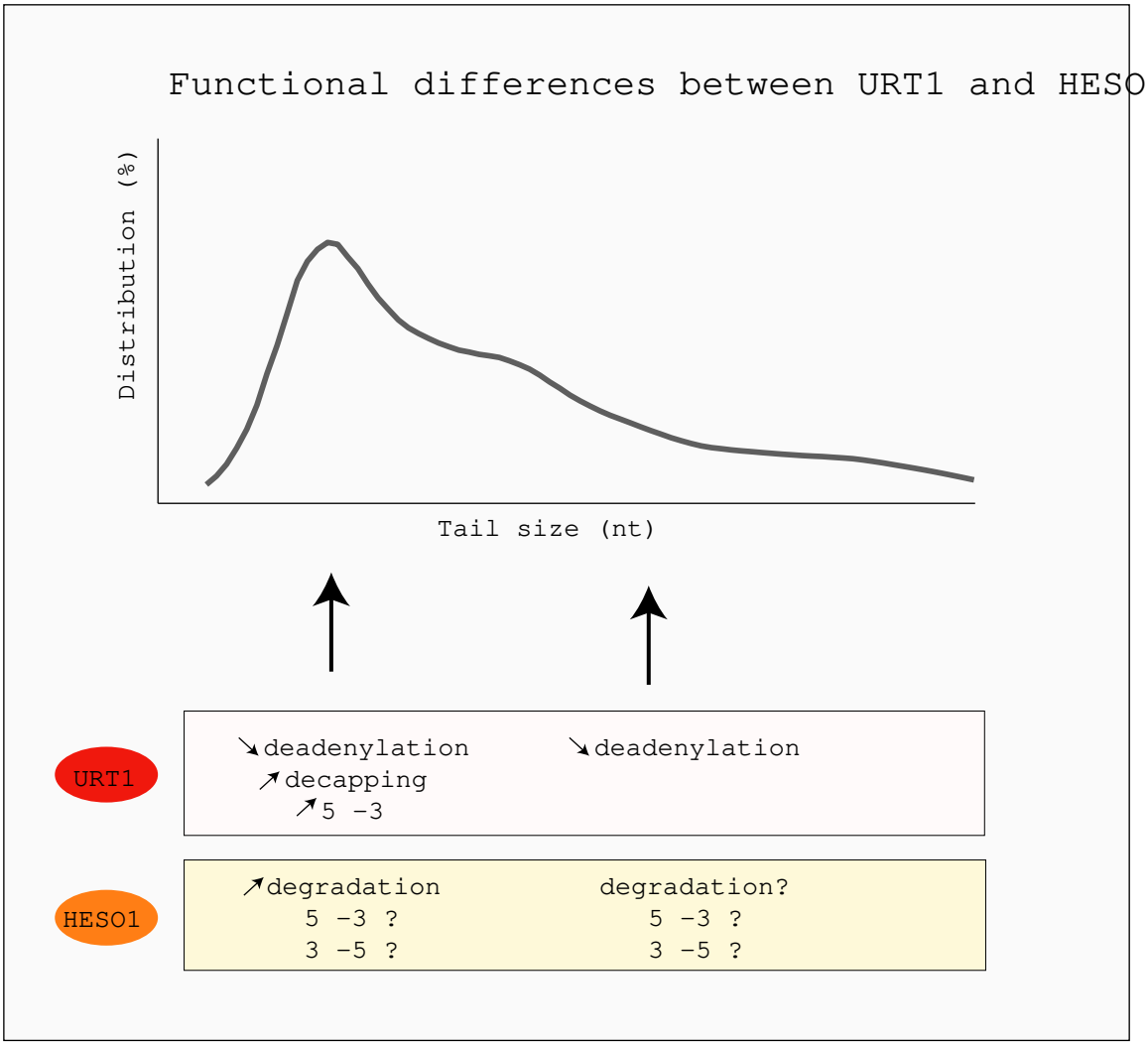


Figure 42: Summary chart of the possible functions of URT1 and HESO1.
URT1 uridylated preferentially short, but also longer polyA tails. The main functio
down deadenylation. It is also promoting the decapping and degradation from 5 to 3
short polyA tails of 10-20 nucleotides. Thus, URT1 prevents the accumulation of ex
lated mRNAs (<10 nucleotides).
HESO1 is preferentially targeting short polyA tails to promote their fast degradati
if HESO1 promotes the degradation from 5 -3 or from 3 -5 .

samples than in the URT1-myc samples. Uridylation by HESO1 *per se* could induce the rapid degradation of transcripts. Because the mRNAs that are most affected are the ones with short polyA tails, HESO1 may target mostly short tails. The elimination of mRNAs with short polyA tails leads to an indirect increase of longer polyA tails in the polyA size distribution.

5.3. Expression of HESO1-myc affects the steady-state of reporter *GFP* mRNAs.

The RNA samples obtained from infiltrated leaf patches were also analysed by northern blots (Figure 41). As observed before, the levels of the *GFP* reporter mRNAs were similar in controls and in samples expressing URT1-myc or URT1$^{D491/3A}$-myc. Equal levels of GFP mRNA were also detected in HESO1$^{D66/8A}$-myc. By contrast, samples expressing catalytically active HESO1-myc presented a clear shift towards longer *GFP* mRNAs as compared to the controls (Figure 41A). These observations support well the working hypothesis that HESO1-mediated uridylation accelerates the turnover of the *GFP* reporter mRNAs with short polyA tails. In contrast to URT1, that increases highly the number of long, uridylated polyA tails when transiently expressed in *N. benthamiana*, expressing HESO1 may lead to the fast degradation of oligoadenylated mRNAs. Either, the addition of longer U-extensions by HESO1 leads to the subsequent degradation of the targeted mRNAs, or HESO1 directly recruits decay factors to its targets and promotes their degradation. For instance, a possible candidate for an exoribonuclease degrading deadenylated reporter mRNA with uridylated short polyA tails quickly from the 3' end is the *N. benthamiana* homolog of SOV/DIS3L2.

NbPR2 transcripts however do not seem to be affected as much by the expression of HESO1-myc in the two tested replicates (Figure 41B). As already seen previously, the polyA size distribution profiles of *NbPR2* seem to be less altered upon HESO1-myc expression. Thus, for a yet unknown reason, *NbPR2* mRNAs may be less targeted and destabilised by HESO1 than *GFP* reporter mRNAs, explaining a better homeostasis of *PR2* mRNAs upon HESO1 overexpression. Of note, slight variations are seen for the *NbPR2* mRNA steady-state level in the URT1$^{D491/3A}$-myc samples. However, *NbPR2* mRNAs are accumulating in 3 out of 4 tested biological samples (compare to Figure 41B to Figure 28).

In conclusion, HESO1 does not lead to increased proportions of uridylated mRNAs and does rather suppress the accumulation of transcripts with short polyA tails in *N. benthamiana*. These results give a first insight into the possible functions of HESO1 in regulation mRNA metabolism in plants and highlight functional differences between URT1 and HESO1 (Figure 42 for summary).

Discussion

1. URT1-mediated uridylation controls the extent of mRNA deadenylation.

Using *N. benthamiana* as transient expression system, I have shown that uridylation by URT1 on reporter *GFP* and *NbPR2* mRNAs leads to a dramatic change in their polyadenylation profile: the decrease of short polyA tails and an increase of long polyA tails. In addition, I provided evidence that URT1 is implicated in the turnover of mRNAs with short polyA tails of 10-20 nucleotides. Two mutually non-exclusive hypotheses were drawn from these results: (1) Uridylation *per se* slows down deadenylation, thereby decreasing the formation of the short polyA tail population. (2) URT1-mediated uridylation promotes the degradation of mRNAs with short tails, which indirectly increases the proportion of mRNAs with longer tails.

The first hypothesis is supported by *in vitro* data. *In vitro* assays with purified hisGST-CAF1b and different RNA substrates showed that 1 and 2 terminal uridines added to the 3' end of an oligoA stretch impede the enzymatic activity of CAF1b. Because the other recombinant proteins I tested did not show any measurable activity under my experimental conditions, I could show this property only for CAF1b. The effects of terminal uridines on other Arabidopsis deadenylases remain to be investigated. Two very recent studies managed to reconstitute the whole CCR4/NOT complexes of *S. pombe* and humans (Webster et al., 2018b; Yi et al., 2018) and showed that the activities of both CCR4 and CAF1 are only slightly modulated upon their integration into the CCR4/NOT complex. Hence, the basic catalytic properties remain likely unchanged upon complex assembly. Therefore, we think that CAF1b activity is probably also inhibited by uridylation upon assembly of Arabidopsis CCR4/NOT complex. Whether this is also the case for CCR4 remains to be demonstrated. Interestingly, a recent study showed that deadenylation by the reconstituted CCR4/NOT complex of *S. pombe* with the two active CAF1 and CCR4 subunits is slowed down by the presence of 3 terminal cytidines, guanosines or uridines (Tang et al., 2019). The impact of uridines is lesser than for cytidines or guanosines. Yet, uridylated mRNAs are frequent *in vivo* substrates, as compared to guanylated or cytidylated mRNAs. Interestingly, the complex with inactive CAF1 and active CCR4 seem to be further inhibited by 3 terminal uridines than the complex with active CAF1 and inactive CCR4 or with both active deadenylases. This suggests that terminal uridines may slow down the activity of CCR4 more efficiently than CAF1. The data by Tang *et al.* fit to our observation that terminal uridines impact directly the activity of deadenylases and suggests that this property is conserved for the full CCR4/NOT complex in Arabidopsis. Hence, the

1 or 2 terminal uridines that are added to the 3' extremities of Arabidopsis mRNAs likely hinder the activity of deadenylases *in vivo*. In line with this hypothesis, I showed that all 16 mRNAs that I analysed by 3'RACE-seq have longer polyA tails in wild-type plants as compared to the *urt1* null mutants, even for transcript with low uridylation frequencies. Hence, our hypothesis that inhibiting deadenylation is an intrinsic property of URT1-mediated uridylation is also supported by *in vivo* data.

1.1. URT1 can uridylate longer polyA tails.

How URT1 is recruited to the polyA tails was not clear. Initially, URT1 was proposed to target only mRNAs with short oligo-adenylated polyA tails (Sement et al., 2013; Zuber et al., 2016). The optimisation of the 3'RACE-seq method in combination with the use of the *N. benthamiana* system that allowed URT1 overexpression revealed that URT1 can also uridylate much longer polyA tails than initially observed in Arabidopsis. Moreover, we also detected long uridylated polyA tails of some Arabidopsis mRNAs (*e.g. AT3G56940*, *AT2G34430* or *AT1G29920*, Figure 35). Taken together, these data demonstrate that URT1's function is not restricted to mRNAs with short polyA tails.

However, long uridylated polyA tails are less frequently detected in Arabidopsis. One explanation for this observation is that the access of URT1 to the 3' extremities of long polyA tails could be blocked by bound PABPs. In addition, the processivity of CCR4/NOT may explain why long polyA tails are less frequently uridylated in Arabidopsis. Because CCR4/NOT is suggested to be highly processive for long polyA tails, such tails are quickly shortened and degraded, and hence less accessible for URT1. It is also possible that terminal uridines could be removed more efficiently by deadenylases on long polyA tails than on short ones.

Nonetheless, Arabidopsis URT1 co-purifies in 8 out of 8 recent co-immunoprecipitation (co-IP) assays with the CCR4/NOT complex (unpublished results). Hence, the CCR4/NOT complex itself could recruit URT1 to mRNA polyA tails during deadenylation. The addition of uridines during the deadenylation process likely directly regulates the rate and/or the extent of deadenylation. As the polyA tails become short, the activity of CCR4/NOT complex could become distributive, which would facilitate the access of URT1 to the tails. URT1 would thus efficiently uridylate short polyA tails and slow down deadenylation, thereby preventing excessive deadenylation and the formation of 3' truncated mRNAs. In accordance with this idea, we detect excessively

deadenylated mRNAs in the *urt1* null mutant, especially for mRNAs that are frequently uridylated in the wild-type (Figure 36).

Slowing down deadenylation could promote the recruitment of other decay factors, such as decapping activators, to favour the 5'-3' over the 3'-5' decay pathway.

1.2. URT1 influences the turnover of mRNAs with short polyA tails.

Compelling evidence for the idea that URT1 helps to recruit decay factors to deadenylated mRNAs comes from the transient expression of mutated versions of URT1 in the *N. benthamiana* system. The overexpression of inactive URT1$^{D491/3A}$-myc lead to the accumulation of transcripts with short tails (uridylated and non-uridylated tails) of 10-20 nucleotides. Importantly, this accumulation was prevented by mutating the M1 motif in the IDR of URT1$^{D491/3A}$-myc, which was shown to participate in the interaction of URT1 with DCP5. Therefore, the compromised turnover of short oligoadenylated mRNAs is most certainly linked to the sequestration of the *N. benthamiana* DCP5 homologue by its binding to the massively overexpressed URT1$^{D491/3A}$-myc protein.

Through arginine/glycine rich motifs (RGG/RG) and FDF motif in its C-terminus, DCP5 interacts with the DCP1 and DCP2 decapping complex (Xu and Chua, 2009). By recruiting DCP5, URT1 could thus trigger decapping and subsequent 5'-3' degradation of mRNAs with polyA tails of 10-20 nucleotides. Uridylation by URT1 has no considerable impact on the steady-state levels of the few mRNA targets that have been tested to date. Nonetheless, delaying excessive deadenylation and recruiting DCP5 could be essential to confer 5'-3' polarity to the degradation of the targets. In line with this idea, the group has shown that deadenylated and uridylated mRNAs accumulate upon loss of XRN4 (Zuber et al., 2016), suggesting that the degradation of such mRNAs is at least partially ensured by decapping and 5'-3' exoribonucleolytic degradation. Such a link between uridylation and the 5'-3' decay pathway has been shown in other eukaryotes. Uridylation positively stimulates the decapping of mRNAs by favouring the binding of LSm1-7 to short uridylated mRNA polyA tails in humans, *S. pombe* and *A. nidulans* (Morozov et al., 2010; Rissland and Norbury, 2009; Song and Kiledjian, 2007). It is yet unknown if the plant homologues of LSm1-7 are implicated in the degradation of mRNAs with uridylated polyA tails, but our data suggest that URT1-mediated uridylation and the recruitment of DCP5 are important for the correct turnover of mRNAs by the 5'-3' decay pathway.

Taken together, these results suggest that URT1 finetunes the turnover of mRNAs by avoiding uncontrolled and excessive deadenylation of mRNAs and instead promotes the regulated

degradation from the 5’ end. The 5’-3’ polarity of degradation is likely essential to prevent the production of truncated proteins in the case of co-translational mRNA decay. Degradation from the 5’ end could also avoid the production of spurious small RNAs against endogenous mRNAs by the RNA-dependant RNA polymerase RDR6. Indeed, RDR6 has been shown to preferentially recognise excessively deadenylated substrates (<4 As) or polyA-less substrates *in vitro* (Baeg et al., 2017).

2. The polyA size distribution is shaped by the binding of PABPs.

In Arabidopsis, the size distribution of uridylated tails peaked at 17-20 nucleotides. This size distribution fits well to the previous observation that Arabidopsis PAB2 binds polyA tails of 16-17 nucleotides and supports the hypothesis that uridylation repairs deadenylated tails to allow binding of a single PABP (Zuber et al., 2016). PAB2 probably stabilises deadenylated mRNAs because it protects from further deadenylation and 3’-5’ exoribonucleases. However, it could also have a role in controlling translation or storage of the deadenylated mRNA.

By contrast, uridylated and non-uridylated polyA tails of less than 10 nucleotides were hardly detected in wild-type plants (Figures 35 and 36). Thus, very short polyA tails (<10 nucleotides) do not accumulate because they are quickly degraded. The drop in the size distribution at 10 nucleotides may indicate that the last PABP needs a minimum of 10 adenosines to bind. Excessive deadenylated tails (<10) that are not bound by PABPs could be accessible to 3’-5’ exoribonucleases. For instance, mRNAs that are excessively deadenylated may be quickly degraded by SOV or the exosome. Very short oligoA and uridylated tails of less than 10 nucleotides are also preferred targets of the LSm1-7 complex. The binding of the LSm1-7 complex promotes the rapid degradation of the mRNA from 5’-3’ (Chowdhury et al., 2007; He and Parker, 2001). Alternatively, excessive deadenylated mRNAs could be recognised by RDR6 and lead to the formation of secondary small RNAs against endogenous mRNAs (De Alba et al., 2015; Baeg et al., 2017).

2.1. Highly uridylated mRNAs show a dispersed and phased polyA tail size distribution *in vivo*.

Both, the genome-wide data obtained by TAIL-seq (Results Chapter 1) and the 3'RACE-seq datasets for specific mRNAs (Results Chapter 4) revealed that different mRNAs have very different uridylation frequencies. As URT1-mediated uridylation is mostly detected on short polyA tails, we expected that mRNA with high uridylation rates have predominantly short polyA tails. Surprisingly however, highly uridylated mRNA showed a rather dispersed polyA tail size distribution, with polyA tails spread across a wide size range (Figure 36). By contrast, mRNAs with low uridylation frequencies (<2%) had polyA tail size distributions mostly below 50 nucleotides. A possible explanation for this observation is that the deadenylases are more active on these mRNAs, leading to an overall short polyA tail size distribution, making the polyA tails of those mRNAs less accessible for URT1. The recruitment and processivity of CCR4/NOT can be affected by RNA binding proteins that recognise specific sequence elements within mRNAs, or by deadenylase activators such as members of the BTG/Tob family in mammals (see Figure 6 and Introduction Chapter 3).

Highly translated mRNAs have many PABPs bound to the polyA tail that block the action of CAF1 and are only deadenylated by CCR4 (Webster, 2018). This leads to a slower shortening of the polyA tail, and terminal uridines could further affect the activity of CCR4/NOT. We did not evaluate the translation rates of the mRNAs with high uridylation frequencies so far. Yet, it would be interesting to assess the possible correlation between dispersed polyA tail size distributions, translation efficiencies and uridylation frequencies. As discussed above, URT1-mediated uridylation could have a prime role in favouring the 5'-3' polarity of co-translational decay. Preventing the production of 3' truncated proteins may be particularly important for highly translated mRNAs.

In good agreement with our results, a recent study in *C. elegans* showed that the polyA tails of highly expressed mRNAs had size distributions that spanned the entire range of detectable tail lengths (Lima et al., 2017). They also noticed that polyA tails of highly expressed genes are shortened to defined lengths, showing a phasing pattern of *circa* 30 nucleotides. Their hypothesis is that these phased profiles resulted from the protection of the polyA tail by PABP and the temporary docking of the CCR4 deadenylase on each PABP before displacing the protein efficiently from the tail. Interestingly, we observe such phased profiles for the polyA tail size distributions of both the reporter *GFP* mRNA and the endogenous *NbPR2* mRNAs in *N. benthamiana*. The size

distribution of *GFP* mRNA polyA tails showed a first peak at 19 nucleotides, and a second peak at 37-41 nucleotides (Figure 19). The *NbPR2* polyA tail size distribution showed a peak at 20 nucleotides and a second peak at 37 nucleotides (Figure 25). Depending on the organism, the footprint of a single PABP with all 4 RRMs is 25-34 nucleotides (Baer and Kornberg, 1980; Lima et al., 2017; Smith et al., 1997; Wang et al., 1999; Yi et al., 2018). The patterns that were observed in the polyA tail size distributions could correspond to the footprint of a PABP. The phasing size of about 20 nucleotides or less indicates that the *N. benthamiana* PABP's are not engaged with all four RRMs on the polyA tail. The first peak at 20 nucleotides could correspond to a PABP that is partially bound to the 3'UTR, and the second peak at 37 a PABP that is bound with all 4RRMs to the polyA tail.

The size distributions of uridylated tails also showed a phased profile for both mRNAs in samples overexpressing URT1-myc. Similarly, to the size distribution of non-modified tails, the size distribution of uridylated tails had peaks at 19 and 41 nucleotides (Figure 19 and 26). Interestingly, we observed an important shift of the 19 nucleotide-peak towards shorter uridylated tails (14 nucleotides for GFP and 17 nucleotides for NbPR2) in control and URT1$^{D491/3A}$-myc samples. This result suggests that a tail size of 14-17 nucleotides is sufficient for the binding of a single PABP. A recent study proposed a model in which deadenylation pushes the last remaining PABP into the 3'UTR. As the RRM4 is not strictly specific for homopolymeric adenosines, it could anchor the PABP to the remaining short tails trough interactions with the 3' UTR when the tails are shortened beyond its conventional binding site (Webster et al., 2018b). Hence, the peaks at 14 and 17 nucleotides could be indeed the footprints of PABPs that were pushed into the 3' UTRs, and the difference between the two mRNAs could be linked to the sequence of the composition of the 3' UTRs. For example, the RRM4 motif of human PABP has been shown to bind especially AU-rich sequences (Sladic et al., 2004), and polyA stretches that are interrupted by other nucleotides (Khanam et al., 2006). I therefore compared the sequences of the 20 nucleotides upstream of the polyadenylation sites in the 3' UTRs of *GFP* and *NbPR2* mRNAs. 65% of the 3' end of *NbPR2* were composed of uridines and adenosines compared to 85% for *GFP* mRNA 3' ends. Thus, the shorter footprint in the size distribution of uridylated tails for *GFP* mRNAs could be due to a more stable interaction of RRM4 of the PABP with the 3'UTR.

The shift between the peak at 19 nucleotides that is observed for URT1-myc samples compared to the peak at 14 nucleotides of the control and inactive URT1$^{D491/3A}$-myc samples can be explained by the fact that the expression of URT1-myc in *N. benthamiana* slowed down the formation of

short polyA tails. Thus, as the uridylated polyA tail are generally longer in these samples, the PABP is most likely entirely bound to the polyA tail and not pushed into the 3'UTR.

The peaks of the size distribution profiles of non-modified tails were also at different positions for the 16 Arabidopsis mRNAs that I analysed by 3'RACE-seq (Figure 36). For example, the size distribution for the tails of *AT2G46820* has a first peak at 16/17 nucleotides and a second peak at 37/38, whereas polyA tails of *AT1G29910* has a peak at 25/26 nucleotides and a second peak at 45 nucleotides (Figure 36, wild-type).

Interestingly, the distance between the first and second peak was *circa* 20 nucleotides for all analysed mRNAs, whereas the first peak was between 17 to 25 nucleotides. Slight differences are also seen for the uridylated tails, for which the maximum of the distribution varies between 13 and 20 nucleotides (Figure 35). To understand whether the composition of the 3'UTRs may explain the pruning profiles of the size distribution for non-modified polyA tails, I compared the polyadenylation sites (up to 20 nucleotides upstream of the cleavage site) for *AT1G29920*, *AT3G56940* and *AT1G29910* all of which had their first peak at 25 nucleotides, with the polyadenylation site of *AT5G42530*, *AT1G24160* and *AT2G64820* which had their first peak at 17 nucleotides (Figure 36). I did not observe a clear correlation between the position of the first peak in the size distribution and the AU content of the 3'UTRs. However, the binding of PABP may not only depend on the sequence but also on the secondary structure of the 3'UTR, or the binding of other RBP to the polyadenylation site.

In summary, the results that we obtained by analysing the 3' ends of Arabidopsis mRNAs resembled the results that we obtained for the 3' ends of *NbPR2* and *GFP* in *N. benthamiana*. Our data fit well with the model that the last PABP can slide alongside the polyA tail and move further in the 3'UTR when the polyA tails are shortened. This likely also protects the remaining oligoA tail from further deadenylation. Thus, URT1-mediated uridylation and PABP cooperate in preventing the production of excessively deadenylated mRNAs. Of note, the poly A tails size distributions of frequently uridylated mRNAs seem to have more pronounced phasing profiles that is widely expanded from 10 to 80 nucleotides (Figure 36). mRNAs with low uridylation frequencies on the other side have a more centred polyA tail size distribution, restricted between 10 and 50 nucleotides. Highly uridylated mRNAs could have thus more PABPs bound to the polyA tails, which are then slowly degraded mainly by the CCR4 deadenylase of the CCR4/NOT complex.

3. HESO1 and URT1 may have overlapping and specific functions in the turnover of mRNAs.

URT1 is the main TUTase that uridylates mRNAs in Arabidopsis, but the detection of residual mRNA uridylation in *urt1* and *urt1 xrn4* null mutants showed that at least one other TUTase can uridylate mRNAs, at least in absence of URT1 (Zuber et al., 2016). This residual uridylation is also observed for the 16 mRNAs that I analysed in this study (Figure 34 and 35). The fact that this residual uridylation is completely abolished in *urt1 heso1* double mutants (unpublished results by Hélène Zuber) reveals that the second TUTase is HESO1. Uridylation by HESO1 was shown to trigger the 3'-5' degradation of both mi- and siRNAs (Ren et al., 2012; Tu et al., 2015; Wang et al., 2015; Zhao et al., 2012b). In addition, U-tailing of 5' RISC cleaved fragments by HESO1 incudes their rapid degradation by both 3'-5' and 5'-3' decay pathways (Branscheid et al., 2015; Orban and Izaurralde, 2005; Zhang et al., 2017a). How uridylation by HESO1 influences mRNA turnover is an open question.

URT1 and HESO1 can act on common RNA substrates as they are both involved in the turnover of miRNAs and RISC-cleaved mRNA fragments (Tu et al., 2015; Wang et al., 2015; Zuber et al., 2018). Yet their activities are not fully redundant. For instance, the residual mRNA uridylation that we observe in Arabidopsis *urt1* mutants does not prevent the formation of excessively deadenylated mRNAs (Figure 36). Especially for mRNAs that are frequently uridylated by URT1, we observe a notable increase of very short polyA tails of less than 10 nucleotides. This results strongly suggests that HESO1 and URT1 have distinct functions in the turnover of mRNAs.

In contrast to URT1, expression of HESO1 did not increase the uridylation levels of the *GFP* reporter and *NbPR2* mRNAs in the *N. benthamiana* system. However, HESO1 expression did impact the polyA tail size distributions (Figure 39 and 40). This is best seen for the *GFP* reporter mRNA. The main effect of HESO1 expression was a reduction of the population of transcripts with short polyA tails of 10-20 nucleotides. The decrease of *GFP* transcripts with short polyA tails was also observed by norther blots (Figure 41A). These results suggest that uridylation by HESO1 actually promotes the rapid degradation of the target mRNA, which also explains why we cannot detect more uridylated mRNAs upon expression of HESO1-myc in *N. benthamiana.* Similar to the effects of HESO1-mediated uridylation on mi- and siRNAs in Arabidopsis, mRNA uridylation by HESO1 could promote degradation by 3'-5' exoribonucleases that specifically recognise such U-tails. To validate

this hypothesis, it would be interesting to test whether the *GFP* transcripts are stabilised upon downregulating *N. benthamiana* SOV or components of the exosome. In addition, mRNA uridylation by HESO1 may also trigger decapping and 5'-3' degradation by XRN4, although an interaction of HESO1 with components of the 5'-3' pathway has not been revealed yet.

In summary, these preliminary results suggest that HESO1 and URT1 have distinct roles in the turnover of mRNAs *in vivo*. HESO1 seems to preferentially targets transcripts with short polyA tails and could induce their rapid decay. By contrast, overexpression of URT1 leads to the accumulation transcripts with very long uridylated tails, probably due to two effects, the inhibition of deadenylases and the recruitment of DCP5 and other components of the 5'-3' decay pathway to mRNAs with short polyA tails of 10-20 nucleotides. As our hypothesis is that uridylation *per se* slows down deadenylation, the distinct specificities and consequences of URT1- or HESO1-mediated uridylation could be linked to their processivity and/ or their interaction networks.

Considering that URT1 and HESO1 seem to impact the turnover of mRNAs with specific polyA tail sizes, it would be interesting to assess and compare the half-lives of mRNAs in wild-type, *urt1* and *heso1* mutant backgrounds. I tried to develop a strategy using actinomycin D to block mRNA transcription and evaluate the half-lives of specific transcripts by qPCR, combined with 3'RACE-seq to assess polyA tail lengths and modifications during their degradation. The combination of the protocols worked well, however, the main issue of this approach was the limited uptake of actinomycin D (or even cordycepin) into the plant material. I tested the actinomycin D uptake of plants grown for 7, 14 or 21 days on solid or liquid medium with various compositions, constant light or dark, and long day conditions (16h light/8h dark). However, in none of the tested conditions we observed efficient and reproducible transcription inhibition. The only protocols that allowed a good uptake of the drugs relied on vacuum infiltration, which turned out to induce major artefacts in our hands as indicated by a strong and irreproducible de-regulation of mRNA levels in mock-treated samples. An efficient protocol, such as metabolic labelling, that would allow to measure mRNA half-lives in different tissues is still to be developed for plants.

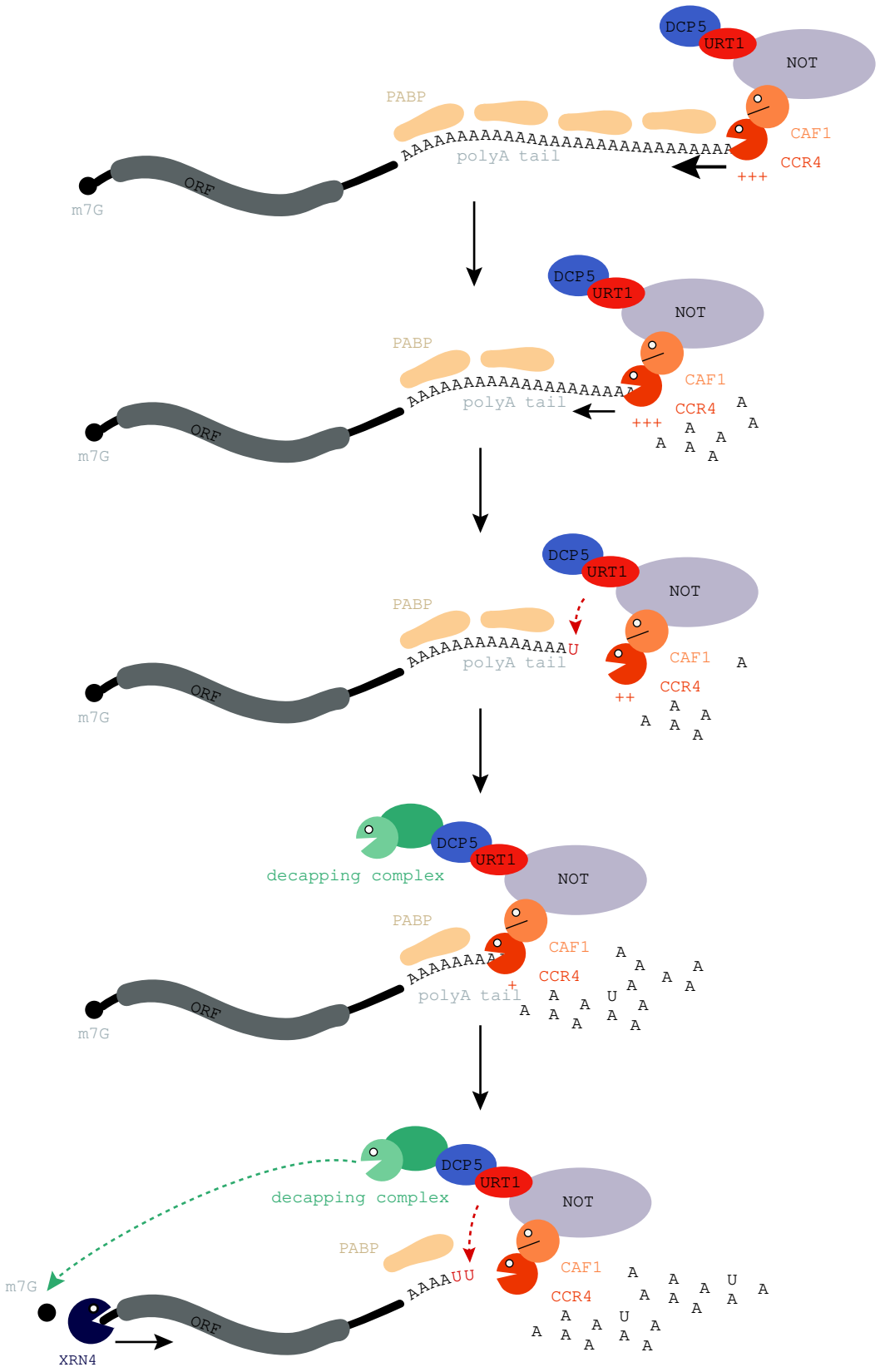
4. General Conclusion

The transient overexpression of URT1-myc in *N. benthamiana* turned out to be an efficient means to manipulate the equilibrium between uridylation and deadenylation. URT1 is a distributive uridylyltransferase for the first added uridines. As most mRNAs have only 1 or 2 terminal uridines at the 3' end of their polyA tails *in vivo*, the *N. benthamiana* expression system was better suited than a tethering approach which artificially forces the binding of URT1 to the polyA tails and may have induced the synthesis of artificial long poly U tails.

The development and improvement of the 3'RACE-seq method allowed us to analyse longer polyA tails and to obtain more accurate polyA tail size distributions for specific mRNAs, but it is not suited for genome-wide analysis. Using TAIL-seq, we now identified a subset of highly uridylated mRNAs to be analysed by 3'RACE-seq in *urt1* and *heso1* mutants. However, due to the biases of the TAIL-seq method against longer polyA tails, we cannot use it for the accurate measurement of polyA size distributions. A recent method to determine the polyA tails size distributions in a genome-wide scale is tail-end displacement sequencing (TED-seq). TED-seq consists in the sequencing of precisely size-selected RNA-seq libraries containing 3' mRNA fragments that include the poly(A) tail (Woo et al., 2018). The sequencing of the fragments allows the identification of the targets and their 3' end polyadenylation site. Thus, the polyA tails can be estimated by the subtraction of the distance of the mapped 5' end to the 3' polyadenylation sites. Although the TED-seq approach avoids the biases linked to the sequencing of long homopolymeric sequences, this method does not allow the identification of 3' modifications. By contrast, new sequencing strategies using Oxford Nanopore Technologies (ONT) to directly sequence native RNA molecules are very promising and could be a powerful tool to measure the polyA tails lengths and modifications (Jain et al., 2016). Complementary, a newly developed R package rectifies the biases linked to the sequencing of homopolymer regions by ONT (Krause et al., 2019). This tailfindr R tool improves the polyA tail size estimations. ONT has a great potential for the genome-wide sequencing of mRNA polyA tails. Comparing genome wide polyA tail size distributions in different mutants and conditions is of great interest to us. *Inter alia*, it would allow us to test the hypothesis that highly uridylated tails have a more dispersed and phased polyA size distribution in Arabidopsis in a genome-wide scale.

The experiments that I realised during my PhD led to a better understanding of how URT1 controls mRNA deadenylation. My results show that URT1-mediated uridylation slows down

Uridylation by URT1 slows down deadenylation and recruits decapping fa



Model: URT1-mediated uridylation slows down deadenylation and promo
decapping.
URT1 interacts with DCP5 via its M1 motif. When the CCR4/NOT complex becomes di
URT1 adds uridines at the 3 ends of the polyA tail, which intrinsically slows
preventing thereby the production of excessively deadenylated mRNAs. When the
starts to be short (< 25 nucleotides approximately), DCP5 could recruit decapp
promote the decapping, leading to degradation from the 5 end of the mRNAs by X

deadenylation. Additionally, my data highlight that the interaction between URT1 and DCP5 is important for the turnover of short polyadenylated reporter mRNAs *GFP* and *NbPR2* in *N. benthamiana*. Finally, my work provided first insights into the function of HESO1 for the turnover of mRNAs.

Our most recent co-immunopurification data revealed that URT1 copurifies with the CCR4/NOT complex. Other factors that are highly enriched in the co-IPs using URT1 as bait are the decapping factor DCP5 as well as homologues of the translation repressors Dhh1/DDX6 (unpublished data). Interestingly, Dhh1/DDX6 have been shown to be recruited to the CCR4/NOT complex in human and yeast (Maillet and Collart, 2002; Mathys et al., 2014; Ozgur et al., 2015). These data support the idea that CCR4/NOT recruits URT1 to polyA tails undergoing deadenylation. Interestingly, DDX6 also regulate the translation of the targeted mRNAs (Kamenska et al., 2016). Thus, molecular links exists between mRNA deadenylation, uridylation, decapping and translation. Our current hypothesis is that uridylation by URT1 controls the extent of mRNA deadenylation by the CCR4/NOT complex in plants (Figure 43). When mRNA polyA tails are deadenylated beyond 20 nucleotides, URT1-mediated uridylation protects them from further shortening and ensures binding of one PABP. In addition, URT1 recruits DCP5 which promotes decapping and degradation by the 5'-3' pathway. Alternatively, the recruitment of DCP5 may promote translation repression. Uridylation, 5'-3' degradation and the binding of translational repressors could help to prevent that excessively deadenylated or 3'-truncated mRNAs are produced and translated. Furthermore, aberrant 3' truncated mRNAs can be recognised by RDR6, a central factor of the post-transcriptional gene silencing pathway. Some new unpublished data of our research group show that URT1 is a silencing suppressor and that the simultaneous loss of URT1 and XRN4 leads to the pronounced accumulation of siRNA produced from endogenous genes. Interestingly, the double mutation *urt1 xrn4* in Arabidopsis causes frequent premature death, and this phenotype is supressed by mutating PTGS components. Taken together these data underline the importance of URT1 mediated uridylation for mRNA homeostasis.

Materials and Methods

Materials.

1. Plant materials.

The Arabidopsis thaliana plants used in this work are of Columbia Accession (Col0). They were grown on soil at 21/18°C for 16h day / 8h night light conditions (long day conditions). The *urt1-1* (SALK_087647) T-DNA insertion mutant lines were ordered from NASC (Nottingham Arabidopsis Stock Center). Homozygote mutants were selected by PCR using genotyping primers designed using the T-DNA Primer Design software available on the SIGnAL website (http://signal.salk.edu/tdnaprimers.2.html)

Nicotiana benthamiana plants used for infiltration experiments were grown for 4 weeks under long day conditions before infiltration. The plant material was harvested 4 days after infiltration.

2. Bacterial strains.

Chemically Competent *Escherichia coli* **TOP10F'** (Invitrogen™) cells were routinely used for plasmid amplification. The strain is characterized by mutations in the *endA1* gene, inactivating intracellular endonuclease activity, and in the *recA* gene, eliminating homologous recombination to increase the amount of plasmid DNA that is produced. The genotype of the used strain is *F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL endA1 nupG*.

Chemically Competent *Escherichia coli* **BL21 DE3** cells of genotype *F- ompT gal dcm lon hsdSB(rB- mB-) λ(DE3[lacI lacUV5-T7 gene 1 ind1 sam7 nin5])* were used for protein production. They contain the phage T7 RNA polymerase gene linked to an IPTG-inducible lacUV5 promoter.

The *Agrobacterium tumefaciens* strain used for the transformation of *Nicotiana benthamiana* plants was **GV3101** (pMP90). This strain has a gentamicin resistance gene on its genome and a rifampicin resistance gene on its tumour-inducing Ti plasmid. The Ti plasmid is a helper plasmid that has been disarmed and contains no functional T-DNA region of its own. The virulence genes of the Ti plasmid ensure the transfer of sequences that are flanked by T-DNA regions in the binary vectors or expression vectors that are introduced into agrobacteria for plant transformation.

3. Vectors.

pDONR207 is a donor Gateway® plasmid with attP1 and attP2 sites, a gentamycin resistance marker and the classical *ccdB* gene.

pGWB17 is a destination Gateway® plasmid with attR1 and attR2 sites, a spectinomycin resistance marker for bacterial expression, and kanamycin and hygromycin resistance markers for plant expression. This plasmid is used for the expression of sequences of interest fused to a C-terminal tag composed of 4 myc units. The construct is expressed under the control of the 35S CaMV promoter and the Nopaline synthase terminator (T-NOS). The pGWB17 plasmid has been used in this study for the expression of the different URT1-myc constructs in the *N. benthamiana* expression system.

pBIN is a binary vector that is used for *Agrobacterium tumefaciens* plant transformation for the expression of sequences under the 35S promoter and terminator. The plasmid confers resistance to kanamycin in bacterial expression systems. This vector is used in this study to express the silencing suppressor P19 in *N. benthamiana*.

pB7FWG2 is a destination Gateway® plasmid with attR1 and attR2 sites. The plasmid confers resistance to spectinomycin in bacteria, and BASTA in plants. It allows the expression of sequences of interest fused to a C-terminal GFP tag, under the control of the 35S promoter and terminator. This plasmid has been used to express the *GFP* reporter mRNA in the expression system *N. benthamiana*.

pHGGWA is a destination Gateway® plasmid with attR1 and attR2 sites that is used in bacterial expression systems. The plasmid has a resistance marker for ampicillin and a *lacI* gene that codes for the repressor of the lactose operon. It allows the expression of sequences of interest fused to a N-terminal 6xhis and GST tag, under the control of the T7 promoter and terminator. This plasmid has been used for the production and purification of hisGST tagged CAF1 proteins.

pHMGWA is a destination Gateway® plasmid with attR1 and attR2 that is used in bacterial expression systems. The plasmid has a resistance marker for ampicillin and a *lacI* gene that codes for the repressor of the lactose operon. It allows the expression of constructs of interest fused to a N-terminal 6xhis and MBP tag, under the control of the T7 promoter and terminator. This plasmid has been used for the production and purification of hisMBP tagged CCR4 proteins.

4. Antibodies.

The @myc antibodies used in this study are mouse monoclonal antibodies produced by Roche. They are used at a dilution of 1/10000. They are recognized with secondary goat antibodies that are coupled to peroxidase (HRP) and that are directed against the heavy and light chains of mouse immunoglobulins (@GAM).

5. Primers.

Table 1: Primers used for Gateway cloning.

Gene	*Primer sequence*	*Description*
CCR4a *AT3G58560*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCTTAGCGTTATACGAG	Gateway (GW) forward primer
	GAAAACCTGTATTTTCAGGGCATGCTTAGCGTTATACGAG	TEV forward primer
	GAAAACCTGTATTTTCAGGGCAATGTGGGCCTTTCTTGTAC	TEV CCR4a EEP forward primer
	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATAAAATATTGTTTCGTCTG	GW reverse primer (with STOP codon)
CCR4b *AT3G58580*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCTGAGCGTGATCCGAG	GW forward primer
	GAAAACCTGTATTTTCAGGGCATGCTGAGCGTGATCCGAG	TEV forward primer
	GAAAACCTGTATTTTCAGGGCACCATATTGACTTCCCGTG	TEV CCR4b EEP forward primer
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAACCGCGTCTTGTCCTAG	GW reverse primer (with STOP codon)
CAF1a *AT3G44260*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCGATCATTAAACCAAAC	GW forward primer
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAAAAACCTCAAGCCCATAC	GW reverse primer (with STOP codon)
CAF1b *AT5G22250*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGATCAAATCAGAAGCAGATC	GW forward primer
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAAAAACCTCTAAACCATAC	GW reverse primer (with STOP codon)
CAF1h *AT1G15920*	GAAAACCTGTATTTTCAGGGCATGTCGCAGGCTCCGAATC	GW forward primer
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAATTTTTGGATCCAGTG	GW reverse primer (with STOP codon)
CAF1i	GAAAACCTGTATTTTCAGGGCATGGCGGAGACTTTGAAAG	GW forward primer

AT4G10960	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAATTAGCAACAGTAGTAG	GW reverse primer (with STOP codon)
CAF1j AT1G80780	GAAAACCTGTATTTTCAGGGCATGTCTCTGTTTCTGAAAG	GW forward primer
	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAGATAGCAACCTGACC	GW reverse primer (with STOP codon)
CAF1k AT2G32070	GAAAACCTGTATTTTCAGGGCATGTCACTGTTTCTAAAAG	GW forward primer
	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAATGAACAATCTGACC	GW reverse primer (with STOP codon)
HESO1 AT2G39740	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGTAGAAACCCTTTC	GW forward primer
	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTACTGCTCATGTCTCGG	GW reverse primer (without STOP codon)
TEV	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAAACCTGTATTTTCAGGGC	GW forward primer

Table 2: Primers used for 3'RACE-seq and TAIL-seq.

Gene	*Primer sequence*	*Description*
AT4G38770	ATTGTGAAGAAGCCGTGTCC	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCATTCACCACCCCAAGTTCG	PCR2 forward primer with Illumina sequence
AT1G29910	GATAGAGAACCTTGCTGACC	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCTTCGCAACCAACTTTGTTC	PCR2 forward primer with Illumina sequence
AT5G42530	TCTCCACGAATCCCCTTTTG	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCAAACCTACCTTCTCCTCCGC	PCR2 forward primer with Illumina sequence
AT2G46820	GGTTACACAGGATGGTTCACTTAC	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCGAGAAGGTCAAGAGCACATAC	PCR2 forward primer with Illumina sequence
AT2G21660	AGGATACGGTGGTGGTGAAG	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCCAATGAATGTTCTCTCTCTCG	PCR2 forward primer with Illumina sequence
AT3G56940	GGTTACTTCTTTAGCCTCTGAG	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCGCCGAGTTTGAGCCTAATC	PCR2 forward primer with Illumina sequence
AT5G19140	AGTTCCTGCTAATGAGGGAG	PCR1 forward primer

	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCAAGGTGGAAGGAGCAACAG	PCR2 forward primer with Illumina sequence
AT2G41430	GAAAAGCCAGCGAAATGGGG	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCAGAAACATCCACCAGCCTCG	PCR2 forward primer with Illumina sequence
AT3G05880	TACGCCATTTATGTCCTCAC	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCTACGCCATTTATGTCCTCAC	PCR2 forward primer with Illumina sequence
AT2G34430	AAAGGGACCGTTGGAGAACC	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCTTCGCTACCAACTTCGTCCC	PCR2 forward primer with Illumina sequence
AT2G30570	GGTTGGATTCTGTTTGGAGTC	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCGAGGAGGATGAAGAATCTG	PCR2 forward primer with Illumina sequence
AT4G22150	CGACATTCGTGGTTTCATTGAG	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCTCAGACCATTGAACAAGCCG	PCR2 forward primer with Illumina sequence
AT4G02770	GTTTCCTAACGGTGAGGTTC	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCAGGAAGAGAAGGTGTTGGTC	PCR2 forward primer with Illumina sequence
AT1G29920	ACCGATAGAGAACCTTGCTG	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCTGCCACCAACTTTGTTCCCG	PCR2 forward primer with Illumina sequence
BAM3	GGAAGAAGGGATAGACATTGTGC	PCR1 forward primer
AT4G20270	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCTGTGGCAATGCTATGTGTGC	PCR2 forward primer with Illumina sequence
AT1G24160	GGTTCTCTTCCAATGCGGA	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCCACCAATGTTGTTCCATCAG	PCR2 forward primer with Illumina sequence
GFP	AACGAAAAGAGAGACCAC	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCCTAGAGTCCGCAAAAATCACC	PCR2 forward primer with Illumina sequence
NbPR2 Niben101Scf04869g03002.1	GCCATGTTTGATGAAAAT	PCR1 forward primer
	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGATCTGAGTAATAAGGAGAACTGC	PCR2 forward primer with Illumina sequence
3'RACE-seq reverse primer	CCTTGGCACCCGAGAATTCCA	PCR1 reverse primer complemented to the RT primer sequence
Illumina RPI reverse primer	CAAGCAGAAGACGGCATACGAGATXXXXXXGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA	PCR2 Illumina RPI reverse primer with index

RT primer	GCCTTGGCACCCGAGAA	RT reverse primer
3'RACE-seq 3' adapter	/5rApp/CTGACNNNNNNNNNNNNNNNTGGAATTCTCGGGTGCCAAGGC/3ddC/	3' adapter for 3'RACE-seq
TAIL-seq 3' adapter	/5App/CTGACNNNNNNNNNNNNNNNTGGAATTCTCGGGTGCCAAGGC/iBiodT//iBiodT//3ddC/	3' adapter for TAIL-seq
5' adapter	5'GUUCAGAGUUCUACAGUCCGACGAU-3'	5' adapter for TAIL-seq

Table 3: Primers used for mutagenesis.

Gene	*Primer sequence*	*Description*
CCR4a AT3G58560	GTTTGTCTACAAGCGGTACAGAATGATC	Forward primer for E299A mutation
	GATCATTCTGTACCGCTTGTAGACAAAC	Reverse primer for E299A mutation
CCR4b AT3G58580	GATGTAGTTTGCCTGCAAGCTGTACAAAGTGATCATTTC	Forward primer for E302A mutation
	GAAATGATCACTTTGTACAGCTTGCAGGCAAACTACATC	Reverse primer for E302A mutation
CAF1a AT3G44260	CCATTTATCTCAATGGCTACAGAATTTCCCGGCG	Forward primer for D47A mutation
	CGCCGGGAAATTCTGTAGCCATTGAGATAAATGG	Reverse primer for D47A mutation
CAF1b AT5G22250	CATTCATCTCAATGGCTACAGAATTTCCCGG	Forward primer for D42A mutation
	CCGGGAAATTCTGTAGCCATTGAGATGAATG	Reverse primer for D42A mutation
CAF1h AT1G15920	CTTACGTAGCCATGGCGACAGAGTTCCCAGG	Forward primer for D44A mutation
	CCTGGGAACTCTGTCGCCATGGCTACGTAAG	Reverse primer for D44A mutation
CAF1i AT4G10960	GCTATATCGCTATGGCGACTGAGTTTCCAGG	Forward primer for D40A mutation
	CCTGGAAACTCAGTCGCCATAGCGATATAGC	Reverse primer for D40A mutation
CAF1j AT1G80780	TTACGTCGCCATGGCAACTGAGTTCCCCG	Forward primer for D40A mutation
	ATGCAGCGGTACCGTTGACTCAAGGGGCC	Reverse primer for D40A mutation
CAF1k AT2G32070	CATTCGTAGCGATGGCGACAGAGTTCCCAGG	Forward primer for D40A mutation

	CCTGGGAACTCTGTCGCCATCGCTACGAATG	Reverse primer for D40A mutation
HESO1 AT2G39740	CACACGATGGGGAGCCTTAGCTATCTCTG	Forward primer for D66/8A mutations
	CAGAGATAGCTAAGGCTCCCCATCGTGTG	Reverse primer for D66/8A mutations
URT1 AT2G45620	GTTTCCCGAAGAGCGCAATCGCCGTTTGCCTTGCAATC	Forward primer for D491/3A mutations
	GATTGCAAGGCAAACGGCGATTGCGCTCTTCGGGAAAC	Reverse primer for D491/3A mutations

Methods.

1. Gateway cloning method.

The Gateway® cloning technology is a universal method based on site-specific lambda bacteriophage recombination properties. The first step of the Gateway cloning method consists of the insertion of a sequence of interest into Gateway® donor plasmids using the BP Clonase® II enzyme (ThermoFisher). This enzyme assures recombination of the *attB*-sites of the PCR products with the *attP* sites of the donor plasmid (p207 in this study). The product of the recombination is an entry vector containing the inserted sequence of interest flanked by the *attL* sites. The second step of this method is completed by the Gateway® LR Clonase® II enzyme that assures recombination of the the *attL* sites of the entry plasmid with *attR* sites of Gateway® destination vectors. The obtained expression vectors can be used for bacterial expression (*e.g.* pHGGWA and pHMGWA) or plant expression (*e.g.* pBin, pB7FWG2 and pGWB17).

1.1. PCR amplification of sequence of interest.

The genomic sequences or coding sequences of proteins of interest were amplified by PCR using the Phusion DNA polymerase (Thermo Scientific) from genomic DNA (gDNA) or complementary DNA (cDNA), respectively. The attB1 and attB2 sequences required for the Gateway cloning are introduced during the polymerization chain reaction (PCR). The reaction is carried out in a volume of 20µl containing 1x Phusion Master Mix, 0.5µM of forward and reverse primers that are specific to the sequence of interest and that contain the attB cloning sites (Table 1), and *circa* 200ng of gDNA or 30ng of cDNA. The reaction consists of a 2-minute denaturation step at 98°C, 35 cycles consisting of 30 seconds of denaturation at 98°C, 45 seconds at the hybridization temperature of

the primers and 1.5 minutes at 72°C, followed by a final elongation step of 5 minutes at 72°C. The PCR products were purified on gel with the NucleoSpin Gel and PCR clean-up kit according to the manufacturer's protocol.

1.2. Gateway BP and LR reactions.

The BP reaction is performed overnight at 25°C, in the presence of 100ng of PCR product, 100ng of the donor vector p207 and 1µl of the BP Clonase ™ II reaction mix (Invitrogen™) in a final volume of 5µl. The reaction is stopped by adding 0.2µg/µl of proteinase K and incubating the reaction 10 minutes at 37°C. The reaction mix is used to transform 50µl from competent TOP10F' *E. coli* cells. The LR reaction is performed during 1 hour at 25°C, in the presence of 100ng of purified entry plasmid, 100ng of destination vector and 1µl of the LR Clonase ™ II reaction mix (Invitrogen™) in a final volume of 5µl. The LR reaction is stopped by adding 0.2µg/µl of proteinase K and incubating the reaction 10 minutes at 37°C. The reaction mix is used to transform 50µl from competent TOP10F' *E. coli* cells.

1.3. TOP10F' bacterial transformation.

50µl of competent TOP10F' cells were added to the 6µl of BP/LR reaction and incubated for 30 minutes on ice. The heat shock was done by incubating the reaction mix for 30 seconds at 42°C. After 1 minute of incubation on ice, 1ml of liquid LB medium was added to the mix. The suspension is then incubated for 1h at 37°C. 100µl of the suspension containing the transformed bacteria are put on LB agar plates supplemented with the antibiotics for the selection of the Gateway entry or expression plasmids. The plasmid DNA preparations were made according to the standard protocol of the Macherey-Nagel "NucleoSpin Plasmid QuickPure" Kit.

1.4. Site-directed mutagenesis.

This method consists of the introduction of a site-directed mutagenesis by using as template a plasmid with the sequence of interest and primers designed to introduce the wanted mutation. Directed mutagenesis is performed in a reaction volume of 25µl in the presence of 2.5units of Pfu polymerase (Promega) in 1x of reaction buffer, 2.5pmoles of forward and reverse primers to introduce the point mutations (see Table 2), and 5nM of dNTPs. Amplification is carried out with the following PCR program: a first denaturation step of 5 minutes at 95°C followed by 18 amplification cycles composed of a denaturation step of 50 seconds at 95°C, a hybridization step of 50 seconds at 60°C, and an elongation step of 1 minute/kb of template at 68°C. The PCR ends

with a final elongation step of 30 seconds at 72°C. PCR products are treated with 10 unites of DpnI (Promega) enzyme for 1 hour at 37°C. The treated PCR products are transformed into competent TOP10F' bacteria. The primers used for the mutagenesis were designed using the primer web program (*http://www.bioinformatics.org/primerx/*). The M1M2-myc, ΔIDR-myc and $\Delta IDR^{D491/3A}$-myc constructs were cloned by Hélène Scheer.

2. *N. benthamiana* leaf infiltration.

2.1. Agrobacterial transformation of GV3101 cells.

50µl of competent GV3101 cells were added to 100ng of expression plasmid (pBIN-P19, pGWB17-URT1/HESO1 or pB7FWG2-GFP plasmids). The reaction mix is incubated 20 minutes on ice before performing the heat shock reaction. The agrobacteria are frozen in liquid nitrogen, and then incubated 5 minutes at 37°C. After 1 minute of incubation on ice, 1ml of liquid LB medium was added to the mix. The suspension was then incubated for 2h at 28°C. 1ml of the bacteria suspension is put on LB agar plates supplemented with the antibiotics for the selection of the Gateway expression plasmids and the resistance markers of the bacterial strain (rifampicin and gentamycin). After 2 days of incubation at 28°C, an isolated colony is used to start a liquid culture of 5ml using liquid LB supplemented with the selection antibiotics. The precultures are incubated 2 days at 28°C in a shaking incubator. These cultures are used to produce bacterial glycerol stocks that are stored at -80°C.

2.2. Agroinfiltration of *N. benthamiana* leaves.

10ml of liquid cultures (supplemented with antibiotics) are inoculated from 75µl of bacterial glycerol stock (or 500µl of preculture) and incubated for 20 hours at 28°C in a shaking incubator. The agrobacteria are then centrifuged at 5000g for 15 minutes. The pellets are resuspended using 5ml of agroinfiltration buffer composed of 10mM MgCl2 and 250µM acetosyringone. The absorbance of the cell suspension is measured by UV spectroscopy at 600nm and adjusted to an optical density (OD) of 1. The cell suspension containing the P19, URT1/HESO1 and GFP constructs are mixed to a 1:1:1 ratio and infiltrated into *N. benthamiana* leaves using needle-less syringes. The plant material is harvested 4 days after infiltration for RNA and protein extraction.

3. Western blot analysis of protein extractions from *N. benthamiana* leaves.

3.1. Protein extractions.

Nicotiana benthamiana plant tissues of infiltrated leaf patches are flash-frozen in liquid nitrogen and placed in Eppendorf tube containing 4mm diameter glass beads (Carl Roth). The samples are ground 8 seconds with the mixing divice Silamat® S6 (Ivoclar Vivadent) and immediately flash-frozen in liquid nitrogen. 100µl of SDS-urea extraction buffer (Tris pH 6.8 62.5 mM; urea 4M; SDS 3% (w/v); glycerol 10% (w/v); bromophenol blue 0.01%) are added to each sample before grinding a second time for 20 seconds with the Silamat mixing device. The samples are heated during 5 minutes at 95°C and centrifuged at 16000g for 5 minutes to remove the cell debris.

3.2. SDS-PAGE.

The protein extractions were separated by SDS-PAGE (polyacrylamide gel electrophoresis containing sodium lauryl sulfate). The 10% denaturating gel consists of a separation gel containing 375mM Tris-HCl (pH 6.8), 0.1% SDS (w/v) and 10% polyacrylamide/N,N' methylene bisacrylamide 37.5/1 (v/v). The concentration gel is composed of 125mM Tris-HCl (pH6.8), 0.1% SDS (w/v) and 4% polyacrylamide/N,N' methylene bisacrylamide 37.5/1 (v/v). 5-10µl of protein extractions were loaded on the gel and the migration was done at 25mA during 1 hour using SDS-PAGE migration buffer (25mM Tris, 250mM glycine, 0.1% SDS (w/v), pH 8.5).

3.3. Western blot.

Once separated by denaturating SDS-PAGE, the proteins are transferred to a 0.45µm Immobilon-P PVDF membrane (Millipore) by immersion in transfer buffer (48mM Tris; 39mM glycine; 15% methanol (v/v)) at 4°C by applying 250mA during 50 minutes (using the Bio-Rad Trans-Blot SD system). The membrane has been previously activated in 100% methanol. After transfer, the membrane is blocked in 50ml of TBS-T solution (20mM Tris-HCl (pH 7.4) ; 150 mM NaCl ; 0.2% Tween-20 (v/v)) with 5% (w/v) of skimmed milk powder for 30 minutes at room temperature. After saturation, the membrane is incubated overnight at 4°C with the primary antibodies diluted to 1/10000 in TBS-T solution with 2% of skimmed milk powder (w/v). The unbound antibodies are then removed by 5 consecutive washes of 5 minutes using TBS-T solution and incubated with the secondary antibodies @GAM for 30 minutes at room temperature. Finally, the membrane is

washed by 5 consecutive washes of 5 minutes before revelation. Muc-tagged proteins were detected by chemiluminescence using Lumi-Light Western Blotting Substrate (Roche®). Luminescence is detected busing autoradiography films or with a Fusion FX camera system.

3.4. Coomassie staining of polyacrylamide gels or PVDF membranes.

SDS-PAGE gels or PVDF membranes were rinsed with clear water and incubated with Coomassie staining solution (0.0025% Coomassie Brilliant Blue R-250, 50% methanol, 10% acetic acid) for 30 minutes. Destaining is performed using a destaining solution (50% methanol, 7% acetic acid). The coomassie dye binds to proteins through hydrophobic bonds and has a detection limit of 0.1µg.

4. Northern blot analysis of total RNA extracts.

4.1. RNA extraction.

100mg of infiltrated leaf material from *Nicotiana benthamiana* are harvested and flash-frozen in Eppendorf tubed containing 4mm diameter glass beads (Carl Roth). The samples are grinded for 8 seconds with the Silamat® mixing device at an average speed of 4500 movements per minute, before adding 750µl of Tri-Reagent® (Sigma-Aldrich). The samples are grinded a second time for 20 seconds. 400µl of chloroform is added to the suspension and the samples are homogenized by vortexing. A centrifugation of 15 minutes at 18000g is performed to separate the aqueous and phenolic phases. The aqueous phase is transferred into a new tube and total RNA is precipitated using 250µl of isopropanol. All samples are well mixed by tube inversion and incubated for 5 minutes at room temperature. Samples are centrifuged for 15 minutes at 18000g, and the pellets are washed twice with 75% ethanol (v/v) and dissolved in 100µl ultra-pure water (Milli-Q, Millipore filtration system). 100µl of a Phenol: Chloroform: Isoamyl alcohol solution (25 :24 :1, v:v, pH4.5) is added to the RNAs. After vortexing, the samples are centrifuged at room temperature for 15 minutes at 18000g. The aqueous phase is transferred in a new tube and the RNAs are precipitated by the addition of NaAc pH 5.2 (ration 1:10, v:v), 0.5µl glycogen (1mg/ml) and 100% ethanol (2.5:1 ratio, v:v). The samples are incubated for precipitation at -80°C for at least 1 hour. A 30-minute centrifugation (4°C) at 18000g is performed to precipitate the purified RNAs. The pellets are washed twice with 75% ethanol (v/v) and dissolved in 30µl of ultra-pure water (Milli-Q, Millipore filtration system). The purified RNAs are then measured by NanoDrop spectrophotometry (Thermo Scientific).

4.2. Agarose gel separation of total RNAs.

For the detection of *NbPR2* and *GFP* mRNAs, 5µg of total RNAs extracted from Benthamiana leaves are separated on a 1.5% agarose gel containing 5.55% (v/v) formaldehyde and 1x MOPS. The migration is carried out in a 1x MOPS solution for 2 hours at 70V under a chemical hood. RNAs are mixed with a 2x RNA loading dye (2x MOPS, 20% formaldehyde (v/v), 8% (v/v) formamide, 10% glycerol, 2mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanide).

4.3. RNA transfer on Hybond N+ membrane

After migration, the gel is washed 3 times for 5 minutes in ultra-pure water, and 5 minutes in a 10x SSC solution (3M sodium chloride and 0.3 sodium citrate, pH 7.0). RNAs are transferred by capillary to an Amersham's Hybond N+ nylon membrane (GE Healthcare). A nylon membrane and four identically-sized pieces of Whatmann filter paper were cut to fit the gel. All blot components were equilibrated in 10x SSC transfer buffer. The components were assembled into a sandwich by placing the membrane and the gel between 2x2 buffer-soaked filter papers. The sandwich is placed on sponges soaked in 10x SSC buffer in a container, with the membrane-side of the sandwich on the upper site. Absorbent paper and a metal plate (*circa* 1kg) are placed on the assembly to allow the diffusion of the SSC salt solution and the transfer of the RNAs to the nylon membrane.

4.4. Methylene blue coloration.

After transfer, the membrane is incubated for 5 minutes in a 2x SSC solution. The RNAs are fixed on the membrane by UV radiation at 100 x 1200 µjoules for 30 seconds with a Stratalinker (Stratagene) device. The membrane is washed for 5 minutes with ultra-pure water and coloured in a methylene blue coloration solution (0.02% of methylene blue, 0.3M sodium acetate pH5). The membrane is washed with water before being scanned. Before hybridization with the specific probes, the membrane needs to be completely destained in a 0.2x SSD, 1% SDS solution.

4.5. Preparation of radiolabeled hybridisation probes.

The template DNA of about 500 base pairs is amplified by PCR using GoTaq polymerase (Promega®), then purified on agarose gel using the NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel™). 100ng of matrix DNA are used for the production of radiolabeled probes with the Decalabel DNA Labeling kit from Thermo Scientific. The probe is marked with [α^{32}P]-dCTP according to the recommended instructions in the kit. This method consists of the incorporation of radioactive [α^{32}P]-dCTP by the Klenow fragment using random primers that can hybridize along the matrix DNA and allow the engagement of the Klenow fragment and the synthesis of complementary DNA strand. The radiolabeled probe is purified by size exclusion chromatography on a Sephadex G-50 matrix before usage. The forward primer TCTCCTATCATTATCCTCG and the reverse primer GTGTCTCCCTCAAACTTG are used for PCR amplification for the preparation of *GFP* specific probes. The forward primer ATCAAATCAAGGTCTCAAC and reverse primer ATTTTCATCAAACATGGC are used for PCR amplification for the preparation of *NbPR2* specific probes.

4.6. Hybridisation with radiolabeled probes.

The membrane is incubated for 30 minutes with the PerfectHyb® hybridization buffer at 65°C. The radiolabaled probe is added to the hybridization solution and incubated overnight at 65°C under rotation. The membrane is washed 3 times in SSC 2x, 0.1% SDS for 20 minutes before being exposed with a photo-sensible PhosphorImager screen at room temperature or to an autoradiography film at -80°C.

5. Protein production and purification.

5.1. Bacterial transformation of BL21 cells.

50µl of competent BL21 cells were transformed with 100ng of plasmid pHGGWA-CAF1 or pHMGWA-CCR4 plasmids. After the thermal shock and incubation (follow TOP10 transformation protocol), 100µl of the bacterial suspension are put on LB agar plates supplemented with ampicillin. Plates were incubated 1 day at 37°C, and an isolated colony was used for the inoculation of a 5ml preculture (LB supplemented with ampicillin and 1% of glucose).

5.2. Batch production of fusion proteins.

After overnight incubation at 37°C, 2ml of the preculture were used to inoculate 200ml of liquid LB (ampicillin + 1% glucose) at 37°C. The absorbance of the cell suspension is measured by UV spectroscopy at 600nm. When the OD_{600} is at *circa* 0.7, the culture is incubated 30 minutes at 4°C before the addition of 1mM of IPTG that induces the expression of the fusion proteins of interest. The bacterial suspension is incubated 5 hours at 20°C. A centrifugation of 15 minutes at 4000g is performed to pellet the cells. The pellet is washed twice with 50ml and 20ml of a 1x PBS solution (phosphate buffered saline solution). The pellets are finally resuspended in 25ml of lysis buffer (20mM MOPS pH 7.2, 250mM KCl, 15% glycerol, 0.1% Tween 20, 0.2mM DTT and 1x cOmplete™ Protease Inhibitor Cocktail (Roche)). Lysis buffer is added until the bacterial cell suspension reaches a DO_{600} of 20. The bacterial cell walls are disrupted by 10 sonication cycles (15s ON/30s OFF) at an amplitude of 60%. All samples are centrifuged at 16000g for 15 minutes at 4°C to remove cell debris. In parallel, 150µl of the 60% amylose suspension (New England BioLAbs Inc) and 150µl of the 60% glutathione-sepharose suspension (Glutathione SepharoseTM 4B, GE healthcare) are washed twice with 1ml lysis buffer to be resuspended in a final lysis buffer volume of 100µL. The 100µL resin mixture is added to the bacterial lysate and the suspension is incubated for 30 minutes at 4°C under rotation. After incubation, the suspension is introduced into proteus Spin columns (Generon). Purification is performed following the manufacturer's protocol (Washing buffer: 20mM MOPS, 250mM KCl, 15% glycerol and 0.1% Tween 20). The elution is done by 2 consecutive centrifugations with 5ml of elution buffer (20mM MOPS pH7.2, 150mM KCl, 15% glycerol, 0.1% Tween 20, 10mM maltose monohydrate (Sigma Aldrich 63418) or reduced glutathione (Sigma Aldrich G4251)). The two elution fractions are assembled and dialysed over-night at 4°C under stirring using dialysis tubing cellulose membranes (Sigma Aldrich D9277) in a dialyse solution (20mM MOPS ph7.2, 100mM KCl, 15% glycerol and 0.1% Tween 20). These tubing membranes have a weight cut-off of 14kDa and allow the removal of glutathione and maltose from the samples. Aliquots of the dialysed elution were instantly frozen in liquid nitrogen and stored at -80°C.

A second round of purification using NiNTA resin (His60 Ni SuperflowTM Clontech) is performed to minimised RNase contaminants (rep1 bis). The same protocol has been followed, using NiNTA Lysis buffer, Washing buffer and elution buffer. The double-purified proteins were dialysed over-night to remove imidazole from the sample.

Lysis buffer: 20mM MOPS pH 7.2, 250mM NaCl, 15% glycerol, 0.1% Tween 20, 0.2mM DTT and 1x cOmplete™ Protease Inhibitor Cocktail (Roche), 30mM Imidazole

Washing buffer: 20mM MOPS pH 7.2, 250mM NaCl, 15% glycerol, 0.1% Tween 20, 30mM Imidazole

Elution buffer: 20mM MOPS pH 7.2, 15% glycerol, 0.1% Tween 20, 300mM Imidazole

The second independent replicate was purified by affinity chromatography on glutathione sepharose coupled to size exclusion chromatography with an AKTA purification system following manufacturer's protocol.

5.3. Sypro Ruby staining.

The protein extractions were separated by SDS-PAGE (see Methods 3.2) to analyse and quantify the purified fusion proteins. The SYPRO Ruby dye (Bio-Rad) is a fluorescent dye capable of binding proteins through non-covalent interactions. This dye is of equal or greater sensitivity than silver nitrate staining and is used for the precise quantification of proteins. The gel is coloured in 50ml of SYPRO Ruby over-night at room temperature. The gel is destained in 100ml of decoloration solution (10% MeOH, 7% acetic acid) for 30 minutes. It is important to avoid contamination by keratin, which can be detected by Sypro Ruby giving more or less important background noises. The SYPRO Ruby has two excitation peaks, at 450nm and 280nm, and a peak emission at 618nm. The gel images are viewed and scanned using a Bio-Rad ChemiDoc system. In order to quantify the purified proteins, a calibration curve was established by quantifying the fluorescence signals of a range of BSA (bovine serum albumin) proteins of known concentrations using ImageJ.

6. *In vitro* deadenylation test.

6.1. 5' end radiolabelling of RNA substrates.

5pmol of RNA substrate is added to 9.5µl of ultra-pure water and heated for 3 minutes at 70°C. Reaction is quenched on ice for 1 minute, then 1.5µl of 10x PNK buffer A (Thermo Scientific), 1µl of T4 PNK (10U, Thermo Scientific) and 25µCi (circa 25pmol) of fresh [γ-^{32}P]ATPs are added to the sample. The reaction is incubated in a water bath for 30 minutes at 37°C. The radiolabeled RNA substate is purified using Sephadex G-50 columns (Sigma Aldrich).

6.2. Deadenylation test.

The deadenylation test of hisGST-CAF1b is done in 20mM MOPS at pH7.2, 5mM of $MgCl_2$, 50mM of KCl, 7% of glycerol and 0.1% of Tween 20 (for other compositions, see Figure S8). Radiolabeled CACCAACCACU-A_{14}, CACCAACCACU-$A_{13}U_1$ and CACCAACCACU-$A_{12}U_2$ RNA substrates were incubated at 25°C for 1 hour with the purified deadenylases. 10-30nM of purified hisGST-CAF1b proteins were incubated with 17.5nM of radiolabelled substrate. 5 µl aliquots were taken at 0, 5, 10, 15, 20, 30 and 1h, and diluted with 5µl of 2xRNA loading buffer (95% (v/v) formamide, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol FF, 5 mM EDTA, 0.025% (w/v) SDS, pH 8.5) before being flash-freezed in liquid nitrogen.

6.3. UREA PAGE separation of substrates.

Aliquots were separated using 17% polyacrylamide/7M UREA gel (7.4g UREA, 8.5ml 40% acrylamide (19:1), 4ml 5x TBE, final volume of 20ml). The gel was pre-runned for 15 minutes in 1x TBE buffer at 25mA. Samples were previously denaturated 5 minutes at 95°C before loading. The migration is done at 25mA for 1 hour. After migration, the radioactive gel was put on filter paper and wrapped with plastic film before being exposed to an autoradiography film at -80°C or using a phosphorimager system.

7. 3'RACE-seq library preparation.

7.1. 3' adapter ligation and reverse transcription.

5µg of total RNAs (*N. benthamiana* infiltrated leaves of Arabidopsis flowers) are incubated with 10pmol of 3' adapter (Table 2) in a final volume of 44µl. The reaction mix is incubated 3 minutes at 65°C, then 2 minutes on ice. 5µl of T4 RNA Ligase Reaction buffer and 1µl of T4 RNA Ligase 1 (ssRNA ligase, (New England BioLAbs Inc) are added to the sample. The reaction mix is incubated 1h at 37°C. In order to remove the salts, unligated 3' adapters and enzymes present in the reaction mix, the sample is purified using purification columns of the "Nucleospin® RNA plant" kit (Macherey Nagel®) according to the manufacturer's recommendations. After purification (elution in 50µl of ultra-pure water), the ligated RNAs are precipitated with NaAc pH 5.2 (ration 1:10, v:v), 0.5µl glycogen (1mg/ml) and 100% ethanol (2.5:1 ratio, v:v). After 1h of incubation at -80°C and a centrifugation of 30 minutes (4°C) at 18000g, the RNA pellet is washed twice with 50ml of 70% ethanol. The pellet is resuspended in 11µl of ultra-pure water. Reverse transcription is performed on 2.5µg of ligated RNA in PCR strip tubes with 50pmol of RT primer (Table 2) complementary to

the 3' adapter sequence. The RT reaction is performed with the SuperScript IV reverse transcriptase (Invitrogen™) following the manufacturer's protocol in a reaction volume of 20µl.

7.2. Nested PCR amplification of 3'RACE-seq libraries.

cDNAs are used as a matrix for nested PCR. The first PCR consists of the amplification of the 3' ends of a given mRNA using a forward primer specific to the sequence of interest and a reverse primer complementary to the 3' adapter sequence (Table 2). The second PCR consists of a nested PCR on the PCR1 products with primers specific to the 5' and 3' ends of the amplicons. These primers contain the Illumina sequences that are indispensable for the hybridisation of the amplicons to the Illumina flow cell surface (Table 2). The PCR2 reverse primers also contain an index sequence of 6 nucleotides (barcode) that allows the sequencing of multiple samples in parallel (1 barcode is attributed to 1 sample). For the first PCR, 1µL of cDNA is incubated with 1 unit of GoTaq® DNA polymerase (Promega), 1x of the GoTaq® reaction buffer, 30nmol of $MgCl_2$, 4nmol of dNTPs and 1U of GoTaq® polymerase in a final reaction volume of 20µl. 10 pmol of RACE PCR1 primers are used for the PCR amplification. The PCR consists of a first denaturation step of 30 seconds at 94°C, followed by 30 amplification cycles composed of a denaturation step of 30 seconds at 94°C, a primer hybridization step of 20 seconds at 50°C and an elongation step of 30 seconds at 72°C. The PCR ends with a final elongation step of 30 seconds at 72°C. For the second PCR, 0.5µL of PCR1 product is placed in the presence of 1 unit of GoTaq® polymerase, 1x GoTaq®, 30nmol of $MgCl_2$, 4nmol of dNTPs and 10pmol of RACE PCR2 primers with Illumina sequences in a final reaction volume of 20µl. The PCR reaction consists of a first denaturation step of 1 minute at 94°C followed by 20 amplification cycles composed of a denaturation step of 30 seconds at 94°C, a hybridization step of 20 seconds at 56°C, and an elongation step of 30 seconds at 72°C. The PCR ends with a final elongation step of 30 seconds at 72°C. The quality of the amplification products is analysed on a 1.5% agarose gel.

7.3. Library purifications on AMPure beads.

PCR products are purified with AMPure XP beads (Agencourt) with a bead/PCR product ratio of 0.8. The beads are first left at room temperature for 10 minutes and then gently mixed with the amplicons of the nested PCR. The suspension is left at room temperature for 5 minutes before being transferred to a magnetic stand (DynaMag-2 Invitrogen™). After 5 minutes, the supernatant is carefully discarded and the beads are washed with 80% fleshly-made ethanol (on magnetic stand), with incubation times of 1 minutes per washing step. The ethanol is finally discarded, the

beads air-dried for 3 minutes, and the PCR products are eluted with 100µl of ultra-pure water. The elution is done by taking the tubes from the magnetic stand, carefully resuspending the beads in water and incubating the suspension 5 minutes before putting the tubes back on the magnetic stand. After 1 minute, the supernatant is carefully collected in a new tube. The elution is performed a second time. The eluted PCR products are then precipitated for 1 hour at -80°C with ethanol (2.5 volumes absolute ethanol), sodium acetate (0.1 volumes of sodium acetate 3M pH 5.2) and 0.5µl glycogen (1mg/ml). After a centrifugation of 30 minutes at 18000g, the DNA pellets are washed twice with 75% of ethanol. Finally, the pellets are resuspended in 6µl of ultra-pure water. The obtained libraries are quantified using the Qubit fluorometer (Invitrogen™) and their quality is assessed using the Bioanalyzer 2100 system (Agilent). The different samples are pooled at an equimolar ratio and 10pM of the final library is used for sequencing. 25-30% of a Phix control v3 library is sequenced in parallel. Spike-in sequences with known polyA tail lengths are added to evaluate the quality of the polyA tails size estimation during the analysis. The libraries are sequenced on a MiSeq sequencer by pair sequencing (pair-end), with 41 and 111 cycles for read 1 and read 2, respectively.

8. TAIL-seq library reparation

8.1. DNase treatment and ribodepletion.

Total RNAs of Arabidopsis flower buds are extracted using Tri-Reagent® (Sigma-Aldrich) (for protocol, see Methods 4.1). After purification, 20µg of total RNAs are treated with 40 unites of DNase I (Thermo Scientific AM2222) in a final volume of 200µl of 1x DNaseI buffer during 30 minutes at 37°C in a water bath. After treatment, total RNAs are purified using the RNeasy MinElute Clean-up Kit (Qiagen, 74204) following the manufacturer's instructions. The RNAs are eluted from the columns with 15µl of RNase-free water. RNAs are then ribodepleted using the RiboMinus Plant kit (Invitrogen™) following the manufacturer's instructions. *Per* sample, two reactions are done on 10µg of total RNAs in a final volume of 10µl. The pellets are resuspended in 8.5µl ultra-pure water (The two replicates are pooled together by resuspending the first pellet with 8.5µl of water and the second pellet with the 8.5µl of the first elution). The ribodepletion efficiency is checked using the Bioanalyzer 2100 system (Agilent).

8.2. 3’ adapter ligation and RNase T1 treatment.

The 3’ adapter ligation is done using the T4 RNA ligase 1 (ssRNA ligase, New England BioLAbs Inc) as in the 3’RACE-seq protocol. 10pmol of 3’ adapter is added to 6.5µl of ribodepleted RNA and incubated 2 minutes at 70°C. After incubation of 1 minute on ice, 1µl of 10x RNA ligase buffer, 1µl of T4 RNA ligase 1 and 0.5µl of Superase-In (Ambion) are added to the reaction and incubated 1 hour at 37°C. After ligation, the sample is treated with RNase T1 (Invitrogen™) to partially digest the ligated RNA molecules. Home-made sequencing buffer containing 20mM of sodium citrate pH5, 1mM of EDTA and 7M of urea is used in the reaction. 9.5µl of the ligation reaction containing the ligated RNA’s are added to 80µl of sequencing buffer in a final volume of 100µl and incubated during 5 minutes at 50°C in a thermocycler. The reaction mix is then cooled to 22°C before adding 2µl of RNase T1. The sample is incubated for 5 minutes at 22°C, stopped with 220µl of inactivation buffer (Invitrogen™) and incubated 1 hour at -80°C. The digested RNAs are then centrifuged for 30 minutes at 18000g at 4°C and the pellet is washed twice with 75% freshly made 75% ethanol. The pellet is resuspended in 50µl of ultra-pure water.

8.3. Biotin pull-down, 5’ phosphorylation and gel purification.

To purify ligated 3’ ends of RNAs, streptavidin beads (Dynabeads M-280, Invitrogen™) were used. *Per* reaction, 50µl of dynabeads were placed in an Eppendorf tube on a magnetic stand (DynaMag-2 Invitrogen™) for at least 1 minute. The supernatant is removed and the beads are washed twice with an equal volume of freshly-made solution A (0.1M NaOH, DEPC-treated 0.05M NaCl). The washing step consists of the resuspension of the dynabeads in the solution A and incubation of 2 minutes at room-temperature before placing the tube back on the magnetic stand. After 1 minute, the supernatant is discarded and an equal volume of solution B (DEPC-treated 0.1M NaCl) is used for a third washing step. Finally, the beads are resuspended in an equal volume of 2x B&W buffer (10mM Tris-HCl pH 7.5, 1mM EDTA, 2M NaCl). 50µl of RNAs are mixed to 5µ of bead suspension and incubated at 25°C for 15 minutes (500rpm on a thermomixer). The mix is put on magnetic stand and incubated 1 minute before carefully discarding the supernatant. The beads are washed twice with 1x B&W buffer and once with 1x PNK buffer (New England BioLAbs Inc). Finally, the beads are resuspended in 50µl of PNK reaction buffer (1x PNK buffer, 2µl of 10mM ATP, 1µl of Superase-In (Ambion), 2µl of PNK and 40µl of ultra-pure water. The reaction is incubated 30 minutes at 37°C in a water bath. After phosphorylation, the reaction is put on magnetic stand and the beads are washed twice with 1x PNK buffer. The elution is done by adding 12µl of 2x RNA loading dye (95% formamide, 2.5 mg bromophenol blue, 2.5mg of xylene blue, 5mM EDTA and

0.25% SDS) and heating the suspension 3 minutes at 95°C. The tube is put on magnetic stand and the supernatant is collected in a new tube. The elution is done twice. The eluted samples are finally separated on a precast 6% TBE-UREA gel (Novex™) at 30mA for 50 minutes using 1x TBE as migration buffer. The gel is stained with SYBR gold (Invitrogen™) for 5 minutes at room temperature (in 50ml of 1x TBE buffer, SYBR gold diluted to 1/10000th). The gel is visualised using a dark reader Transilluminator (Avantor®) and the samples are cut between 300 and 1200 nucleotides. The gel slices are placed into gel breaker tubes (CliniSciences), centrifuged for 10 seconds and the grinded gel slices places in a new tube with 800µl of 0.3M NaCl. Elution is done over-night at 4°C under rotation. The mix is then placed into gel filter tubes (CliniSciences) and centrifuged 10 seconds to get rid off the gel slices. The RNAs are then purified using ethanol precipitation by dividing the elution into 2 Eppendorf tubes. The pellets resuspended in 7µ of ultra-pure water.

8.4. 5’ adapter ligation and reverse transcription.

4.2µl of purified RNAs are added to 5pmol of 5’ adapter and incubated 2 minutes at 70°C. 0.8µl of 10x RNA ligase buffer, 0.4µl of Superase-In (Ambion), 0.8µl of 10mM ATP and 0.8µl of T4 RNA ligase 1 (ssRNA ligase, New England BioLAbs Inc) are added and incubated 1 hour at 37°C. The reverse transcription is done using the Superscript III reverse transcriptase (Invitrogen™, 18080085) following the manufacturer’s protocol using 8µl of purified RNA and 50pmol of RT primer.

8.5. PCR amplification of librairies.

The libraries were amplified from 17µl of cDNAs using the DNA Phusion Polymerase (ThermoScientific). 25µl of Phusion master mix, 2.5µl of 10µM RP1 and 2.5µl of Illumina RPI reverse primer are mixed and divided into 2 tubes. The PCR consists of a first denaturation step of 30 seconds at 98°C, followed by 19 amplification cycles composed of a denaturation step of 10 seconds at 98°C, a primer hybridization step of 30 seconds at 60°C and an elongation step of 45 seconds at 72°C. The PCR ends with a final elongation step of 5 minutes at 72°C. The final libraries are combined into one tube and precipitated using ethanol. Pellets are resuspended in 10µl of ultra-pure water and 2µl of 6x DNA loading dye (Thermo Scientific) is added. The libraries are migrated on a precast 6% PAGE gel (Novex™) at 140V for 45 minutes using 1x TBE as migration buffer. The gel is coloured using SYBR gold and visualised as describes before (see Methods 8.4). The libraries are cut from 300-1000 nucleotides and eluted as describe before (Methods 8.4). The

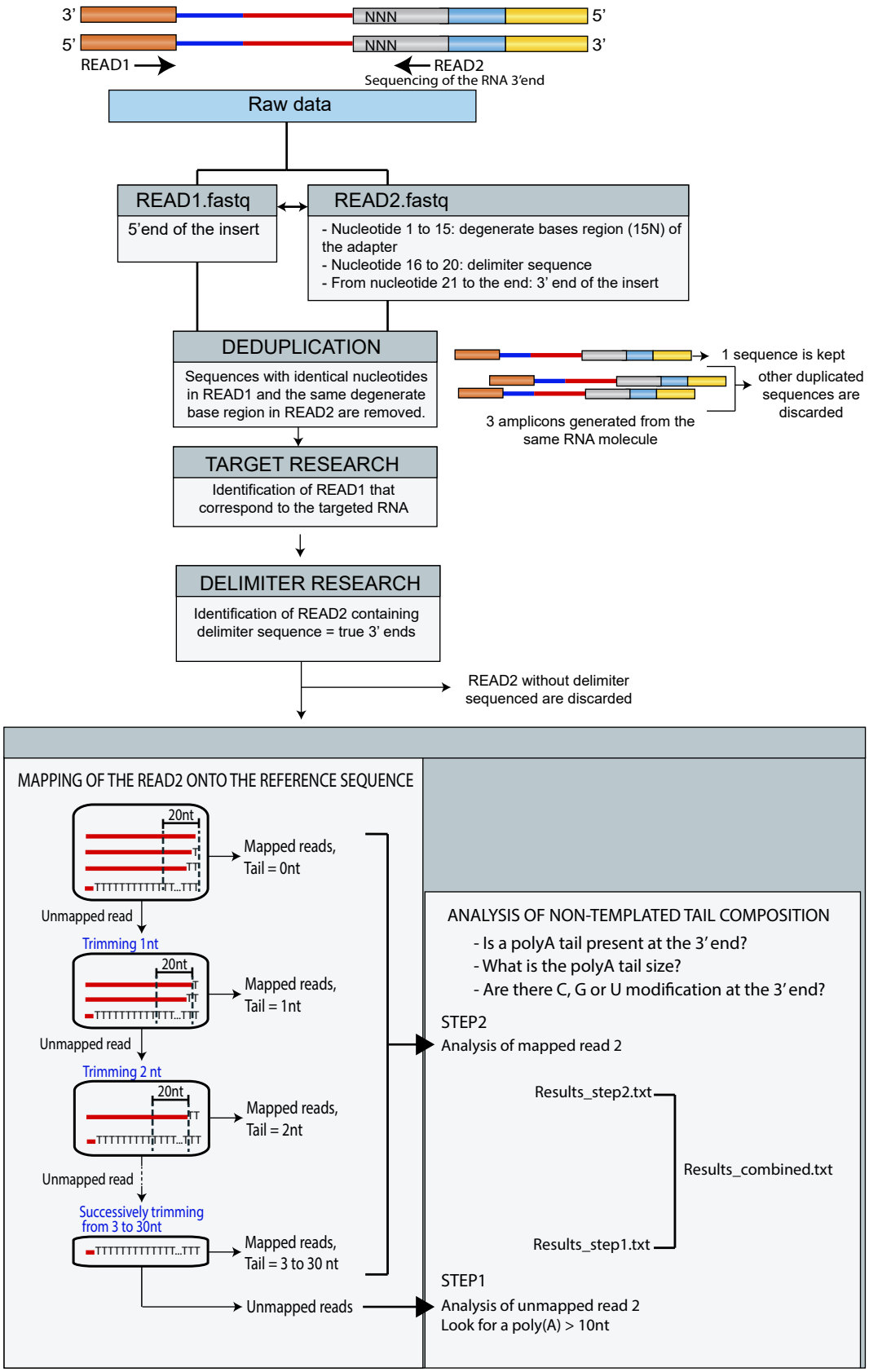


Figure 44: Flowchart of the 3'RACE-seq data analysis.

final libraries are resuspended in 7µl of ultra-pure water. 10pM of the final library is used for sequencing. 25% of a Phix control v3 library is sequenced in parallel. Spike-in sequences with known polyA tail lengths are added to evaluate the quality of the polyA tails size estimation during the analysis. The libraries are sequenced on a MiSeq sequencer by pair sequencing (pair-end), with 41 and 111 cycles for read 1 and read 2, respectively.

9. RACE-seq data analysis

After initial data processing by the MiSeq Control Software v 2.5. (Illumina), base calls were extracted and further analysed by a set of homemade scripts adapted from Sikorska et *al.* 2017 using python (v2.7), biopython (v1.63) and regex (v2.4) libraries. Reads with low quality bases (=< Q10) within the 15-base random sequence of the read 2 or within the 30 bases downstream the delimiter sequence were filtered out. Sequences with identical nucleotides in 15-base random sequence were deduplicated (Figure 44 for an overview of the flowchart). Next, the 20 first nucleotides that are sequenced downstream of the forward PCR2 primer of READ1 were used to identify the corresponding target mRNAs. One mismatch was tolerated. Matched reads 1 and their corresponding reads 2 were extracted and annotated. Reads 2 that contain the delimiter sequence were selected and subsequently trimmed from their random and delimiter sequences. Then, the analysis was divided into two steps. The aim of the first step was to identify the position of mRNA 3' extremities and to detect untemplated nucleotides. To do this, the 30 nucleotide sequences downstream of the read 2 delimiter sequence were mapped to the corresponding reference sequence, which goes from the first nucleotide of the transcript that maps the forward PCR2 primer to the end of the mRNA. Two mismatches were tolerated, with the exception of the first 10 nucleotides downstream of the mapping site that had to perfectly map. To map the 3' end position of reads 2 with untemplated tails, the sequences of the unmatched reads 2 were successively trimmed from their 3' end, with a 1 nucleotide trimming step, until they could be mapped to the reference sequence or until a maximum of 30 nucleotide has been removed. For each successfully mapped read 2, untemplated nucleotides at the 3' end were extracted. The goal of the second step was to analyse long mRNA polyA tail. Sequencing of long homopolymeric stretches causes a rapid decrease of sequencing quality, making it impossible to exactly map the 3' end of mRNA with long polyA. We thus looked for long T stretches of at least 10 Ts in the read 2 that failed to map the reference sequence. PolyA tails were searched in the first 30 cycles, which means that the maximal length of the added 3' end modification is limited to 29 nucleotides. Finally, results from step 1 and 2 were compiled and the 3' extension were analysed.

10. TAIL-seq data analysis

10.1. Base calling-based pipeline

The base calling-based pipeline was adapted from Zuber et al. 2016. As for 3'RACE-seq, the base calls were extracted after initial data processing by the MiSeq Control Software v 2.5 and the sequences having identical nucleotides in the 1st to 15th cycle in read 2 (degenerate bases in 3' adapter) were deduplicated. In order to identify the transcripts, read 1 sequences were mapped onto the *Arabidopsis thaliana* reference genome (TAIR 10) using Hisat2 (v2.1.0) (ref) and Bowtie2 (Langmead and Salzberg, 2012). The resulting alignment was annotated using the intersectBed tool from the BEDTools suite (v2.17.5) (Quinlan and Hall, 2010) and the Arabidopsis annotation file TAIR10_GFF3_genes.gff (http://www.arabidopsis.org). Reads 1 that map to cytosolic mRNAs and their corresponding reads 2 were extracted and used for further analyses. Reads 2 that contain the delimiter sequence were selected and subsequently trimmed from their random and delimiter sequences. A homemade python script was then used to extract polyA tails (>5nt) and their potential 3' end modification from the remaining read 2. PolyA tails were searched in the first 30 cycles, which means that the maximal length of the added 3' end modification is limited to 29 nucleotides. Detected polyA tail sizes and their 3' extensions were finally analysed.

10.2. PolyA length estimation using Tailseeker software

In addition to the base calling-based analysis, the polyA sizes were also estimated using the TAILseeker software (v3.1, https://github.com/hyeshik/tailseeker, Chang et al. 2014) developed by Hyeshik Chang in the lab of Narry kim (Seoul, Korea). We used the level 1 of the software which performs:

- PolyA length measurement ($\geq$ 5nt)
- Non-A additions to polyA tails
- PCR duplicate removal
- Quality check for polyA length measurement, *i.e.* estimation of the polyA size for spike-in

Default setting were used with the following parameters: "max_cctr_scan_left_space", "max_cctr_scan_right_space" and "polya_boundary_pos" that were set to 20, 20 and 120, respectively. The "third party basecaller" option was also disactivated.

10.3. Spike-in analysis

To evaluate the quality of the polyA tail size estimation during the analysis, the polyA length of the spike-in sequence was estimated using both Tailseeker software (cf above) and base calling-based pipeline. For base calling-based analysis, the 15-nucleotide sequence 'ACAGTAGCTCGTCAG' corresponding to a fixed sequence present in all spike-in were searched into reads 2 to extract true spike-in sequences. T stretches of at least 6 successive Ts (corresponding to polyA in the read 2) were then searched into the remaining read 2 and analysed for their size.

Supplementary data

Chapter 8

High-Resolution Mapping of 3' Extremities of RNA Exosome Substrates by 3' RACE-Seq

Hélène Scheer, Caroline De Almeida, Natalia Sikorska, Sandrine Koechler, Dominique Gagliardi, and Hélène Zuber

Abstract

The main 3′-5′ exoribonucleolytic activity of eukaryotic cells is provided by the RNA exosome. The exosome is constituted by a core complex of nine subunits (Exo9), which coordinates the recruitment and the activities of distinct types of cofactors. The RNA exosome cofactors confer distributive and processive 3′-5′ exoribonucleolytic, endoribonucleolytic, and RNA helicase activities. In addition, several RNA binding proteins and terminal nucleotidyltransferases also participate in the recognition of exosome RNA substrates.

To fully understand the biological roles of the exosome, the respective functions of its cofactors must be deciphered. This entails the high-resolution analysis of 3′ extremities of degradation or processing intermediates in different mutant backgrounds or growth conditions. Here, we describe a detailed 3′ RACE-seq procedure for targeted mapping of exosome substrate 3′ ends. This procedure combines a 3′ RACE protocol with Illumina sequencing to enable the high-resolution mapping of 3′ extremities and the identification of untemplated nucleotides for selected RNA targets.

Key words Exosome, rRNA maturation, Rapid amplification of cDNA 3′ end, 3′ RACE-seq, 3′ Adapter ligation, Illumina sequencing, MiSeq, Untemplated nucleotides

1 Introduction

The RNA exosome provides all eukaryotic cells with a 3′-5′ exoribonucleolytic activity which plays a central role in the processing and the degradation of many nuclear and cytosolic RNAs. Nine subunits compose the exosome core, which is also called Exo9. Exo9 is structurally related to bacterial polynucleotide phosphorylases (PNPases) and archaeal exosomes (*see* Chapters 2–4). These prokaryotic exoribonucleases are processive enzymes whose central channel accommodates three phosphorolytic active sites. By contrast, in mammals and yeast, Exo9 have lost the original phosphorolytic activity and the ribonucleolytic activity of the exosome relies on ribonucleases associated to Exo9 [1–3]. In yeast, Rrp6 confers a

John LaCava and Štěpánka Vaňáčová (eds.), *The Eukaryotic RNA Exosome: Methods and Protocols*, Methods in Molecular Biology, vol. 2062, https://doi.org/10.1007/978-1-4939-9822-7_8, © Springer Science+Business Media, LLC, part of Springer Nature 2020

distributive 3′-5′ exoribonuclease activity while Dis3/Rrp44 provides both processive 3′-5′ exoribonuclease and endonucleolytic activity [1–3]. In addition to Rrp6, the human Exo9 associates with two Dis3 homologs, Dis3 and Dis3L, the latter lacking an endoribonucleolytic active site. In Arabidopsis, homologs of RRP6 and DIS3 also contribute to exosome activity [4–6]. However the most striking difference with the mammalian and yeast Exo9 is that Arabidopsis Exo9 has retained a single site conferring a distributive and phosphorolytic activity [6].

A complex set of ribonucleolytic activities are therefore coordinated by Exo9 in eukaryotes. In addition, RNA helicases, various RNA binding proteins and terminal nucleotidyl transferases (TNTases) also assist the exosome in recognizing, degrading or maturing its RNA substrates. The respective functions of all these cofactors and the coordination of their associated activities must be determined to fully appreciate the biological functions of the RNA exosome.

One of the key aspects toward understanding the roles of each activity linked to exosome function is to analyze the degradation or trimming of exosome RNA substrates. One way to do that is to map 3′ extremities of exosome RNA substrates at high density by determining their precise position at nucleotide level and by identifying eventual untemplated nucleotides added to processed 3′ extremities. We present here a high throughput sequencing-based strategy called 3′ RACE (3′ Rapid Amplification of cDNA End)-seq. The 3′ RACE-seq method combines a modified 3′ RACE-PCR method for 3′ end analysis of cDNA [7, 8] and the Illumina sequencing technology. The classical 3′ RACE-PCR is a low-throughput method that implies cloning PCR amplicons and Sanger sequencing of individual clones. By contrast, the 3′ RACE-seq procedure allows for the simultaneous analysis of millions of amplicons for multiple samples. The 3′ RACE-seq procedure is summarized in Fig. 1. It comprises the ligation of an adapter at the 3′ end of each molecule in a total RNA sample and subsequent cDNA synthesis by using a reverse transcriptase (RT) primer complementary to the ligated 3′ adapter. Importantly, the ligated 3′ adapter, described in TAIL-seq protocol [9–12], contains a random region that allows for the removal of PCR duplicates during bioinformatics analysis and thus each final sequence corresponds to a unique RNA molecule. Chosen targets are then amplified using forward and reverse primers that bind to the target sequence and the 3′ adapter, respectively, and that comprise the Illumina sequences required for flow cell hybridization and sequencing. Finally, amplicon libraries are sequenced using MiSeq paired-end sequencing for an average yield per run of 40 million of reads: 20 million of read 1 and 20 million of read 2. The use of barcoded Illumina adapters allows for sequencing in parallel more than 30 conditions or replicates in a single MiSeq run.

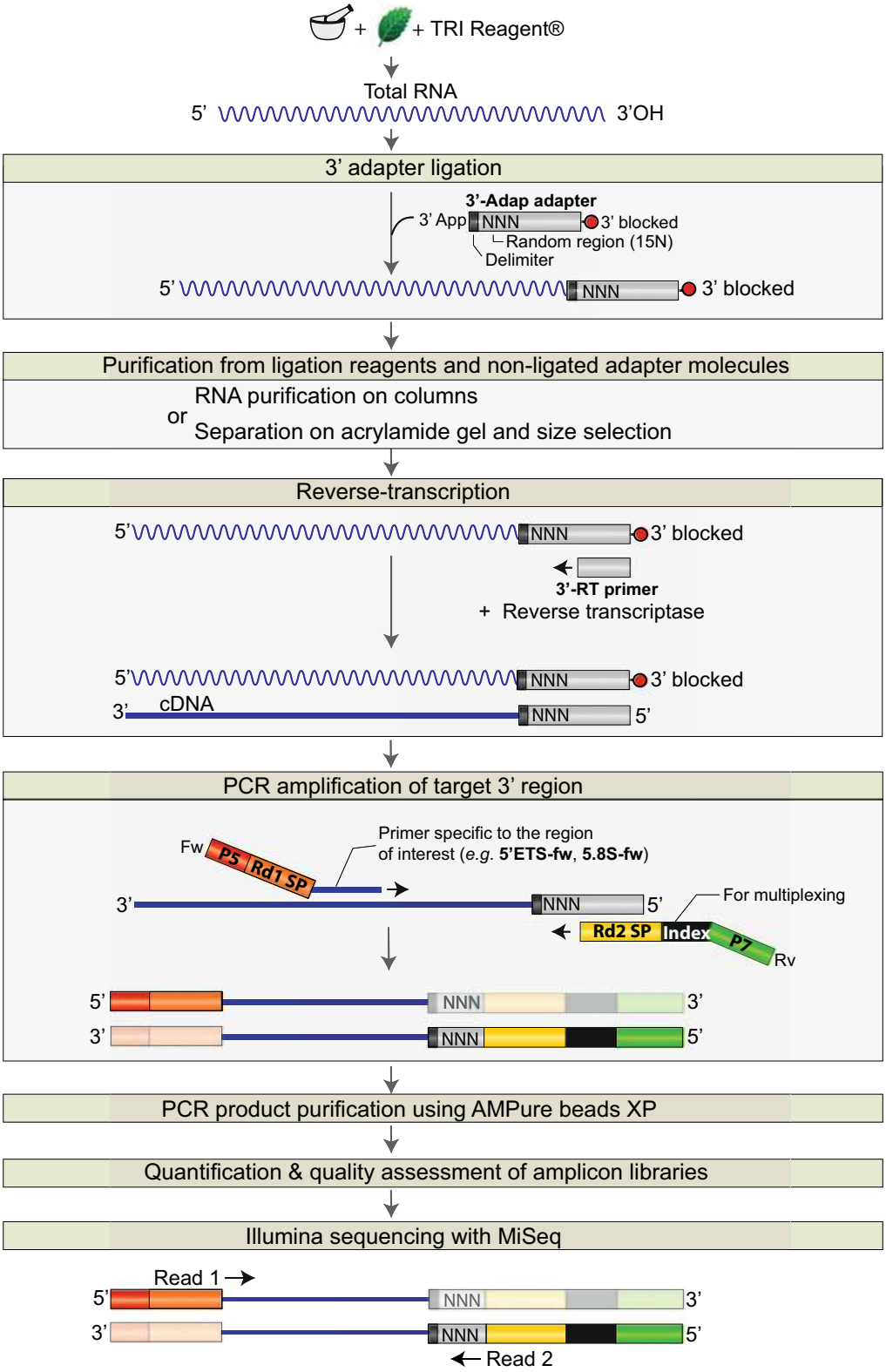


Fig. 1 Flowchart steps for 3′ RACE-seq procedure. After total RNA purification, the 3′ hydroxyl (3′ OH) end of each RNA molecule is ligated to the 3′-Adap adapter (*see* Fig. 2, Subheading 3.2 in the text). To remove nonligated 3′ adapters, ligated RNAs are purified using RNA purification columns or separated on an acrylamide gel when size-selection is possible (Subheading 3.3 in the text, *see* **Note 6**). cDNA synthesis is then initiated using the 3′-RT primer complementary to the 3′-Adap adapter (Subheading 3.4 in the text, Table 1). To specifically analyze 3′ regions of targets of interest, cDNA are PCR amplified using a forward

To illustrate the 3′ RACE-seq procedure, we present the detailed protocol adapted for two classical RNA substrates of the exosome during ribosomal RNA (rRNA) maturation. In eukaryotes, three out of four ribosomal RNAs, the 18S, 5.8S, and 25/28S rRNAs, are transcribed as a common polycistronic precursor. The 18S, 5.8S, and 25/28S rRNAs are separated by internal transcribed spacers (ITS1 and 2) and flanked by two external transcribed spacers (5′ and 3′ ETS). The production of mature rRNAs requires endonucleolytic cleavages and exoribonucleolytic processing steps to remove internal and external transcribed spacers. Two of the archetypical RNA substrates of the exosome in eukaryotes are the 5′ external transcribed spacer (5′ ETS) of the 18S-5.8S-25/28S rRNA primary transcript and the 5.8S rRNA precursors. Using 3′ RACE-seq, we have recently shown the complexity of ribonucleolytic and tailing activities that contribute to these rRNA maturation steps in Arabidopsis [6]. Of note, the 3′ RACE-seq procedure can easily be adapted to any other RNA targets, such as poly(A) tailed transcripts or RISC-cleaved fragments, with slight modifications of the protocol.

2 Materials

2.1 RNA Extraction

1. TRI Reagent® (Molecular Research Center).
2. Acid phenol (Biophenol water saturated, pH 4)–chloroform–isoamyl alcohol solution (25:24:1, v/v/v).
3. Absolute ethanol.
4. 3 M sodium acetate, pH 5.2.
5. 20 mg/ml glycogen.
6. Refrigerated microcentrifuge reaching 16,000 × *g*.
7. 75% ethanol.
8. Nuclease-free water.
9. Ultraviolet (UV) spectrophotometer (e.g., Thermo Scientific NanoDrop 2000).

Fig. 1 (continued) primer that binds specifically to the regions of interest and a reverse primer that is complementary to the ligated adapter (Subheading 3.5 in the text, Table 1). Forward and reverse primers contain P5/Rd1 SP and P7/Rd2 SP Illumina sequences, respectively. P5 and P7 sequences are used for the hybridization to the flow cell. Rd1 SP and Rd2 SP correspond to the binding site of read 1 and read 2 sequencing primers. Reverse primer also contains an index sequence, which allows for multiplexed sequencing. In order to remove primer-dimers, salts and other PCR reagents, PCR products are then purified using AMPure XP beads (Subheading 3.6 in the text). Amplicon libraries are quantified and analyzed with Bioanalyzer for quality assessment (Subheading 3.7 in the text). Libraries are paired-end sequenced on the Miseq Illumina system. Read 1 allows for the identification of the target and read 2 for the identification of the RNA 3′ end

Table 1
Oligonucleotides used in 3′ RACE-seq procedure to analyze Arabidopsis 5′ ETS P-P1 intermediates and 5.8S precursors

Oligonucleotide name	5'-3' sequence
3'-Adap	/5rApp/CTGACNNNNNNNNNNNNNNNNTGGAATTCTCGGGTGCCAAGGC/3ddC/
3'-RT	GCCTTGGCACCCGAGAA
5' ETS-fw	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGAT **CATCTCGCGCTTGTACGGCTTTG**
5.8S-fw	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGACGAT **CTCTGCCTGGGTGTCACAAATC**
Illumina RPI	CAAGCAGAAGACGGCATACGAGAT**XXXXXX**GTGACTGGAGTTCCTTGGCACCC GAGAATTCCA

All oligonucleotides are listed in the 5′ to 3′ orientation. 3′-Adap adapter contains two modifications: 5rApp = 5′, 5′-adenyl pyrophosphoryl moiety, 3ddC = 3′-dideoxy-C. 3′-Adap should be HPLC purified in an RNase-free environment. For 5′ ETS-fw and 5.8S-fw primers, bolded nucleotides correspond to the target specific sequence, while the rest of the sequence is used for hybridization to the Illumina flow cell and for sequencing. Red bold nucleotides in Illumina RPI PCR primer correspond to the index sequence (for further details *see* Illumina manufacturer's instruction for TruSeq Small RNA RPI primers [23])

10. Heating block that can heat to 65 °C.
11. 2 × RNA loading buffer: 95% (v/v) formamide, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol FF, 5 mM EDTA, 0.025% (w/v) SDS, pH 8.5.
12. Agarose.
13. 0.5× TBE (10× stock solution: 1 M Tris base, 1 M boric acid and 0.02 M EDTA, pH 8).
14. Gel system for agarose electrophoresis (well combs, casting tray, gel box) and electrophoresis power supply.
15. 10 mg/ml ethidium bromide (EtBr).
16. (*Optional*) Agilent 2100 Bioanalyzer.

2.2 3′ Adapter Ligation

1. 3′-Adap oligonucleotide (Table 1).
2. Nuclease-free water.
3. Water bath or heating block for 37 °C and 65 °C incubation.
4. 10× T4 RNA Ligase Reaction Buffer (NEB): 10 mM $MgCl_2$, 50 mM Tris–HCl, 1 mM DTT, pH 7.5.
5. 10,000 U/ml T4 ssRNA Ligase 1 (NEB).

2.3 Electrophoresis and RNA Isolation

1. 2× RNA loading buffer: 95% (v/v) formamide, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol FF, 5 mM EDTA, 0.025% (w/v) SDS, pH 8.5.
2. 40% acrylamide (19:1) solution.
3. 1× TBE (10× stock solution: 1 M Tris base, 1 M boric acid and 0.02 M EDTA).

4. Urea.
5. 10% (w/v) ammonium persulfate solution (APS).
6. *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED).
7. Gel system for PAGE (gel combs, gel cassettes and spacers) and electrophoresis power supply.
8. Water bath or heating block that can heat to 65 °C.
9. Syringe with a needle.
10. 10 mg/ml ethidium bromide (EtBr).
11. Scalpel.
12. Elution buffer: 500 mM ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA and 0.1% (w/v) SDS.
13. Rotating wheel.
14. UV light Transilluminator.
15. Acid phenol (Biophenol water saturated, pH 4)–chloroform–isoamyl alcohol solution (125:24:1).
16. Absolute ethanol.
17. 3 M sodium acetate, pH 5.2.
18. 20 mg/ml glycogen.
19. Refrigerated microcentrifuge reaching 16,000 × *g*.
20. 75% ethanol.
21. Nuclease-free water.
22. UV spectrophotometer (e.g., Thermo Scientific NanoDrop 2000).
23. Dry ice.

2.4 cDNA Synthesis

1. 3′-RT primer (Table 1).
2. 10 mM dNTP mix (dATP, dGTP, dCTP, and dTTP, each at 10 mM).
3. Nuclease-free water.
4. PCR thermal cycler.
5. 0.2 ml strip PCR tubes.
6. 5× SuperScript™ IV buffer (Invitrogen™).
7. 0.1 M DTT.
8. 40 U/μl RNaseOUT™ (Invitrogen™).
9. 200 U/μl SuperScript™ IV Reverse transcriptase (Invitrogen™).

2.5 PCR and Quality Assessment

1. Primers: forward PCR primer (target specific) and reverse PCR primer (TruSeq RNA PCR index primer, RPI, Table 1).
2. 5 U/μl DreamTaq DNA Polymerase (Thermo Fisher Scientific) supplied with 10× DreamTaq buffer.
3. 10 mM dNTP mix (dATP, dGTP, dCTP and dTTP, each at 10 mM).
4. Nuclease-free water.
5. PCR thermal cycler.
6. 0.2 ml strip PCR tubes.
7. 6× DNA loading buffer: 10 mM Tris–HCl (pH 7.6), 60% (v/v) glycerol, 0.03% (w/v) bromophenol blue, 0.03% (w/v) xylene cyanol FF, 60 mM EDTA.
8. 0.5× TBE (10× stock solution: 1 M Tris base, 1 M boric acid and 0.02 M EDTA, pH 8).
9. Agarose.
10. Gel system for agarose electrophoresis (well combs, casting tray, gel box) and electrophoresis power supply.
11. 10 mg/ml ethidium bromide (EtBr).

2.6 PCR Product Purification

1. AMPure XP beads (Agencourt).
2. Benchtop minicentrifuge.
3. Magnetic stand compatible with 1.5 ml microtubes (e.g., DynaMag-2 Invitrogen™).
4. 80% ethanol.
5. Nuclease-free water.
6. Absolute ethanol.
7. Refrigerated microcentrifuge reaching 16,000 × *g*.
8. 75% ethanol.
9. 3 M sodium acetate, pH 5.2.
10. 20 mg/ml glycogen.
11. UV spectrophotometer (e.g., Thermo Scientific NanoDrop 2000).

2.7 Qubit and Bioanalyzer Analysis of Purified Amplicons

1. Qubit fluorometric quantitation system (Invitrogen™).
2. Agilent 2100 Bioanalyzer.
3. DNA chip kit (*see* **Note 1**).
4. Microcentrifuge.

2.8 Preparing Libraries for Sequencing on MiSeq

1. Illumina MiSeq system.
2. 1.0 N NaOH.
3. PhiX control v3 library (Illumina, FC-110-3001).

4. HT1 (Hybridization Buffer provided by Illumina).
5. Benchtop microcentrifuge.

2.9 MiSeq Run and Analysis

1. MiSeq Reagent Kit v3 (Illumina, MS-102-3001) that contains:
   - Reagent Cartridge.
   - HT1 (Hybridization Buffer).
   - PR2 (Incorporation Buffer).
   - MiSeq Flow Cell.

3 Methods

3.1 RNA Extraction

1. Extract total RNA using TRI Reagent® (Molecular Research Center) according to the manufacturer's protocol.
2. A second round of purification using acid phenol–chloroform–isoamyl alcohol and a subsequent RNA precipitation can be performed in order to remove residual contaminants. Add 1 volume of acid phenol–chloroform–isoamyl alcohol solution (25:24:1) (*see* **Note 2**).
3. Vortex well and centrifuge for 15 min at 16,000 × *g*.
4. Transfer supernatant into new tube and precipitate RNA by adding 0.1 volume of 3 M sodium acetate pH 5.2, 0.5 μl of glycogen (20 mg/ml) and 2.5 volumes of absolute ethanol.
5. Mix by tube inversion.
6. Incubate for at least 1 h at −80 °C.
7. Centrifuge for 30 min at 16,000 × *g* (4 °C).
8. Discard supernatant.
9. Wash pellet with 75% ethanol (500 μl) to remove residual salt. Centrifuge 5 min at 16,000 × *g*.
10. Discard supernatant thoroughly, dry the RNA pellet and dissolve it in 20 μl of nuclease-free water.
11. Measure the RNA quantity and purity of your samples with an UV spectrophotometer (*see* **Note 2**).
12. To assess the integrity of your total RNA preparation in a quick and cheap manner, you can check the profile(s) of your sample (s) on a 1% agarose gel (*see* **Note 3**). Take a volume containing between 500 ng and 1 μg of RNA and add 1 volume of 2× RNA denaturing loading dye.
13. Heat 5 min at 65 °C and chill on ice prior to loading into the wells of agarose gel.

    Alternatively, RNA quality can be assessed on Agilent Bioanalyzer system.

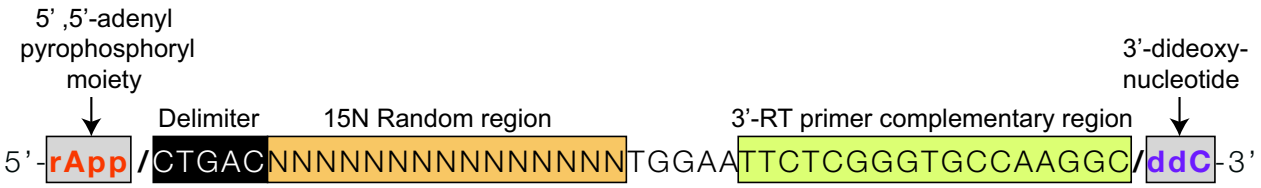


Fig. 2 Sequence details for the 3′-Adap adapter. 3′-Adap is a preadenylated oligonucleotide containing a 5′,5′-adenyl pyrophosphoryl moiety (*see* **Note 4**). The delimiter sequence allows for the demarcation between the sequence corresponding to the 3′ end of the RNA and the adapter sequence. This sequence is also used during bioinformatics analysis to discriminate read 2 containing adapter sequence from other sequences that could arise from artifacts of reverse transcription priming (Fig. 4). The 15 random bases (15 N) allow for deduplicating and therefore for eliminating PCR duplicates. The 3′-RT primer complementary region is used as template during reverse-transcription reaction. Finally, the 3′-dideoxynucleotide prevents the 5′ adenylated oligo from self-ligation

3.2 3′ Adapter Ligation

In order to analyze RNA 3′ extremities, the 5′-riboadenylated DNA oligonucleotide (3′-Adap, Table 1) is ligated at the RNA 3′ end. This primer is as described in [9] except that it is not biotinylated. The features of 3′-Adap are shown in Fig. 2.

1. Take 20 μg of total RNA for each sample.
2. Add 5 pmol of the 3′-Adap.
3. Add nuclease-free water to a final volume of 44 μl.
4. Denature sample for 3 min at 65 °C.
5. Put on ice for at least 2 min.
6. Add 5 μl of 10× T4 RNA Ligase Reaction Buffer.
7. Add 1 μl of T4 ssRNA Ligase 1 (10,000 U/ml) (*see* **Note 4**).
8. Incubate for 1 h at 37 °C (*see* **Note 5**).

3.3 RNA Separation by Denaturing PAGE and Isolation of RNA Fragments from Polyacrylamide Gel

Before proceeding to cDNA synthesis, the ligation reaction needs to be stopped and the ligation products purified from reagents and nonligated adapter molecules. This can be achieved using RNA purification columns (*see* **Note 6**). Here, RNAs are separated by denaturing PAGE and RNA molecules of 100–400 nucleotides are eluted from the gel in order to enrich for desired targets, that is, 5.8S rRNA precursors and 5′ ETS fragments in Arabidopsis (Fig. 3).

1. Cast a 6% urea–polyacrylamide gel (6% polyacrylamide, 7 M urea, 1× TBE), *see* example protocol in [13].
2. Prerun the gel in 1× TBE at 15 W for 15–20 min.
3. During the prerun, add 1 volume of 2× RNA loading buffer to your samples. Heat the samples at 65 °C for 3 min and chill on ice. Spin down briefly before loading.
4. Wash the residual urea from the wells using a syringe with a needle.

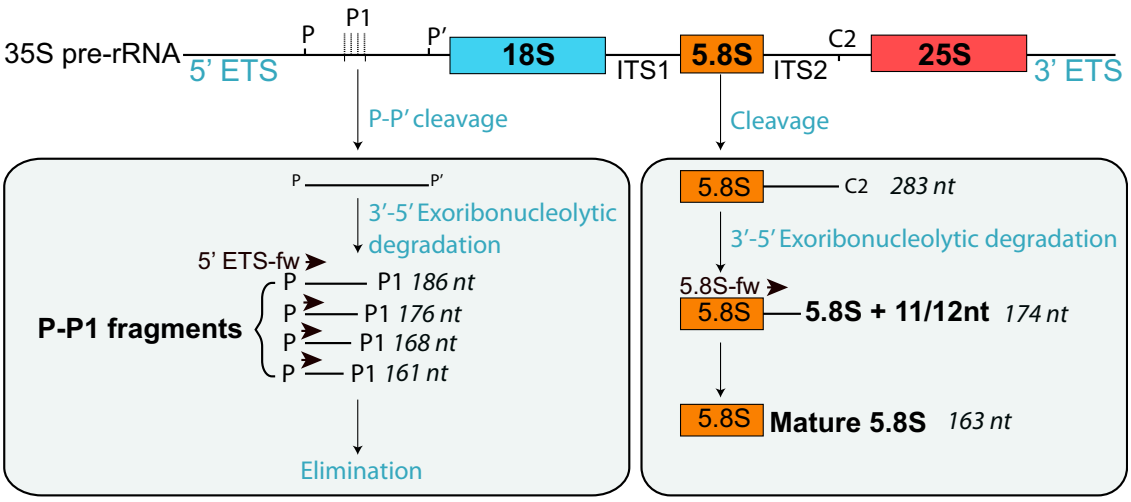


Fig. 3 Scheme of Arabidopsis 5′ ETS maturation by-products and 5.8S precursors analyzed by 3′ RACE-seq. 3′ RACE-seq procedure described in this chapter was used to map at high density the 3′ extremities of P-P1 fragments and of 5.8S precursors. P-P1 fragments result from endonucleolytic cleavage of the 5′ ETS at P and P′ sites followed by 3′-5′ exonucleolytic degradation of the P-P′ fragment by the RNA exosome. 5.8S precursors result from endonucleolytic cleavage at the 5′ end of 5.8S and at the C2 site. This 5.8S-C2 fragment is then further processed by 3′-5′ exoribonucleolytic activity, including by the RNA exosome. The shortest 5.8S rRNA precursor is extended by 11/12 nt in Arabidopsis. The 5′ ETS-fw primer was used to analyze 3′ extremities of P-P1 fragments. The 5.8S-fw primer was used to map 3′ extremities of 5.8S rRNA precursors

5. Load the samples on the gel and run at 15 W until the bromophenol blue tracking dye reaches three quarters of the gel.
6. Stain the gel in an ethidium bromide solution (0.5 μg/ml of EtBr in 1× TBE solution) for approximately 5 min and visualize using a UV transilluminator.
7. Excise RNA molecules of 100–400 nucleotides using a clean scalpel, put the gel slices in 1.5 ml tubes. Place tubes on dry ice to freeze the gel slices.
8. Fragment slices using a sterile 1 ml tip.
9. Elute RNA by adding 1:1 (v:w) volume of elution buffer to gel slices, for example 100 μl of elution buffer to 100 mg of gel, and incubate tubes overnight at 4 °C on a rotating wheel.
10. Centrifuge for 10 min at 16,000 × *g* to pellet the gel pieces.
11. Collect the supernatant and add 1 volume of acid phenol–-chloroform–isoamyl alcohol solution (25:24:1, v/v/v).
12. Vortex well and centrifuge for 15 min at 16,000 × *g*.
13. Transfer supernatant into new tube and precipitate RNA by adding 0.1 volume of 3 M sodium acetate pH 5.2, 0.5 μl of glycogen (20 mg/ml), and 2.5 volumes of absolute ethanol (*see* **Note 7**).
14. Mix by tube inversion.

15. Incubate for at least 1 h at −80 °C.
16. Centrifuge for 30 min at 16,000 × *g* (4 °C).
17. Discard the supernatant.
18. Wash pellet with 75% ethanol (500 μl) to remove residual salt. Centrifuge for 5 min at 16,000 × *g*.
19. Discard the supernatant thoroughly, dry the pellet and dissolve it in 20 μl of nuclease-free water.
20. Determine the RNA concentration and purity of your samples using UV spectrophotometer (*see* **Note 2**).

3.4 cDNA Synthesis

cDNA synthesis is initiated using a primer complementary to the last 17 nucleotides of the adapter sequence ligated at the 3′ end (3′-RT; Table 1, *see* **Note 8**). All steps of the cDNA synthesis are performed in a PCR thermal cycler using 0.2 ml strip PCR tubes.

1. Take 500 ng of adapter-ligated and size-selected RNAs and add 50 pmol of 3′-RT primer and 1 μl of 10 mM dNTP mix.
2. Add nuclease-free water to a final volume of 13 μl.
3. Denature for 5 min at 65 °C.
4. Chill samples on ice for at least 2 min.
5. Add 7 μl of RT Mix comprising 4 μl of 5× SuperScript™ IV buffer, 1 μl of 0.1 M DTT, 1 μl of RNaseOUT™ (40 U/μl), and 1 μl of SuperScript™ IV (200 U/μl).
6. Incubate for 10 min at 50 °C (*see* **Note 9**).
7. Inactivate the reaction by incubating for 10 min at 80 °C.

3.5 PCR Amplification and Quality Assessment by Electrophoresis

In order to prepare the libraries for MiSeq sequencing, the cDNA molecules of interest are amplified by PCR using forward PCR primers that comprise the Illumina P5 sequence and 21 nucleotides of the sequence of interest (here the 5′ ETS downstream of the P processing site or 3′ extended 5.8 S rRNA precursors) and a TruSeq RNA PCR index primer (RPI) complementary to the 3′ end of the 3′ adapter sequence that comprise the Illumina P7 sequence (*see* **Note 10,** Table 1). PCR reactions are performed in a PCR thermal cycler in 0.2 ml strip PCR tubes.

1. Set up PCR reaction by mixing the following components for each reaction:
   (a) 2.5 μl of 10× DreamTaq buffer.
   (b) 0.5 μl of 10 mM dNTP.
   (c) 0.5 μl of 10 μM forward primer (Table 1).
   (d) 0.5 μl of 10 μM reverse primer (Table 1).
   (e) 0.25 μl DreamTaq Polymerase (5 U/μl).
   (f) Add nuclease-free water to a final volume of 25 μl.

2. Distribute PCR mix in strip PCR tubes and add 1 μl of template cDNA.
3. Spin down and place reaction in thermal cycler.
4. Run PCR reaction with the following settings:
   Initial denaturation step: 2 min at 94 °C.
   30 cycles composed of:
   (a) Denaturation step: 30 s at 94 °C.
   (b) Hybridization step: 30 s at 55 °C.
   (c) Elongation step: 30 s at 72 °C.
   Final extension: 2 min at 72 °C.
5. Visualize PCR products by loading a 3 μl aliquot with DNA loading dye on a 2% agarose gel (0.5× TBE) (*see* **Note 11**).

3.6 PCR Product Purification

After amplification, the PCR products are purified using AMPure XP beads (Agencourt). This system uses magnetic beads that can bind PCR amplicons of at least 100 bp, thereby purifying amplicons from nucleotides, primer dimers, salts or other reagents. The protocol has been adapted from the manufacturer's protocol. AMPure XP beads purification is performed in 1.5 ml microtubes that are compatible with the magnetic stand.

1. Warm the AMPure XP beads to room temperature for at least 10 min and shake the Agencourt AMPure XP bottle before pipetting to resuspend magnetic particles.
2. Transfer each PCR reaction to individual 1.5 ml tubes.
3. For each tube, add 1 volume of beads to 1 volume of PCR reaction and mix well by pipetting or gentle vortexing (*see* **Note 12**).
4. Incubate the mixture for 5 min at room temperature.
5. Transfer tubes to a magnetic stand.
   For the next steps (**6–11**), tubes are kept on the magnetic stand.
6. Let sit for about 5 min or until solution appears clear.
7. Carefully discard the supernatant without disturbing the beads.
8. Keep the tubes on the magnetic stand and wash beads carefully with 200 μl of 80% ethanol.
9. Incubate for 1 min and carefully remove the ethanol.
10. Repeat the washing step.
11. Air-dry beads for a maximum of 3 min (*see* **Note 13**).
12. Add 100 μl of nuclease-free water to beads, remove tubes from the magnetic stand and mix gently by pipetting.
13. Incubate for 5 min at room temperature.

14. Put tubes back on stand and let sit for about 5 min or until solution is clear.
15. Transfer the eluate to a new, clean tube by paying attention not to take beads (*see* **Note 14**).
16. Perform a second elution with 100 μl of nuclease-free water.
17. Precipitate purified amplicons with 5 volumes of absolute ethanol, 0.1 volume of sodium acetate 3 M, pH 5.2, and 0.5 μl of glycogen (20 mg/ml).
18. Mix by tube inversion.
19. Incubate for at least 1 h at −80 °C.
20. Centrifuge for 30 min at 16,000 × *g* (4 °C).
21. Discard the supernatant.
22. Wash pellet with 75% ethanol (500 μl) to remove residual salt. Centrifuge at 16,000 × *g* during 5 min.
23. Discard the supernatant thoroughly, dry the pellet and dissolve it in 11 μl of nuclease-free water.
24. Measure the DNA quantity and purity of your samples using UV spectrophotometer (*see* **Note 2**).

3.7 Qubit and Bioanalyzer Analysis of Purified Amplicons

Before Illumina sequencing, we quantify amplicon libraries by a fluorometric method, (i.e., Qubit fluorometric quantitation system). We also control the size and the quality of our amplicon profiles after AMPure XP beads purification using Agilent 2100 Bioanalyzer, notably to check the complete removal of primer dimers (*see* **Note 15**).

1. Determine Qubit concentration of each library (Qubit fluorometric quantitation system; see the manufacturer's protocol [14]) (*see* **Note 16**).
2. Check library profiles using Agilent 2100 Bioanalyzer according to the manufacturer's instructions [15]. Choose the reagent kit according to the range of concentration of your samples (*see* **Note 17**).
3. Use Qubit concentration and Bioanalyzer size estimation to calculate the molarity of each sample library (*see* **Note 18**).

3.8 Preparing Libraries for Sequencing on MiSeq

Here, we prepare libraries for sequencing with v3 MiSeq chemistry (*see* **Note 19**). You need to prepare a final amplicon library according to the depth wanted for each sample and to denature the library. We usually allocate from 0.5% to 3% of the flow cell per condition/genotype and target. Ten to twenty percentage of PhiX control v3 library are also included to compensate for the low-diversity of the samples (*see* **Notes 20** and **21**).

1. Prepare at least 5 μl of a final 4 nM amplicon library with all individual sample libraries to be sequenced.
2. Combine 5 μl of 4 nM final amplicon library and 5 μl of a fresh 0.2 N NaOH dilution.
3. Vortex briefly and centrifuge for 1 min at 280 × *g*.
4. Incubate for 5 min at room temperature.
5. Stop denaturation reaction by adding 990 μl prechilled HT1 to the 10 μl denatured library. This results in a 20 pM denatured library.
6. Dilute the 20 pM amplicon library to 15 pM by adding 150 μl prechilled HT1 to 450 μl of the 20 pM denatured amplicon library.
7. Mix by inversion and quickly centrifuge the resulting 15 pM amplicon.
8. Final 15 pM library should be kept on ice.
9. Repeat **steps 2–8** with the 4 nM PhiX control library as described above for amplicon library to get a 15 pM PhiX library.
10. For MiSeq sequencing with 15% of PhiX, combine 90 μl of 15 pM denature PhiX control library and 510 μl of 15 pM denature amplicon library (*see* **Notes 20** and **21**). Keep tubes on ice until loading on the reagent cartridge.

3.9 MiSeq Run and Analysis

Here we use a MiSeq Reagent kit v3 150 cycles. The final library is paired-end sequenced with a 76 × 76 bp cycle setting. Cycle setting may be adjusted according to the type of analyzed RNA target (*see* **Note 22**). Read 1 and read 2 will be used during bioinformatics analysis for RNA target identification and 3′ end analysis, respectively.

1. Thaw Reagent Cartridge and mix according to Illumina manufacturer's instructions.
2. In the Illumina Experiment Manager (IEM) software, create a custom library prep kit as indicated in the IEM software guide [16]. Take the "TruSeq Small RNA.txt" template as model (model with RPI barcodes) and change the setting section to allow for paired-end sequencing (*see* **Note 23**).
3. Use the IEM software to create sample sheet. Select "MiSeq," "other," and "fastq only" in the instrument, category, and application sections, respectively. For workflow parameters, select as option the new custom library prep kit and set cycle setting as 76 × 76 bp (*see* **Note 22**).
4. Fill the sample sheet wizard as indicated in the IEM software guide.

5. Start MiSeq Control Software and follow steps indicated by the software to start MiSeq sequencing.
6. When asked by the MiSeq Control Software, load your sample (combination of PhiX and sample libraries as prepared in Subheading 3.8) in the reservoir labeled "Load Sample" of the reagent cartridge.
7. Use the Sequence Analysis Viewer (SAV) to monitor sequencing during run.
8. When sequencing is finished, check quality control metrics using SAV and control quality of read 1 and read 2 fastq files using the quality control tool FASTQC [17].
9. Finally process fastq files using the pipeline available in [6]. An overview of the bioinformatics workflow is shown in Fig. 4.

4 Notes

1. NanoDrop spectrophotometer measures the sample absorbance across a wide spectrum that spans UV and visible light. Nucleotides, RNA and DNA, have an absorbance peak at 260 nm. By contrast, proteins have a peak of absorbance at 280 nm, while other usual RNA contaminants, such as carbohydrates, EDTA and phenol have an absorbance maximum at 230 nm or less. 260/280 and 260/230 ratios can thus be used to assess RNA purity. Values around 2.0 are usually considered as acceptable for 260/280 and 260/230 ratios [18].
2. During phenol extraction of nucleic acid molecules, the partition between aqueous and organic phase is pH-dependent. At acidic pH conditions, RNA molecules are highly soluble and retained in the aqueous phase, while DNA molecules are retained in the organic phase and interphase. Acid phenol is thereby used for the isolation or RNA molecules, whereas DNA isolation is best performed with buffer-saturated phenol equilibrated to pH >7.4.
3. The assessment of RNA integrity by electrophoresis on agarose gel and ethidium bromide staining is a basic and cheap technique that gives a first indication of the quality of your RNA preparation. Sharp bands corresponding to rRNAs should be visible on the gel. Partially degraded RNA will appear as smeared bands.
4. T4 RNA ligase 1 catalyzes the ligation of 5′-phosphoryl terminated DNA or RNA to 3′-hydroxyl terminated single strand DNA or RNA. Here, we use a preadenylated adapter (3′-Adap) containing a 5′,5′-adenyl pyrophosphoryl moiety, which can be directly ligated to the RNA without the addition of ATP during the ligation reaction (Fig. 2). This strategy prevents that

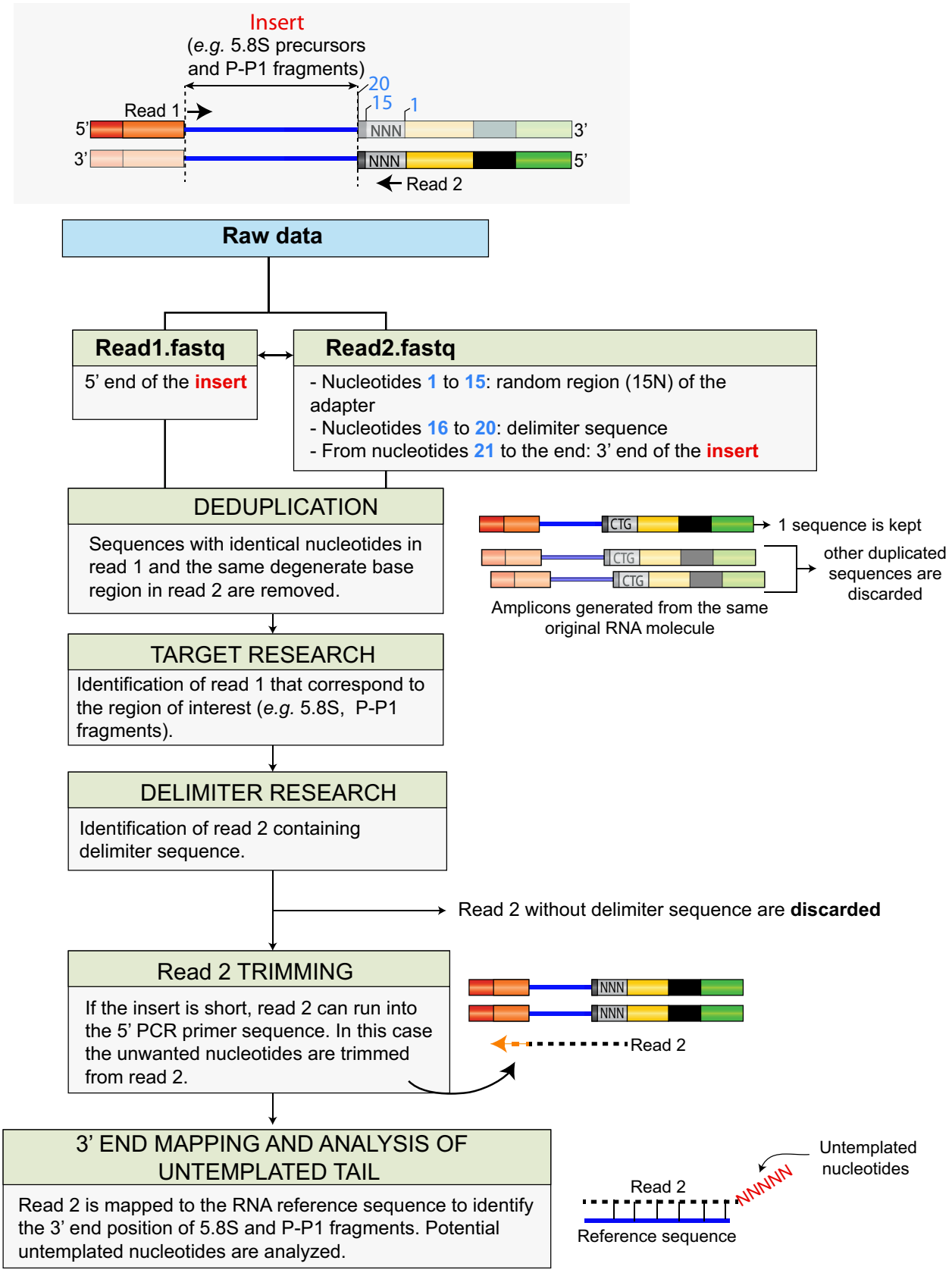


Fig. 4 Schematic overview of the bioinformatic pipeline for 3′ RACE-seq analysis of 5.8S and P-P1 fragments. Color code as in Fig. 1. Scripts are available in [6]

endogenous RNA with 5′ phosphate extremities compete with the adapter for 3′ ligation and ensures that only the preadenylated adapter is ligated to the 3′ hydroxylated end of endogenous transcripts.

5. We usually use a water bath for the incubation at 37 °C during the ligation step. Water bath provides a better contact surface area for heat transfer as compared to dry bath, allowing for better reproducibility.
6. Separation and purification of ligated RNA on PAGE is not required for all applications. Here, the molecules of interest are small RNA fragments (<400 nt). Consequently, we purified ligated RNA on PAGE to remove larger RNAs and to enrich for desired targets before proceeding to cDNA synthesis. Alternatively, to purify the ligated RNAs from reagents and nonligated adapter, you can perform fast purification of nucleic acids using RNA purification kits.
7. GlycoBlue™ (15 mg/ml, Thermo Fisher Scientific) can be used instead of glycogen to increase pellet visibility and is recommended if you are working with low amounts of RNA.
8. During the setting up of 3′ RACE-seq procedure, we detected in our first amplicon libraries a strong proportion of sequences that did not contain the delimiter sequence (Figs. 2 and 4) and thus did not correspond to real RNA 3′ ends. These artifacts were likely caused by nonspecific binding of the primer used in the reverse transcription reaction. Consequently, to specifically amplify cDNA 3′ ends, we designed a shortened RT oligonucleotide (3′-RT, Table 1) that is complementary to the last 17 nt of the 3′ ligated adapter and lacks five nucleotides contained by both 3′ ligated adapter and reverse PCR primers. Thus, cDNA resulting from nonspecific reverse transcription cannot be amplified.
9. The reverse transcriptase SuperScript™ IV (Invitrogen™) is very effective and robust, allowing for efficient cDNA synthesis in only 10 min. If you use other reverse transcriptases, adapt the incubation time according to the manufacturer's protocol.
10. RPI primers used for cDNA amplification contain index sequences, also called barcodes, which allow for the sequencing of a large number of samples in a single run (i.e., multiplexing sequencing). Individual index sequences can be assigned to each genotype and/or conditions and then be used to distinguish and sort samples during data analysis. The use of individual index sequences for each different target, here 5.8S and P-P1 fragments, is not necessary as target sequences can be distinguished during bioinformatic analysis.

11. Separation of smaller DNA molecules and fragments is improved with TBE buffer, whereas TAE buffer is well suited for larger fragments [19].
12. The concentration of PEG and NaCl in the AMPure XP bead solution is crucial for size selection of the DNA fragments that are purified. The size of purified fragments is determined by the ratio of beads/sample: the lower is the chosen ratio, the larger are the eluted fragments. To adapt the ratio according to the size of your amplicon library, refer to [20].
13. Take care to not over dry beads. This would significantly decrease the elution efficiency.
14. The carryover of magnetic beads results in an additional peak in Bioanalyzer electropherograms and could lead to inaccurate estimation of the library size. Trace amounts of beads may also affect the performance of Illumina sequencing.
15. As we usually analyze a large number of samples, replicates, and genotypes, we are used to pool PCR products for each analyzed target prior to Qubit and Bioanalyzer analysis. We first adjust each sample to the same concentration, verify the dilution by checking the profiles on a 2% agarose gel (at least 75 ng DNA for ethidium bromide staining), and then pool the samples.
16. The Invitrogen™ Qubit is a fluorometric quantitation system that allows for a more specific and sensitive quantification of RNA and DNA than using the NanoDrop spectrophotometer. Select the Qubit assay kit according to your sample. The dsDNA HS Assay kit is well suited for the quantification of the prepared libraries and is designed to measure sample with initial concentration from 10 pg/μl to 100 ng/μl.
17. DNA kits for Bioanalyzer analysis may be chosen according to the size of the amplification product and the range of concentration of the sample. DNA 12000, DNA 7500, and DNA 1000 kits allow for the analysis of dsDNA fragments from 100 to 12,000 bp, 100 to 7500 bp and 25 to 1000, respectively. All of them offer a 0.5–50 ng/μl sensitivity [21]. For samples with a low concentration, DNA High Sensitivity kit allows for the analysis of 50–7000 bp fragments and offers a 5–500 pg/μl sensitivity [22].
18. The molarity in nM of amplicon library is calculated using the following formula ([ng/μl conc.] $\times$ 10^6)/([bp length] $\times$ 607.4 + 157.9). See Illumina manufacturer's protocol for more information.
19. Here, we sequenced amplicon libraries using MiSeq Reagent Kit v3 (150 cycles). Reagent kits with v2 chemistry are also

available for MiSeq. v2 chemistry enables smaller depth compared to v3: for single-end run, up to 15 and 25 millions of output reads are obtained for v2 and v3, respectively. Final concentration of denatured library need to be adjusted according to the selected MiSeq chemistry. v2 and v3 chemistry support a maximum of 10 pM and 20 pM concentration, respectively.

20. To compensate for the low-diversity of amplicon libraries, PhiX Control v3 Library should be sequenced alongside samples. Illumina recommends spiking-in a minimum of 5% of PhiX control library. This percentage may be adjusted according to experiments and may be increased if the sample library clusters more efficiently than the PhiX library. When analyzing rRNA maturation intermediates, we usually spike-in from 10% to 20% of PhiX control v3 library.
21. The procedure detailed here can be extended to other types of RNA targets, such as poly(A) tailed transcripts. Illumina sequencing of poly(A) stretches requires spiking-in a particular high amount of PhiX control library, at least 20% of the flow cell. Indeed, sequencing a highly diversified library alongside samples is necessary to counteract the strong negative impact of the base composition bias toward A of poly(A) tails on sequencing quality.
22. Amplicon libraries are paired-end sequenced: read 1 is use to identify target, whereas read 2 is used to determine RNA 3′ end position and to identify added untemplate nucleotides (Fig. 4). Reagents provided in MiSeq 150-Cycle kit are sufficient to perform 152 sequencing cycles. The number of cycles for read 2 sequencing can be adjusted according to the type of analyzed RNA targets and the expected length of 3′ untemplated tails. A cycle setting of 76 × 76 nt enables the analysis of untemplated tails up to 56 nt (76 nt – 15 nt of the random sequence – 5 nt of the delimiter sequence). For longer 3′ tail, such as mRNA poly(A) tail, cycle setting can be desynchronized and cycle number for read 1 sequencing can be reduced in favor of read 2 sequencing. For example in [6], we sequenced in parallel amplicon libraries for rRNA intermediates and poly(A) tailed mRNA and we thus set cycle setting as 41 × 111 nt, 41 nt being sufficient for transcript identification.
23. In this protocol, we use the TruSeq RNA PCR index primers, classically used for the preparation of TruSeq Smal RNA library. However, default options for TruSeq Small sequencing

in the IEM software does not allow for paired-end sequencing. Therefore, a custom template, based on the "TruSeq Small RNA.txt" template, needs to be created.

Acknowledgments

This work was supported by the Centre National de la Recherche Scientifique (CNRS, France) and research grants from the French National Research Agency as part of the "Investments for the Future" program in the frame of LABEX ANR-10-LABX-0036_NETRNA and ANR-15-CE12-0008-01 to D.G, and in the frame of the IdEx Unistra to H.Z.

References

1. Januszyk K, Lima CD (2014) The eukaryotic RNA exosome. Curr Opin Struct Biol 24:132–140
2. Schneider C, Tollervey D (2014) Looking into the barrel of the RNA exosome. Nat Struct Mol Biol 21:17–18
3. Zinder JC, Lima CD (2017) Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev 31:88–100
4. Kumakura N, Otsuki H, Tsuzuki M et al (2013) Arabidopsis AtRRP44A is the functional homolog of Rrp44/Dis3, an exosome component, is essential for viability and is required for RNA processing and degradation. PLoS One 8:e79219
5. Lange H, Holec S, Cognat V et al (2008) Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in Arabidopsis thaliana. Mol Cell Biol 28:3038–3044
6. Sikorska N, Zuber H, Gobert A et al (2017) RNA degradation by the plant RNA exosome involves both phosphorolytic and hydrolytic activities. Nat Commun 8:2162
7. Sement FM, Ferrier E, Zuber H et al (2013) Uridylation prevents 3′ trimming of oligoadenylated mRNAs. Nucleic Acids Res 41:7115–7127
8. Sement FM, Gagliardi D (2014) Detection of uridylated mRNAs. Methods Mol Biol 1125:43–51
9. Chang H, Lim J, Ha M, Kim VN (2014) TAIL-seq: genome-wide determination of poly(a) tail length and 3′ end modifications. Mol Cell 53:1044–1052
10. Lim J, Ha M, Chang H et al (2014) Uridylation by TUT4 and TUT7 marks mRNA for degradation. Cell 159:1365–1376
11. Zuber H, Scheer H, Ferrier E et al (2016) Uridylation and PABP cooperate to repair mRNA deadenylated ends in Arabidopsis. Cell Rep 14:2707–2717
12. Morgan M, Much C, DiGiacomo M et al (2017) MRNA 3′ uridylation and poly(A) tail length sculpt the mammalian maternal transcriptome. Nature 548:347–351
13. Summer H, Grämer R, Dröge P (2009) Denaturing urea polyacrylamide gel electrophoresis (urea PAGE). J Vis Exp:3–5
14. Thermo Fisher Scientific (2017) Qubit ™ 4 Fluorometer guide
15. Agilent (2001) Agilent 2100 bioanalyzer user guide
16. Illumina (2018) Illumina Experiment Manager User Guide. Accessed 19 Jan 2018
17. Babraham Bioinformatics (2018) FastQC a quality control tool for high throughput sequence Data. Accessed 19 Jan 2018
18. Desjardins P, Conklin D (2010) NanoDrop microvolume quantitation of nucleic acids. J Vis Exp pii:2565
19. Sanderson BA, Araki N, Lilley JL et al (2014) Modification of gel architecture and TBE/TAE buffer composition to minimize heating during agarose gel electrophoresis. Anal Biochem 454:44–52

20. Bronner IF, Quail MA, Turner DJ, Swerdlow H (2009) Improved protocols for illumina sequencing. Curr Protocol Human Genet 79:18.2.1–18.2.42
21. Agilent (2018) DNA analysis kits & reagents - details & specifications. Accessed 19 Jan 2018
22. Agilent (2018) High sensitivity DNA analysis kits - details & specifications. Accessed 19 Jan 2018
23. Illumina (2018) Illumina adapter sequences document. Accessed 26 Jan 2018

Core Subunits of the CCR4/NOT complex.

H. sapiens	*S. cerevisiae*	*D. melanogaster*	*A. thaliana*
CNOT1	NOT1	Not1	NOT1
CNOT2	NOT2	Regena (Rga)	NOT2
CNOT3	NOT5	Not3	NOT3
CNOT4	NOT4	Not4	NOT4
CNOT7/8 (CAF)	CAF1/POP2	Pop2	CAF1
CNOT6/6L(CCR4)	CCR4	twin/Ccr4	CCR4
CNOT9	CAF40	Rcd1/Caf40	NOT9
CNOT10		Not10	NOT10
CNOT11		Not11	NOT11
	CAF130		
	NOT3		

Table S1: Name of the core subunits of the CCR4/NOT complex in human, yeast, flies and Arabidopsis.
The subunits of the CCR4/NOT complex are conserved among eukaryotes.
Yeast has no orthologue to NOT10 and NOT11, but instead an additional CAF130 subunit. The closest orthologue of NOT3 in yeast is NOT5.
Adapted from Collart, 2016

Number of reads (*GFP* mRNAs)

Sample	total	uri.
ctrl rep1	64249	2291
ctrl rep2	43314	1025
ctrl rep3	51868	1432
ctrl rep4	77824	2886
ctrl rep5	58488	2083
ctrl rep6	110253	3320
URT1 rep1	75600	8830
URT1 rep2	49474	6786
URT1 rep3	58318	7875
URT1 rep4	112835	17278
URT1 rep5	100338	15681
URT1 rep6	62754	8393
URT1$^{D491/3A}$ rep1	89177	5785
URT1$^{D491/3A}$ rep2	45981	2583
URT1$^{D491/3A}$ rep3	59463	3409
URT1$^{D491/3A}$ rep4	8791	7890
URT1$^{D491/3A}$ rep5	70675	5841
URT1$^{D491/3A}$ rep6	131498	10876

Figure S1: Number of reads for all and only uridylated polyA tails for *GFP* mRNAs.

A

Examples of A-rich tails (*GFP* mRNAs)

AAAAAAAAAAAAAAAAAAAAAAAAATAA
AAAAAAAAAAAAAAAAAAAAAAAAAGAA
AAAAAAAAAAAAAAAAAAAAAAAAATAAA
AAAAAAAAAAAATG
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA
AAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAA
AAAAAAAAAAAAAAAAAAAAGAAT
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA
AAAAAAAAAAAAAAAACAAAAAAATAGAAAAAAAAAAGAA
AAAAAAAAAAAATAAAAAAAAAA
AAAAAAAAAAAAAAAAACATAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAATT
AAAAAAAAAAGAAACACACTACAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA
AAAAAAAAAAAAAAGAAAAAAAAATA
AAAAAAAAAAAAAATA

B

Number of reads (*GFP* mRNAs)

Sample	total	uri	A-rich
ctrl rep1	64249	2291	4638
ctrl rep2	43314	1025	3119
ctrl rep3	51868	1432	3775
URT1 rep1	75600	8830	10161
URT1 rep2	49474	6786	6093
URT1 rep3	58318	7875	7134
URT1$^{D491/3A}$ rep1	89177	5785	17794
URT1$^{D491/3A}$ rep2	45981	2583	10514
URT1$^{D491/3A}$ rep3	59463	3409	13373
M1M2 rep1	87258	11820	7019
M1M2 rep2	48776	7648	5013
ΔIDR rep1	60884	5497	5878
ΔIDR rep3	63992	3078	4550
ΔIDR$^{D491/3A}$ rep1	61302	2635	4301
ΔIDR$^{D491/3A}$ rep3	38248	1060	2966

Figure S2: A-rich tails with heteropolymeric extensions for *GFP* mRNAs.
A) Examples of A-rich tails.
B) Number of reads for all, uridylated or A-rich tails.

A

Number of reads for *PR2* mRNAs

Sample	total	uri	A-rich
ctrl rep1	48998	898	4440
ctrl rep2	53753	845	3655
ctrl rep3	57237	1279	4531
URT1 rep1	101120	5437	10046
URT1 rep2	70233	2497	5939
URT1 rep3	100540	4418	8209
URT1$^{D491/3A}$ rep1	106613	3137	13810
URT1$^{D491/3A}$ rep2	46539	1246	7064
URT1$^{D491/3A}$ rep3	81613	2312	12647
M1M2 rep1	88413	4515	7903
M1M2 rep2	60029	2501	4086
ΔIDR rep1	84306	3188	6486
ΔIDR rep3	57330	2235	4574
ΔIDR$^{D491/3A}$ rep1	65469	1666	5205
ΔIDR$^{D491/3A}$ rep3	48730	906	3822

B

Number of uridines added to the *PR2* mRNAs

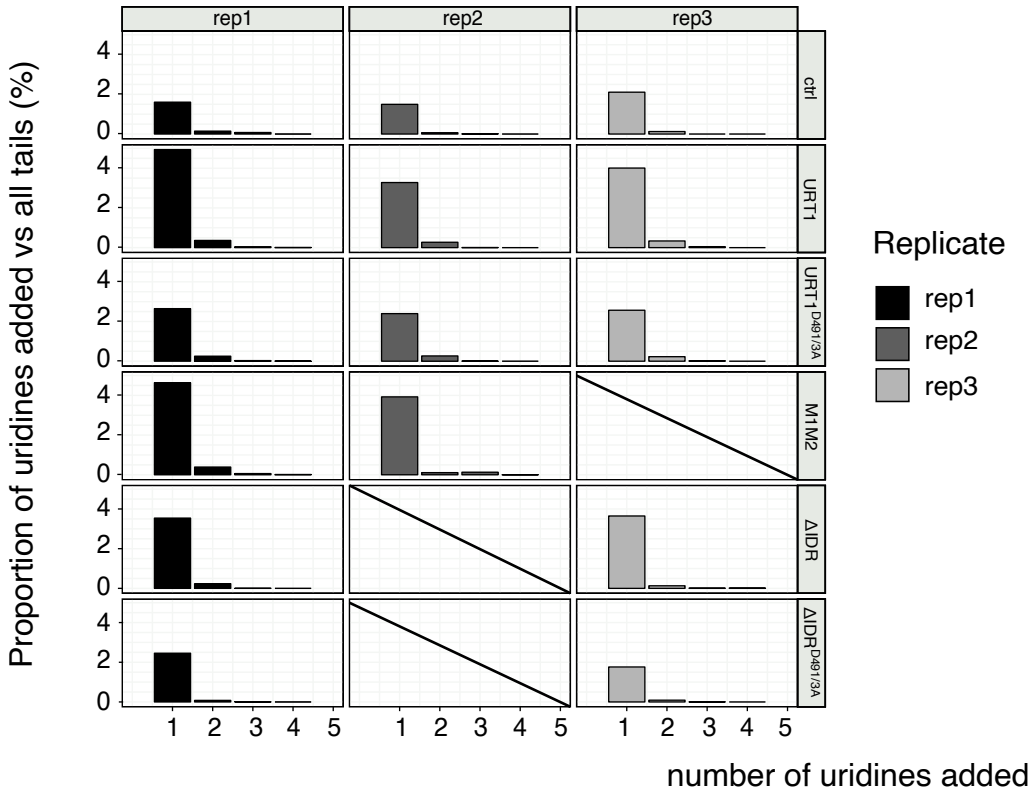


Figure S3: Number of reads and number of uridines added to *NbPR2* mRNAs.
A) Number of reads for each sample and replicate of *NbPR2* mRNAs.
B) Number of uridines added to the polyA tails of *NbPR2* mRNAs.

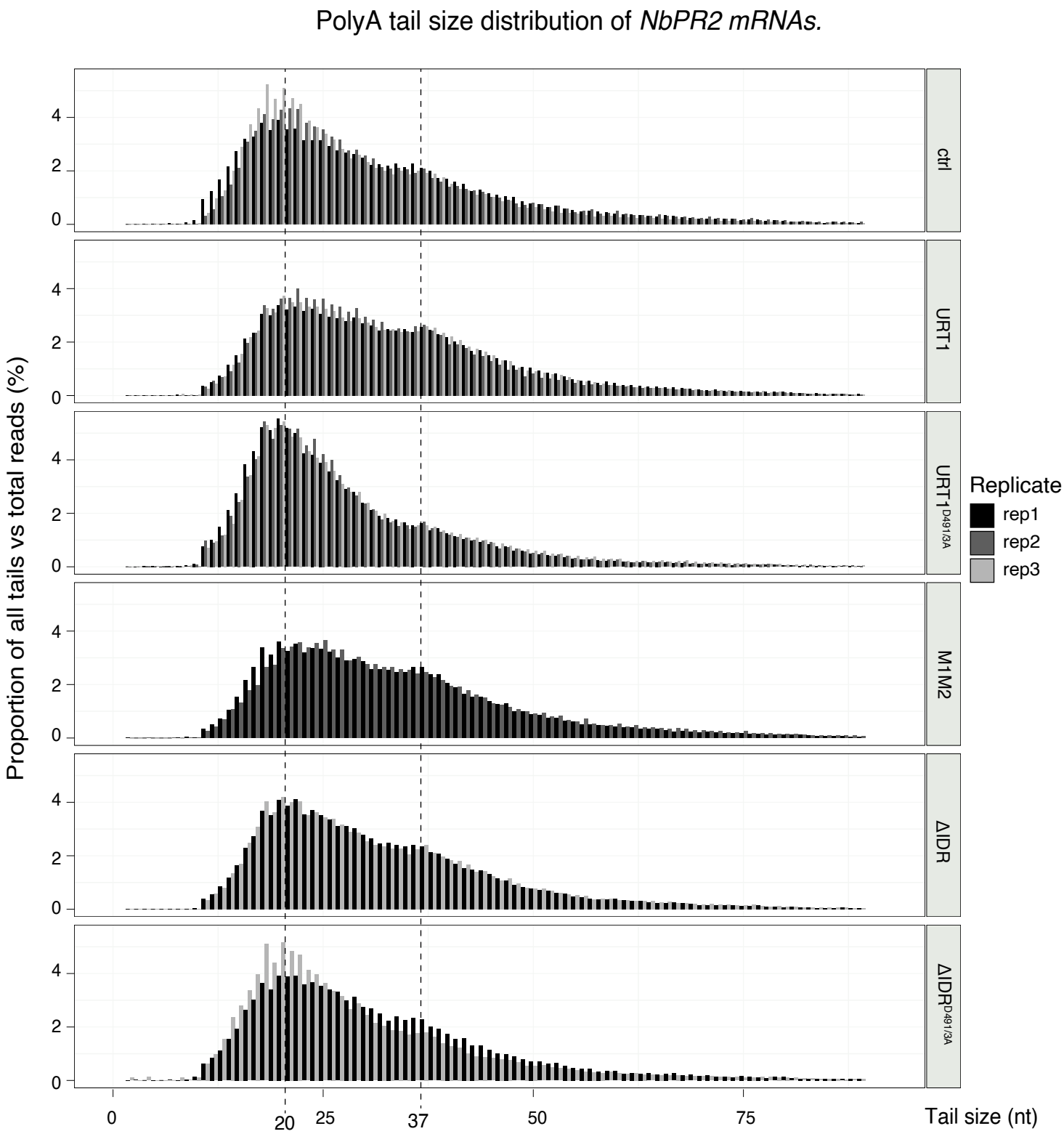


Figure S4: Size distribution of all tails of *NbPR2* mRNAs.
NbPR2 mRNAs show a rather short polyA tail size distribution with peaks at 20 and 37 nucleotides.

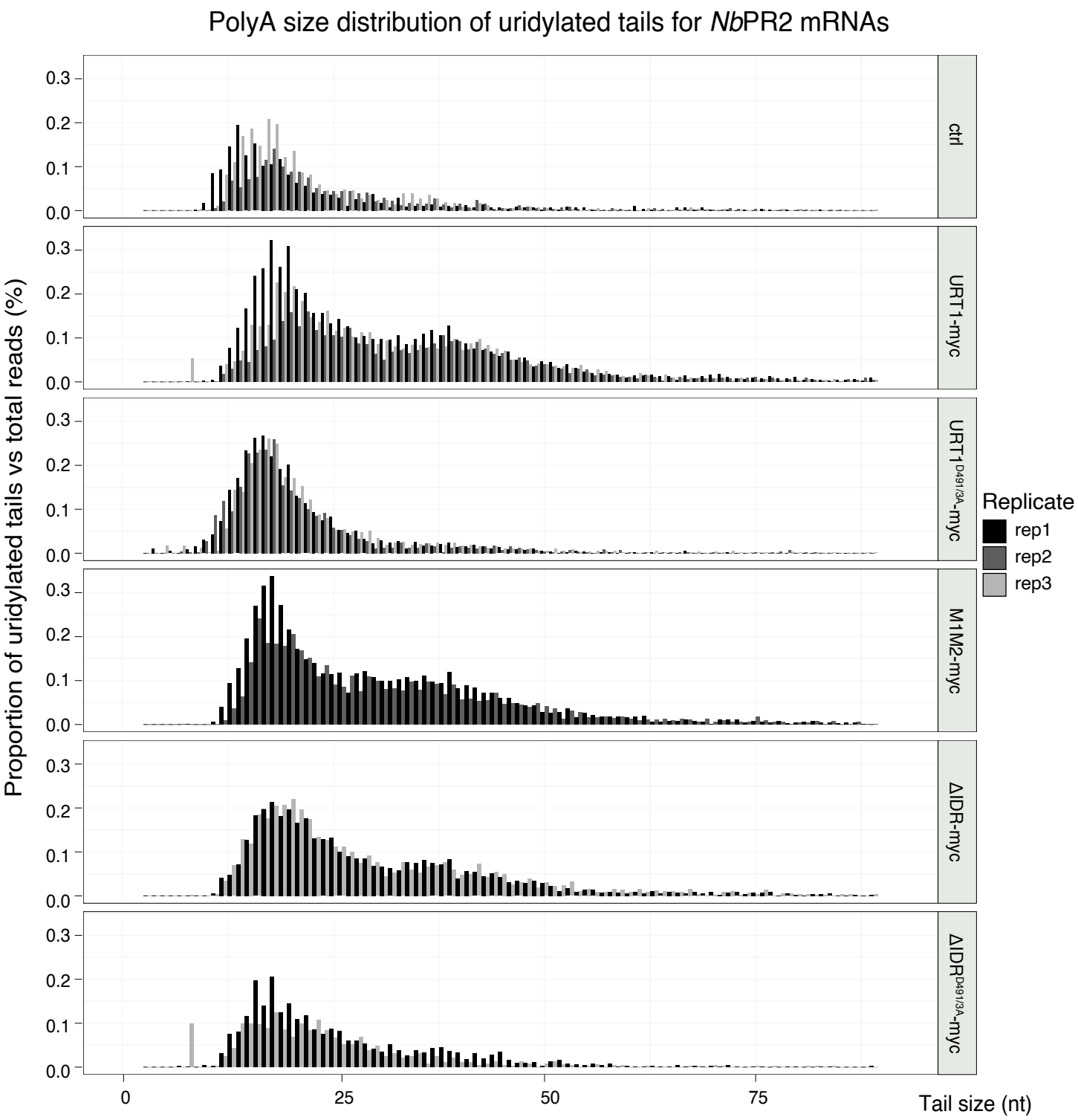


Figure S5: Size distribution of uridylated tails of *NbPR2* mRNAs.
Expression of active URT1-myc and M1M2-myc leads to the accumulation of longer uridylated polyA tails (circa 38 nt). The three biological replicates show slight differences, the line-chart that shows the median polyA tail length of the replicates (Figure 26B) was used to count the peak sizes.

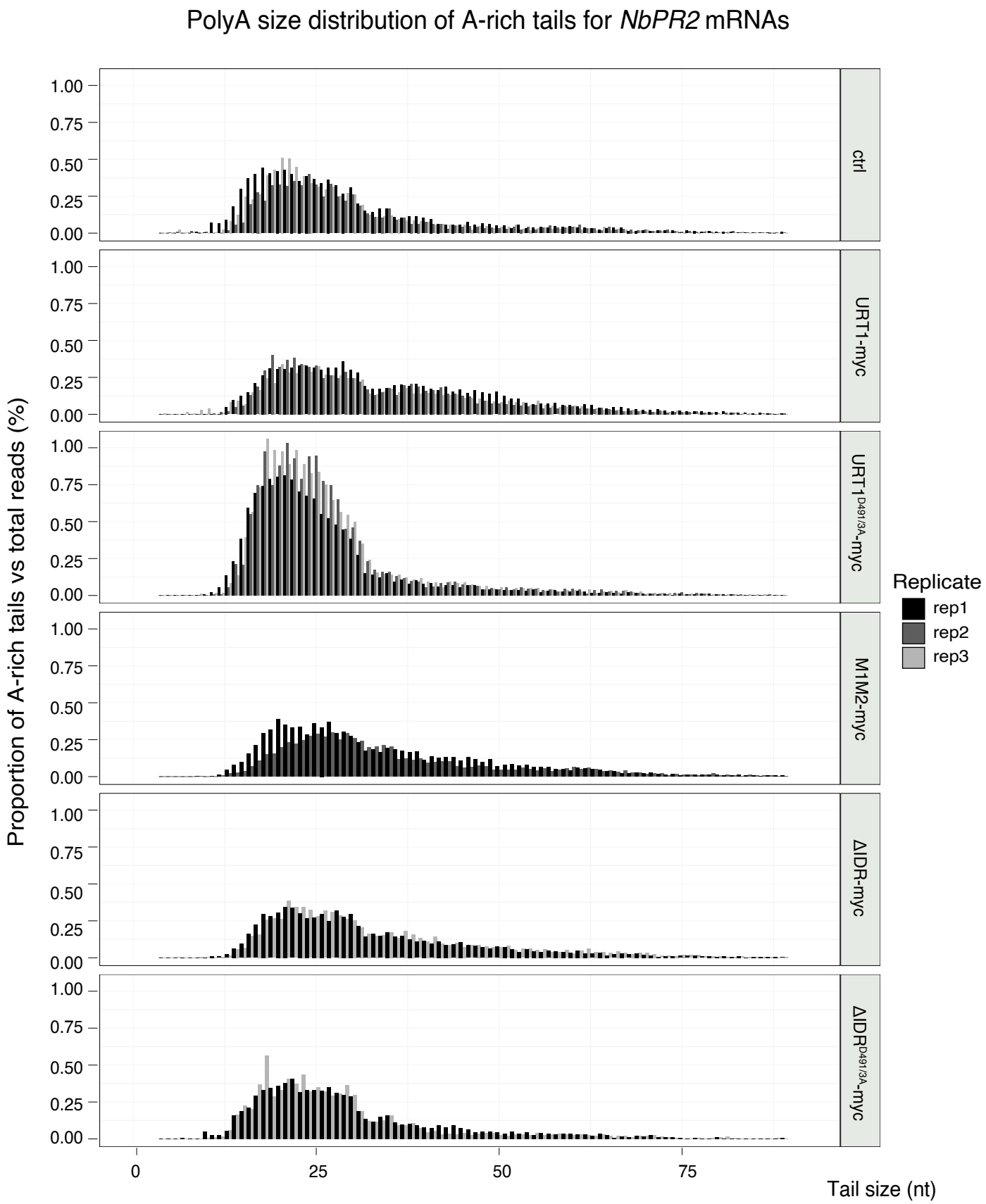


Figure S6: Size distribution of A-rich tails of *NbPR2* mRNAs.
The three biological replicates show slight differences, the line-chart that shows the median polyA tail length of the replicates (Figure 27B) was used to count the peak sizes.

Number of reads Run 3 (*GFP* mRNAs).

Sample	total	uri	A-rich
ctrl rep1	99611	3395	6853
ctrl rep2	196789	7162	10910
ctrl rep3	196065	7811	10542
URT1 rep1	81155	7590	8581
URT1 rep2	140096	18600	13460
URT1 rep3	125383	15655	11589
URT1$^{D491/3A}$ rep1	158377	10951	31778
URT1$^{D491/3A}$ rep2	312914	27622	87351
URT1$^{D491/3A}$ rep3	323740	27887	75000
URT1$^{D491/3A}$M1 rep1	127136	5801	9912
URT1$^{D491/3A}$M1 rep2	216762	12728	15840
URT1$^{D491/3A}$M1 rep3	203086	11858	13795
URT1$^{D491/3A}$M2 rep1	115393	10552	26121
URT1$^{D491/3A}$M2 rep2	237760	18628	64051
URT1$^{D491/3A}$M2 rep3	318515	28317	73526
URT1$^{D491/3A}$M1M2 rep1	124585	5307	9938
URT1$^{D491/3A}$M1M2 rep2	234210	10970	15023
URT1$^{D491/3A}$M1M2 rep3	262098	13304	16213

Figure S7: Number of reads for the *GFP* mRNAs in the 3'RACE-seq run3 dataset.
Number of reads for all, uridylated or A-rich tails.

Conditions tested with purified deadenylases.

Constructs	0mM Mg^{2+}	2mM Mg^{2+}	5mM Mg^{2+}	15% gly.	7% gly.	100mM KCl	50mM KCl	TEV cleav.
hisGST-CAF1a	-	-	-	-	-	-	-	/
hisGST-CAF1b	-	+	++	+	++	+	++	/
hisMBP-CCR4a	-	-	-	-	-	-	-	-
hisMBP-CCR4b	-	-	-	-	-	-	-	-
hisMBP-EEPa	-	-	-	/	-	/	-	-
hisMBP-EEPb	-	-	-	/	-	/	-	-
hisGST-CAF1h	/	/	-	/	-	/	-	-
hisGST-CAF1i	/	/	-	/	-	/	-	-
hisGST-CAF1j	/	/	-	/	-	/	-	-
hisGST-CAF1k	/	/	-	/	-	/	-	-

Figure S8: I*n vitro* buffer conditions tested with the purified deadenylases.
All the reaction buffers had 20mM MOPS, pH 7.2, 0.1% of Tween 20, and the indicated concentrations of $MgCl_2$, KCl and glycerol. For some of the purified constructs, additional activity tests were done with the N-terminal tag cleaved using TEV protease (TEV cleav.). hisGST-CAF1b was mostly active in the presence of 5mM of Mg^{2+}, 7% of glycerol and 50mM of KCl.

- “no activity”
+ “activity”
/ “not tested”

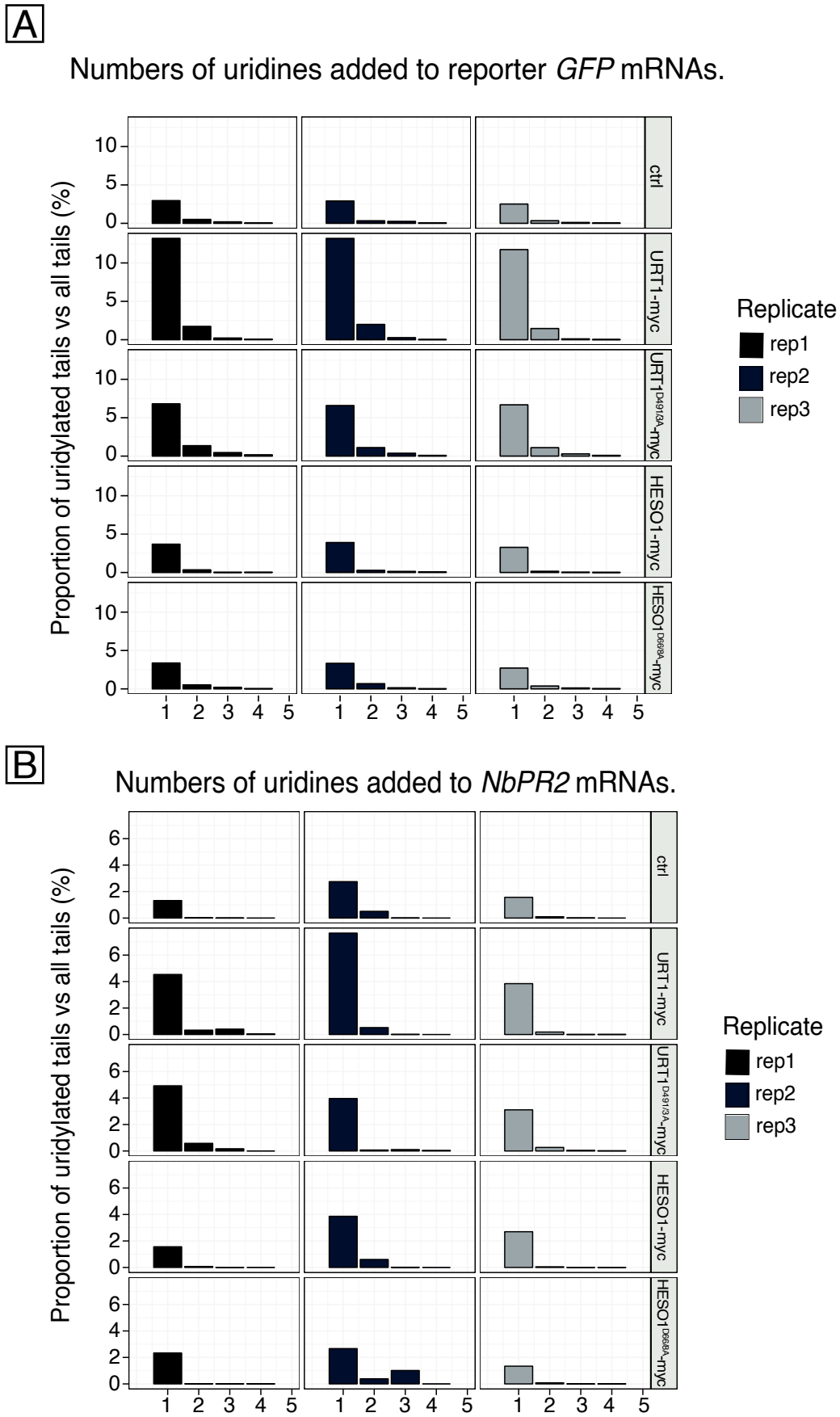


Figure S9: Number of uridines added to the reporter *GFP* and *NbPR2* mRNAs.
A) Number of uridines added to uridylated polyA tails for *GFP* mRNAs.
B) Number of uridines added to uridylated polyA tails for *NbPR2* mRNAs.

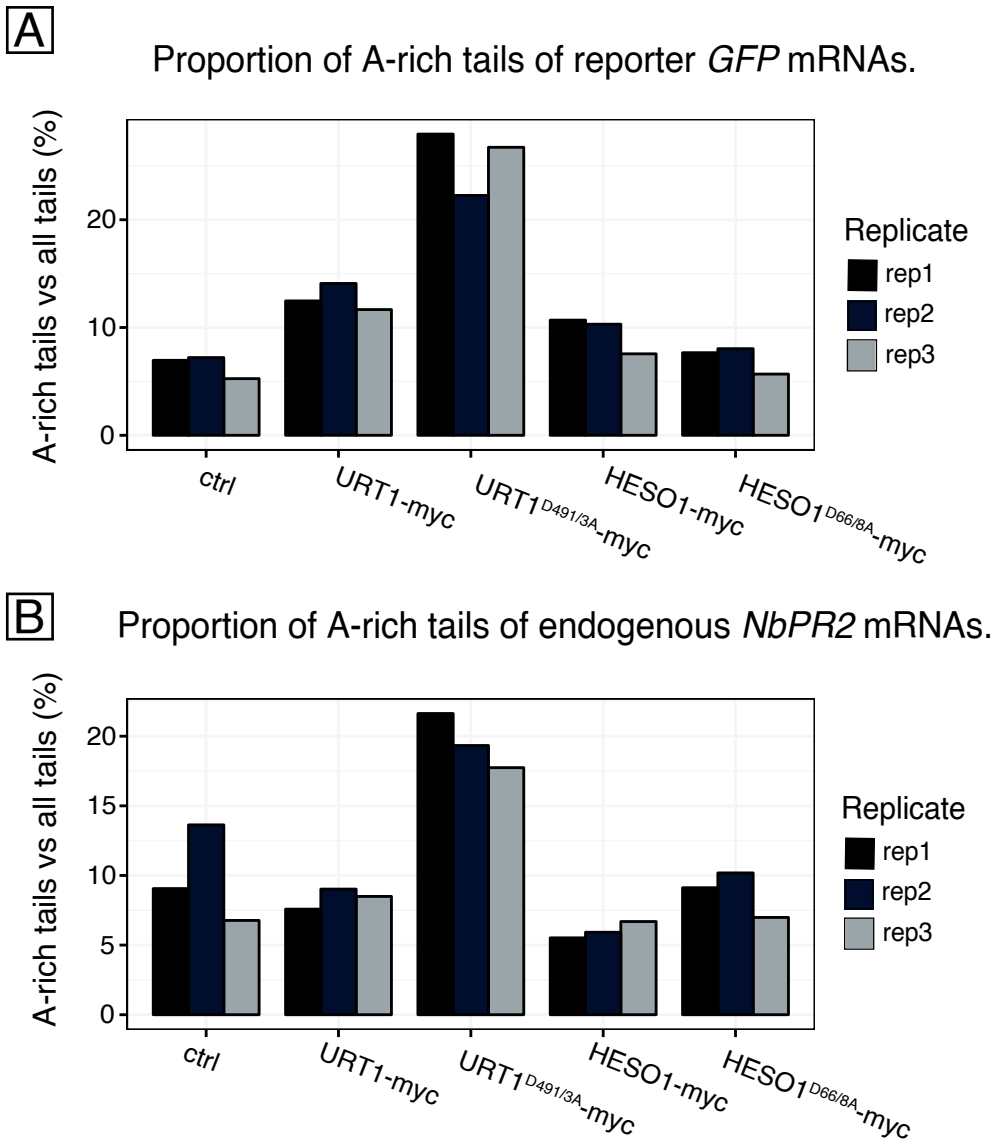


Figure S10: Proportion of A-rich tails of reporter *GFP* and endogenous *NbPR2* mRNAs.
A) Proportion of A-rich tails of reporter *GFP* mRNAs.
B) Proportion of A-rich tails of endogenous *NbPR2* mRNAs.

Bibliography

Achsel, T., Brahms, H., Kastner, B., Bachi, A., Wilm, M., and Lührmann, R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. EMBO J. *18*, 5789–5802.

Adivarahan, S., Livingston, N., Nicholson, B., Rahman, S., Wu, B., Rissland, O.S., and Zenklusen, D. (2018). Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway. Mol. Cell *72*, 727-738.e5.

De Alba, A.E.M., Moreno, A.B., Gabriel, M., Mallory, A.C., Christ, A., Bounon, R., Balzergue, S., Aubourg, S., Gautheret, D., Crespi, M.D., et al. (2015). In plants, decapping prevents RDR6-dependent production of small interfering RNAs from endogenous mRNAs. Nucleic Acids Res. *43*, 2902–2913.

De Almeida, C. and Scheer, H., Gobert, A., Fileccia, V., Martinelli, F., Zuber, H., and Gagliardi, D. (2018). RNA uridylation and decay in plants. Philos. Trans. R. Soc. B Biol. Sci. *373*.

De Almeida, C., Scheer, H., Zuber, H., and Gagliardi, D. (2018). RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. Wiley Interdiscip. Rev. RNA *9*.

Amaya Ramirez, C.C., Hubbe, P., Mandel, N., and Béthune, J. (2018). 4EHP-independent repression of endogenous mRNAs by the RNA-binding protein GIGYF2. Nucleic Acids Res. *46*, 5792–5808.

Amorim, M. de F., Willing, E.-M., Francisco-Mangilet, A.G., Droste-Borel, I., Macek, B., Schneeberger, K., and Laubinger, S. (2018). The U1 snRNP subunit LUC7 controls plant development and stress response through alternative splicing regulation. BioRxiv 150805.

Aphasizhev, R., Suematsu, T., Zhang, L., and Aphasizheva, I. (2016). Constructive edge of uridylation-induced RNA degradation. RNA Biol. *13*, 1078–1083.

Arae, T., Morita, K., Imahori, R., Suzuki, Y., Yasuda, S., Sato, T., Yamaguchi, J., and Chiba, Y. (2019). Identification of Arabidopsis CCR4-NOT complexes with Pumilio RNA binding proteins, APUM5 and APUM2. Plant Cell Physiol.

Astuti, D., Morris, M.R., Cooper, W.N., Staals, R.H.J., Wake, N.C., Fews, G.A., Gill, H., Gentle, D., Shuib, S., Ricketts, C.J., et al. (2012). Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat. Genet. *44*, 277–284.

Ayache, J., Bénard, M., Ernoult-Lange, M., Minshall, N., Standart, N., Kress, M., and Weil, D. (2015). P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes. *26*.

Badis, G., Saveanu, C., Fromont-Racine, M., and Jacquier, A. (2004). Targeted mRNA degradation by deadenylation-independent decapping. Mol. Cell *15*, 5–15.

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

Baer, B.W., and Kornberg, R.D. (1980). Repeating structure of cytoplasmic poly(A)-ribonucleoprotein. Proc. Natl. Acad. Sci. *77*, 1890–1892.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in Xenopus laevis retina: A mechanism for posttranscriptional control of circadian-related mRNA. Curr. Biol. *13*, 189–198.

Bai, Y., Salvadore, C., Chiang, Y.C., Collart, M.A., Liu, H.Y., and Denis, C.L. (1999). The CCR4 and CAF1 proteins of the CCR4-NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. Mol. Cell. Biol. *19*, 6642–6651.

Balatsos, N.A.A., Nilsson, P., Mazza, C., Cusack, S., and Virtanen, A. (2006). Inhibition of mRNA deadenylation by the nuclear cap binding complex (CBC). J. Biol. Chem. *281*, 4517–4522.

Balzeau, J., Menezes, M.R., Cao, S., and Hagan, J.P. (2017). The LIN28/let-7 pathway in cancer. Front. Genet. *8*.

Barbee, S.A., Estes, P.S., Cziko, A.M., Hillebrand, J., Luedeman, R.A., Coller, J.M., Johnson, N., Howlett, I.C., Geng, C., Ueda, R., et al. (2006). Staufen- and FMRP-Containing Neuronal RNPs Are Structurally and Functionally Related to Somatic P Bodies. Neuron *52*, 997–1009.

Basquin, J., Roudko, V. V., Rode, M., Basquin, C., Séraphin, B., and Conti, E. (2012). Architecture of the nuclease module of the yeast ccr4-Not complex: The not1-caf1-ccr4 interaction. Mol. Cell *48*, 207–218.

Bawankar, P., Loh, B., Wohlbold, L., Schmidt, S., and Izaurralde, E. (2013). NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. RNA Biol. *10*, 228–244.

Beelman, C.A., Stevens, A., Caponigro, G., LaGrandeur, T.E., Hatfield, L., Fortner, D.M., and Parker, R. (1996). An essential component of the decapping enzyme required for normal rates of mRNA turnover. Nature *382*, 642–646.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev. *20*, 1885–1898.

Beilharz, T.H., and Preiss, T. (2007). Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. RNA *13*, 982–997.

Belliot, G., Sosnovtsev, S. V, Chang, K.O., McPhie, P., and Green, K.Y. (2008). Nucleotidylylation of the VPg protein of a human norovirus by its proteinase-polymerase precursor protein. Virology *374*, 33–49.

Belostotsky, D.A. (2003). Unexpected complexity of poly(A)-binding protein gene families in flowering plants: Three conserved lineages that are at least 200 million years old and possible auto- and cross-regulation. Genetics *163*, 311–319.

Berndt, H., Harnisch, C., Rammelt, C., Stohr, N., Zirkel, A., Dohm, J.C., Himmelbauer, H., Tavanez, J.P., Huttelmaier, S., Wahle, E., et al. (2012). Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. RNA *18*, 958–972.

Bhaskar, V., Basquin, J., and Conti, E. (2015). Architecture of the ubiquitylation module of the yeast Ccr4-Not complex. Structure *23*, 921–928.

Blahna, M.T., Jones, M.R., Quinton, L.J., Matsuura, K.Y., and Mizgerd, J.P. (2011). Terminal uridyltransferase enzyme Zcchc11 promotes cell proliferation independent of its uridyltransferase activity. J. Biol. Chem. *286*, 42381–42389.

Boccaletto, P., MacHnicka, M.A., Purta, E., Pitkowski, P., Baginski, B., Wirecki, T.K., De Crécy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A., et al. (2018). MODOMICS: A database of RNA modification pathways. 2017 update. Nucleic Acids Res. *46*, D303–D307.

Boeck, R., Tarun, S., Rieger, M., Deardorff, J.A., Müller-Auer, S., and Sachs, A.B. (1996). The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. J. Biol. Chem. *271*, 432–438.

Boland, A., Chen, Y., Raisch, T., Jonas, S., Kuzuoğlu-Öztürk, D., Wohlbold, L., Weichenrieder, O., and Izaurralde, E. (2013). Structure and assembly of the NOT module of the human CCR4-NOT complex. Nat. Struct. Mol. Biol. *20*, 1289–1297.

Borja, M.S., Piotukh, K., Freund, C., and Gross, J.D. (2011). Dcp1 links coactivators of mRNA decapping to Dcp2 by proline recognition. RNA *17*, 278–290.

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

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

Braun, J.E., Tritschler, F., Haas, G., Igreja, C., Truffault, V., Weichenrieder, O., and Izaurralde, E. (2010). The C-terminal α-α superhelix of Pat is required for mRNA decapping in metazoa. EMBO J. *29*, 2368–2380.

Braun, J.E., Huntzinger, E., and Izaurralde, E. (2013). The role of GW182 proteins in miRNA-mediated gene silencing. Adv. Exp. Med. Biol. *768*, 147–163.

Brown, C.E., and Sachs, A.B. (1998). Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. Mol. Cell. Biol. *18*, 6548–6559.

Browning, K.S., and Bailey-Serres, J. (2015). Mechanism of Cytoplasmic mRNA Translation. Arab. B. *13*, e0176.

Burd, C.G., Matunis, E.L., and Dreyfuss, G. (1991). The multiple RNA-binding domains of the mRNA poly(A)-binding protein have different RNA-binding activities. Mol. Cell. Biol. *11*, 3419–3424.

Callahan, K.P., and Butler, J.S. (2010). TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. J. Biol. Chem. *285*, 3540–3547.

Caponigro, G., and Parker, R. (1995). Multiple functions for the poly(A)binding protein in mRNA decapping and deadenylation in yeast. Genes Dev. *9*, 2421–2432.

Castellano, L.A., and Bazzini, A.A. (2017). Poly(A) tails: longer is not always better. Nat. Struct. Mol. Biol. *24*, 1010–1011.

Cautain, B., Hill, R., De Pedro, N., and Link, W. (2015). Components and regulation of nuclear transport processes. FEBS J. *282*, 445–462.

Cevher, M.A., Zhang, X., Fernandez, S., Kim, S., Baquero, J., Nilsson, P., Lee, S., Virtanen, A., and Kleiman, F.E. (2010). Nuclear deadenylation/polyadenylation factors regulate 3' processing in response to DNA damage. EMBO J. *29*, 1674–1687.

Chang, H.-M., Triboulet, R., Thornton, J.E., and Gregory, R.I. (2013). A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature *497*, 244–248.

Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014). TAIL-seq: Genome-wide determination of poly(A) tail length and 3' end modifications. Mol. Cell *53*, 1044–1052.

Chapat, C., and Corbo, L. (2014). Novel roles of the CCR4-NOT complex. Wiley Interdiscip. Rev. RNA *5*, 883–901.

Charenton, C., Taverniti, V., Gaudon-Plesse, C., Back, R., Séraphin, B., and Graille, M. (2016). Structure of the active form of Dcp1-Dcp2 decapping enzyme bound to m 7 GDP and its Edc3 activator. Nat. Struct. Mol. Biol. *23*, 982–986.

Charlesworth, A., Meijer, H.A., and De Moor, C.H. (2013). Specificity factors in cytoplasmic polyadenylation. Wiley Interdiscip. Rev. RNA *4*, 437–461.

Chen, J., Rappsilber, J., Chiang, Y.-C., Russell, P., Mann, M., and Denis, C.L. (2001). Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. J. Mol. Biol. *314*, 683–694.

Chen, X.J., Zhang, X.H., Hu, L.D., Zhang, J.Q., Jiang, Y., Yang, Y., and Yan, Y. Bin (2016). DsCaf1 is involved in environmental stress response of Dunaliella salina. Int. J. Biol. Macromol. *82*, 369–374.

Chen, Y., Boland, A., Kuzuo, D., Bawankar, P., Loh, B., Chang, C. Te, Weichenrieder, O., Izaurralde, E., Kuzuoğlu-Öztürk, D., Bawankar, P., et al. (2014). DDX6-CNOT1 Complex and W-Binding Pockets in CNOT9 Reveal Direct Links between miRNA Target Recognition and Silencing. Mol. Cell *54*, 737–750.

Cheng, S., and Gallie, D.R. (2013). Eukaryotic initiation factor 4B and the poly(A)-binding protein bind eIF4G competitively. Translation *1*, e24038.

Cheng, H., Dufu, K., Lee, C.S., Hsu, J.L., Dias, A., and Reed, R. (2006). Human mRNA Export Machinery Recruited to the 5′ End of mRNA. Cell *127*, 1389–1400.

Chi, B., Wang, Q., Wu, G., Tan, M., Wang, L., Shi, M., Chang, X., and Cheng, H. (2013). Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. Nucleic Acids Res. *41*, 1294–1306.

Chiba, Y., Johnson, M.A., Lidder, P., Vogel, J.T., van Erp, H., and Green, P.J. (2004). AtPARN is an essential poly(A) ribonuclease in Arabidopsis. Gene *328*, 95–102.

Chicois, C., Scheer, H., Garcia, S., Zuber, H., Mutterer, J., Chicher, J., Hammann, P., Gagliardi, D., and Garcia, D. (2018). The UPF1 interactome reveals interaction networks between RNA degradation and translation repression factors in Arabidopsis. Plant J. *96*, 119–132.

Chiu, S.Y., Lejeune, F., Ranganathan, A.C., and Maquat, L.E. (2004). The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex. Genes Dev. *18*, 745–754.

Choi, Y.H., and Hagedorn, C.H. (2003). Purifying mRNAs with a high-affinity eIF4E mutant identifies the short 3′ poly(A) end phenotype. Proc. Natl. Acad. Sci. *100*, 7033–7038.

Chou, W.-L., Chung, Y.-L., Fang, J.-C., and Lu, C.-A. (2016). Novel interaction between CCR4 and CAF1 in rice CCR4–NOT deadenylase complex. Plant Mol. Biol. *93*, 79–96.

Chou, W.L., Huang, L.F., Fang, J.C., Yeh, C.H., Hong, C.Y., Wu, S.J., and Lu, C.A. (2014). Divergence of the expression and subcellular localization of CCR4-associated factor 1 (CAF1) deadenylase proteins in Oryza sativa. Plant Mol. Biol. *85*, 443–458.

Chowdhury, A., Mukhopadhyay, J., and Tharun, S. (2007). The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. RNA *13*, 998–1016.

Collart, M.A. (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip. Rev. RNA *7*, 438–454.

Collart, M.A., and Panasenko, O.O. (2012). The Ccr4 – Not complex. Gene *492*, 42–53.

Collart, M.A., and Struhl, K. (1994). NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization. Genes Dev. *8*, 525–537.

Conti, E., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: Molecular insights and mechanistic variations across species. Curr. Opin. Cell Biol. *17*, 316–325.

Copeland, P.R., and WORMINGTON, M. (2001). The mechanism and regulation of deadenylation: Identification and characterization of Xenopus PARN. RNA *7*, 875–886.

Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.Y., Schneller, S., Zust, R., Dong, H., et al. (2010). 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members.

Nature *468*, 452–456.

Daszkowska-Golec, A. (2018). Emerging Roles of the Nuclear Cap-Binding Complex in Abiotic Stress Responses. Plant Physiol. *176*, 242–253.

Decroly, E., Ferron, F., Lescar, J., and Canard, B. (2012). Conventional and unconventional mechanisms for capping viral mRNA. Nat. Rev. Microbiol. *10*, 51–65.

Dehecq, M., Decourty, L., Namane, A., Proux, C., Kanaan, J., Le Hir, H., Jacquier, A., and Saveanu, C. (2018). Nonsense-mediated mRNA decay involves two distinct Upf1-bound complexes. EMBO J. *37*, e99278.

Dehlin, E., Wormington, M., G. Körner, C., Wahle, E., and Anderson, J. (2000). Cap-dependent deadenylation of mRNA after poly(A) tail removal (but the subsequent degradation events remain.

Delis, C., Krokida, A., Tomatsidou, A., Tsikou, D., Beta, R.A.A., Tsioumpekou, M., Moustaka, J., Stravodimos, G., Leonidas, D.D., Balatsos, N.A.A.A., et al. (2015). AtHESPERIN: A Novel Regulator of Circadian Rhythms with Poly(A)-degrading Activity in Plants. RNA Biol. *6286*, 00–00.

Denis, C.L. (1984). Identification of new genes involved in the regulation of yeast alcohol dehydrogenase II. Genetics *108*, 833–844.

Deo, R.C., Bonanno, J.B., Sonenberg, N., and Burley, S.K. (1999). Recognition of polyadenylate RNA by the poly(A)-binding protein. Cell *98*, 835–845.

Devarkar, S.C., Wang, C., Miller, M.T., Ramanathan, A., Jiang, F., Khan, A.G., Patel, S.S., and Marcotrigiano, J. (2016). Structural basis for m7G recognition and 2′-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. Proc. Natl. Acad. Sci. *113*, 596–601.

Dias, S.M.G., Wilson, K.F., Rojas, K.S., Ambrosio, A.L.B., and Cerione, R.A. (2009). The molecular basis for the regulation of the cap-binding complex by the importins. Nat. Struct. Mol. Biol. *16*, 930–937.

Domingues, M.N., Sforça, M.L., Soprano, A.S., Lee, J., De Souza, T.D.A.C.B., Cassago, A., Portugal, R.V., De Mattos Zeri, A.C., Murakami, M.T., Sadanandom, A., et al. (2015). Structure and Mechanism of Dimer-Monomer Transition of a Plant Poly(A)-Binding Protein upon RNA Interaction: Insights into Its Poly(A) Tail Assembly. J. Mol. Biol. *427*, 2491–2506.

Dong, S., Jacobson, A., and He, F. (2010). Degradation of YRA1 pre-mRNA in the cytoplasm requires translational repression, multiple modular intronic elements, Edc3p, and Mex67p. PLoS Biol. *8*, e1000360.

Draper, M.P., Liu, H.Y., Nelsbach, A.H., Mosley, S.P., and Denis, C.L. (2015). CCR4 is a glucose-regulated transcription factor whose leucine-rich repeat binds several proteins important for placing CCR4 in its proper promoter context. Mol. Cell. Biol. *14*, 4522–4531.

Dreyfus, M., and Régnier, P. (2004). The Poly(A) Tail of mRNAs. Cell *111*, 611–613.

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m 6 A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat. Commun. *7*, 12626.

Dupressoir, A., Morel, A.-P., Barbot, W., Loireau, M.-P., Corbo, L., and Heidmann, T. (2001). Identification of four families of yCCR4- and Mg2+-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. BMC Genomics *2*, 9.

Dyson, H.J. (2016). Making Sense of Intrinsically Disordered Proteins. Biophys. J. *110*, 1013–1016.

Dziembowski, A., Lorentzen, E., Conti, E., and Séraphin, B. (2007). A single subunit, Dis3, is essentially

responsible for yeast exosome core activity. Nat. Struct. Mol. Biol. *14*, 15–22.

Edmonds, M. (2002). A history of poly A sequences: from formation to factors to function. Prog. Nucleic Acid Res. Mol. Biol. *71*, 285–389.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. Mol. Cell. Biol. *27*, 3970–3981.

Evguenieva-Hackenberg, E., Hou, L., Glaeser, S., and Klug, G. (2014). Structure and function of the archaeal exosome. Wiley Interdiscip. Rev. RNA *5*, 623–635.

Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., et al. (2011). MiRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. Nat. Struct. Mol. Biol. *18*, 1211–1217.

Faehnle, C.R., Walleshauser, J., and Joshua-Tor, L. (2014). Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. Nature *514*, 252–256.

Faehnle, C.R., Walleshauser, J., and Joshua-Tor, L. (2017). Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis. Nat. Struct. Mol. Biol. *24*, 658–665.

Falk, S., Weir, J.R., Hentschel, J., Reichelt, P., Bonneau, F., and Conti, E. (2014). The Molecular Architecture of the TRAMP Complex Reveals the Organization and Interplay of Its Two Catalytic Activities. Mol. Cell *55*, 856–867.

Ferrier, E. (2013) Rôle et mode d'action de l'UTP:RNA uridylyltransferase URT1 dans l'uridylation et la dégradation des ARNm chez Arabidopsis thaliana (doctoral dissertation). University of Strasbourg, Strasbourg, France.

Flaherty, S.M., Izaurralde, E., Mattaj, I.W., and Gilmartin, G.M. (1997). Porticipation of the nuclear cap binding complex in pre-mRNA 3’ processing.

Fortes, P., Kufel, J., Fornerod, M., Polycarpou-Schwarz, M., Lafontaine, D., Tollervey, D., and Mattaj, I.W. (2015). Genetic and Physical Interactions Involving the Yeast Nuclear Cap-Binding Complex. Mol. Cell. Biol. *19*, 6543–6553.

Funakoshi, Y., Doi, Y., Hosoda, N., Uchida, N., Osawa, M., Shimada, I., Tsujimoto, M., Suzuki, T., Katada, T., and Hoshino, S.I. (2007). Mechanism of mRNA deadenylation: Evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. Genes Dev. *21*, 3135–3148.

Furuichi, Y., Morgan, M., Shatkin, A.J., Jelinek, W., Salditt-Georgieff, M., and Darnell, J.E. (1975). Methylated, blocked 5 termini in HeLa cell mRNA. Proc. Natl. Acad. Sci. U. S. A. *72*, 1904–1908.

Gallie, D.R. (2002). The 5’-leader of tobacco mosaic virus promotes translation through enhanced recruitment of eIF4F. Nucleic Acids Res. *30*, 3401–3411.

Gallie, D.R. (2014). The role of the poly(A) binding protein in the assembly of the Cap-binding complex during translation initiation in plants. Translation *2*, e959378.

Gallie, D.R. (2017). Class II members of the poly(A) binding protein family exhibit distinct functions during Arabidopsis growth and development. Translation *5*, e1295129.

Gallie, D.R., and Browning, K.S. (2001). eIF4G Functionally Differs from eIFiso4G in Promoting Internal Initiation, Cap-independent Translation, and Translation of Structured mRNAs. J. Biol. Chem. *276*, 36951–36960.

Gallie, D.R., and Liu, R. (2014). Phylogenetic analysis reveals dynamic evolution of the poly(A)-binding protein gene family in plants.

Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C., and Wilson, T. michael A. (1987). The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts in vitro and in vivo. Nucleic Acids Res. *15*, 3257–3273.

Gao, M., Fritz, D.T., Ford, L.P., and Wilusz, J. (2000). Interaction between a poly(A)-specific ribonuclease and the 5' cap influences mRNA deadenylation rates in vitro. Mol. Cell *5*, 479–488.

Garbarino-Pico, E., Niu, S., Rollag, M.D., Strayer, C.A., Besharse, J.C., and Green, C.B. (2007). Immediate early response of the circadian polyA ribonuclease nocturnin to two extracellular stimuli. RNA *13*, 745–755.

Ghosh, A., and Lima, C.D. (2010). Enzymology of RNA cap synthesis. Wiley Interdiscip. Rev. RNA *1*, 152–172.

Glaunsinger, B.A., and Lee, Y.J. (2010). How tails define the ending: Divergent roles for polyadenylation in RNA stability and gene expression. RNA Biol. *7*, 13–17.

Godwin, A.R., Kojima, S., Green, C.B., and Wilusz, J. (2013). Kiss your tail goodbye: The role of PARN, Nocturnin, and Angel deadenylases in mRNA biology. Biochim. Biophys. Acta - Gene Regul. Mech. *1829*, 571–579.

Goldstrohm, A.C., and Wickens, M. (2008). Multifunctional deadenylase complexes diversify mRNA control. Nat. Rev. Mol. Cell Biol. *9*, 337–344.

Goss, D.J., and Kleiman, F.E. (2013). Poly(A) binding proteins: Are they all created equal? Wiley Interdiscip. Rev. RNA *4*, 167–179.

Green, C.B., and Besharse, J.C. (2002). Identification of a novel vertebrate circadian clock-regulated gene encoding the protein nocturnin. Proc. Natl. Acad. Sci. *93*, 14884–14888.

Gregory, B.D., O'Malley, R.C., Lister, R., Urich, M.A., Tonti-Filippini, J., Chen, H., Millar, A.H., and Ecker, J.R. (2008). A link between RNA metabolism and silencing affecting Arabidopsis development. Dev. Cell *14*, 854–866.

Guo, C., Spinelli, M., Liu, M., Li, Q.Q., and Liang, C. (2016). A Genome-wide Study of "non-3UTR" Polyadenylation Sites in Arabidopsis thaliana. Sci. Rep. *6*, 28060.

Gy, I., Gasciolli, V., Lauressergues, D., Morel, J.-B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C. (2007). Arabidopsis FIERY1, XRN2, and XRN3 Are Endogenous RNA Silencing Suppressors . Plant Cell *19*, 3451–3461.

Haas, G., Braun, J.E., Igreja, C., Tritschler, F., Nishihara, T., and Izaurralde, E. (2010). HPat provides a link between deadenylation and decapping in metazoa. J. Cell Biol. *189*, 289–302.

Hajnsdorf, E., and Kaberdin, V.R. (2018). RNA polyadenylation and its consequences in prokaryotes. Philos. Trans. R. Soc. B Biol. Sci. *373*.

Halter, D., Collart, M.A., and Panasenko, O.O. (2014). The Not4 E3 ligase and CCR4 deadenylase play distinct roles in protein quality control. PLoS One *9*.

Hanson, G., and Coller, J. (2018). Translation and Protein Quality Control: Codon optimality, bias and usage in translation and mRNA decay. Nat. Rev. Mol. Cell Biol.

He, W., and Parker, R. (2001). The yeast cytoplasmic lsm1/pat1p complex protects mRNA 3' termini from partial degradation. Genetics *158*, 1445–1455.

He, F., Li, C., Roy, B., and Jacobson, A. (2014). Yeast Edc3 Targets RPS28B mRNA for Decapping by Binding to a 3' Untranslated Region Decay-Inducing Regulatory Element. Mol. Cell. Biol. *34*, 1438–1451.

Hector, R.E., Nykamp, K.R., Dheur, S., Anderson, J.T., Non, P.J., Urbinati, C.R., Wilson, S.M., Minvielle-Sebastia, L., and Swanson, M.S. (2002). Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. EMBO J. *21*, 1800–1810.

Hématy, K., Bellec, Y., Podicheti, R., Bouteiller, N., Anne, P., Morineau, C., Haslam, R.P., Beaudoin, F., Napier, J.A., Mockaitis, K., et al. (2016). The Zinc-Finger Protein SOP1 Is Required for a Subset of the Nuclear Exosome Functions in Arabidopsis. PLoS Genet. *12*.

Hilleren, P., McCarthy, T., Rosbash, M., Parker, R., and Jensen, T.H. (2001). Quality control of mRNA 3′-end processing is linked to the nuclear exosome. Nature *413*, 538–542.

Hirayama, T. (2014). A unique system for regulating mitochondrial mRNA poly(A) status and stability in plants. Plant Signal. Behav. *9*, e973809.

Hirayama, T., Matsuura, T., Ushiyama, S., Narusaka, M., Kurihara, Y., Yasuda, M., Ohtani, M., Seki, M., Demura, T., Nakashita, H., et al. (2013). A poly(A)-specific ribonuclease directly regulates the poly(A) status of mitochondrial mRNA in Arabidopsis. Nat. Commun. *4*, 1–9.

Horvathova, I., Voigt, F., Kotrys, A. V., Zhan, Y., Artus-Revel, C.G., Eglinger, J., Stadler, M.B., Giorgetti, L., and Chao, J.A. (2017). The Dynamics of mRNA Turnover Revealed by Single-Molecule Imaging in Single Cells. Mol. Cell *68*, 615-625.e9.

Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., and Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. Nature *461*, 225–229.

Hu, W., Petzold, C., Coller, J., and Baker, K.E. (2010). Nonsense-mediated mRNA decapping occurs on polyribosomes in Saccharomyces cerevisiae. Nat. Struct. Mol. Biol. *17*, 244–247.

Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J.B., Munier, A., Fradet, M., et al. (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. Mol. Cell *68*, 144-157.e5.

Hughes, K.L., Abshire, E.T., and Goldstrohm, A.C. (2018). Regulatory roles of vertebrate Nocturnin: insights and remaining mysteries. RNA Biol. *15*, 1255–1267.

Hunt, A.G. (2008). Messenger RNA 3′ end formation in plants. Curr. Top. Microbiol. Immunol. *326*, 151–177.

Hunt, A.G., Xing, D., and Li, Q.Q. (2012). Plant polyadenylation factors: Conservation and variety in the polyadenylation complex in plants. BMC Genomics *13*, 1.

Hunter, R.W., Liu, Y., Manjunath, H., Acharya, A., Jones, B.T., Zhang, H., Chen, B., Ramalingam, H., Hammer, R.E., Xie, Y., et al. (2018). Loss of Dis3l2 partially phenocopies perlman syndrome in mice and results in upregulation of Igf2 in nephron progenitor cells. Genes Dev. *32*, 903–908.

Ibrahim, F., Rohr, J., Jeong, W.J., Hesson, J., and Cerutti, H. (2006). Untemplated oligoadenylation promotes degradation of RISC-cleaved transcripts. Science (80-.). *314*, 1893.

Ibrahim, F., Rymarquis, L.A., Kim, E.-J., Becker, J., Balassa, E., Green, P.J., and Cerutti, H. (2010). Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. PNAS *107*, 3906–3911.

Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: Safeguarding cells from abnormal mRNA function. Genes Dev. *21*, 1833–1856.

Izaurralde, E., McGuigan, C., Mattaj, I.W., Lewis, J.D., and Jarmolowski, A. (2007). A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5′ splice site.

Izumi, N., Shoji, K., Sakaguchi, Y., Honda, S., Kirino, Y., Suzuki, T., Katsuma, S., and Tomari, Y. (2016).

Identification and Functional Analysis of the Pre-piRNA 3' Trimmer in Silkworms. Cell *164*, 962–973.

Jackson, R.J., Hellen, C.U.T., and Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. *11*, 113–127.

Jain, M., Olsen, H.E., Paten, B., and Akeson, M. (2016). The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. Genome Biol. *17*.

Januszyk, K., and Lima, C.D. (2014). The eukaryotic RNA exosome. Curr. Opin. Struct. Biol. *24*, 132–140.

Järvelin, A.I., Noerenberg, M., Davis, I., and Castello, A. (2016). The new (dis)order in RNA regulation. Cell Commun. Signal. *14*, 9.

Jiao, X., Chang, J.H., Kilic, T., Tong, L., and Kiledjian, M. (2013). A Mammalian Pre-mRNA 5' End Capping Quality Control Mechanism and an Unexpected Link of Capping to Pre-mRNA Processing. Mol. Cell *50*, 104–115.

Kamenska, A., Simpson, C., Vindry, C., Broomhead, H., Bénard, M., Ernoult-Lange, M., Lee, B.P., Harries, L.W., Weil, D., and Standart, N. (2016). The DDX6-4E-T interaction mediates translational repression and P-body assembly. Nucleic Acids Res. *44*, 6318–6334.

Kastenmayer, J.P., and Green, P.J. (2002). Novel features of the XRN-family in Arabidopsis: Evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. Proc. Natl. Acad. Sci. *97*, 13985–13990.

Kaufmann, I., Martin, G., Friedlein, A., Langen, H., and Keller, W. (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. EMBO J. *23*, 616–626.

Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fitzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. J. Cell Biol. *169*, 871–884.

Kerwitz, Y., Kühn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., and Wahle, E. (2003). Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. EMBO J. *22*, 3705–3714.

Khanam, T., Sondekoppa Muddashetty, R., Kahvejian, A., Muddashetty, R.S., Kahvejian, A., Sonenberg, N., and Brosius, J. (2006). Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1+2 and 3+4. RNA Biol. *3*, 170–177.

Kierzkowski, D., Kmieciak, M., Piontek, P., Wojtaszek, P., Szweykowska-Kulinska, Z., and Jarmolowski, A. (2009). The Arabidopsis CBP20 targets the cap-binding complex to the nucleus, and is stabilized by CBP80. Plant J. *59*, 813–823.

Kilchert, C., Wittmann, S., and Vasiljeva, L. (2016). The regulation and functions of the nuclear RNA exosome complex. Nat. Rev. Mol. Cell Biol. *17*, 227–239.

Kim, B., Ha, M., Loeff, L., Chang, H., Simanshu, D.K., Li, S., Fareh, M., Patel, D.J., Joo, C., and Kim, V.N. (2015). TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. EMBO J. *34*, 1801–1815.

Kim, S., Yang, J.Y., Xu, J., Jang, I.C., Prigge, M.J., and Chua, N.H. (2008). Two cap-binding proteins CBP20 and CBP80 are involved in processing primary microRNAs. Plant Cell Physiol. *49*, 1634–1644.

Kini, H.K., Silverman, I.M., Ji, X., Gregory, B.D., and Liebhaber, S.A. (2016). Cytoplasmic poly(A) binding protein-1 binds to genomically encoded sequences within mammalian mRNAs. RNA *22*, 61–74.

Kolesnikova, O., Back, R., Graille, M., and Séraphin, B. (2013). Identification of the Rps28 binding motif from yeast Edc3 involved in the autoregulatory feedback loop controlling RPS28B mRNA decay. Nucleic

Acids Res. *41*, 9514–9523.

Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. *14*, 2452–2460.

Körner, C.G., and Wahle, E. (1997). Poly(A) tail shortening by a mammalian poly(A)-specific 3'-exoribonuclease. J. Biol. Chem. *272*, 10448–10456.

Körner, C.G., Wormington, M., Muckenthaler, M., Schneider, S., Dehlin, E., and Wahle, E. (1998). The deadenylating nuclease (DAN) is involved in poly(A) tail removal during the meiotic maturation of Xenopus oocytes. EMBO J. *17*, 5427–5437.

Krause, M., Niazi, A.M., Labun, K., Torres Cleuren, Y.N., Müller, F.S., and Valen, E. (2019). taillndr: Alignment-free poly(A) length measurement for Oxford Nanopore RNA and DNA sequencing. BioRxiv 1–21.

Kroupova, A., Ivasçu, A.I., Reimão, M.M., Reimão-Pinto, R., Ameres, S.L., and Jinek, M. (2019). Structural basis for acceptor RNA substrate selectivity of the 3 terminal uridylyl transferase Tailor. Nucleic Acids Res. *47*, 1030–1042.

Kuai, L., Fang, F., Butler, J.S., and Sherman, F. (2004). Polyadenylation of rRNA in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. *101*, 8581–8586.

Kühn, U., and Pieler, T. (1996). Xenopus poly(A) binding protein: Functional domains in RNA binding and protein-protein interaction. J. Mol. Biol. *256*, 20–30.

Kühn, U., and Wahle, E. (2004). Structure and function of poly(A) binding proteins. Biochim. Biophys. Acta - Gene Struct. Expr. *1678*, 67–84.

Kühn, U., Gündel, M., Knoth, A., Kerwitz, Y., Rüdel, S., and Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear Poly(A)-binding protein regulating the interaction between Poly(A) polymerase and the cleavage and polyadenylation specificity factor. J. Biol. Chem. *284*, 22803–22814.

Kurihara, Y. (2017). Activity and roles of Arabidopsis thaliana XRN family exoribonucleases in noncoding RNA pathways. J. Plant Res. *130*, 25–31.

Kurihara, Y., Schmitz, R.J., Nery, J.R., Schultz, M.D., Okubo-Kurihara, E., Morosawa, T., Tanaka, M., Toyoda, T., Seki, M., and Ecker, J.R. (2012). Surveillance of 3' Noncoding Transcripts Requires FIERY1 and XRN3 in *Arabidopsis*. G3: Genes|Genomes|Genetics *2*, 487–498.

łabno, A., Warkocki, Z., Kulínski, T., Krawczyk, P.S., Bijata, K., Tomecki, R., and Dziembowski, A. (2016). Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. Nucleic Acids Res. *44*, 10437–10453.

Łabno, A., Tomecki, R., and Dziembowski, A. (2016). Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim. Biophys. Acta - Mol. Cell Res. *1863*, 3125–3147.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell *121*, 713–724.

Lahudkar, S., Shukla, A., Bajwa, P., Durairaj, G., Stanojevic, N., and Bhaumik, S.R. (2011). The mRNA cap-binding complex stimulates the formation of pre-initiation complex at the promoter via its interaction with Mot1p in vivo. Nucleic Acids Res. *39*, 2188–2209.

Lai, W.S., Kennington, E.A., and Blackshear, P.J. (2003). Tristetraprolin and Its Family Members Can Promote the Cell-Free Deadenylation of AU-Rich Element-Containing mRNAs by Poly(A) Ribonuclease.

Mol. Cell. Biol. *23*, 3798–3812.

Lange, H., Holec, S., Cognat, V., Pieuchot, L., Le Ret, M., Canaday, J., and Gagliardi, D. (2008). Degradation of a Polyadenylated rRNA Maturation By-Product Involves One of the Three RRP6-Like Proteins in Arabidopsis thaliana. Mol. Cell. Biol. *28*, 3038–3044.

Lange, H., Sement, F.M., Canaday, J., and Gagliardi, D. (2009). Polyadenylation-assisted RNA degradation processes in plants. Trends Plant Sci. *14*, 497–504.

Lange, H., Zuber, H., Sement, F.M., Chicher, J., Kuhn, L., Hammann, P., Brunaud, V., Bérard, C., Bouteiller, N., Balzergue, S., et al. (2014). The RNA Helicases AtMTR4 and HEN2 Target Specific Subsets of Nuclear Transcripts for Degradation by the Nuclear Exosome in Arabidopsis thaliana. PLoS Genet. *10*.

Lange, H., Ndecky, S.Y.A., Gomez-Diaz, C., Pflieger, D., Butel, N., Zumsteg, J., Kuhn, L., Piermaria, C., Chicher, J., Christie, M., et al. (2019). RST1 and RIPR connect the cytosolic RNA exosome to the Ski complex in Arabidopsis. BioRxiv 617894.

Lapointe, C.P., and Wickens, M. (2013). The nucleic acid-binding domain and translational repression activity of a Xenopus terminal uridylyl transferase. J. Biol. Chem. *288*, 20723–20733.

Lässig, C., and Hopfner, K.P. (2017). Discrimination of cytosolic self and non-self RNA by RIG-I-like receptors. J. Biol. Chem. *292*, 9000–9009.

Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Ratsch, G., and Weigel, D. (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., et al. (2014). Classification of Intrinsically Disordered Regions and Proteins. Chem. Rev. *114*, 6589–6631.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol. Cell *12*, 675–687.

Lejeune, F., Ranganathan, A.C., and Maquat, L.E. (2004). eIF4G is required for the pioneer round of translation in mammalian cells. Nat. Struct. Mol. Biol. *11*, 992–1000.

Lellis, A.D., Allen, M.L., Aertker, A.W., Tran, J.K., Hillis, D.M., Harbin, C.R., Caldwell, C., Gallie, D.R., and Browning, K.S. (2010). Deletion of the eIFiso4G subunit of the Arabidopsis eIFiso4F translation initiation complex impairs health and viability. Plant Mol. Biol. *74*, 249–263.

Levy, S., and Schuster, G. (2016). Polyadenylation and degradation of RNA in the mitochondria. Biochem. Soc. Trans. *44*, 1475–1482.

Li, Y., and Kiledjian, M. (2010). Regulation of mRNA decapping. Wiley Interdiscip. Rev. RNA *1*, 253–265.

Li, Y., and Maine, E.M. (2018). The balance of poly(U) polymerase activity ensures germline identity, survival and development in Caenorhabditis elegans . Development *145*, dev165944.

Li, R., Yue, J., Zhang, Y., Zhou, L., Hao, W., Yuan, J., Qiang, B., Ding, J.M., Peng, X., and Cao, J.M. (2008). CLOCK/BMAL1 regulates human nocturnin transcription through binding to the E-box of nocturnin promoter. Mol. Cell. Biochem. *317*, 169–177.

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

Liao, S., Sun, H., and Xu, C. (2018). YTH Domain: A Family of N 6 -methyladenosine (m 6 A) Readers.

Genomics, Proteomics Bioinforma. *16*, 99–107.

Licht, K., Medenbach, J., Luhrmann, R., Kambach, C., Bindereif, A., Lührmann, R., Kambach, C., and Bindereif, A. (2008). 3'-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. Rna *14*, 1532–1538.

Lim, J., Ha, M., Chang, H., Kwon, S.C., Simanshu, D.K., Patel, D.J., and Kim, V.N. (2014). Uridylation by TUT4 and TUT7 marks mRNA for degradation. Cell *159*, 1365–1376.

Lim, J., Lee, M., Son, A., Chang, H., and Kim, V.N. (2016). MTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. Genes Dev. *30*, 1671–1682.

Lim, J., Kim, D., Lee, Y.S., Ha, M., Lee, M., Yeo, J., Chang, H., Song, J., Ahn, K., and Kim, V.N. (2018). Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. Science (80-.). *361*, 701–704.

Lima, S.A., Chipman, L.B., Nicholson, A.L., Chen, Y.H., Yee, B.A., Yeo, G.W., Coller, J., and Pasquinelli, A.E. (2017). Short poly(A) tails are a conserved feature of highly expressed genes. Nat. Struct. Mol. Biol. *24*, 1057–1063.

Lin, C.-J.J., Wen, J., Bejarano, F., Hu, F., Bortolamiol-Becet, D., Kan, L., Sanfilippo, P., Kondo, S., and Lai, E.C. (2017). Characterization of a TUTase/RNase complex required for Drosophila gametogenesis. RNA *23*, 284–296.

Liu, X., Zheng, Q., Vrettos, N., Maragkakis, M., Alexiou, P., Gregory, B.D., and Mourelatos, Z. (2014). A MicroRNA Precursor Surveillance System in Quality Control of MicroRNA Synthesis. Mol. Cell *55*, 868–879.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402–408.

Loke, J.C. (2005). Compilation of mRNA Polyadenylation Signals in Arabidopsis Revealed a New Signal Element and Potential Secondary Structures. PLANT Physiol. *138*, 1457–1468.

Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. EMBO J. *32*, 1855–1868.

Lutz, C.S., and Moreira, A. (2011). Alternative mRNA polyadenylation in eukaryotes: An effective regulator of gene expression. Wiley Interdiscip. Rev. RNA *2*, 22–31.

Machín, Á., Martín Alonso, J.M., and Parra, F. (2001). Identification of the Amino Acid Residue Involved in Rabbit Hemorrhagic Disease Virus VPg Uridylylation. J. Biol. Chem. *276*, 27787–27792.

Maillet, L., and Collart, M.A. (2002). Interaction between Not1p, a component of the Ccr4-Not complex, a global regulator of transcription, and Dhh1p, a putative RNA helicase. J. Biol. Chem. *277*, 2835–2842.

Maillet, L., Tu, C., Hong, Y.K., Shuster, E.O., and Collart, M.A. (2000). The essential function of Not1 lies within the Ccr4-Not complex. J. Mol. Biol. *303*, 131–143.

Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M.G., and Arraiano, C.M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. TL - 32. EMBO J. *32 VN-r*, 1842–1854.

Mangus, D.A., Evans, M.C., and Jacobson, A. (2003). Poly(A)-binding proteins: Multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol. *4*, 223.

Mangus, D.A., Evans, M.C., Agrin, N.S., Smith, M., Gongidi, P., and Jacobson, A. (2004). Positive and Negative Regulation of Poly(A) Nuclease. Mol. Cell. Biol. *24*, 5521–5533.

Maragozidis, P., Papanastasi, E., Scutelnic, D., Totomi, A., Kokkori, I., Zarogiannis, S.G., Kerenidi, T., Gourgoulianis, K.I., and Balatsos, N.A.A. (2015). Poly(A)-specific ribonuclease and Nocturnin in squamous cell lung cancer: Prognostic value and impact on gene expression. Mol. Cancer *14*.

Martin, G., and Keller, W. (2007). RNA-specific ribonucleotidyl transferases. RNA *13*, 1834–1849.

Martînez, J., Ren, Y.G., Nilsson, P., Ehrenberg, M., and Virtanen, A. (2001). The mRNA Cap Structure Stimulates Rate of Poly(A) Removal and Amplifies Processivity of Degradation. J. Biol. Chem. *276*, 27923–27929.

Mathys, H., Basquin, J.Ô., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., and Filipowicz, W. (2014). Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in MicroRNA Repression. Mol. Cell *54*, 751–765.

Mauxion, F., Faux, C., and Séraphin, B. (2008). The BTG2 protein is a general activator of mRNA deadenylation. EMBO J. *27*, 1039–1048.

Mauxion, F., Prève, B., and Séraphin, B. (2013). C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. RNA Biol. *10*, 267–276.

Mayberry, L.K., Allen, M.L., Nitka, K.R., Campbell, L., Murphy, P.A., and Browning, K.S. (2011). Plant Cap-binding Complexes Eukaryotic Initiation Factors eIF4F and eIFISO4F. J. Biol. Chem. *286*, 42566–42574.

Meaux, S.A., Holmquist, C.E., and Marzluff, W.F. (2018). Role of oligouridylation in normal metabolism and regulated degradation of mammalian histone mRNAs. Philos. Trans. R. Soc. B Biol. Sci. *373*.

Meijer, H.A., Bushell, M., Hill, K., Gant, T.W., Willis, A.E., Jones, P., and de Moor, C.H. (2007). A novel method for poly(A) fractionation reveals a large population of mRNAs with a short poly(A) tail in mammalian cells. Nucleic Acids Res. *35*.

Melo, E.O., Dhalia, R., De Sa, C.M., Standart, N., and De Melo Neto, O.P. (2003). Identification of a C-terminal poly(A)-binding protein (PABP)-PABP interaction domain: Role in cooperative binding to poly(A) and efficient cap distal translational repression. J. Biol. Chem. *278*, 46357–46368.

Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A., et al. (2016). Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. Mol. Cell *64*, 520–533.

Merret, R., Nagarajan, V.K., Carpentier, M.-C., Park, S., Favory, J.-J., Descombin, J., Picart, C., Charng, Y.-Y., Green, P.J., Deragon, J.-M., et al. (2015). Heat-induced ribosome pausing triggers mRNA co-translational decay in Arabidopsis thaliana. Nucleic Acids Res. *43*, 4121–4132.

Miki, T.S., and Großhans, H. (2013). The multifunctional RNase XRN2. Biochem. Soc. Trans. *41*, 825–830.

Miller, J.E., and Reese, J.C. (2012). Ccr4-Not complex: The control freak of eukaryotic cells. Crit. Rev. Biochem. Mol. Biol. *47*, 315–333.

Millevoi, S., and Vagner, S. (2009). Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. Nucleic Acids Res. *38*, 2757–2774.

Moon, D.H., Segal, M., Boyraz, B., Guinan, E., Hofmann, I., Cahan, P., Tai, A.K., and Agarwal, S. (2015). Poly(A)-specific ribonuclease (PARN) mediates 3'-end maturation of the telomerase RNA component. Nat. Genet. *47*, 1482–1488.

Moreno, A.B., De Alba, A.E.M., Bardou, F., Crespi, M.D., Vaucheret, H., Maizel, A., and Mallory, A.C. (2013). Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-

transcriptional gene silencing in plants. Nucleic Acids Res. *41*, 4699–4708.

Morgan, M., Much, C., DiGiacomo, M., Azzi, C., Ivanova, I., Vitsios, D.M., Pistolic, J., Collier, P., Moreira, P.N., Benes, V., et al. (2017). mRNA 3' uridylation and poly(A) tail length sculpt the mammalian maternal transcriptome. Nature *548*, 347–351.

Morgan, M., Kabayama, Y., Much, C., Ivanova, I., Di Giacomo, M., Auchynnikava, T., Monahan, J.M., Vitsios, D.M., Vasiliauskaitė, L., Comazzetto, S., et al. (2019). A programmed wave of uridylation-primed mRNA degradation is essential for meiotic progression and mammalian spermatogenesis. Cell Res. *29*, 221–232.

Morozov, I.Y., Jones, M.G., Razak, A.A., Rigden, D.J., and Caddick, M.X. (2010). CUCU modification of mRNA promotes decapping and transcript degradation in Aspergillus nidulans. Mol. Cell. Biol. *30*, 460–469.

Morozov, I.Y., Jones, M.G., Gould, P.D., Crome, V., Wilson, J.B., Hall, A.J.W., Rigden, D.J., and Caddick, M.X. (2012). mRNA 3' Tagging Is Induced by Nonsense-Mediated Decay and Promotes Ribosome Dissociation. Mol. Cell. Biol. *32*, 2585–2595.

Morris, M.R., Astuti, D., and Maher, E.R. (2013). Perlman syndrome: overgrowth, Wilms tumor predisposition and DIS3L2. Am. J. Med. Genet. C. Semin. Med. Genet. *163C*, 106–113.

Mugridge, J.S., Tibble, R.W., Ziemniak, M., Jemielity, J., and Gross, J.D. (2018). Structure of the activated Edc1-Dcp1-Dcp2-Edc3 mRNA decapping complex with substrate analog poised for catalysis. Nat. Commun. *9*, 1152.

Müller, W.E.G., Schröder, H.C., Pisignano, D., Markl, J.S., and Wang, X. (2013). Metazoan circadian rhythm: Toward an understanding of a light-based zeitgeber in sponges. In Integrative and Comparative Biology, (Narnia), pp. 103–117.

Munoz-tello, P., Rajappa, L., Coquille, S., Thore, S. xe9, and phane (2015). Polyuridylation in Eukaryotes : A 3' -End Modification Regulating RNA Life. Biomed Res. Int. *2015*, 1–12.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'→3' exoribonucleases: Structure, mechanisms and functions. Biochim. Biophys. Acta - Gene Regul. Mech. *1829*, 590–603.

Nasertorabi, F., Batisse, C., Diepholz, M., Suck, D., and Böttcher, B. (2011). Insights into the structure of the CCR4-NOT complex by electron microscopy. FEBS Lett. *585*, 2182–2186.

Nguyen, A.H., Matsui, A., Tanaka, M., Mizunashi, K., Nakaminami, K., Hayashi, M., Iida, K., Toyoda, T., Nguyen, D. Van, and Seki, M. (2015). Loss of Arabidopsis 5'–3' Exoribonuclease AtXRN4 Function Enhances Heat Stress Tolerance of Plants Subjected to Severe Heat Stress. Plant Cell Physiol. *56*, pcv096.

Nicholson, A.L., and Pasquinelli, A.E. (2019). Tales of Detailed Poly(A) Tails. Trends Cell Biol. *29*, 191–200.

Nietfeld, W., Mentzel, H., and Pieler, T. (1990). The Xenopus laevis poly(A) binding protein is composed of multiple functionally independent RNA binding domains. EMBO J. *9*, 3699–3705.

Nilsson, P., Henriksson, N., Niedzwiecka, A., Balatsos, N.A.A., Kokkoris, K., Eriksson, J., and Virtanen, A. (2007). A multifunctional RNA recognition motif in poly(A)-specific ribonuclease with cap and poly(A) binding properties. J. Biol. Chem. *282*, 32902–32911.

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

Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010). Decapping Activators in Saccharomyces cerevisiae Act by Multiple Mechanisms. Mol. Cell *39*, 773–783.

Nojima, T., Hirose, T., Kimura, H., and Hagiwara, M. (2007). The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. J. Biol. Chem. *282*, 15645–15651.

Norbury, C.J. (2013). Cytoplasmic RNA: a case of the tail wagging the dog. Nat Rev Cancer *13*, 643–653.

Ogami, K., Hosoda, N., Funakoshi, Y., and Hoshino, S. (2014). Antiproliferative protein Tob directly regulates c-myc proto-oncogene expression through cytoplasmic polyadenylation element-binding protein CPEB. Oncogene *33*, 55–64.

Ogami, K., Chen, Y., and Manley, J. (2018). RNA Surveillance by the Nuclear RNA Exosome: Mechanisms and Significance. Non-Coding RNA *4*, 8.

Oh, N., Kim, K.M., Choe, J., and Kim, Y.K. (2007). Pioneer round of translation mediated by nuclear cap-binding proteins CBP80/20 occurs during prolonged hypoxia. FEBS Lett. *581*, 5158–5164.

Olmedo, G., Guo, H., Gregory, B.D., Nourizadeh, S.D., Aguilar-Henonin, L., Li, H., An, F., Guzman, P., and Ecker, J.R. (2006). ETHYLENE-INSENSITIVE5 encodes a 5'->3' exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. Proc. Natl. Acad. Sci. *103*, 13286–13293.

Orban, T.I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. RNA *11*, 459–469.

Ozgur, S., Chekulaeva, M., and Stoecklin, G. (2010). Human Pat1b Connects Deadenylation with mRNA Decapping and Controls the Assembly of Processing Bodies. Mol. Cell. Biol. *30*, 4308–4323.

Ozgur, S., Basquin, J., Kamenska, A., Filipowicz, W., Standart, N., and Conti, E. (2015). Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex. Cell Rep. *13*, 703–711.

Pabis, M., Neufeld, N., Steiner, M.C., Bojic, T., Shav-Tal, Y., and Neugebauer, K.M. (2013). The nuclear cap-binding complex interacts with the U4/U6·U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. RNA *19*, 1054–1063.

Park, J.-E., Yi, H., Kim, Y., Chang, H., and Kim, V.N. (2016). Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. Mol Cell *62*, 462–471.

Parker, R., and Sheth, U. (2007). P Bodies and the Control of mRNA Translation and Degradation. Mol. Cell *25*, 635–646.

Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. Nat. Struct. Mol. Biol. *11*, 121–127.

Patrick, R.M., and Browning, K.S. (2012). The eIF4F and eIFiso4F Complexes of Plants: An Evolutionary Perspective. Comp. Funct. Genomics *2012*, 1–12.

Pavlopoulou, A., Vlachakis, D., Balatsos, N.A.A., and Kossida, S. (2013). A comprehensive phylogenetic analysis of deadenylases. Evol. Bioinforma. *2013*, 491–497.

Pelechano, V., Wei, W., and Steinmetz, L.M.M. (2015). Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. Cell *161*, 1400–1412.

Le Pen, J., Di Domenico, T., Kneuss, E., Kosałka, J., Rudolph, K.L.M., Miska, E.A., Jiang, H., Leung, C., Wang, D., Morgan, M., et al. (2018). Terminal uridylyltransferases target RNA viruses as part of the innate immune system. Nat. Struct. Mol. Biol. *25*, 778–786.

Petit, A.P., Wohlbold, L., Bawankar, P., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2012). The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. Nucleic Acids Res. *40*, 11058–11072.

Picard-Jean, F., Brand, C., Tremblay-Létourneau, M., Allaire, A., Beaudoin, M.C., Boudreault, S., Duval, C., Rainville-Sirois, J., Robert, F., Pelletier, J., et al. (2018). 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. PLoS One *13*, e0193804.

Pierron, G., and Weil, D. (2018). Re-viewing the 3D Organization of mRNPs. Mol. Cell *72*, 603–605.

Pirouz, M., Du, P., Munafò, M., and Gregory, R.I. (2016). Dis3l2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs. Cell Rep. *16*, 1861–1873.

Pisacane, P., and Halic, M. (2017). Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi. Nat. Commun.

Potuschak, T., Vansiri, A., Binder, B.M., Lechner, E., Vierstra, R.D., and Genschik, P. (2006). The Exoribonuclease XRN4 Is a Component of the Ethylene Response Pathway in Arabidopsis . Plant Cell *18*, 3047–3057.

Raczynska, K.D., Stepien, A., Kierzkowski, D., Kalak, M., Bajczyk, M., McNicol, J., Simpson, C.G., Szweykowska-Kulinska, Z., Brown, J.W.S., and Jarmolowski, A. (2014). The SERRATE protein is involved in alternative splicing in Arabidopsis thaliana. Nucleic Acids Res. *42*, 1224–1244.

Radhakrishnan, A., Chen, Y.H., Martin, S., Alhusaini, N., Green, R., and Coller, J. (2016). The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. Cell *167*, 122-132.e9.

Rajappa-Titu, L., Suematsu, T., Munoz-Tello, P., Long, M., Demir, Ö., Cheng, K.J., Stagno, J.R., Luecke, H., Amaro, R.E., Aphasizheva, I., et al. (2016). RNA Editing TUTase 1: structural foundation of substrate recognition, complex interactions and drug targeting. Nucleic Acids Res. *44*, gkw917.

Rajecka, V., Skalicky, T., and Vanacova, S. (2019). The role of RNA adenosine demethylases in the control of gene expression. Biochim. Biophys. Acta - Gene Regul. Mech. *1862*, 343–355.

Rajyaguru, P., and Parker, R. (2012). RGG motif proteins : Modulatirs if mRNA functional states. Cell Cycle *11*.

Rajyaguru, P., She, M., and Parker, R. (2012). Scd6 Targets eIF4G to Repress Translation : RGG Motif Proteins as a Class of eIF4G-Binding Proteins. Mol. Cell *45*, 244–254.

Ramanathan, A., Robb, G.B., and Chan, S.H. (2016). mRNA capping: Biological functions and applications. Nucleic Acids Res. *44*, 7511–7526.

Read, L.K., Zimmer, S.L., and Ammerman, M.L. (2011). Marked for Translation: Long A/U Tails as an Interface between Completion of RNA Editing and Ribosome Recruitment. Mol. Cell *42*, 6–8.

Rebbapragada, I., and Lykke-Andersen, J. (2009). Execution of nonsense-mediated mRNA decay: what defines a substrate? Curr. Opin. Cell Biol. *21*, 394–402.

Reimão-Pinto, M.M., Manzenreither, R.A., Burkard, T.R., Sledz, P., Jinek, M., Mechtler, K., and Ameres, S.L. (2016). Molecular basis for cytoplasmic RNA surveillance by uridylation-triggered decay in *Drosophila*. EMBO J. *35*, e201695164.

Ren, G., Chen, X., and Yu, B. (2012). Uridylation of miRNAs by HEN1 SUPPRESSOR1 in arabidopsis. Curr. Biol. *22*, 695–700.

Ren, G., Xie, M., Zhang, S., Vinovskis, C., Chen, X., and Yu, B. (2014). Methylation protects microRNAs from an AGO1-associated activity that uridylates 5' RNA fragments generated by AGO1 cleavage. Proc.

Natl. Acad. Sci. U. S. A. *111*, 6365–6370.

Reverdatto, S. V, Dutko, J.A., Chekanova, J.A., Hamilton, D.A., and Belostotsky, D.A. (2004). mRNA deadenylation by PARN is essential for embryogenesis in higher plants. RNA *10*, 1200–1214.

Ries, R.J., Zaccara, S., Klein, P., Olarerin-George, A., Namkoong, S., Pickering, B.F., Patil, D.P., Kwak, H., Lee, J.H., and Jaffrey, S.R. (2019). m6A enhances the phase separation potential of mRNA. Nature.

Rissland, O.S., and Norbury, C.J. (2009). Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. TL - 16. Nat. Struct. Mol. Biol. *16 VN-r*, 616–623.

Rodrigues, J.P., Rode, M., Gatfield, D., Blencowe, B.J., Carmo-Fonseca, M., and Izaurralde, E. (2012). REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. Proc. Natl. Acad. Sci. *98*, 1030–1035.

Rohayem, J., Robel, I., Jager, K., Scheffler, U., and Rudolph, W. (2006). Protein-Primed and De Novo Initiation of RNA Synthesis by Norovirus 3Dpol. J. Virol. *80*, 7060–7069.

Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. Cell *169*, 1187–1200.

Roy, D., and Rajyaguru, P.I. (2018). Suppressor of clathrin deficiency (Scd6)—An emerging RGG-motif translation repressor. Wiley Interdiscip. Rev. RNA *9*, e1479.

Rymarquis, L.A., Souret, F.F., and Green, P.J. (2011). Evidence that XRN4, an Arabidopsis homolog of exoribonuclease XRN1, preferentially impacts transcripts with certain sequences or in particular functional categories. RNA *17*, 501–511.

Sachs, A.B., and Deardorff, J.A. (1992). Translation initiation requires the PAB-dependent poly(A) ribonuclease in yeast. Cell *70*, 961–973.

Sachs, A.B., Davis, R.W., and Kornberg, R.D. (1987). A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. Mol. Cell. Biol. *7*, 3268–3276.

Sarowar, S., Oh, H.W., Cho, H.S., Baek, K.H., Seong, E.S., Joung, Y.H., Choi, G.J., Lee, S., and Choi, D. (2007). Capsicum annuum CCR4-associated factor CaCAF1 is necessary for plant development and defence response. Plant J. *51*, 792–802.

Sawazaki, R., Imai, S., Yokogawa, M., Hosoda, N., Hoshino, S.I., Mio, M., Mio, K., Shimada, I., and Osawa, M. (2018). Characterization of the multimeric structure of poly(A)-binding protein on a poly(A) tail. Sci. Rep. *8*, 1455.

Schäfer, I.B., Yamashita, M., Schuller, J.M., Schüssler, S., Reichelt, P., Strauss, M., and Conti, E. (2019). Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase. Cell 1619–1631.

Scheer, H. (2018) Rôle complexe de l'uridylation dans le métabolisme des ARNm chez Arabidopsis thaliana (doctoral dissertation). University of Strasbourg, Strasbourg, France.

Scheer, H., Zuber, H., De Almeida, C., and Gagliardi, D. (2016). Uridylation Earmarks mRNAs for Degradation… and More. Trends Genet. *32*, 607–619.

Schmid, M., Poulsen, M.B., Olszewski, P., Pelechano, V., Saguez, C., Gupta, I., Steinmetz, L.M., Moore, C., and Jensen, T.H. (2012). Rrp6p Controls mRNA Poly(A) Tail Length and Its Decoration with Poly(A) Binding Proteins. Mol. Cell *47*, 267–280.

Schmidt, M.J., and Norbury, C.J. (2010). Polyadenylation and beyond: Emerging roles for noncanonical poly(A) polymerases. Wiley Interdiscip. Rev. RNA *1*, 142–151.

Schmidt, C., Kowalinski, E., Shanmuganathan, V., Defenouillère, Q., Braunger, K., Heuer, A., Pech, M., Namane, A., Berninghausen, O., Fromont-Racine, M., et al. (2016). The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. Science (80-.). *354*, 1431–1433.

Schneider, C., and Tollervey, D. (2013). Threading the barrel of the RNA exosome. Trends Biochem. Sci. *38*, 485–493.

Schulze, W.M., Stein, F., Rettel, M., Nanao, M., and Cusack, S. (2018). Structural analysis of human ARS2 as a platform for co-transcriptional RNA sorting. Nat. Commun. *9*.

Schutz, S., Noldeke, E.R., and Sprangers, R. (2017). A synergistic network of interactions promotes the formation of in vitro processing bodies and protects mRNA against decapping. Nucleic Acids Res. *45*, 6911–6922.

Scott, D.D., and Norbury, C.J. (2013). RNA decay via 3' uridylation. Biochim. Biophys. Acta - Gene Regul. Mech. *1829*, 654–665.

Sement, F.M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J.-M.J.-M., Bousquet-Antonelli, C., Lange, H., and Gagliardi, D. (2013). Uridylation prevents 3' trimming of oligoadenylated mRNAs. Nucleic Acids Res. *41*, 7115–7127.

Sen, R., Barman, P., Kaja, A., Ferdoush, J., Lahudkar, S., Roy, A., and Bhaumik, S.R. (2019). Distinct Functions of the Cap-Binding Complex in Stimulation of Nuclear mRNA Export. Mol. Cell. Biol. *39*.

Sharif, H., and Conti, E. (2013). Architecture of the Lsm1-7-Pat1 Complex: A Conserved Assembly in Eukaryotic mRNA Turnover. Cell Rep. *5*, 283–291.

Shen, Y., Ji, G., Haas, B.J., Wu, X., Zheng, J., Reese, G.J., and Li, Q.Q. (2008a). Genome level analysis of rice mRNA 3'-end processing signals and alternative polyadenylation. Nucleic Acids Res. *36*, 3150–3161.

Shen, Y., Liu, Y., Liu, L., Liang, C., and Li, Q.Q. (2008b). Unique features of nuclear mRNA poly(A) signals and alternative polyadenylation in Chlamydomonas reinhardtii. Genetics *179*, 167–176.

Shirafuji, H., Kida, M., Asako, T., and Yoneda, M. (1979). Nucleotidylation of aminoglycoside antibiotics by bacillus brevis. Agric. Biol. Chem. *43*, 2579–2584.

Shirai, Y.T., Suzuki, T., Morita, M., Takahashi, A., and Yamamoto, T. (2014). Multifunctional roles of the mammalian CCR4-NOT complex in physiological phenomena. Front. Genet. *5*.

Shoemaker, C.J., and Green, R. (2012). Translation drives mRNA quality control. Nat. Struct. Mol. Biol. *19*, 594–601.

Shukla, S., and Parker, R. (2017). PARN Modulates Y RNA Stability and Its 3'-End Formation. Mol. Cell. Biol. *37*, e00264-17.

Shukla, S., Schmidt, J.C., Goldfarb, K.C., Cech, T.R., and Parker, R. (2016). Inhibition of telomerase RNA decay rescues telomerase deficiency caused by dyskerin or PARN defects. Nat. Struct. Mol. Biol. *23*, 286–292.

Shukla, S., Bjerke, G.A., Muhlrad, D., Yi, R., and Parker, R. (2019). The RNase PARN Controls the Levels of Specific miRNAs that Contribute to p53 Regulation. Mol. Cell *73*, 1204-1216.e4.

Siddiqui, N., Mangus, D.A., Chang, T.C., Palermino, J.M., Shyu, A. Bin, and Gehring, K. (2007). Poly(A) nuclease interacts with the C-terminal domain of polyadenylate-binding protein domain from poly(A)-binding protein. J. Biol. Chem. *282*, 25067–25075.

Sikorska, N., Zuber, H., Gobert, A., Lange, H., and Gagliardi, D. (2017). RNA degradation by the plant RNA exosome involves both phosphorolytic and hydrolytic activities. Nat. Commun. *8*, 2162.

Sikorski, P.J., Zuber, H., Philippe, L., Sement, F.M., Canaday, J., Kufel, J., Gagliardi, D., and Lange, H. (2015). Distinct 18S rRNA precursors are targets of the exosome complex, the exoribonuclease RRP6L2 and the terminal nucleotidyltransferase TRL in Arabidopsis thaliana. Plant J. *83*, 991–1004.

Siwaszek, A., Ukleja, M., and Dziembowski, A. (2014). Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. RNA Biol. *11*, 1122–1136.

Sladic, R.T., Lagnado, C.A., Bagley, C.J., and Goodall, G.J. (2004). Human PABP binds AU-rich RNA via RNA-binding domains 3 and 4. Eur. J. Biochem. *271*, 450–457.

Slomovic, S., Fremder, E., Staals, R.H.G., Pruijn, G.J.M., and Schuster, G. (2010). Addition of poly(A) and poly(A)-rich tails during RNA degradation in the cytoplasm of human cells. Proc. Natl. Acad. Sci. *107*, 7407–7412.

Smith, B.L., Gallie, D.R., Le, H., and Hansma, P.K. (1997). Visualization of poly(A)-binding protein complex formation with poly(A) RNA using atomic force microscopy. J. Struct. Biol. *119*, 109–117.

Son, A., Park, J.E., and Kim, V.N. (2018). PARN and TOE1 Constitute a 3' End Maturation Module for Nuclear Non-coding RNAs. Cell Rep. *23*, 888–898.

Song, J., and Yi, C. (2017). Chemical Modifications to RNA: A New Layer of Gene Expression Regulation. ACS Chem. Biol. *12*, 316–325.

Song, M.-G.G.M.-G., and Kiledjian, M. (2007). 3' Terminal oligo U-tract-mediated stimulation of decapping. RNA *13*, 2356–2365.

Song, J.B., Song, J., Mo, B.X., and Chen, X.M. (2015). Uridylation and adenylation of RNAs. Sci. China Life Sci. *58*, 1057–1066.

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

Souret, F.F., Kastenmayer, J.P., and Green, P.J. (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. Mol. Cell *15*, 173–183.

Staals, R.H.J., Bronkhorst, A.W., Schilders, G., Slomovic, S., Schuster, G., Heck, A.J.R., Raijmakers, R., and Pruijn, G.J.M. (2010). Dis3-like 1: A novel exoribonuclease associated with the human exosome. EMBO J. *29*, 2358–2367.

Standart, N., and Weil, D. (2018). P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. Trends Genet. *34*, 612–626.

Steiger, M., Carr-Schmid, A., Schwartz, D.C., Kiledjian, M., and Parker, R. (2003). Analysis of recombinant yeast decapping enzyme. RNA *9*, 231–238.

Stoecklin, G., and Kedersha, N. (2013). Relationship of GW/P-bodies with stress granules. Adv. Exp. Med. Biol. *768*, 197–211.

Stubblefield, J.J., Terrien, J., and Green, C.B. (2012). Nocturnin: At the crossroads of clocks and metabolism. Trends Endocrinol. Metab. *23*, 326–333.

Stubblefield, J.J., Gao, P., Kilaru, G., Mukadam, B., Terrien, J., and Green, C.B. (2018). Temporal Control of Metabolic Amplitude by Nocturnin. Cell Rep. *22*, 1225–1235.

Stupfler, B., Birck, C., Séraphin, B., and Mauxion, F. (2016). BTG2 bridges PABPC1 RNA-binding domains and CAF1 deadenylase to control cell proliferation. Nat. Commun. *7*, 10811.

Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly(A)-tail profiling reveals

an embryonic switch in translational control. Nature *508*, 66–71.

Suzuki, Y., Arae, T., Green, P.J., Yamaguchi, J., and 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 Physiol. *56*, 863–874.

Swisher, K.D., and Parker, R. (2011). Interactions between Upf1 and the decapping factors Edc3 and pat1 in saccharomyces cerevisiae. PLoS One *6*, e26547.

Tang, T.T.L., Stowell, J.A.W., Hill, C.H., and Passmore, L.A. (2019). The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. Nat. Struct. Mol. Biol. *26*, 433–442.

Tang, W., Tu, S., Lee, H.C., Weng, Z., and Mello, C.C. (2016). The RNase PARN-1 Trims piRNA 3' Ends to Promote Transcriptome Surveillance in C. elegans. Cell *164*, 974–984.

Tarun, S.Z., and Sachs, A.B. (1995). A common function for mRNA 5' and 3' ends in translation initiation in yeast. Genes Dev. *9*, 2997–3007.

Teixeira, D., and Parker, R. (2007). Analysis of P-Body Assembly in Saccharomyces cerevisiae. Mol. Biol. Cell *18*, 2274–2287.

Teixeira, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M., and Parker, R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. RNA *11*, 371–382.

Temme, C., Zhang, L., Kremmer, E., Ihling, C., Chartier, A., Sinz, A., Simonelig, M., and Wahle, E. (2010). Subunits of the Drosophila CCR4-NOT complex and their roles in mRNA deadenylation. RNA *16*, 1356–1370.

Tharun, S., and Parker, R. (2001). 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.

Tharun, S., Tharun, S., He, W., Mayes, A.E., Mayes, A.E., Lennertz, P., Lennertz, P., Beggs, J.D., and Parker, R. (2000). Yeast sm-like proteins function in mRNA decapping and decay. Nature *404*, 515–518.

Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Drazkowska, K., Pastula, A., Andersen, J.S., Stepien, P.P., et al. (2010). The human core exosome interacts with differentially localized processive RNases: HDIS3 and hDIS3L. EMBO J. *29*, 2342–2357.

Towler, B.P., Jones, C.I., Harper, K.L., Waldron, J.A., and Newbury, S.F. (2016). A novel role for the 3'-5' exoribonuclease Dis3L2 in controlling cell proliferation and tissue growth. RNA Biol.

Tseng, C.K., Wang, H.F., Schroeder, M.R., and Baumann, P. (2018). The H/ACA complex disrupts triplex in hTR precursor to permit processing by RRP6 and PARN. Nat. Commun. *9*, 1–12.

Tu, B., Liu, L., Xu, C., Zhai, J., Li, S.S., Lopez, M.A., Zhao, Y., Yu, Y., Ramachandran, V., Ren, G., et al. (2015). Distinct and Cooperative Activities of HESO1 and URT1 Nucleotidyl Transferases in MicroRNA Turnover in Arabidopsis. PLoS Genet. *11*, e1005119.

Tudek, A., Lloret-Llinares, M., and Heick Jensen, T. (2018). The multitasking polyA tail: Nuclear RNA maturation, degradation and export. Philos. Trans. R. Soc. B Biol. Sci. *373*.

Tutucci, E., Vera, M., Biswas, J., Garcia, J., Parker, R., and Singer, R.H. (2018). An improved MS2 system for accurate reporting of the mRNA life cycle. Nat. Methods *15*, 81–89.

Uchida, N., Hoshino, S.I., and Katada, T. (2004). Identification of a Human Cytoplasmic Poly(A) Nuclease Complex Stimulated by Poly(A)-binding Protein. J. Biol. Chem. *279*, 1383–1391.

Ustianenko, D., Hrossova, D., Potesil, D., Chalupnikova, K., Hrazdilova, K., Pachernik, J., Cetkovska, K.,

Uldrijan, S., Zdrahal, Z., and Vanacova, S. (2013). Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. RNA *19*, 1632–1638.

Ustianenko, D., Pasulka, J., Feketova, Z., Bednarik, L., Zigackova, D., Fortova, A., Zavolan, M., and Vanacova, S. (2016). TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. EMBO J. *35*, 2179–2191.

Ustyantsev, I.G., Golubchikova, J.S., Borodulina, O.R., and Kramerov, D.A. (2017). Canonical and noncanonical RNA polyadenylation. Mol. Biol. *51*, 262–273.

Vanacova, S., Stefl, R., and Stef, R. (2007). The exosome and RNA quality control in the nucleus (European Molecular Biology Organization).

Vaňáčová, Š., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol. *3*, 0986–0997.

Villanyi, Z., and Collart, M.A. (2016). Building on the Ccr4-Not architecture. BioEssays *38*, 997–1002.

Vindry, C., Marnef, A., Broomhead, H., Twyffels, L., Ozgur, S., Stoecklin, G., Llorian, M., Smith, C.W., Mata, J., Weil, D., et al. (2017). Dual RNA Processing Roles of Pat1b via Cytoplasmic Lsm1-7 and Nuclear Lsm2-8 Complexes. Cell Rep. *20*, 1187–1200.

Virtanen, A., Henriksson, N., Nilsson, P., and Nissbeck, M. (2013). Poly(A)-specific ribonuclease (PARN): An allosterically regulated, processive and mRNA cap-interacting deadenylase. Crit. Rev. Biochem. Mol. Biol. *48*, 192–209.

Viswanathan, P., Chen, J., Chiang, Y.C., and Denis, C.L. (2003). Identification of multiple RNA features that influence CCR4 deadenylation activity. J. Biol. Chem. *278*, 14949–14955.

Wahle, E. (1991). A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. Cell *66*, 759–768.

Wahle, E., and Winkler, G.S. (2013). RNA decay machines: Deadenylation by the Ccr4-Not and Pan2-Pan3 complexes. Biochim. Biophys. Acta - Gene Regul. Mech. *1829*, 561–570.

Walley, J.W., Coughlan, S., Hudson, M.E., Covington, M.F., Kaspi, R., Banu, G., Harmer, S.L., and Dehesh, K. (2007). Mechanical stress induces biotic and abiotic stress responses via a novel cis-element. PLoS Genet. *3*, 1800–1812.

Walley, J.W., Kelley, D.R., Nestorova, G., Hirschberg, D.L., and Dehesh, K. (2010). Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. Plant Physiol. *152*, 866–875.

Wang, X., Zhang, S., Dou, Y., Zhang, C., Chen, X., Yu, B., and Ren, G. (2015). Synergistic and Independent Actions of Multiple Terminal Nucleotidyl Transferases in the 3' Tailing of Small RNAs in Arabidopsis. PLOS Genet. *11*, e1005091.

Wang, Z., Day, N., Trifillis, P., and Kiledjian, M. (1999). An mRNA Stability Complex Functions with Poly(A)-Binding Protein To Stabilize mRNA In Vitro. Mol. Cell. Biol. *19*, 4552–4560.

Warkocki, Z., Liudkovska, V., Gewartowska, O., Mroczek, S., and Dziembowski, A. (2018a). Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism. Philos. Trans. R. Soc. B Biol. Sci. *373*.

Warkocki, Z., Krawczyk, P.S., Adamska, D., Bijata, K., Garcia-Perez, J.L., and Dziembowski, A. (2018b). Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s. Cell *174*, 1537-1548.e29.

Webster, M.W., Stowell, J.A.W., Tang, T.T.L., and Passmore, L.A. (2017). Analysis of mRNA deadenylation by multi-protein complexes. Methods *126*, 95–104.

Webster, M.W., Stowell, J.A.W., and Passmore, L.A. (2018a). The mechanism of Ccr4-Not recruitment to specific mRNAs involves sequence-selective tethering by RNA-binding proteins. BioRxiv 478008.

Webster, M.W., Chen, Y.H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Coller, J., and Passmore, L.A. (2018b). mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. Mol. Cell *70*, 1089-1100.e8.

Wei, C.M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5′ terminus of HeLa cell messenger RNA. Cell *4*, 379–386.

Wells, S.E., Hillner, P.E., Vale, R.D., and Sachs, A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. Mol. Cell *2*, 135–140.

Werner, M., Purta, E., Kaminska, K.H., Cymerman, I.A., Campbell, D.A., Mittra, B., Zamudio, J.R., Sturm, N.R., Jaworski, J., and Bujnicki, J.M. (2011). 2′-O-ribose methylation of cap2 in human: Function and evolution in a horizontally mobile family. Nucleic Acids Res. *39*, 4756–4768.

West, S., Gromak, N., Norbury, C.J., and Proudfoot, N.J. (2006). Adenylation and exosome-mediated degradation of cotranscriptionally cleaved pre-messenger RNA in human cells. Mol. Cell *21*, 437–443.

Wolf, J., and Passmore, L.A. (2014). mRNA deadenylation by Pan2–Pan3. Biochem. Soc. Trans. *42*, 184–187.

Wolf, J., Valkov, E., Allen, M.D., Meineke, B., Gordiyenko, Y., McLaughlin, S.H., Olsen, T.M., Robinson, C. V, Bycroft, M., Stewart, M., et al. (2014). Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. EMBO J. *33*, 1514–1526.

Woo, Y.M., Kwak, Y., Namkoong, S., Kristjánsdóttir, K., Lee, S.H., Lee, J.H., and Kwak, H. (2018). TED-Seq Identifies the Dynamics of Poly(A) Length during ER Stress. Cell Rep. *24*, 3630-3641.e7.

Wu, M., Reuter, M., Lilie, H., Liu, Y., Wahle, E., and Song, H. (2005). Structural insight into poly(A) binding and catalytic mechanism of human PARN. EMBO J. *24*, 4082–4093.

Wu, X., Liu, M., Downie, B., Liang, C., Ji, G., Li, Q.Q., and Hunt, A.G. (2011). Genome-wide landscape of polyadenylation in Arabidopsis provides evidence for extensive alternative polyadenylation. Proc. Natl. Acad. Sci. *108*, 12533–12538.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.C., Dufour, M.E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., et al. (2005). Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell *121*, 725–737.

Xie, J., Kozlov, G., and Gehring, K. (2014). The “tale” of poly(A) binding protein: The MLLE domain and PAM2-containing proteins. Biochim. Biophys. Acta - Gene Regul. Mech. *1839*, 1062–1068.

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

Xu, K., Bai, Y., Zhang, A., Zhang, Q., and Bartlam, M.G. (2014). Insights into the structure and architecture of the CCR4-NOT complex. Front. Genet. *5*, 1–12.

Yamashita, S., Takagi, Y., Nagaike, T., and Tomita, K. (2017). Crystal structures of U6 snRNA-specific terminal uridylyltransferase. Nat. Commun. *8*, 15788.

Yan, Y. Bin (2014). Deadenylation: enzymes, regulation, and functional implications. Wiley Interdiscip. Rev. RNA *5*, 421–443.

Yang, A., Bofill-De Ros, X., Shao, T.-J., Jiang, M., Li, K., Villanueva, P., Dai, L., and Gu, S. (2019). 3′ Uridylation Confers miRNAs with Non-canonical Target Repertoires. Mol. Cell 1–12.

Yashiro, Y., and Tomita, K. (2018). Function and regulation of human terminal uridylyltransferases. Front. Genet. *9*, 538.

Yates, L.A., Fleurdépine, S., Rissland, O.S., De Colibus, L., Harlos, K., Norbury, C.J., and Gilbert, R.J.C. (2012). Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. Nat. Struct. Mol. Biol. *19*, 782–787.

Yi, H., Park, J., Ha, M., Lim, J., Chang, H., and Kim, V.N. (2018). PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. Mol. Cell *70*, 1081-1088.e5.

Yu, X., Willmann, M.R., Anderson, S.J., and Gregory, B.D. (2016). Genome-Wide Mapping of Uncapped and Cleaved Transcripts Reveals a Role for the Nuclear mRNA Cap-Binding Complex in Cotranslational RNA Decay in Arabidopsis. Plant Cell *28*, 2385–2397.

Zakrzewska-Placzek, M., Souret, F.F., Sobczyk, G.J., Green, P.J., and Kufel, J. (2010). Arabidopsis thaliana XRN2 is required for primary cleavage in the pre-ribosomal RNA. Nucleic Acids Res. *38*, 4487–4502.

Zeidan, Q., He, F., Zhang, F., Zhang, H., Jacobson, A., and Hinnebusch, A.G. (2018). Conserved mRNA-granule component Scd6 targets Dhh1 to repress translation initiation and activates Dcp2-mediated mRNA decay in vivo.

Zhang, E., Khanna, V., Dacheux, E., Namane, A., Doyen, A., Gomard, M., Turcotte, B., Jacquier, A., and Fromont-Racine, M. (2019). A specialised SKI complex assists the cytoplasmic RNA exosome in the absence of direct association with ribosomes. EMBO J. e100640.

Zhang, J.Q., He, G.J., and Yan, Y. Bin (2013). Biochemical and Biophysical Characterization of the Deadenylase CrCaf1 from Chlamydomonas reinhardtii. PLoS One *8*, e69582.

Zhang, W., Murphy, C., and Sieburth, L.E. (2010). Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. Proc. Natl. Acad. Sci. *107*, 15981–15985.

Zhang, X., Devany, E., Murphy, M.R., Glazman, G., Persaud, M., and Kleiman, F.E. (2015). PARN deadenylase is involved in miRNA-dependent degradation of TP53 mRNA in mammalian cells. Nucleic Acids Res. *43*, 10925–10938.

Zhang, X., Koiwa, H., Hu, F., Sung, M.W., Tang, G., Castillo-González, C., Li, P., Zhang, Z., Shu, C., and Dickman, M. (2017a). RISC-interacting clearing 3'- 5' exoribonucleases (RICEs) degrade uridylated cleavage fragments to maintain functional RISC in Arabidopsis thaliana. Elife *6*, 1–29.

Zhang, Z., Hu, F., Sung, M.W., Shu, C., Castillo-González, C., Koiwa, H., Tang, G., Dickman, M., Li, P., Zhang, X., et al. (2017b). RISC-interacting clearing 3'- 5' exoribonucleases (RICES) degrade uridylated cleavage fragments to maintain functional RISC in Arabidopsis thaliana. Elife *6*, 1–29.

Zhao, H., Xing, D., and Li, Q.Q. (2009). Unique Features of Plant Cleavage and Polyadenylation Specificity Factor Revealed by Proteomic Studies. PLANT Physiol. *151*, 1546–1556.

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B., and Chen, X. (2012a). The arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. Curr. Biol. *22*, 689–694.

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., and Theresa, T. (2012b). HESO1, a nucleotidyl transferase in Arabidopsis, uridylates unmethylated miRNAs and siRNAs to trigger their degradation. Curr Biol. *22*, 689–694.

Zhou, L., Hang, J., Zhou, Y., Wan, R., Lu, G., Yin, P., Yan, C., and Shi, Y. (2014). Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. Nature *506*, 116–120.

Zhu, Y., and Vaughn, J.C. (2018). Experimental Verification and Evolutionary Origin of 5'-UTR Polyadenylation Sites in Arabidopsis thaliana. Front. Plant Sci. *9*, 969.

Zigáčková, D., and Vaňáčová, Š. (2018). The role of 3' end uridylation in RNA metabolism and cellular physiology. Philos. Trans. R. Soc. B Biol. Sci. *373*.

Zimmer, S.L., Schein, A., Zipor, G., Stern, D.B., and Schuster, G. (2009). Polyadenylation in Arabidopsis and Chlamydomonas organelles: The input of nucleotidyltransferases, poly(A) polymerases and polynucleotide phosphorylase. Plant J. *59*, 88–89.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev. *31*, 88–100.

Zuber, H., Scheer, H., Ferrier, E., Sement, F.M., Mercier, P., Stupfler, B., and Gagliardi, D. (2016). Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in Arabidopsis. Cell Rep. *14*, 2707–2717.

Zuber, H., Scheer, H., Joly, A.-C., and Gagliardi, D. (2018). Respective Contributions of URT1 and HESO1 to the Uridylation of 5' Fragments Produced From RISC-Cleaved mRNAs. Front. Plant Sci. *9*, 1438.

Zuo, Y. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res. *29*, 1017–1026.

Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., et al. (2011). Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. Nat. Immunol. *12*, 137–143.

Influence de l'uridylation sur la déadénylation des ARN messagers chez *Arabidopsis thaliana*.

Introduction

L'uridylation des ARN est une modification post-transcriptionnelle très courante qui a un rôle fondamental dans la régulation de l'expression des gènes. L'uridylation est conservée dans une grande variété d'eucaryotes, à l'exception de la levure *Saccharomyces cerevisiae*, et consiste en l'ajout d'uridines aux extrémités 3' des ARN messagers et ARN non-codants. De plus, l'uridylation a été détectée de manière extensive sur divers ARN viraux infectant des champignons, plantes et animaux. L'ajout d'uridines est catalysée par des terminal uridylytransférases (TUTases) et peut influencer positivement ou négativement la destinée d'un ARN, en fonction de la nature du transcrit ciblé, de la processivité de la TUTase impliquée et de l'environnement cellulaire. La fonction principale décrite pour l'uridylation est la stimulation de la dégradation des ARN messagers et des ARN non-codants matures et de leurs précurseurs (De Almeida et al., 2018; Aphasizhev et al., 2016; Łabno et al., 2016; Lim et al., 2014; Munoz-tello et al., 2015; Scheer et al., 2016; Scott and Norbury, 2013; Ustianenko et al., 2016; Zigáčková and Vaňáčová, 2018). Chez l'homme, *Schizosaccharomyces pombe* et *Aspergillus nidulans*, l'uridylation des ARNm déadénylés permet notamment le recrutement de LSm1-7, un complexe qui se lie aux queues polyA courtes et qui active l'enlèvement de la coiffe et la dégradation de 5' en 3'. Alternativement, l'uridylation a été montrée comme pouvant rapidement mener à la dégradation de 3' en 5' par Dis3L2 ou l'exosome.

Les membres de la famille TUTase sont composés d'un domaine catalytique (CCD) (Martin and Keller, 2007) et d'éventuels domaines supplémentaires qui peuvent influencer la reconnaissance des substrats, leurs activités et le recrutement de facteurs auxiliaires (De Almeida et al., 2018; Warkocki et al., 2018a; Yashiro and Tomita, 2018; Zigáčková and Vaňáčová, 2018). Un trait fréquent des TUTases est la présence de régions intrinsèquement désorganisées (IDRs) qui sont pauvres en acides aminés hydrophobes et qui n'ont pas de structure secondaire (Dyson, 2016; Järvelin et al., 2016). Ces IDRs comportent souvent des motifs linéaires courts qui peuvent entre autres assurer le recrutement de protéines partenaires aux TUTases (van der Lee et al., 2014). La TUTase URT1, qui uridyle principalement les ARNm chez *Arabidopsis thaliana,* contient une telle IDR dans sa région N-terminale. Cette partie de la protéine a été montrée par

l'équipe comme étant impliquée dans le recrutement de DCP5 via un motif conservé M1. DCP5 et ses orthologues Scd6 chez la levure et LSM14 chez l'homme interagissent avec le complexe d'élimination de la coiffe DCP1/2 (Barbee et al., 2006; Nissan et al., 2010; Xu and Chua, 2009). Ces facteurs DCP5/Scd6/LSM14 se comportent également comme des inhibiteurs de la traduction (Ayache et al., 2015; Rajyaguru et al., 2012; Zeidan et al., 2018). Une interaction entre URT1 et DCP5 pourrait donc constituer un lien moléculaire enter uridylation et élimination de la coiffe ou inhibition de traduction chez Arabidopsis.

L'équipe de recherche dans laquelle j'ai effectué ma thèse a montré que chez *Arabidopsis thaliana*, une fonction supplémentaire peut être associée à l'uridylation des ARN messagers. En effet, la TUTase URT1 est capable de réparer les extrémités 3' des ARN messagers déadénylés par l'ajout de quelques uridines pour restaurer un site de fixation pour une protéine de liaison aux queues polyA (PABP) (Zuber et al., 2016). Cette activité prévient la production d'ARNm excessivement déadénylés, ce qui permet de déplacer la polarité de la dégradation de 5' en 3'. En outre, l'uridylation par URT1 pourrait avoir un rôle dans le stockage ou dans l'inhibition de la traduction des ARN messager uridylés. Il existe une deuxième TUTase chez Arabidopsis, HESO1, qui uridyle principalement les ARN non-codants pour accélérer leurs dégradations (Ren et al., 2012; Tu et al., 2015; Wang et al., 2015; Zhao et al., 2012). De plus, HESO1 pourrait également être responsable de l'uridylation résiduelle observée en absence de URT1 (Sement et al., 2013; Zuber et al., 2016). De manière intéressante, cette uridylation résiduelle dans le mutant *urt1* ne semble pas être capable de réparer les extrémités 3' des ARNm. Le rôle moléculaire associé à l'activité d'uridylation des ARNm par HESO1 reste à être défini.

Objectifs de thèse

L'uridylation et la déadénylation sont intimement liées. Cependant, les mécanismes moléculaires qui régulent les interactions entre ces deux processus sont encore peu connus. Mon objectif de thèse était d'étudier la relation entre uridylation et déadénylation pendant la dégradation des ARN messagers, afin de mieux comprendre le rôle moléculaire de la TUTase URT1. J'ai également réalisé des premières expériences visant à étudier la fonction de l'uridylation des ARNm par HESO1.

Afin d'analyser la taille et les modifications des queues polyA d'ARNm, j'ai participé à l'optimisation du protocole de TAIL-seq (Chang et al., 2014) visant à analyser à l'échelle du transcriptome les queues polyA des ARNm chez *Arabidopsis thaliana*. J'ai également participé au

perfectionnement de la méthode 3'RACE-seq qui consiste en l'amplification par PCR des régions 3' d'ARN messagers cibles (RACE : rapid amplification of cDNA ends) couplée au séquençage à haut débit de type Illumina. La méthode de 3'RACE-seq permet d'obtenir une profondeur importante de séquençage et d'analyser de manière extensive la taille et composition de la queue polyA pour une molécule d'ARN donnée. Grâce à cette méthode, j'ai pu analyser les queues polyA d'ARNm rapporteurs dans le système d'expression transitoire dans des feuilles de *Nicotiana benthamiana* et des ARNm endogènes dans des plantes sauvages et mutantes d'*Arabidopsis thaliana.* L'ensemble des principales expériences représentées ci-dessous ont participé à une meilleure compréhension de l'équilibre qui existe entre déadénylation et uridylation des queues polyA et des fonctions moléculaires qui régulent la taille des queues polyA chez la plante.

Résultats

- **Mise au point des méthodes TAIL-seq et 3'RACE-seq pour l'analyse des extrémités 3' des ARN messagers chez la plante.**

Durant ma thèse, j'ai participé à l'optimisation de la méthode TAIL-seq qui a initialement été développée pour l'analyse des extrémités 3' des ARNm chez l'homme et la souris (Chang et al., 2014). J'ai notamment généré un premier jeu de données TAIL-seq d'ARNm de fleurs de plantes sauvages d'accession Columbia. L'analyse des données de séquençage montre que nous avons obtenu une profondeur de séquençage suffisante pour une analyse gène-à-gène de la taille et la composition des queues polyA. 2 044 ARNm ont pu être détectés avec au moins 20 lectures (Figure 13A). Néanmoins, il faudra augmenter la couverture de séquençage et améliorer le protocole de préparation des banques TAIL-seq pour avoir des estimations de distributions de tailles de queues polyA *in vivo* plus fiables. Malgré ces limitations, ce jeu de donnée nous permet de faire une sélection d'ARNm ayant des taux d'uridylation variés et qu'il serait intéressant d'analyser par 3'RACE-seq dans différent contextes génétiques. La méthode 3'RACE-seq, quant à elle, s'est avérée être une procédure efficace pour l'analyse des tailles et compositions des queues polyA. En effet, l'amplification et le séquençage d'ARNm cibles à haut débit permet d'augmenter considérablement la couverture de séquençage et ainsi la qualité de l'analyse des extrémités 3' des ARNm (Figure 14B et C). Ceci représente une méthode performante pour l'analyse de variations minimes dans les profils de distributions des queues polyA entre

différents échantillons. Pour la suite des expériences, j'ai utilisé la méthode 3'RACE-seq pour l'analyse d'ARNm rapporteurs exprimés de manière transitoire et d'ARNm endogènes.

- **La surexpression transitoire de URT1 dans *N. benthamiana* induit un allongement important des queues polyA.**

L'utilisation de feuilles de *N. benthamiana* pour l'expression transitoire de la TUTase URT1 et de différentes constructions mutantes par la technique d'agro-infiltration nous a permis de surexprimer de manière efficace nos protéines d'intérêt et d'étudier par 3'RACE-seq l'effet de cette surexpression sur les extrémités 3' de l'ARNm endogène *NbPR2* et de l'ARNm rapporteur GFP. Les agro-infiltrations sur feuilles de tabac consistaient en la co-infiltration des différentes constructions avec le suppresseur de silencing P19 et le gène rapporteur *GFP*. Les contrôles utilisés dans ces expériences consistaient en l'analyse des extrémités 3' des ARNm *GFP* et *NbPR2* dans les patchs de feuilles agro-infiltrées seulement avec P19 et le rapporteur *GFP* (ctrl).

Le premier résultat marquant que nous avons pu observer est le changement important des profils d'uridylation et de polyadénylation pour les ARNm rapporteurs *GFP* et pour l'ARNm endogène *NbPR2* lorsque la construction active de URT1-myc est surexprimée. En effet, l'expression transitoire de URT1 résulte en une accumulation remarquable de longues queues polyA uridylées (Figure 18). De manière surprenante, la distribution des queues polyA non-uridylées est également fortement affectée, et montre une diminution des queues courtes en dessous de 25 nucléotides et une importante augmentation des queues plus longues (Figure 19). Ces phénotypes moléculaires sont liés en partie à l'activité catalytique de URT1, puisque la mutation de résidus du site actif de URT1 ($URT1^{\Delta D491/3A}$ -myc) mène à une importante diminution de queue plus longues et une augmentation des queues courtes en dessous de 25 nucléotides, similaire au profil observé dans les contrôles.

Ces premiers résultats suggèrent que l'activité de URT1 a un effet direct sur la taille des queues polyA lorsque URT1 est transitoirement surexprimée dans *N. benthamiana*. Le changement de profil a deux explications possibles qui ne sont pas nécessairement exclusives :

1- La surexpression de URT1 mène à la dégradation des queues courtes, déplaçant ainsi la distribution vers des queues plus longues.

2- La surexpression de URT1 permet l'uridylation des queues plus longues, qui pourrait inhiber l'activité des déadénylases, ayant comme conséquence la diminution de la formation de la population d'ARN à queues courtes.

Ces deux hypothèses se sont révélées fondées et sont détaillées à la suite de ce résumé.

- **Le motif M1 de URT1 est impliqué dans la régulation de l'accumulation d'ARNm avec des queues polyA courtes.**

Nous avons remarqué que dans les feuilles exprimant URT1$^{\Delta D491/3A}$-myc, la population d'ARN à queues polyA courtes (<25 nucléotides) était plus importante que celles observées dans les contrôles pour *GFP* et *NbPR2* (Figure 21 et S4). De manière intéressante, en enlevant l'IDR de URT1$^{\Delta D491/3A}$ -myc (ΔIDR$^{\Delta D491/3A}$-myc), la population d'ARN avec des queues polyA courtes diminue et le profil observé est semblable à celui des contrôles. Ces résultats suggèrent que la partie N-terminale de la protéine URT1 est requise pour la dégradation d'ARN avec des queues polyA courtes, indépendamment de son activité. Afin d'évaluer si la dégradation des ARN à queues polyA courtes dépend des motifs conservés M1 ou M2 qui sont intégrés dans la région IDR, j'ai muté les motifs M1 et M2 dans URT1$^{\Delta D491/3A}$-myc. Seule la mutation du motif M1 mène à une importante réduction de la population d'ARNm *GFP* à queues courtes et restaure le profil observé dans les contrôles (Figure 30 et 31). Ces résultats suggèrent que URT1 est impliquée dans la dégradation de cette population d'ARN via son motif M1. La surexpression de URT1$^{\Delta D491/3A}$-myc dans les feuilles de *N. benthamiana* pourrait en effet mener à une déplétion du facteur DCP5 et de la machinerie d'élimination de la coiffe. Ainsi, les ARNm avec des queues polyA courtes qui sont les substrats idéaux de cette voie de dégradation ne seront pas correctement décoiffé et dégradé, ce qui mène à une augmentation de cette population d'ARNm à queues courtes.

Ces expériences suggèrent qu'un lien moléculaire existe entre la TUTase URT1 et le complexe d'enlèvement de la coiffe via le recrutement de DCP5. Ainsi, l'uridylation d'ARNm avec des queues polyA de moins de 25 nucléotides pourrait jouer un rôle important pour assurer l'enlèvement de la coiffe et la dégradation de 5'-3' de ces ARNm.

- **La présence d'uridines en 3' des queues polyA entraine un ralentissement de l'activité de CAF1b *in vitro*.**

Afin de vérifier si l'accumulation de queues longues est due à un effet direct de l'uridylation par URT1 sur l'activité des déadénylases, j'ai surexprimé chez *Escherichia coli* et purifié des versions actives et inactives d'une des déadénylases principales de *A. thaliana*, CAF1b. Différents ARNs d'intérêt contenant quelques adénosines en 3' suivies d'aucune, d'une ou de deux uridines terminales ont été incubés en présence de la déadénylase. Les résultats de ces tests *in vitro* montrent clairement que la présence d'une seule uridine en 3' est suffisante pour ralentir

l'activité de CAF1b, et ce pour les différents réplicas testés (Figure 33). De plus, deux uridines terminales semblent davantage freiner l'activité de déadénylation.

Ces résultats révèlent que l'uridylation peut intrinsèquement freiner la dégradation des queues polyA par les déadénylases *in vitro*.

- **L'absence de URT1 dans les mutants *urt1* d'*A. thaliana* résulte en un raccourcissement excessif des queues polyA pour des ARN cibles.**

En plus des expériences *in planta* dans le système d'expression transitoire de *N. benthamiana,* j'ai préparé des banques 3'RACE-seq pour 16 ARNm d'intérêt à partir de plantes sauvages et mutantes *urt1*. L'analyse des extrémités 3' de ces ARN montre que pour toutes les cibles analysées, nous avons accumulation de queues polyA plus longues dans le sauvage par rapport au mutant (Figure 36). L'effet est d'autant plus visible que les ARNm présentent un fort taux d'uridylation. En effet, les ARNm qui sont les plus uridylés dans le sauvage montrent une importante accumulation d'ARNm avec des queues polyA excessivement déadénylés (<10 nucleotides). Ainsi, la TUTase URT1 pourrait jouer un rôle crucial pour éviter la formation de ces ARNm excessivement déadénylés, d'autant plus que cette population d'ARNm ne peut pas être détectée en condition sauvage.

Ces résultats et les expériences de surexpression transitoire menées dans *N. benthamiana* suggèrent que l'expression de la TUTase URT1 permet de freiner le raccourcissement des queues polyA afin d'éviter la formation de queues excessivement déadénylées, et ce par deux mécanismes : l'ajout d'uridines terminales freine l'activité des déadénylases *per se*, et le recrutement de DCP5 par URT1 aux queues courtes de moins de 25 nucléotides déplace la polarité de la dégradation des ARNm vers la voie 5'-3'. Mes résultats ont ainsi permis d'établir de nouveaux mécanismes moléculaires par lesquels l'uridylation régule la dégradation d'ARNm chez un eucaryote.

- **L'expression transitoire de HESO1 dans *N. benthamiana* mène à une réduction d'ARNm avec des queues polyA courtes.**

La fonction de l'uridylation des ARN non-codants par HESO1 est d'induire la dégradation rapide de ses ARN cibles. Ainsi, HESO1 pourrait également être impliqué dans la dégradation rapide des ARN messagers. En vue d'évaluer la fonction moléculaire de HESO1 dans l'uridylation des ARNm, j'ai également réalisé des expériences en utilisant le système d'expression transitoire utilisant *N. benthamiana*. Des versions actives et inactives de HESO1-myc ont été co-infiltrées

avec P19 et le rapporteur *GFP*. L'analyse par RACE-seq des extrémités 3' de l'ARNm de la *GFP* et de l'ARNm endogène *NbPR2* montrent que HESO1 n'affecte pas les queues polyA des ARNm *GFP NbPR2* de la même manière que URT1. Nous observons une diminution des queues polyA courtes (<25 nucléotides) (Figure 40), mais nous ne détectons aucune augmentation dans le pourcentage d'uridylation des queues (Figure 38). Cependant, le changement du profil est dû à l'activité d'uridylation de HESO1, puisque la mutation de résidus du site active mène à la restauration du profil de distribution des queues polyA tels qu'il est observé dans les échantillons contrôles. De manière intéressante, nous ne pouvons pas observer une augmentation importante de queues polyA longues comme lorsque URT1-myc est exprimé. Ces résultats suggèrent que l'uridylation par HESO1 mène plutôt à la dégradation rapide des ARNm cibles, les rendant difficilement détectables par 3'RACE-seq. De plus, en vue de l'importante diminution des queues polyA courtes en dessous de 25 nucléotides, HESO1 semble principalement viser la population d'ARN avec des queues courtes. En accord avec cette hypothèse, l'analyse des ARNm *GFP* et *NbPR2* par northern blot révèle que l'expression de HESO1 mène à une diminution de l'accumulation des ARNm dans les feuilles infiltrées (Figure 41). Ces résultats donnent une première appréciation des éventuelles fonctions de HESO1 dans la régulation du métabolisme de l'ARNm chez les plantes et suggèrent une fonction moléculaire différente entre URT1 et HESO1 dans la dégradation des ARNm.

Conclusion et perspectives

Mes expériences menées avec le système d'expression transitoire dans des feuilles de *N. benthamiana* et les expériences *in vitro* et *in vivo* chez Arabidopsis ont permis de mieux comprendre comment URT1 pouvait influencer la taille des queues polyA et les rôles moléculaires liés à cette uridylation. Mes résultats révèlent que URT1 peut freiner les déadénylases directement par l'ajout d'uridines en 3' des queues polyA *in vivo*. De plus, le lien moléculaire entre URT1 et l'activateur de decapping DCP5 semple être important pour la dégradation d'ARNm avec des queues polyA de 10-25 nucléotides (Figure 43). Ces rôles peuvent être décisifs pour éviter la formation d'ARN excessivement déadénylés. Ce processus revête une importance particulière sur le plan biologique car les ARNm excessivement déadénylés deviennent des substrats de la RNA polymérase RNA dépendante RDR6, menant à la formation de petits ARN dans la plante (Baeg et al., 2017). Ces siRNAs une fois produits ciblent les ARNm endogènes, conduisant à la mort de la plante. D'autre part, le ralentissement du

raccourcissement des queues polyA permet de recruter d'autres facteurs, tels que la PABP, qui pourrait avoir un rôle dans la traduction ou le stockage de l'ARN.

Mes expériences ont également permis d'avoir un premier aperçu sur le rôle éventuel de l'uridylation par HESO1 des ARNm. Contrairement à URT1, qui semble pouvoir uridyler des queues courtes et longues, HESO1 semble être spécifique pour les queues polyA courtes de moins de 25 nucléotides et son activité mène à une déstabilisation plus importante des ARNm cibles. Ces résultats suggèrent que l'uridylation par URT1 ou HESO1 a des fonctions spécifiques dans la régulation de la dégradation des ARNm chez *Arabidopsis thaliana* (Figure 42). HESO1 pourrait par exemple induire la rapide dégradation des ARNm par la voie 5'-3' grâce au recrutement de LSm1-7 sur les queues uridylées. L'uridylation par HESO1 pourrait également induire la rapide dégradation des ARNm de 3'-5' par l'exosome ou les exoribonucléases SOV, qui reconnaissent préférentiellement des ARNm ayant quelques uridines terminales en 3' des queues polyA.

L'ensemble de ces résultats montrent que l'uridylation des ARNm joue des rôles complexes permettant une dégradation efficace des ARN tout en évitant que des intermédiaires de dégradation ne deviennent des substrats de voies de silencing. Des résultats récents de l'équipe ont montré en accord avec mes résultats que URT1 serait un suppresseur de silencing et que les ARNm fortement uridylés en condition sauvage produisent une quantité importante de petits ARN dans le double mutant *urt1 xrn4*.

Les perspectives immédiates à ce projet seraient d'analyser par 3'RACE-seq des ARNm fortement uridylés dans des fond mutants *urt1* et *heso1* afin d'évaluer le rôle de HESO1 dans la régulation de la taille des queues polyA *in vivo*. Il serait également intéressant d'améliorer la profondeur de séquençage de la méthode TAIL-seq afin d'avoir de meilleures données de séquençage pour la bonne estimation des tailles et du taux d'uridylation des queues polyA des ARNm à l'échelle du transcriptome. Une alternative à ceci serait l'utilisation de méthodes alternatives qui permettrait de réduire les biais dans l'estimation de la taille des queues polyA, telles que le tail-end displacement sequencing (TED-seq) ou l'utilisation de la technologie de séquençage direct des ARN d'Oxford Nanopore Technologies (ONT). De manière intéressante, des expériences récentes de co-immunoprécipitations montrent que URT1 co-purifie avec plusieurs sous-unités du principal complexe de déadénylation CCR4/NOT chez *Arabidopsis thaliana*. Ainsi, il serait intéressant d'étudier comment la TUTase URT1 est recrutée aux ARNm lors de la déadénylation par CCR4/NOT.

To Snarky Puppy, Joe Hisaishi, Hans Zimmer, Marcus Miller, Nina Simon, Seeed, Gramatik, Ibrahim Maalouf, John Powell, Miles Davis, Paul Desmond, Jacob Collier, Knower, Harry Gregson-Williams, Bob Marley, Michael Giacchino, The Cat Empire, Stevie Wonder, Yael Naim, Ibrahim Ferrer, Cesaria Evora, James Newton Howard and many more to whom I've been listening especially (and continuously) this last 4 months: Music is the medicine of the mind!

Université de Strasbourg

Caroline de Almeida

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

Interplay between uridylation and deadenylation during mRNA degradation in *Arabidopsis thaliana*

Résumé

L'uridylation des ARN est une modification post-transcriptionnelle qui consiste en l'ajout d'uridines aux extrémités 3' des ARN codants et non-codants. La fonction principale décrite pour l'uridylation est la stimulation de la dégradation des ARN cibles. Cependant, les processus moléculaires responsables de cette stimulation sont encore peu connus. Mes travaux de thèse ont eu pour but la caractérisation de la fonction moléculaire de URT1, la principale uridylyltransferase terminale qui uridyle les ARN messagers (ARNm) chez *Arabidopsis thaliana*. Mes résultats démontrent que URT1 module le profil de polyadenylation des ARNm de deux manières. Premièrement, l'ajout d'uridines en 3' des ARNm par URT1 freine la déadénylation et entraîne l'accumulation de queues polyadénylées (polyA) plus longues. Deuxièmement, URT1 est impliqué dans la dégradation d'ARNm uridylés ayant une queue polyA courte de 10-20 nucléotides. Un motif peptidique M1 dans la région N-terminale de URT1 est impliqué dans la dégradation de cette population d'ARNm. Ce motif recrute DCP5, un activateur de l'élimination de la coiffe, qui pourrait stimuler l'élimination de la coiffe et favoriser la dégradation de 5'-3' de ces ARN cibles. Ainsi, les fonctions d'URT1 sont essentielles pour favoriser la dégradation de 5'-3' et prévenir l'accumulation délétère d'ARNm excessivement déadenylés et tronqués en 3'.

Mots-clés : uridylation, déadénylation, ARNm, TUTases, URT1, dégradation des ARN, *Arabidopsis thaliana*.

Abstract

RNA uridylation is a post-transcriptional modification of 3' ends of coding and non-coding RNAs. The best known function of uridylation is to stimulate RNA degradation, but the underlying molecular mechanisms are poorly understood. In this work, I investigate the functions of URT1, the main terminal uridylyltransferase (TUTase) that uridylates messenger RNA (mRNA) in *Arabidopsis thaliana*. My results indicate a dual role of uridylation in mRNA turnover. First, the addition of terminal uridines by URT1 directly slows down deadenylation and leads to the accumulation of mRNAs with longer polyA tails. Second, URT1 promotes the degradation of deadenylated mRNAs with tails of 10-20 nucleotides, likely by recruiting the decapping activator DCP5 via the M1 motif that is present in the N-terminal intrinsically disordered region of URT1. Both effects are essential to favour the 5'-3' polarity of mRNA degradation and to prevent the deleterious accumulation of excessively deadenylated and 3' truncated mRNAs.

Keywords: mRNA uridylation, TUTases, URT1, deadenylation, RNA degradation, *Arabidopsis thaliana*.