UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

Institut de Biologie Moléculaire des Plantes – CNRS – UPR2357

THÈSE

présentée par :

Hélène SCHEER

soutenue le : **28 septembre 2018**

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

Discipline/ Spécialité : Science de la vie et de la Santé

Rôle complexe de l'uridylation dans le métabolisme des ARNm chez *Arabidopsis thaliana*

Directeur de Thèse	Dr Dominique GAGLIARDI
Rapporteur Externe	Dr Dominique WEIL
Rapporteur Externe	Pr Jean-Marc DERAGON
Examinateur Interne	Dr Sébastien PFEFFER

Table des matières

Liste des Figures

INTRODUCTION

Introduction générale

Le contrôle quantitatif, qualitatif et spatio-temporel de l'expression des gènes permet aux cellules de s'adapter aux changements environnementaux et rend possible, pour les organismes multicellulaires, l'expression d'un grand nombre de cellules phénotypiquement distinctes à partir d'une information génétique unique. L'expression des gènes est régulée tout au long d'un continuum débutant par la synthèse d'un ARN lors de sa transcription, suivie d'étapes de maturation complexe au cours desquelles l'ARN est la cible de nombreuses modifications, de son export possible du noyau et du contrôle de sa localisation finale, de sa traduction dans le cas d'un ARN messager (ARNm) et de sa dégradation pour tous les types d'ARN. Les processus de **dégradation des ARN** opèrent à chacune de ces étapes et permettent le contrôle du niveau d'expression des ARN mais également de leur qualité. En effet, les voies de dégradation des ARN permettent d'ajuster continuellement le temps de demi-vie des ARN mais également de détecter et d'éliminer les transcrits défectueux. Contrôler la stabilité ou l'instabilité des ARNm peut également affecter leur engagement polysomal et par conséquent leur traduction rendant les processus de dégradation et de traduction des ARN étroitement liés.

Les **modifications post-transcriptionnelles** constituent l'**outil moléculaire** principal pour contrôler les processus de maturation, de dégradation et de traduction des ARN. Historiquement, les mieux décrites de ces modifications post-transcriptionnelles ont été les modifications chimiques affectant les ARN non codants (ARNnc) abondants tels que les ARN ribosomiques (ARNr) et les ARN de transferts (ARNt). Nos connaissances sur les modifications des ARN codants se résumaient quant à elles au coiffage et la polyadénylation canonique des ARNm. L'amélioration des techniques de détection ont, depuis lors, permis la cartographie et la découverte de nouvelles formes de modifications et démontré de l'étendu de ces dernières. La fonctionnalité d'un ARN n'est plus exclusivement dépendant de l'information contenue dans sa séquence mais est façonnée par les modifications post-transcriptionnelles auxquelles il peut être soumis. A l'instar des marques épigénétiques qui modulent la structure et l'activité de la chromatine, les modifications façonnent les ARN après transcription, de cette dernière analogie découle la notion d'**épitranscriptomique** (Wang and He, 2014; Roundtree *et al.*, 2017; Roundtree and He, 2016; Helm and Motorin, 2017; Schaefer *et al.*, 2017).

Les modifications se distinguent en deux classes principales ; **(1)** les modifications chimiques des nucléosides. Dans cette première catégorie on compte des modifications telles que la méthylation des adénosines (N6-methyladenosine (m6A) (Luo *et al.*, 2014; M., Lee *et al.*, 2014; Fu *et al.*, 2014; Meyer and Jaffrey, 2014; Deng *et al.*, 2015; Huang and Yin, 2018), N1-methyladenosine (m1A), (Dominissini *et al.*, 2016; Li *et al.*, 2017), la pseudouridylation (Zaringhalam and Papavasiliou, 2016; Li *et al.*, 2016), ou la 5'hydroxyméthylation des cytosines (5hmC) (Delatte *et al.*, 2016) et **(2)** le **marquage des ARN en 5' ou 3'**. Cette seconde catégorie, qui connaît également un regain d'attention, comprend **l'ajout de nucléotides à l'extrémité 3' des ARN codants et non codants**. Outre l'adénylation canonique, les ARN peuvent être la cible d'autres types de marquage

en 3' tels que l'adénylation non canonique, l'uridylation (M., Lee *et al.*, 2014), la cytidylation et la guanylation.

L'adénylation non canonique est une modification présente chez l'ensemble des organismes, dans lesquels elle promeut la dégradation des ARN non codants. Elle favorise également la dégradation des ARNm chez les bactéries et les archées et dans les mitochondries notamment chez la plante, ainsi que dans les chloroplastes (Lange *et al.*, 2009; Norbury, 2013). Enfin elle constitue un signal d'activation de la traduction dans le cytoplasme des cellules neuronales et ovocytaire (Charlesworth *et al.*, 2013). La cytidylation et la guanylation des ARNm n'ont été décrites que chez l'homme, les plantes et *Aspergillus* nidulans (Chang *et al.*, 2014; Morozov *et al.*, 2012; Chang *et al.*, 2014; Zuber *et al.*, 2016). Une étude récente indique que l'incorporation de C, G ou U pendant la synthèse de la queue poly(A de certains ARNm chez l'homme permettrait de réguler leur vitesse de déadénylation, une étape clé de la dégradation des ARNm (Lim *et al.*, 2018).

Le rôle joué par **l'uridylation** est quant à lui plus largement étudié et consiste en l'ajout d'uridines monophosphates à l'extrémité 3' de substrats ARN, une action catalysée par des **Terminal UridylylTransférase** (TUTases). L'ajout de U par les TUTases est un mécanisme conservé et a été décrit chez divers organismes, incluant l'homme, la souris, la drosophile, la levure à fission, les trypanosomes, les nématodes et les plantes (Kwak and Wickens, 2007; Kwak and Wickens, 2007; Scott and Norbury, 2013; Aphasizhev and Aphasizheva, 2014; Munoz-Tello *et al.*, 2015; Reimão-Pinto *et al.*, 2015; Scheer *et al.*, 2016; Łabno *et al.*, 2016; Morgan *et al.*, 2017; De Almeida *et al.*, 2018). L'uridylation affecte un panel varié d'ARN codants et non codants et les conséquences de cette modification sont diverses. Si le rôle premier de l'uridylation des ARN est celui de promouvoir la dégradation de 3' en 5' et de 5' en 3', l'uridylation est également impliquée dans la maturation de transcrits primaires et pourrait jouer un rôle dans la traduction des ARNm. Au laboratoire nous nous intéressons tout particulièrement au rôle de l'uridylation chez *Arabidopsis thaliana* et tentons de comprendre son impact sur le métabolisme des ARN.

Mon travail de thèse s'est principalement porté sur le rôle de l'uridylation des ARN messagers par la **terminal uridylyltransférase URT1**, la TUTase majoritairement responsable de l'uridylation des ARNm chez Arabidopsis (Sement *et al.*, 2013; Zuber *et al.*, 2016).

Afin de mieux appréhender mes travaux de thèse, l'introduction aura pour ambition de donner un aperçu global des processus principaux impliqués dans la traduction et la dégradation des ARN messagers. Il y sera également traité le lien entre dégradation et traduction et le rôle joué par la formation des complexes ARN-protéines que sont les P-bodies. *In fine* l'ensemble de ces connaissances servira de support pour la description du rôle complexe joué par l'uridylation des ARN.

Structure, traduction et dégradation des ARNm eucaryotiques

1. La structure des ARN messagers chez les eucaryotes

Lors de la transcription par l'ARN polymérase II (Pol II), les ARNm précurseurs (pré-ARNm) subissent deux modifications déterminantes pour leur maturation, leur stabilité et leur traduction. Ces modifications sont l'ajout d'une **structure coiffe en 5'** et d'une **queue poly(A) en 3'**. Ces deux modifications sont nécessaires pour permettre l'export du noyau vers le cytosol, l'initiation de la traduction des ARNm et joue un rôle crucial dans la stabilité des ARNm. A noter que ces processus sont également influencés par des modifications des nucléotides (par exemple, une méthylation des adénosines en position 6 ou m6A) et les étapes d'épissage de séquences introniques qui affectent nombre d'ARNm eucaryotiques. En effet, lors de l'épissage des introns, un complexe protéique nommé EJC pour Exon Junction Complex est déposé à la jonction des exons. Ce complexe influence l'épissage lui-même, l'export, la traduction et la dégradation de certains ARNm (Hir *et al.*, 2016). Nous nous focaliserons pour la suite de cette partie sur la description des rôles de la coiffe et de la queue poly(A). Ces deux éléments sont en effet retrouvés chez tous les ARNm eucaryotiques à l'exception notable des ARNm d'histones qui ne sont pas polyadénylés chez les mammifères (Marzluff and Koreski, 2017).

1.1 La coiffe

La transcription par Pol II génère un ARN ayant une extrémité 5' triphosphate libre, cette dernière est modifiée co-transcriptionnellement de manière covalente afin d'incorporer une structure particulière, **la coiffe.** La structure de la coiffe est caractérisée par la présence d'une guanosine méthylée en position N7 reliée au premier nucléotide du pré-ARNm par une liaison 5'-5 triphosphate inversée (7mGpppN, avec N étant le premier nucléotide de l'ARN). Une telle coiffe, sans méthylation dans le corps de l'ARN est la coiffe 0. Si le premier nucléotide de l'ARNm est également méthylé (monométhylation) ou les deux premiers nucléotides (diméthylation) on parle de coiffe 1 et 2, respectivement. Ces deux dernières structures sont celles qui prévalent chez les eucaryotes supérieurs. Le profil de méthylation de la coiffe est dynamique et réversible et constitue une modification épitranscriptomique impliquée dans la stabilité des ARN (Mauer *et al.*, 2017). A noter que récemment, une coiffe alternative des ARNm due à la liaison en 5' de nicotinamide adenine dinucleotide (NAD^+) a été décrite chez les bactéries (Jäschke *et al.*, 2016), mais également chez la levure *Saccharomyces cerevisiae* et les mammifères (Kiledjian, 2018). Au contraire de la coiffe canonique m^7G qui stabilisent les ARNm, la coiffe NAD+ promeut la dégradation de ces cibles (Kiledjian, 2018).

Dans le noyau, la coiffe canonique m^7G est requise pour l'épissage, la polyadénylation et l'export des ARNm matures du noyau vers le cytoplasme. Dans le cytoplasme elle constitue une **structure protectrice** contre l'action prématurée d'**exoribonucléases 5'3'**, est impliquée dans la pseudo-

circularisation de l'ARNm et le recrutement de facteurs nécessaires à l'**initiation de la traduction** (Ramanathan *et al.*, 2016). La fonction protectrice de la coiffe tient du fait que les exoribonucléases cytoplasmiques ne reconnaissent pas cette structure en tant que substrat, et la dégradation de l'ARNm coiffé de 5' en 3' nécessite donc l'action des facteurs d'enlèvement de la coiffe. La stimulation de la traduction par la coiffe résulte de l'interaction entre cette dernière et le complexe tertiaire d'initiation eucaryote **eIF4F** (eukaryotic initiation factor 4F) qui comprend la protéine de liaison à la coiffe **eIF4E** connectée à l'hélicase DEAD-box **eiF4A** via le facteur d'initiation de la traduction **eIF4G** . Chez les plantes, il existe également un complexe **eIF4iso4F** composé de eiF4iso4E, eiF4A, et eIFiso4G (Browning and Bailey-Serres, 2015).

1.2 La queue polyadénylique

En plus de l'addition d'une coiffe, les transcrits de Pol II sont polyadénylés en leur extrémité 3'. L'ajout d'une **queue poly(A)** sur les transcrits Pol II est permise par un clivage endonucléolytique au site de polyadénylation et par l'ajout d'une longue queue poly(A) (~150-250 nucléotides chez l'homme et ~70-80 nucléotides chez la levure) par une **poly(A) polymérase canonique** (cPAP). L'ajout de la queue poly(A) permet le recrutement de protéines de liaison aux queues poly(A), les PABP, (**P**oly(**A**) **B**inding **P**rotein) (Eckmann *et al.*, 2011).

1.2.1 Structure et rôles des PABP

De manière générale, chez les eucaryotes il existe une PABP nucléaire qui permet de contrôler la taille de l'extension poly(A) ajouté à l'ARNm et qui contribue à l'export des ARNm du noyau vers le cytoplasme (Goss and Kleiman, 2013). Dans le cytoplasme il peut y avoir une ou plusieurs PABP selon les organismes (Goss and Kleiman, 2013; Mangus *et al.*, 2003). Si le génome de la levure *S. cerevisiae* code pour une seule et unique PABP cytoplasmique, celui de l'homme et de la plante *Arabidopsis thaliana* en expriment 4 et 8, respectivement. Les **8 PABP cytoplasmiques** d'*Arabidopsis thaliana* forment 3 classes distinctes : la classe I (PABP3, PABP5, PABP1) et la classe III (PAB6/PABP7) dont l'expression est faible ou spécifique de certains tissus, et la classe II (**PAB2/4 et 8)** dont l'expression est abondante, notamment dans les tissus végétatifs (Belostotsky, 2003; Gallie and Liu, 2014; Gallie, 2017).

Les PABP sont composées d'une région NH2-terminale (N-terminale) comprenant **4 domaines de liaison à l'ARN** de type **RRM** (RNA recognition motif) et d'une région COOH-terminale (C-terminale) comprenant une région riche en proline riche suivi d'un domaine composé de 5 hélices α, le **domaine MLLE**. Les domaines RRMs 1 et 2 sont suffisants pour la liaison avec une haute affinité à un ARN oligo(A) (Deo *et al.*, 1999). L'association de la PABP avec la queue poly(A) nécessite **un site de liaison minimal de 12 adénosines** (Deo *et al.*, 1999) et plusieurs PABP peuvent se lier sur une même queue poly(A) avec **1 PABP chaque ~27 nucléotides** (Baer and Kornberg, 1980). Chez l'homme, il a été montré, que la multimérisation des PABP sur la queue poly(A) est permise par l'interaction intermoléculaire entre le RRM2 d'une PABP et la région riche

en proline de la PABP voisine (Sawazaki *et al.*, 2018) . De surcroit le domaine RRM2 est également impliqué dans l'interaction de la PABP avec eIF4G (Imataka, 1998; Safaee *et al.*, 2012) et cette interaction est nécessaire pour la stimulation de la traduction par la PABP (Kahvejian *et al.*, 2005). L'interaction eIF4G-PABP a également été décrite chez la levure et les plantes (Le *et al.*, 1997; Tarun and Sachs, 1996; Cheng and Gallie, 2007). Chez la plante et les mammifères, la PABP interagit également avec un autre facteur d'initiation de la traduction, eIF4B qui stimule la liaison de la PABP à l'ARN et l'activité hélicase de eiF4A (Bushell *et al.*, 2001; Cheng and Gallie, 2007).

1.2.2 La queue poly(A) et la coiffe agissent de manière synergique pour promouvoir l'initiation de la traduction

Plusieurs travaux ont comparé l'efficacité de traduction de 4 types d'ARNm rapporteurs, un ARNm avec une coiffe mais sans queue poly(A), un ARNm sans coiffe mais avec une queue poly(A), un ARN avec les deux modifications, et enfin un ARNm sans aucune des deux modifications. Ces expériences ont été menées *via* l'infiltration de protoplasmes de cellules de tabac, de cellules ovariennes de hamster et de levure mais également avec des extraits de cellules humaines HeLa. Les résultats montrent que la coiffe et la queue poly(A) promeuvent la traduction de façon individuelle, et que la présence des deux modifications simultanément améliore l'efficacité de traduction d'un facteur supérieur à la somme des effets individuels. Ces expériences démontrent l'effet synergique de la coiffe et de la queue poly(A) sur la traduction des ARNm chez la levure, l'homme et la plante (Bergamini *et al.*, 2000). Cette effet synergique de la coiffe et de la queue poly(A) sur la traduction est dépendante de l'interaction simultanée entre la PABP et eIF4G et de l'interaction entre eIF4G et le facteur de liaison à la coiffe, eIF4E (Tarun *et al.*, 1997; Tarun and Sachs, 1995; Tarun and Sachs, 1997; Kessler and Sachs, 1998) dont la conséquence est la formation d'une boucle fermée (Wells *et al.*, 1998; Gallie, 1998). L'interaction eIF4G-PABP-eIF4E a également pour conséquence de stabiliser l'interaction de eIF4F avec la coiffe (Ptushkina, 1998; Borman *et al.*, 2000; Haar *et al.*, 2000; Wei *et al.*, 1998) et contribue à favoriser l'initiation de la traduction.

Des papiers récents montrent que la liaison d'une PABP à une queue poly(A) entraine un changement drastique de la structure de la protéine qui adopte une conformation sévèrement courbée au niveau de la région de liaison entre le RRM2 et le RRM3. Ce changement de structure rabouti au rapprochement spatial du RRM1 et 4 (S., H., Lee *et al.*, 2014) et favorise l'interaction de la PABP avec eiF4G. La liaison préférentiel de eIF4G à des PABP qui adopte cette structure courbée pourrait constituer un mécanisme pour distinguer les PABP libres des PABP liées à des queues poly(A) (Hong *et al.*, 2017).

2. Mécanisme de traduction des ARNm

L'acteur majeur de la traduction des ARNm eucaryotes est le **ribosome 80S**. Ce dernier est composé deux sous-unités ; la **grande sous-unité 60S** qui comprend 3 ARN (l'ARN 28S, 5S et 5,8S) et la **petite sous-unité 40S** qui contient l'ARN 18S. Le ribosome permet l'appariement de

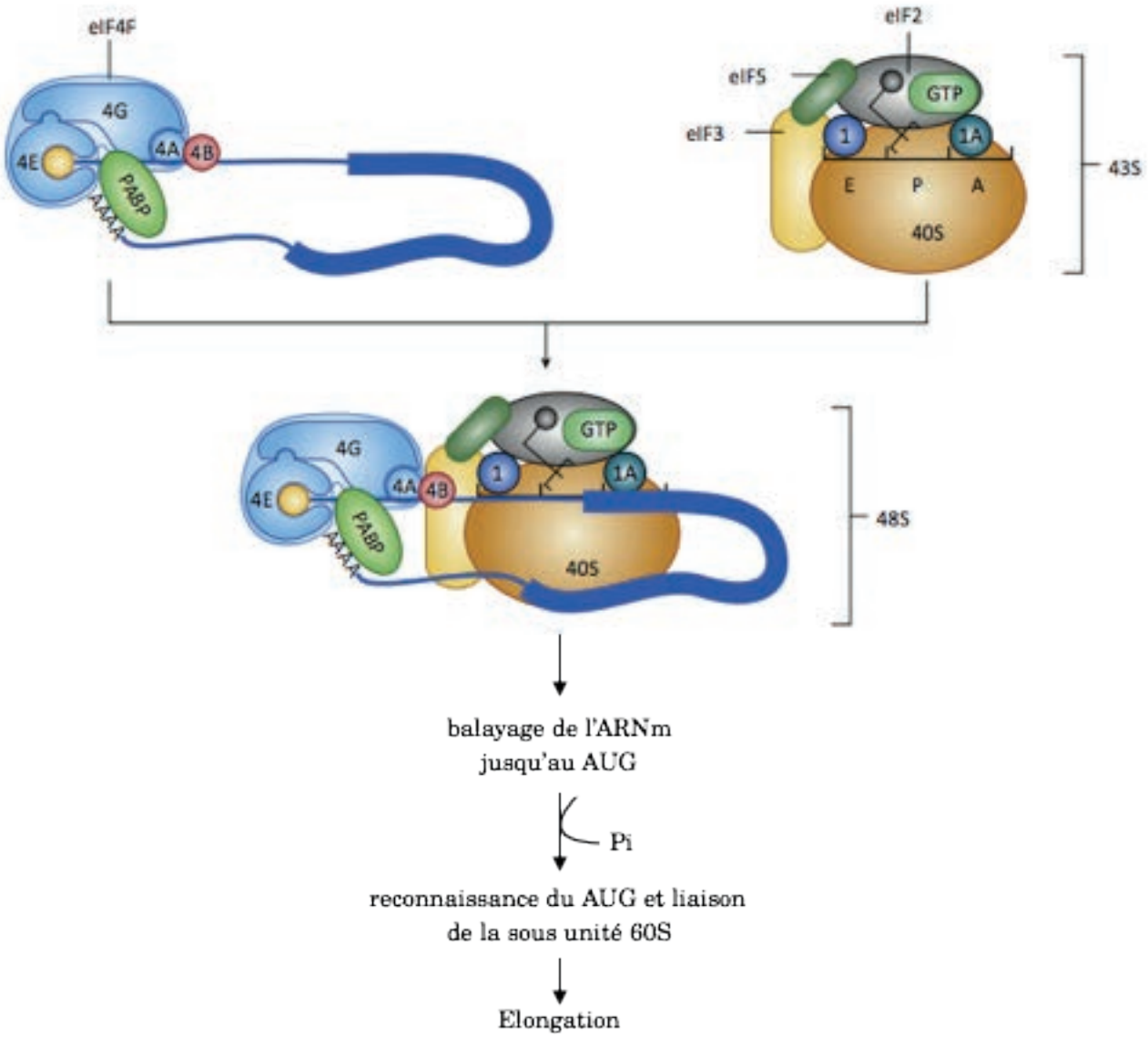


Figure 1 | Représentation schématique de l'assemblage de eIF4F et du ribosome pour former le complexe 48S

Le complexe de pré-initiation 43S est formé par l'association du ribosome 40S avec les facteurs d'initiations de la traduction eIF1, eIF1A, eIF3 et eIF5 et le complexe ternaire (eiF2 : GTP :ARNt méthionine initiateur). L'assemblage du complexe eIF4F sur la coiffe de l'ARNm permet le recrutement du ribosome à l'ARN de part l'interaction eIF4G:eIF3. L'interaction de eIF4G avec la PABP permet la circularisation de l'ARNm. L'association du 43S à l'ARNm conduit à la formation du complexe d'initiation 48S et le balayage de l'ARNm jusqu'à la rencontre du codon initiateur et élongation. Schéma adapté de Chu et al. 2016.

l'anticodon ARNt avec le codon de l'ARNm et la synthèse protéique. La traduction se déroule en 3 étapes : l'initiation, l'élongation et la terminaison. A noter que l'initiation de la traduction est l'étape la plus complexe et celle qui requiert le plus de facteurs, elle est également la première cible de la régulation des gènes liés au contrôle de la traduction (Aitken and Lorsch, 2012; Jackson *et al.*, 2010). L'ensemble des facteurs et des mécanismes impliqués dans ces 3 processus sont conservés chez les eucaryotes.

2.1 L'initiation de la traduction

L'initiation de la traduction débute par l'association du **complexe ternaire** (eIF2 : GTP :methionyl-ARNt initiateur) avec la sous-unité du ribosome 40S associé à eIF1, eIF1A, eIF2, eIF3 et eIF5 pour former le **complexe de pré-initiation 43S** (43S PIC). L'ARNm est ensuite recruté au niveau du complexe 43S grâce à l'interaction de ce dernier avec eIF4G/eIFiso4G. En effet, en plus de son interaction avec la PABP et eIF4A, eIF4G est également capable d'interagir avec le composant du complexe 43S, la protéine **eIF3** (Korneeva *et al.*, 2000; Lamphear *et al.*, 1995; Morino *et al.*, 2000; Villa *et al.*, 2013). L'interaction du 43S PIC avec l'ARNm et ses facteurs associés forme le **complexe de pré-initiation 48S** (48S PIC) (Fig. 1). Par la suite, l'activité hélicase de eIF4A, stimulée par eIF4B et eIF4G (Özeş *et al.*, 2011; Rozen *et al.*, 1990; Rogers *et al.*, 2001), contribue au déroulement des structures secondaires de l'ARNm et permet au 48S PIC de scanner l'ARNm de 5' en 3' jusqu'à reconnaissance du codon d'initiation AUG. Durant ce processus de balayage, le complexe 48S est dans une conformation dite « ouverte » avec le tRNA initiateur partiellement engagé dans le site P (P_{OUT}) (Llácer *et al.*, 2015). eIF1 et eIF1A sont impliqués dans la formation et la stabilisation de cette conformation propice au processus de balayage (Pestova and Kolupaeva, 2002; Hinnebusch and Lorsch, 2012; Aylett and Ban, 2017). La reconnaissance du codon AUG par l'anticodon ARNt initiateur induit le positionnement complet de ce dernier dans le site P (P_{IN}). Ce réarrangement entraine le relargage de eIF1 et l'hydrolyse par eIF5 du GTP lié à eIF2. Le départ de eIF1 a pour conséquence la « fermeture » du complexe 48S PIC et l'arrêt du balayage, ce qui entraine le relargage de eIF5 et de eIF2-GDP. Enfin la liaison de eIF5B-GTP et l'hydrolyse de sa molécule de GTP en GDP conduit à la dissociation des facteurs restants et à la fixation de la **sous-unité 60S** pour former le **ribosome 80S**.

2.2 Élongation et terminaison de la traduction

L'élongation de la traduction consiste en l'addition séquentielle d'acides aminés à une chaine polypeptidique naissante. Ceci implique l'utilisation des 3 sites présents dans le ribosome 80S, le site **E** (site de sortie, Exit), **P** (site de liaison du Peptidyl ARNt) et **A** (site de liaison de l'aminoacyl ARNt) et les acteurs clés de cette étape sont les **facteurs d'élongation de la traduction** (**eEF)**. A la fin de l'initiation l'anticodon de l'ARNt initiateur est apparié au codon d'initiation AUG situé dans le site P du ribosome 80S. L'élongation débute avec l'arrivée, dans le site A, d'un aminoacyl-ARNt (aa-ARNt) apparenté au second codon et lié au complexe binaire eEF1A-GTP. L'hydrolyse du GTP contenu sur ce dernier entraine le relargage de eEF1A-GDP et **l'accommodation de l'aa-**

ARNt dans le site A. S'ensuit la **catalyse d'une liaison peptidique** entre le nouvel acide-animé apporté au site A et l'extrémité carboxyle du polypeptide en cours de synthèse au site P. Cette action est catalysée dans le centre peptidyltransférase (PeptidylTransferase Center, PTC) à l'aide du facteur eEF5A qui participe au positionnement correct des substrats. Le résultat de ce processus est le transfert du peptide en cours de synthèse de l'aa-ARNt du site P sur l'acide aminé de l'aa-ARNt positionné dans le site A. Le peptidyl ARNt occupe ainsi le site A et l'ARNt dans le site P est déacétylé. L'élongation s'achève avec une étape de **translocation** qui conduit au déplacement de l'ARNt déacétylé du site P vers le site de sortie, du peptidyl ARNt du site A vers le site P et du mouvement de l'ARNm d'exactement 3 nucléotides (1 codon). Cette action est promue par eEF2. Une fois le site A libre l'ensemble des étapes précédemment décrites se répètent jusqu'à synthèse complète de la chaine polypeptidique. (Parsyan, 2014; Browning and Bailey-Serres, 2015; Dever *et al.*, 2018).

Lorsque qu'un codon stop atteint le site A du ribosome, cela entraine le recrutement des facteurs de terminaison de la traduction eucaryotes eRF (eukaryotic Release Factor) 1 et 3 qui agissent sous la forme d'un complexe ternaire eRF1: eRF3-GTP. Lorsque le complexe ternaire se lie au site A, eRF1 reconnaît le codon stop et eRF3 hydrolyse son GTP. S'ensuit l'accommodation de eRF1 dans le centre peptidyltransférase du ribosome et la terminaison de la traduction qui aboutit au **relargage de la chaine polypeptidique** (Hellen, 2018).

2.3 Les polysomes

Lorsqu'un ribosome quitte le site d'initiation de la traduction, il laisse la place libre à un autre ribosome qui, à son tour, peut commencer la traduction de la chaine polypeptidique. Ceci implique qu'il y ait plusieurs ribosomes sur un même ARNm engagé dans le processus de synthèse protéique. Ce complexe macromoléculaire formé par l'ARNm et plusieurs ribosomes associés est appelé **polysomes** (ou polyribosomes). Si un seul ribosome est présent sur un ARNm on parle de **monosome**. Les polysomes sont associés à la traduction active des ARNm (Slayter *et al.*, 1963), et les monosomes correspondent à des ribosomes en cours d'initiation de la traduction au codon start. Des travaux récents montrent cependant qu'une traduction active peut avoir également avoir lieu dans les monosomes et notamment celle de transcrits particuliers incluant les transcrits ayant une phase de lecture courte (short ORFs), les cibles de la voie du NMD (nonsense mRNA mediated decay, décrit ultérieurement), des uORFs (upstream ORFs) et des ARNm encodant des protéines régulatrices de faibles abondances (Heyer and Moore, 2016).

Grâce à l'avancée des techniques de séquençages haut-débit et le développement de la technique de profilage du ribosome (ribosome profiling ou Ribo-Seq) il est désormais possible de cartographier la position des ribosomes sur les ARNm en cours de traduction *in vivo*. Le Riboseq se base sur la propension qu'ont les ribosomes à protéger les ARNm sur une zone de ~30 nucléotides, ce qui a pour conséquence de rendre ces régions insensibles à l'action de ribonucléases. Le séquençage haut-débit des fragments protégés permet de déterminer le positionnement des ribosomes sur leur ARNm associé et ainsi de révéler le translatome, c'est à dire l'identification globale des ARNm

traduits. En connaissant le nombre de ribosomes positionnés sur un ARNm donné, il est possible d'apprécier de manière quantitative l'efficacité de traduction (Ingolia *et al.*, 2009; Ingolia *et al.*, 2012; Ingolia, 2014). Ces derniers années ont été également marquées par l'avancée des technologies d'imagerie permettant la visualisation en temps réel de la traduction d'un ARNm unique *in vivo* (Chekulaeva and Landthaler, 2016).

2.4 L'effet de la taille de la queue poly(A) sur la traduction

L'idée qui prélavait jusqu'alors était que la présence d'une queue poly(A) longue entrainait une meilleure traductibilité de l'ARNm. Cette idée a vraisemblablement été inspirée par le fait que l'allongement des queues poly(A) peut activer la traduction de transcrits dans des contextes spécifiques (Eckmann *et al.*, 2011; Weill *et al.*, 2012). A titre d'exemple, la polyadénylation par des poly(A) polymérases cytoplasmiques active la traduction d'ARN maternels lors de la maturation méiotique chez les oocytes de Xénope (Barkoff *et al.*, 1998) et la traduction de transcrits impliqués dans la formation de l'axe antéro-postérieur au cours des stades précoces de l'embryogenèse chez la drosophile (Salles *et al.*, 1994). La polyadénylation cytoplasmique promeut également la traduction d'ARNm dans les neurones en réponse à des signaux synaptiques (Zukin *et al.*, 2009) et la synthèse de la poly(A) binding protein cytoplasmique humaine (PABPC1) dans les cardiomyocytes (Chorghade *et al.*, 2017).

Des travaux récents, permis par l'avènement de nouvelles méthodes d'analyse de la taille des queues poly(A) à l'échelle du génome (TAIL-seq, PAL-seq) en combinaison avec des mesures de l'efficacité de traduction par Ribo-seq révèlent qu'il n'existe, en dehors de ces contextes spécifiques, pas de corrélation ou une corrélation inverse entre la taille des queues poly(A) et leur traduction. En effet, si Subtelny et al. confirment la présence d'un couplage entre la taille de la queue poly(A) et la traduction lors des stades précoces de développement chez le Xénope et le poisson-zèbre, la taille de la queue poly(A) ne semble, en revanche, n'avoir que très peu d'influence sur la traduction chez les adultes (Subtelny *et al.*, 2014). Les auteurs n'observent également aucune corrélation taille de queues poly(A)/traduction chez *S. cerevisiae*, *S. pombe* et dans des cellules hépatiques de souris. Ces travaux concordent avec les observations de Park et al. montrant que la variation de la taille de la queue poly(A) entre la phase S et la phase M du cycle cellulaire ne résulte pas sur une modification de l'efficacité de traduction des ARNm considérés dans les cellules somatiques humaines et que seuls des ARNm ayant une queue poly(A) inférieure à 20 nucléotides semblent présenter une efficacité de traduction réduite (Park *et al.*, 2016) .

De manière surprenante, il existerait même une corrélation inverse entre la taille de la queue poly(A) d'un ARNm et sa traductibilité. En effet, des travaux récents menée chez *Caenorhabditis elegans* révèlent une corrélation négative entre la taille de la queue poly(A) et la traduction des ARNm (Lima *et al.*, 2017). En effet ces auteurs montrent qu'une taille de poly(A) réduite, pouvant accommoder de 1 à 2 PABP, est caractéristique des transcrits abondants, enrichis en codons optimaux et fortement associés aux ribosomes et que cette caractéristique est conservée chez d'autres eucaryotes comme la souris (Lima *et al.*, 2017).

3. Dégradation des ARNm eucaryotes

L'abondance cellulaire d'un ARNm résulte de l'action conjuguée de la régulation de la transcription et des processus de **régulation de la stabilité des ARN** (Fan *et al.*, 2002; Grigull *et al.*, 2004; Marguerat *et al.*, 2014; Vogel and Marcotte, 2012).

La stabilité d'un ARNm peut être appréciée grâce à la mesure de son **temps de demi-vie**. La mesure du temps de demi-vie des ARNm sur l'ensemble du génome a été réalisée chez de nombreuse espèces eucaryotes, dont la levure, l'homme et chez la plante et manifeste d'une grande hétérogénéité dans la stabilité des ARNm (Wang *et al.*, 2002; Chen *et al.*, 2008; Narsai *et al.*, 2007; Mata and Wise, 2017). En effet, le temps de demi-vie d'un ARNm donné peut varier de quelques minutes à plusieurs heures et est fonction des conditions physiologiques.

Il existe notamment une corrélation entre la stabilité d'un ARNm et la fonction de la protéine pour laquelle il code, ainsi des transcrits qui codent pour des protéines ayant des fonctions régulatrices ont un temps de demi-vie court *a contrario* de ceux codant pour des protéines ayant des fonctions dans le métabolisme général (transcription, cycle cellulaire, traduction…) (Wang *et al.*, 2002; Narsai *et al.*, 2007; Yang *et al.*, n.d.; Sorenson *et al.*, 2018). De surcroît la présence de motifs caractéristiques, l'optimalité des codons, la présence de sites de reconnaissance pour les microARN et la présence ou l'absence d'introns sont également des facteurs influençant la vitesse de dégradation des ARN chez les eucaryotes (Narsai *et al.*, 2007; Sorenson *et al.*, 2018; Presnyak *et al.*, 2015). Le contrôle de la stabilité des ARNm est un outil indispensable à la cellule pour répondre aux stimuli de l'environnement et notamment en réponse à des stress chez l'homme (Fan *et al.*, 2002), la levure (Gowrishankar *et al.*, 2005; Greatrix and van Vuuren, 2006; Romero-Santacreu *et al.*, 2009) et la plante (Romero-Santacreu *et al.*, 2009; Merret *et al.*, 2013; Merret *et al.*, 2015; Perea-Resa *et al.*, 2016; Kawa and Testerink, 2017).

La dégradation des ARNm est un processus précisément orchestré par l'action d'une série de réactions enzymatiques qui peuvent amener à la dégradation dans le sens 3'5' ou 5'3'. L'ensemble de ces mécanismes ont pour premier objectif la déprotection de l'ARNm, à savoir l'enlèvement de la coiffe, le raccourcissement de la queue poly(A) ou par clivage endonucléolytique dans le corps même de l'ARNm. Une variété énorme de mécanismes est impliquée dans le contrôle de la dégradation de ARNm, des modifications épitranscriptomiques comme le m6A, au recrutement de protéines de liaison à l'ARN qui elles-mêmes recrutent des facteurs d'élimination de la coiffe ou de déadénylation, ou encore par la reconnaissance par les petits ARN (miRNA ou siRNA) qui assurent la spécificité de reconnaissance des ARNm par le complexe RISC (RNA inducing silencing complex). Je ne traite dans la suite de cette introduction que des voies générales de dégradation des ARNm.

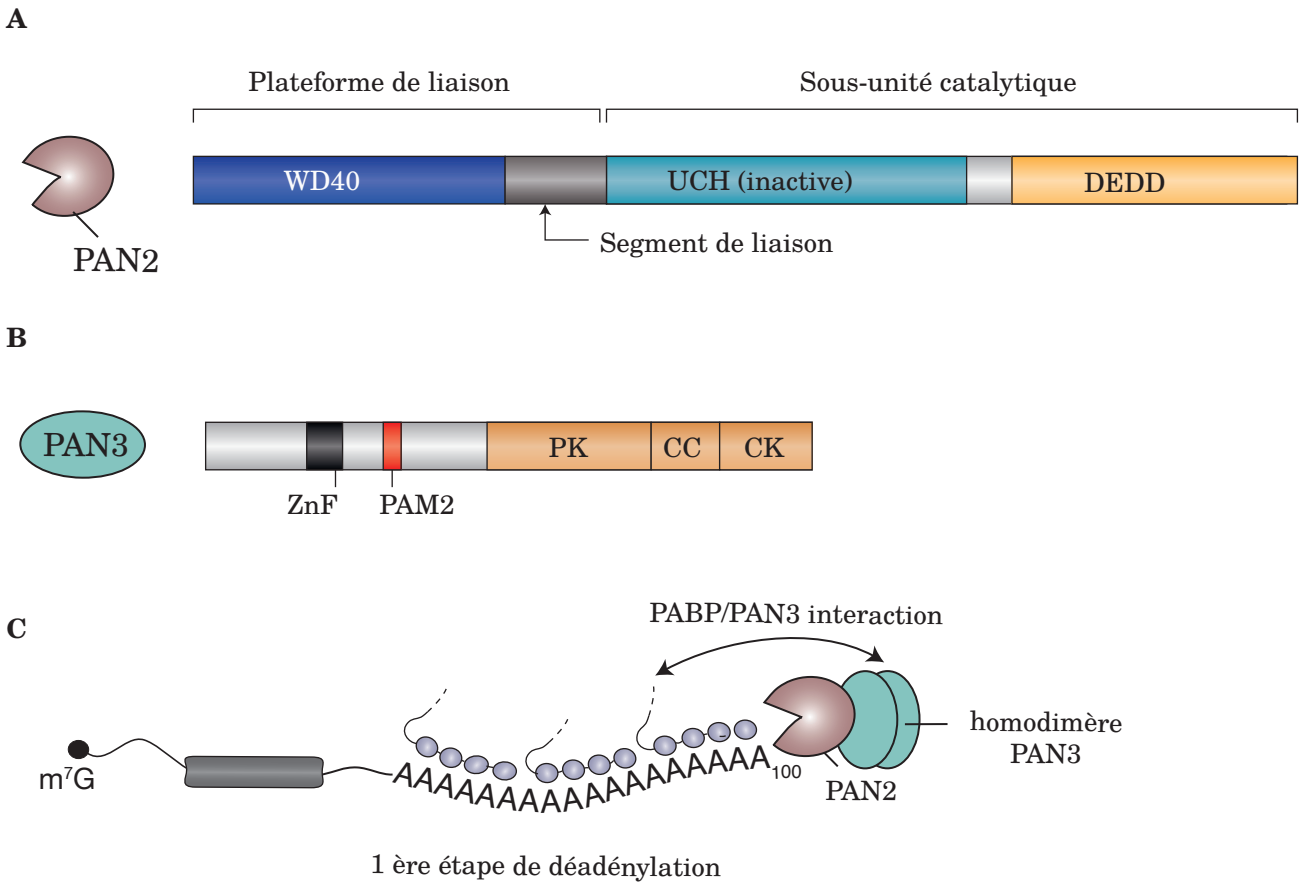


Figure 2 | Organisation générique des protéines PAN2 et PAN3 chez les eucaryotes. Le complexe PAN2-PAN3 est formé de l'association de l'exonucléase PAN2 et d'un homodimère de PAN3. **A |** PAN2 contient 3 domaines globulaires: le domaine WD40 connecté à un domaine ubiquitin C-terminal hydrolase inactif (UCH) par un segment de liaison, et un domaine exonucléase. **B |** PAN3 contient un domaine en doigt de zinc (ZnF), un domaine d'interaction avec la PABP (PAM2), un domaine pseudokinase (PK), un domaine superhélice (coil-coil, CC) et un domaine C-terminal Knob (CK). L'interaction entre PAN2 et PAN3 est permise par les domaines compris dans la plateforme de liaison de PAN2 et le domaine CK de PAN3. Le domaine superhélice est requis pour la l'homodimérisation de PAN3. Le domaine ZnF sert de site de liaison à l'ARN et domaine PAM2 de PAN3 permet l'interaction de PAN3 avec la PABP. **C |** Le complexe PAN2-PAN3 catalyse la première étape de déadenylation chez les eucaryotes supérieurs, exception faite pour les plantes à fleurs qui ne possèdent pas d'orthologue aux protéines PAN2/PAN3. Schéma adapté de : Schäfer et al., 2014; Wolf and Passmore, 2014; Jonas and Izaurralde, 2015.

3.1 La déadénylation, première étape et étape limitante de la dégradation des ARNm

Chez les eucaryotes la stabilité des ARNm est intimement reliée à la présence de la queue poly(A) qui, en plus de son rôle dans l'initiation de la traduction, peut servir de structure protectrice en 3'. La majorité des processus de dégradation implique dès lors une première étape de **déadénylation** consistant au raccourcissement de cette dernière (Chen and Shyu, 2011). La vision classique de la déadénylation, apportée par des travaux menés chez l'homme et la levure, est celle d'un processus **biphasique** qui requiert une première étape de déadénylation distributive par le complexe **PAN2-PAN3** (PAN2/3) et l'action consécutive du complexe **CCR4-NOT** (CNOT). Ce dernier opère de façon processive jusqu'à obtention d'une queue poly(A) de ~45 adénosines puis de façon distributive pour la dégradation des dernières adénosines (Yamashita *et al.*, 2005).

3.1.1 Le complexe PAN2-:PAN3

Le complexe PAN2-PAN3 est un hétérodimère formé par la déadenylase PAN2 en interaction forte avec un homodimère PAN3 (Schäfer *et al.*, 2014; Christie *et al.*, 2013). L'organisation en domaine des deux protéines est représentée en Figure 2. PAN2 est une exonucléase 3'5' distributive qui contient un domaine exonucléase C-terminal de type DEDD (acide aspartique, acide glutamique, acide aspartique, acide aspartique). De ce fait son activité est dépendante de la présence de cations divalents et une mutation dans un des résidus du motif DEDD, responsable de la coordination des ions Mg^{2+}, inactive PAN2 (Zuo and Deutscher, 2001; Wahle and Winkler, 2013).

Le recrutement de PAN2-PAN3 à l'ARNm est permise par l'interaction entre le motif PAM2 de la sous-unité PAN3 et le motif MLLE de PABP, cette interaction stimule l'activité de PAN2 chez l'homme et la levure (Lowell *et al.*, 1992; Boeck *et al.*, 1996; Uchida *et al.*, 2004). Outre son interaction avec la PABP, PAN3 est également capable de se lier directement à l'ARNm grâce à la présence, dans sa région N-terminale, d'un domaine en doigt de zinc (Wolf *et al.*, 2014). PAN3 est donc la sous-unité régulatrice du complexe PAN2-PAN3 puisqu'elle apporte la spécificité de substrat et module l'activité de PAN2.

De manière intéressante, le complexe formé par PAN2-PAN3 est conservé chez la plupart des eucaryotes supérieurs à l'exception des plantes à fleurs. En effet bien qu'un orthologue à PAN2/PAN3 a été identifié chez l'algue verte *Chlamydomanas reinhardtii* ce complexe semble être absent chez *Arabidopsis thaliana.* Ce faisant, il est probable que le modèle de dégradation biphasique ne s'applique pas à ces organismes et que les processus de déadenylation soit la conséquence exclusive de l'activité du complexe CCR4-NOT. L'existence du modèle biphasique est également remise en cause chez l'homme, puisque l'abolition de PAN2-PAN3 n'a qu'un effet minoritaire sur la taille des queues poly(A) et le temps de demi-vie des ARN, contrairement à l'inactivation de CCR4-NOT. De surcroit le complexe CNOT semblerait également capable de dégrader des longues queues poly(A). Ces résultats suggèrent que la première phase de déadenylation par PAN-PAN3 n'est probablement pas une étape critique du processus de

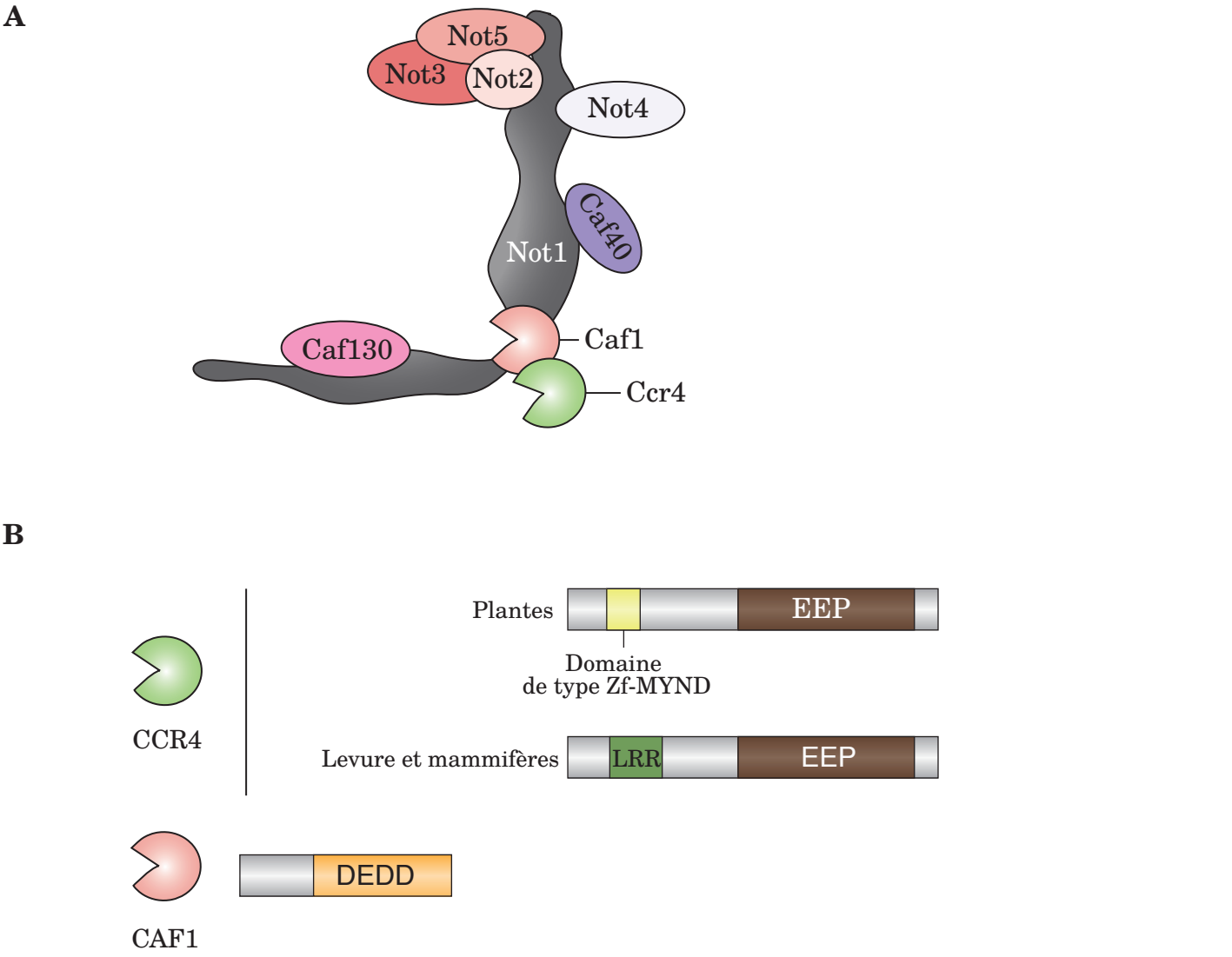


Figure 3 | Le complexe CCR4-NOT
A | Organisation du complexe CCR4-NOT de la levure (schéma adapté de Collard et al. 2016). La protéine Not1 et sa structure en L sert de protéine d'échafaudage pour la liaison des autres facteurs du complexe CCR4-NOT. Les déadenylases Ccr4 et Caf1 interagissent avec le complexe de par l'interaction entre Caf1 et Not1 B | Organisation en domaine des déadénylases CCR4 et CAF1. CCR4 est composé d'un domaine exonucléase de type EEP, et d'un domaine composé de répétition en leucine (Leucine Rich Region, LRR) qui permet son interaction avec CAF1. Le domaine LRR est remplacé par un domaine Zf-MYND. CAF1 est uniquement composé d'un domaine exonucléase de type DEDD.

déadenylation et que la majorité de l'activité déadenylase observée dans les cellules humaines est la due à l'activité du complexe CCR4-NOT (Yi *et al.*, 2018).

3.1.2 Le complexe CCR4-NOT (CNOT)

Le complexe CNOT joue un rôle important dans de nombreux processus cellulaires, il est ainsi impliqué dans les processus de transcription, de contrôle qualité et d'export, de traduction, de contrôle qualité durant la traduction et bien évident il est crucial pour la dégradation des ARNm dans le cytoplasme (Collart, 2016). Il est conservé chez les eucaryotes et est composé des deux déadénylases **CCR4** (Carbon Catabolic Repressor 4) et **CAF1** (CCR4 associated factor) ainsi que des sous-unités NOT1, NOT2, NOT5 et CAF40. On trouve des facteurs additionnels spécifiques à certaines espèces, CAF130 chez *S. cerevisiae*, Mmi1 chez *S. pombe* et NOT10 et NOT11 chez l'homme et la drosophile. Le complexe est construit autour d'une protéine de 200kDa, NOT1, qui sert de protéine d'échafaudage pour les autres facteurs du complexe (Collart, 2016). La compréhension de l'organisation structurelle et du réseau d'interaction du complexe CCR4-NOT1 a notamment été permise par la résolution du complexe chez la levure, représenté schématiquement en Figure 3A et illustre bien le rôle clé de NOT1 dans l'association du complexe CNOT.

L'activité du complexe est apportée par l'activité de CCR4 et CAF1. CCR4 appartient à la famille des exonucléases-endonucléases-phosphatases (EEP) et CAF1 à la famille des nucléases de type DEDD, au même titre que PAN2 (Zuo and Deutscher, 2001; Wahle and Winkler, 2013). L'activité nucléases spécifique des poly(A) de ces deux nucléases a été montrée *in vitro* à l'aide de protéines purifiées chez l'homme (Bianchin *et al.*, 2005; Wang *et al.*, 2010), la souris (Viswanathan *et al.*, 2004), la levure (Daugeron *et al.*, 2001; Tucker *et al.*, 2002; Thore *et al.*, 2003), la drosophile (Temme *et al.*, 2004; Temme *et al.*, 2010) et la plante (Liang *et al.*, 2009; Chou *et al.*, 2014; Chou *et al.*, 2017). Chez la levure et les mammifères CCR4 est porteur d'un motif répété riche en leucine (LRR) dans sa partie N-terminale (Fig. 3B) qui permet son interaction avec CAF1. Ce motif est absent chez les deux homologues de CCR4 présent chez Arabidopsis ou le riz (CCR4a et CCR4b) et remplacé par un domaine de type zf-MYND (Fig. 3B) (Chou *et al.*, 2017). CAF1 interagit directement avec NOT1 et sert d'intermédiaire entre CCR4 et NOT1 dans la formation du complexe CCR4-NOT.

3.1.3 Le rôle complexe des PABP dans le contrôle de la déadenylation

Comme nous l'avions décrit antérieurement les PABP recouvrent les queues poly(A) des ARNm et forment un complexe tripartite avec la coiffe 5' ce qui favorise l'initiation de sa traduction. Les PABP jouent également un rôle complexe ; voire paradoxal, dans les processus de déadenylation.

En effet la déplétion de la PABP résulte en une diminution de la stabilité des ARNm rapporteurs analysés dans des extraits de cellules humaines (Bernstein *et al.*, 1989) et sur un raccourcissement des queues poly(A) chez *S. cerevisiae* (Sachs and Davis, 1989) suggérant que le rôle de la PABP serait celui d'occulter les ARNm de l'activité des déadénylases. Conformément à ces résultats la

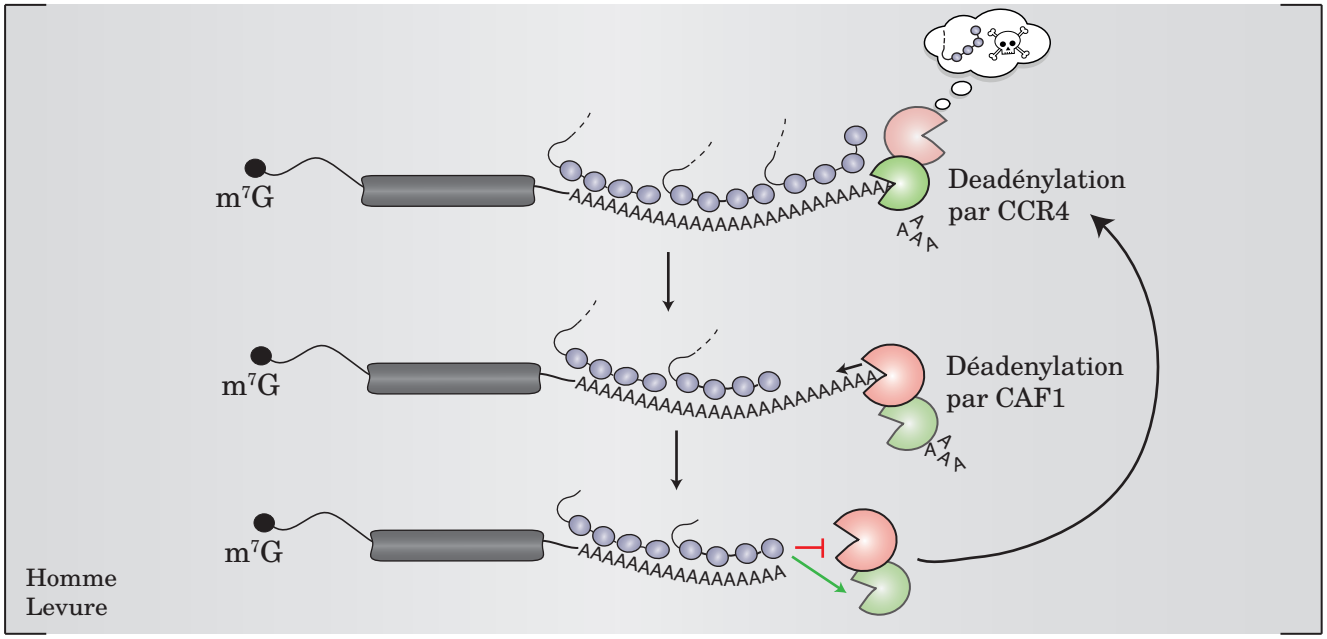


Figure 4 | Modèle de dégradation des ARNm par CCR4-NOT
La queue poly(A) des ARNm chez la levure est liée par plusieurs PABP. La présence de la PABP favorise l'activité de la déadenylase CCR4 chez la levure et l'homme, chez la levure cette stimulation se fait de part une interaction physique entre CCR4 et la PABP. La déadénylation de la queue poly(A) par CCR4 provoque l'éviction de la PABP. La queue poly(A) est alors libre de toute liaison et constitue un bon subtrat pour l'activité de CAF1. CAF1 déadenyle la queue poly(A) jusqu'à ce que son activité soit freinée par la présence de la PABP suivante. Figure adaptée de Bresson et Tollervey, 2018.

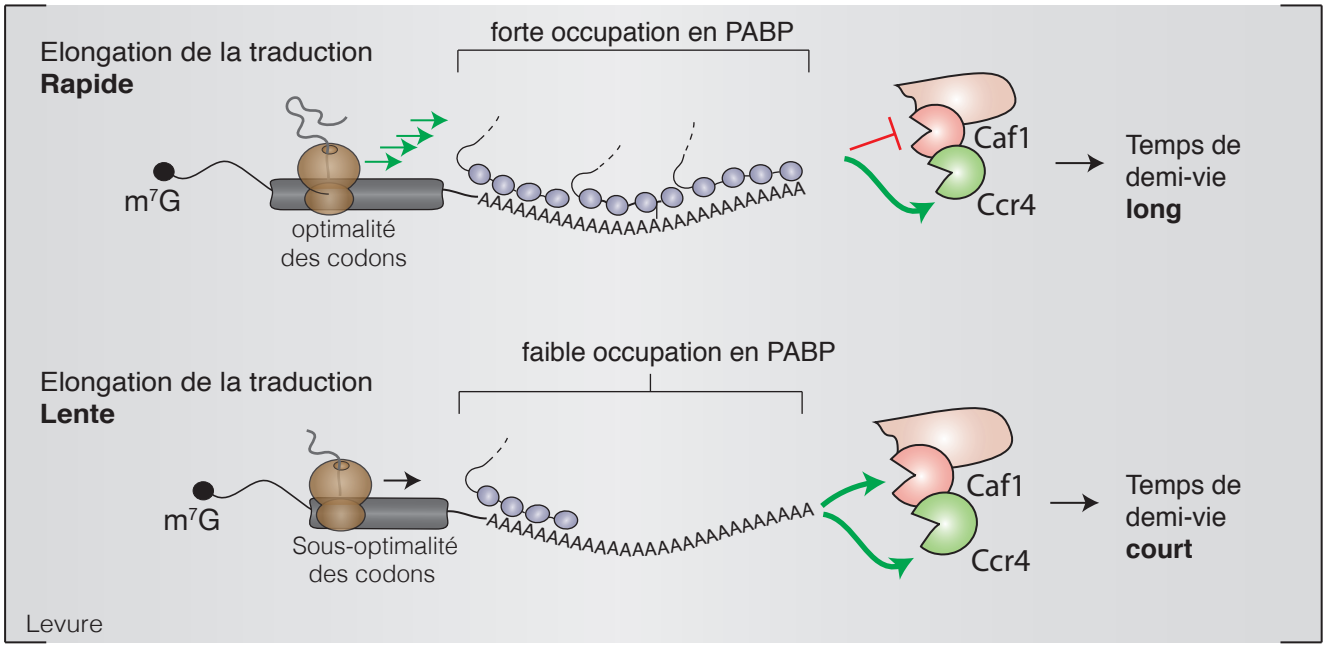


Figure 5 | **Couplage entre la traduction et la vitesse déadenylation par CAF1**
Les travaux de Webster et al. montre qu'il existe une correlation entre l'optimalité des codons d'un transcrit et la fréquence de liaison des PABPs sur les queues poly(A). L'activité de Caf1 est spécifique des transcrits sous optimaux.
Figure adaptée de Webster et al., 2018.

surexpression de la PABP protège les ARNm de la déadénylation dans les oocytes de Xenopus et chez la levure (Tucker *et al.*, 2002; Wormington *et al.*, 1996).

Paradoxalement à son rôle protecteur, la PABP stimule l'activité déadénylase de PAN2-PAN3 et des expériences contradictoires menées chez la levure montrent que la présence de la PABP augmente la vitesse de déadenylation des ARNm rapporteurs testés (Caponigro and Parker, 1995). Chez les mammifères la PABP est également capable de stimuler le complexe CCR4-NOT en interagissant avec les facteurs TOB1 et BTG2 (Ezzeddine *et al.*, 2007; Stupfler *et al.*, 2016).

La liaison des PABP aux queues poly(A) semble donc avoir un rôle ambivalent, elle peut à la fois bloquer l'action des nucléases 3'5' pour protéger les extrémités 3' de la dégradation et à l'inverse recruter et stimuler l'activité des déadénylases. Récemment, les travaux de deux équipes corroborent l'effet antagoniste de la PABP sur la déadenylation (Yi *et al.*, 2018; Webster *et al.*, 2018). En utilisant le complexe CCR4-NOT purifié chez *S. pombe*, Webster et al. montrent que la PABP interagit avec CCR4 et stimule son activité déadenylase mais qu'*a contrario* elle inhibe l'activité de CAF1. Ces résultats sont confortés par les expériences similaires menées avec les versions humaines de CCR4 et CAF1 (Yi *et al.*, 2018). Le raccourcissement de la queue poly(A) pourrait être la conséquence de l'action consécutive de CCR4 sur les queues poly(A) liées à la PABP, et celle de CAF1 sur les queues poly(A) libres (Fig. 4). De manière intéressante des expériences *in vivo* menées chez la levure *S. cerevisiae* montrent que, contrairement à CCR4 dont l'activité déadénylase cible l'ensemble des transcrits, l'activité de CAF1 semble, quant-à-elle, restreinte à la déadenylation de transcrits ayant une composition en codon sous-optimale. Une faible **optimalité de codon** corrèle avec une faible efficacité de traduction, une plus grande instabilité (Presnyak *et al.*, 2015) et une plus faible occupation en PABP (Webster *et al.*, 2018). Ces résultats suggèrent que la connexion entre efficacité de traduction et dégradation pourrait se faire par l'intermédiaire de la PABP et de CAF1 (Fig. 5).

L'ensemble de ces observations dénote d'un statut pivot de la PABP dans les processus de déadénylation et de la traduction des ARNm chez les eucaryotes.

3.1.4 Les autres déadénylases chez les eucaryotes

En plus des complexes PAN2-PAN3 et CCR4-NOT ils existent plusieurs autres protéines ayant une activité déadenylase chez les eucaryotes. Parmi les mieux étudiées, on trouve **PARN** (Poly(A)-specific ribonuclease) et **Nocturnin**. PARN est une exonucléase de type DEDD, retrouvée chez les plantes et l'homme. L'enzyme semble en revanche absente chez *S. cerevisiae* et la drosophile. En plus de se lier à la queue poly(A), PARN est capable d'interagir de manière simultanée avec la coiffe en 5' ce qui favorise son activité enzymatique et sa processivité (Godwin *et al.*, 2013). Chez l'homme la mutation de PARN est associée à des retards développementaux et à de l'insuffisance médullaire (Dhanraj *et al.*, 2015). Chez Arabidopsis PARN est requise pour les stades précoces de développement et est essentielle pour la viabilité (Chiba *et al.*, 2004). Elle joue également un rôle dans la réponse au stress (Nishimura *et al.*, 2005). PARN pourrait être impliquée dans le raccourcissement de la taille de la queues poly(A) de transcrits embryonnaires (Reverdatto *et al.*,

2004). PARN semble être aussi localiser dans les mitochondries chez Arabidopsis et réguler la taille des queues poly(A) des ARNm mitochondriaux en conjonction avec la poly(A) polymérase AGS1 (Hirayama *et al.*, 2013).

La seconde enzyme, Nocturnin appartient à la famille des protéines CCR4, et de ce fait, possède un domaine nucléase de type EEP, elle est en revanche dépourvue du domaine LRR qui permet à CCR4 d'interagir avec CAF1 chez la levure et l'homme. Nocturnin est régulée par le cycle circadien, un trait qui la distingue des autres déadénylases. Chez la plante l'altération de la protéine orthologue à Nocturnin, HESPERIN affecte la rythmicité des transcrits TOC1 (Timing of Cab Expression1) et CCA1 (Circadian Clock Associated 1) deux oscillateurs clés de l'horloge circadienne (Delis *et al.*, 2015).

3.2 Mécanisme de dégradation 5'3'

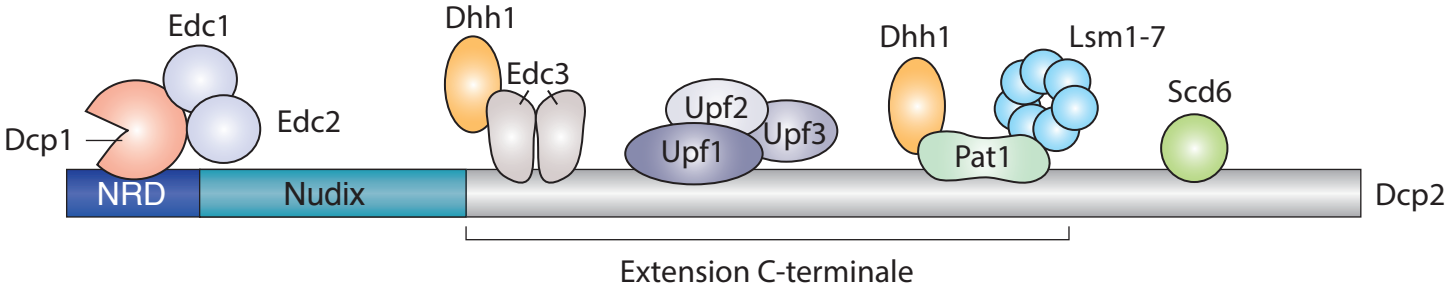
3.2.1 Le complexe d'enlèvement de la coiffe

La coiffe est un déterminant important de la stabilité des ARNm, en effet des transcrits porteurs d'une coiffe en 5' sont plus stables que ceux qui en sont dépourvus (Furuichi *et al.*, 1977). A l'instar de la queue poly(A) cette protection se doit d'être outrepassée afin d'engager la dégradation du messager en réponse à des stimuli internes ou externes. L'enlèvement de la coiffe a pour conséquence l'inhibition de la traduction et la dégradation complète et irréversible des ARNm déadénylés.

Le complexe majeur d'enlèvement de la coiffe est une holoenzyme formée par la sous-unité catalytique DCP2 et le cofacteur DCP1. DCP2 (Decapping mRNA 2) est une protéine de la superfamille conservée des hydrolases Nudix capable de cliver la liaison diphosphate entre le phosphate α et le phosphate β de la structure coiffe. Ce clivage entraine le relargage d'un ARNm avec une extrémité 5' monophosphate, un substrat idéal des exoribonucléases 5'3', et d'une guanosine diphosphate méthylée en position N7 (m^7GDP). A noter qu'il existe plusieurs enzymes additionnelles capables d'éliminer la coiffe dans des cellules humaines, comme Nudt16 et Nudt3 (Grudzien-Nogalska and Kiledjian, 2017). De plus, DXO exerce un contrôle de qualité de la coiffe en hydrolysant préférentiellement les coiffes incomplètes ou coiffe NAD+ (Grudzien-Nogalska and Kiledjian, 2017; Jiao *et al.*, 2017).

In vivo, le complexe DCP1 -DCP2 fonctionne de concert avec des facteurs qui promeuvent son activité, son accessibilité aux substrats et coordonnent l'enlèvement de la coiffe avec les processus de traduction et de déadénylation. Parmi ces facteurs figurent Pat1, Lsm1-7, Dhh1/DDX6 et Scd6/LSM14 et des protéines stimulatrices de l'enlèvement de la coiffe EDCs (Enhancer of decapping).

A



B

Levure	Homme	Plantes
Dcp1	DCP1	DCP2
Dcp2	DCP1	**DCP2**/TDT
Edc1	absent	absent
Edc2	absent	absent
Edc3	**EDC3**/LSM16	absent
Edc4 *	**EDC4**/HEDLS/Ge-1	**VCS**/ EDC4
Dhh1	**DDX6**/RCK/P54	RH6, 8, 12
Scd6 (S. cerevisiae) / Sum2 (S. pombe)	**LSM14A**/RAP55	DCP5
Lsm1-7	Lsm1-7	LSM1-7
Pat1	PAT1B	PAT1
XRN1	XRN1	XRN4
* Edc4 est présent chez la levure à fission *S. pombe* mais absent chez *S. cerevisiae*		

Figure 6 | Le complexe DCP1-DCP2 et ses co-facteurs associés
A | Représentation schématique de Dcp1 associé à Dcp2 et à ses co-facteurs associés chez la levure.
B | Tableau de correspondance des noms des différents facteurs associés au complexe de la coiffe chez l'homme, la levure et chez les plantes.

3.2.2 *Le réseau complexe d'interaction de Dcp1 -Dcp2*

Le complexe d'enlèvement de la coiffe et ses cofacteurs associés est très bien décrit chez la levure et illustre parfaitement la complexité du réseau d'interaction de Dcp1 : Dcp2 (Figure 6A).

Dcp2 interagit avec Dcp1 via son domaine NRD (She *et al.*, 2006; Charenton *et al.*, 2016) et l'activité du complexe ainsi formé est stimulée par le recrutement des stimulateurs de décapping **Edc1** et **Edc2** (Enhancer of decapping 1 & 2) par Dcp1 (Dunckley *et al.*, n.d.; Borja *et al.*, 2011). En outre, des **motifs en hélice α riches en leucines** (**HLM**, helical leucine-rich motif) présents sur une longue extension C-terminale de Dcp2 permettent le recrutement de **Pat1** (Protein associated with topoisomerase I), **Edc3** (enhancer of decapping 4), et **Scd6** (Suppressor of Clathrin Deficiency 6) (Fromm *et al.*, 2012; Charenton *et al.*, 2017; He and Jacobson, 2015a). Cette extension C-terminale assure également le recrutement du régulateur de la voie du NMD (Nonsense Mediated Decay) **Upf1** (Up Frameshift1) (He and Jacobson, 2015a). Les interactants directs de Dcp2 sont eux-mêmes impliqués dans des liaisons avec d'autres partenaires protéiques contribuant à la formation d'un réseau complexe d'interaction. Ainsi, Pat1 sert de pivot pour le recrutement de l'exoribonucléase 5'3' **Xrn1** et du complexe **Lsm1-7**, deux facteurs impliqués dans la dégradation des ARN (Bouveret *et al.*, 2000; Tharun *et al.*, 2000; Nissan *et al.*, 2010; Sharif *et al.*, 2013; Sharif and Conti, 2013) et se lie également, au même titre que Edc3 (Sharif *et al.*, 2013), au répresseur de la traduction et stimulateur de décapping, **Dhh1** (Nissan *et al.*, 2010) . Enfin, Upf1 s'associe à Upf2 et Upf3 (Up frameshift 1 & 2) (He *et al.*, 1997, p.1) pour promouvoir le décapping et la subséquente dégradation des ARNm ayant un codon stop prématuré (He and Jacobson, 2015b).

A l'exception de Edc1 et Edc2, l'ensemble des facteurs associés à DCP1 : DCP2 sont retrouvés conservés chez les eucaryotes supérieurs (Figure 6B), bien qu'il existe quelques différences notables. Chez les eucaryotes multicellulaires, DCP2 est notamment dépourvue de l'extension C-terminale présente chez la levure et c'est DCP1 qui contient un unique motif HLM, ceci implique un remodelage du réseau d'interaction de DCP1-DCP2 et de ses facteurs associés. En outre, chez les métazoaires et les plantes l'activité d'enlèvement de la coiffe requiert la présence additionnelle de **EDC4** (HEDLS/Ge-1), connu sous le nom de **VCS** (Varicose) chez la plante (Xu *et al.*, 2006; Xu and Chua, 2011; Arribas-Layton *et al.*, 2013).

La plupart des cofacteurs associés à DCP1 : DCP2 chez les eucaryotes joue un rôle crucial dans la dégradation et la traduction. La déplétion de ces facteurs entraine, dans la plupart des cas, un fort phénotype ou la létalité, à l'image de leur importance dans le métabolisme des ARN. Ces cofacteurs peuvent avoir plusieurs fonctions, ils peuvent recruter le complexe DCP1 :DCP2 sur des substrats spécifiques, à l'instar du rôle de Upf1, ou comme Edc1/Edc2/Edc3 stimuler directement son activité. D'autres cofacteurs, comme Pat1, Lsm1-7, Dhh1 (DDX6) et Scd6 (LSM14 chez l'homme, DCP5 chez Arabidopsis ont des fonctions plus complexes tel que le couplage avec la déadénylation et l'inhibition de la traduction.

3.2.3 *Couplage entre enlèvement de la coiffe et la déadénylation par Pat1*

Pat1 est une protéine centrale de la dégradation des ARN. Sa déplétion a de lourdes conséquences, puisqu'elle entraine la létalité embryonnaire chez la drosophile (Pradhan *et al.*, 2012) et un phénotype de nanisme et de feuilles dentelés chez *Arabidopsis thaliana* (Roux *et al.*, 2015). De même, l'inactivation du gène *PAT1* chez la levure conduit à un ralentissement de croissance à 30°C et à la létalité à 37°C (Bouveret *et al.*, 2000; Tharun *et al.*, 2000). Pat1 est avant tout un activateur de l'activité de DCP1-DCP2 qui **connecte le processus de déadénylation au processus d'enlèvement de la coiffe**. L'une des évidences en faveur du couplage décapping/déadénylation est l'accumulation, dans les mutants *pat1*, de transcrits coiffés et déadénylés (Bouveret *et al.*, 2000; Tharun *et al.*, 2000). De surcroît, Pat1 interagit avec Lsm1-7, un complexe protéique qui s'associe préférentiellement, *in vivo* et *in vitro* aux ARNm oligoadénylés (Tharun *et al.*, 2000; Tharun and Parker, 2001; Chowdhury *et al.*, 2007) et uridylés, (Song and Kiledjian, 2007) plutôt qu'aux ARNm polyadénylés et qui protège les extrémités 3' de la dégradation (He and Parker, n.d.). Enfin chez les métazoaires, Pat1 interagit également avec le complexe de déadénylation CCR4-Not (Haas *et al.*, 2010; Ozgur *et al.*, 2010). L'ensemble de ces éléments convergent vers un modèle où les ARNm déadénylés par CCR4-Not sont les substrats préférentiels de Lsm1-7 qui recrute Pat1. Pat1 stimule l'enlèvement de la coiffe de par son interaction avec DCP1 : DCP2 et permet également le recrutement de XRN1 pour initier la dégradation de 5' en 3.' Nous verrons également ultérieurement, que l'association de Pat1 : Lsm1-7 contribue à la dégradation des ARNm uridylés par Cid1 (Scheer *et al.*, 2016; De Almeida *et al.*, 2018).

Le rôle de Pat1 dans le métabolisme des ARN ne se restreint pas uniquement à l'activation du décapping, en effet chez *S. cerevisiae*, il a également été montré que Pat1 peut inhiber l'initiation de la traduction en réprimant la formation du complexe 48S (Nissan *et al.*, 2010). La répression de la traduction par Pat1 a été également mis en évidence chez les oocytes de Xénope (Marnef and Standart, 2010).

3.2.4 *Cofacteurs de décapping impliqués dans l'inhibition de la traduction*

Parmi les cofacteurs associés au complexe de décapping DCP1 : DCP2, deux protéines jouent également un rôle dans l'inhibition de la traduction des ARNm : DDX6/Dhh1 et LSM14/Scd6, chez l'homme et la levure, respectivement.

3.2.4.1 Dhh1/DDX6

Les protéines de la famille DDX6/Dhh1 sont des hélicases à ARN de type DEAD extrêmement bien conservées chez les eucaryotes (Ostareck *et al.*, n.d.). DDX6/Dhh1 joue un rôle prépondérant dans la régulation des ARN cytoplasmiques en étant impliquées dans l'inhibition de la traduction des ARNm, dans la formation des P bodies et de par son association avec le complexe de décapping DCP1 :DCP2.

Dans les cellule humaines, DDX6 est remarquablement abondantes puisqu'elles est en large excès molaire par rapport aux ARNm sur lesquels elles se lient sans spécificité de séquence et avec une haute affinité (Ernoult-Lange *et al.*, 2012). Elle est 7 fois plus que abondante que les ARNm dans les cellules humaines (Ernoult-Lange *et al.*, 2012). Cette caractéristique est conservée chez le Xénope (Smillie and Sommerville, 2002) et chez le trypanosome (Kramer *et al.*, 2010)

A l'instar des autres membres de la famille des hélicases à ARN, les protéines de la famille DDX6/Dhh1 possèdent deux domaines globulaires de type RecA (RecA-like, RecA1 et RecA2) connectés via une région flexible. Ces derniers sont composés de différents motifs conservés impliqués dans la liaison à l'ARN, l'hydrolyse de l'ATP et la coordination de l'activité hélicase (Sloan and Bohnsack, 2018).

Chez la levure Dhh1 a été initialement décrit comme un facteur de dégradation des ARNm dont la déplétion conduisait à la stabilisation d'intermédiaires de dégradation dépourvus de queue poly(A) et coiffés (Coller *et al.*, 2001; Fischer and Weis, 2002). Cette idée était soutenue par le fait que Dhh1 fait partie du complexe de décapping et interagit avec le complexe de déadénylation Ccr4-not (Maillet and Collart, 2002).

Des études plus tardives ont montrés que Dhh1 est un inhibiteur de la traduction et que son impact sur la dégradation des ARNm était en réalité dépendant du statut traductionnel des transcrits considérés (Coller and Parker, 2005). Le rôle de Dhh1 en tant que répresseur de la traduction a également été confirmé par la fonction de ses orthologues chez les eucaryotes multicellulaire. Ainsi, chez le Xénope, Xp54, a été identifiée comme un composant de complexes ribonucléoprotéiques (mRNPs) traductionnellement inactifs (Ladomery *et al.*, 1997) et chez la drosophile Me31B a été impliquée dans la répression traductionnelle des ARNm *Oskar* lors de leur transport des cellules nourricières vers l'oocyte durant l'oogenèse (Nakamura *et al.*, 2001). Chez l'homme, l'inactivation de DDX6 entraine une dérépression générale de la traduction des ARNm (Chu and Rana, 2006) . Il a, de ce fait, été rapidement suggéré que le rôle de DDX6/Dhh1 dans la dégradation des ARNm, était une conséquence indirecte de sa fonction dans l'inhibition de la traduction.

Des expériences d'ancrage ou « tethering » révèlent que la capacité de Dhh1 à inhiber la traduction d'un ARNm cible est indépendante du décapping, de la déadénylation ou de la présence de eIF4E et eIF3 et qu'elle intervient après l'initiation de la traduction en provoquant une augmentation drastique du nombre de ribosome sur le messager cible (Sweet *et al.*, 2012). De manière intéressante Radhakrishnan et al. montrent également que la protéine s'associe préférentiellement aux ARNm ayant une faible optimalité de codons et provoque leur dégradation. Les auteurs suggèrent que Dhh1 serait le médiateur entre vitesse d'élongation de la traduction et dégradation (Radhakrishnan *et al.*, 2016).

DDX6 est également impliquée dans le silencing génique par le complexe RISC (RNA-induced silencing Complex) chez les animaux (Iwakawa and Tomari, 2013).

A noter que chez *Arabidopsis thaliana* il existe 3 protéines orthologues à DDX6, RH6, 8 et 12. Leur fonction n'a en revanche pas été caractérisée chez la plante, mais la forte homologie de séquence avec la protéine humaine (87%, 86% et 86% de similarité, respectivement) suggère également une

homologie de fonction. Ces hélicases sont également localisées dans les P-bodies chez Arabidopsis (Bhullar *et al.*, 2017; Chicois *et al.*, 2018).

3.2.4.2 Scd6/LSM14

Les protéines de la famille Scd6/LSM14 et leur organisation en domaines sont extrêmement bien conservées chez les organismes eucaryotes. Cette organisation est caractérisée par la présence en région N-terminale, d'un domaine LSm (« Sm-like ») et en région C-terminale d'un domaine FDF (phénylalanine, acide aspartique, phénylalanine) et de 1 ou 2 domaines composés de motifs répétés RG/RGG (argine-glycine ou arginine-glycine-glycine) (Roy and Rajyaguru, 2018).

A l'instar de Dhh1/DDX6, Scd6/LSm a été initialement décrit comme un activateur de décapping du fait de son interaction avec le complexe DCP1-DCP2 chez *S. pombe* et *S. cerevisiae.* Cette interaction est également observée chez la plante entre DCP5 et DCP1-DCP2 (Xu and Chua, 2011) et chez la drosophile entre Tral (Trailer Hitch) et DCP1. Il existe cependant peu d'évidences montrant que Scd6/LSM14 stimulent directement l'activité du complexe décapping *in vitro* ou *in vivo* (Roy and Rajyaguru, 2018) et la protéine semble principalement impliquée dans l'inhibition de la traduction, à l'instar de Dhh1.

Le rôle principal de Scd6/LSM14 dans l'inhibition de la traduction est soutenu par des expériences *in vitro* et *in vivo* menée chez plusieurs organismes. Chez *S. cerevisiae* Scd6 se lie à eIF4G via son domaine RGG et réprime la traduction en inhibant la formation du complexe 48S *in vitro* (Rajyaguru *et al.*, 2012). Chez le Xénope, des expériences de « tethering » montre que xRAP55 inhibe la traduction d'un ARN reporter *in vivo* (Tanaka *et al.*, 2006). Enfin l'homologue chez Arabidopsis DCP5 inhibe la traduction de l'ARNm OLEO1 *in vivo* et *in vitro* (Xu and Chua, 2009).

Une autre évidence forte en faveur d'un rôle majeur de Scd6/LSM14 dans l'inhibition de la traduction est son interaction conservée avec Dhh1/DDX6 (voir paragraphe précédent). Cette interaction est retrouvée conservée chez S. cerevisiae (Decourty *et al.*, 2008), *C. elegans* (Ko *et al.*, 2013), la drosophile (Wilhelm *et al.*, 2005), le Xénope (Tanaka *et al.*, 2006), la plante (Chicois *et al.*, 2018) et l'homme (Brandmann *et al.*, 2018). Chez la drosophile, le Xénope et l'homme cette interaction est permise par la domaine FDF ou FDF et RGG.

Scd6/LSM14 est localisée au niveau des P-bodies chez l'ensemble des organismes eucaryotes à l'exception de *S. pombe* chez qui la localisation de Scd6 n'a pas été démontrée (Roy and Rajyaguru, 2018). Chez l'homme la protéine est requise, au même titre que DDX6 pour leur formation (Ayache *et al.*, 2015).

3.2.4.3 Couplage entre silencing, inhibition de la traduction et dégradation des ARNm par DDX6

Les miARN peuvent réprimer un ARN cible grâce à deux mécanismes indépendants, le blocage de la traduction et la dégradation. La dégradation de l'ARNm peut être induite par un clivage endonucléolytique via RISC. Ce mode d'action de RISC est principalement observé chez les plantes, mais l'inhibition de traduction induite par miARN est également fréquente (Yu *et al.*, 2017), Chez

l'animal les petits ARN reconnaissent leur cible en s'hybridant de manière imparfaite ce qui prévient le clivage par RISC. Le silencing post-transcriptionnel de RISC s'opère donc par le recrutement de protéines effectrices qui conduisent d'une part à l'inhibition de traduction du messager, d'autre part à sa déadénylation aboutissant à sa dégradation (Iwakawa and Tomari, 2015).

La liaison du miARN à sa cible entraine la dissociation des PABP et la déadénylation par les complexes PAN2-PAN3 et CCR4-NOT. Les deux complexes sont recrutés à l'ARN via TNRC6/GW182 (Fabian *et al.*, 2011; Braun *et al.*, 2011; Chekulaeva *et al.*, 2011; Zekri *et al.*, 2013), une protéine clé de ce mécanisme de silencing puisqu'elle interagît directement avec Ago. La sous-unité CNOT1 du complexe CCR4-NOT interagît également avec DDX6 (Chen *et al.*, 2014, p.6; Mathys *et al.*, 2014; Rouya *et al.*, 2014), ce qui permet le recrutement des autres membres du complexe de décapping PAT1, LSM14, EDC4, DCP1 et DCP2. De par son domaine d'interaction avec eIF4E, 4E-T participe également au silencing des ARNm en connectant le complexe de dégradation formé en 3' à la coiffe en 5'. La protéine est en effet capable d'interagir avec CNOT1, PATL1, LSM14, DDX6, LSM2 et DCP1. *In fine*, la présence du complexe de décapping et de ces cofacteurs associés favorise l'enlèvement de la coiffe et la dégradation de 5' en 3' (Iwakawa and Tomari, 2015).

La liaison imparfaite du miARN à sa cible entraine également l'inhibition de la traduction. Plusieurs éléments pourraient contribuer à ce phénomène de répression. D'une part la dissociation des PABP par GW182 pourrait empêcher la liaison PABP-EIF4G ce qui compromettrait l'initiation de la traduction. D'autre part elle pourrait être la conséquence du recrutement des répresseurs de la traduction 4E-T et DDX6. Certaine études semblent également suggérer que la liaison du miARN à sa cible favoriserait la dissociation de eIF4A (Fukao *et al.*, 2014).

3.2.5 La dégradation exoribonucléolytique 5'3' par XRN

3.2.5.1 Les protéines XRN

Le clivage de la coiffe par DCP2 produit des ARNm avec une extrémité libre 5' monophosphate, un substrat idéal pour les exoribonucléases 5'3'. Chez la plupart des eucaryotes la dégradation des ARN avec une extrémités 5'P est permise par deux exoribonucléases , **XRN1** (Pacman chez la drosophile) qui est principalement cytoplasmique et XRN2 (Rat1 chez la levure) qui est nucléaire (Jones *et al.*, n.d.). XRN1 est responsable de la dégradation des ARNm décoiffés ou des fragments 3' issus de clivages endonucléolytiques. La structure de XRN1 chez la drosophile révèle que le groupement 5' phosphate de l'ARNm est reconnu grâce à des interactions électrostatiques avec une poche électropositive, contenue dans le domaine nucléase, et formée de 4 résidus hautement conservés. Cette poche étroite exclut les groupements 5' terminal plus larges comme les groupements triphosphates ou les extrémités coiffés (Jinek *et al.*, 2011). Les extrémités 5'OH

constituent de mauvais substrats pour XRN1, ce qui suggère que le phosphate en lui-même est également important pour ce mécanisme de reconnaissance par XRN1 (Pellegrini *et al.*, 2008).
XRN1 est une enzyme extrêmement processive qui est uniquement freinée par la présence de structures secondaires particulières. A titre d'exemple, le virus de l'hépatite C (HCV) contient une structure dans sa 5'UTR qui est résistante à l'activité de XRN1 (Moon *et al.*, 2015). De même, XRN1 est également bloquée par la présence d'un pseudonoeud formant une jonction à 3 hélices dans la 3'UTR des flavivirus (Chapman *et al.*, 2014) et par la séquence coremin en 3'UTR chez le virus des nervures jaunes et nécrotique de la betterave sucrière (BNYVV) (Flobinus *et al.*, 2018).
Chez Arabidopsis il n'existe aucun orthologue à XRN1 mais ces dernières possèdent 3 homologues à XRN2. XRN2 et XRN3 sont localisés dans le noyau, et c'est XRN4 qui fait office d'exoribonucléase cytoplasmique (Kastenmayer and Green, 2000). Tout comme XRN1, XRN4 dégrade les fragments 3' issus du clivage par RISC (Rymarquis *et al.*, 2011) et les transcrits décoiffés (Gregory *et al.*, 2008). Les plantes mutantes *xrn4* ne montrent que très peu d'altérations morphologiques dans des conditions standards de culture, mais présentent cependant un phénotype d'insensibilité à l'éthylène. En effet, la mutation de XRN4 conduit à une accumulation des messagers EBF1 et EBF2 (EIN3-binding F-box protein1) qui encodent pour des protéines F-BOX impliquées dans la dégradation de la protéine EIN3, un régulateur transcriptionnel de la voie de signalisation de l'éthylène (Potuschak *et al.*, 2006; Olmedo *et al.*, 2006). De plus, la dégradation par XRN4 est impliquée dans la dormance chez la graine (Basbouss-Serhal *et al.*, 2017), la réponse et la tolérance au stress thermique (Merret *et al.*, 2013; Nguyen *et al.*, 2015) et la réponse antivirale (Cheng *et al.*, 2007; Jaag and Nagy, 2009; Peng *et al.*, 2011; Lee *et al.*, 2016; Jiang *et al.*, 2018).

3.2.5.2 Évidence d'une dégradation co-traductionnelle par les protéines XRN

Il apparaît de plus en plus clairement que la dégradation d'une large proportion des ARNm par les protéines XRN a en réalité lieu co-traductionnellement. Cette dégradation co-traductionnelle 5'3' mise en évidence chez *Saccharomyces cerevisiae* (Hu *et al.*, 2009; Pelechano *et al.*, 2015) existe aussi chez *S. pombe* ou Arabidopsis (Pelechano *et al.*, 2015; Yu *et al.*, 2016). Le séquençage des intermédiaires de dégradation ayant une extrémité 5'P révèlent une périodicité de 3 nucléotides chez *S. cerevisiae* et indique que l'exoribonucléase XRN1 suit le dernier ribosome en cours de traduction (Pelechano *et al.*, 2015). Des travaux similaires menées chez Arabidopsis, le riz et le soja révèlent également l'omniprésence d'intermédiaires de dégradation avec une périodicité de 3 nucléotides dans l'ORF, ce qui suggère qu'une partie de la dégradation effectuée par XRN4 chez les plantes a lieu co-traductionnellement (Yu *et al.*, 2016). La dégradation co-traductionnelle des ARNm pourrait notamment contribuer à l'ajustement rapide du transcriptome en cas de bouleversements environnementaux. Lors d'un stress thermique chez Arabidopsis, XRN4 contribue à la dégradation des ARNm engagés et non engagés dans les polysomes. Une partie de la dégradation des ARNm engagé dans les polysomes est permise par la protéine LARP1 (LA-related protein) qui s'associe à XRN4 et contribue à son association dans les polysomes (Merret *et al.*, 2013).

La dégradation co-traductionnelle par XRN4 cible notamment les ARNm sur lesquels les ribosomes sont bloqués en traduction suite au stress thermique (Merret *et al.*, 2015).

3.3 La dégradation dans le sens 3'5'

Après déadenylation la dégradation des ARNm dans le cytoplasme peut également se faire dans le sens 3'5' par l'exosome ou l'exoribonucléase DIS3L2/SOV.

3.3.1 L'exosome et le complexe SKI

L'exosome est un **complexe multiprotéique** ayant une **activité endoribonucléolytique et exoribonucléolytique 3'5'** et est conservé chez les eucaryotes. Le cœur de l'exosome est composé de 9 sous-unités (**Exo9**): 6 protéines de type RNase PH (Rrp41, Rrp42, Rrp43, Rrp45, Rrp46, et Mtr3) et 3 protéines de liaison à l'ARN (Rrp4, Rrp40, et Csl4) (Lange and Gagliardi, 2011). L'Exo9 forme une structure en tonneau dont le centre renferme un canal assurant le passage des ARN simple brin. L'architecture de l'exosome chez les eucaryotes est similaire à celle de l'exosome présent chez les archées ainsi qu'à la RNase PH et la polynucléotide phosphorylase (PNPase), deux enzymes phosphorolytiques bactériennes. Ces dernières sont des exoribonucléases ayant une **activité phosphorolytique processive** assurée par 3 sites catalytiques situés au sein du canal central (Makino *et al.*, 2013; Schneider and Tollervey, 2013; Januszyk and Lima, 2014; Evguenieva-Hackenberg *et al.*, 2014; Zinder and Lima, 2017). Chez l'homme et la levure des mutations d'acides aminés du site de coordination des phosphates rend **l'Exo9 catalytiquement inactif** et l'activité de l'exosome est assurée par l'activité hydrolytique de RNases qui lui sont physiquement associées, ces enzymes comprennent **hRRP6/ Rrp6p, et hDIS3/Rp44p**. hDIS3/Rp44p possède une activité endoribonucléolytique et 3'5' exoribonucleolytique processive, et hRRP6/Rrp6p une activité 3'5' exoribonucléolytique distributive, les deux protéines associées à Exo9 forment **Exo11**.

La composition générale de l'exosome est conservée chez *Arabidopsis thaliana* (Chekanova *et al.*, 2007). A noter cependant la duplication de deux gènes codant pour les sous-unités de Exo9, RRP40 et RRP45, ainsi que l'existence de 3 homologues à RRP6 (RRP6L1, RRP6L2, RRP6L3) (PMID:18160042) Lange et al 2008). RRP6L1 et RRP6L2 ont une localisation nucléaire et RRP6L3 est quant à elle retrouvée dans le cytosol (Lange *et al.*, 2008). Aucun de ces homologues RRP6L n'a pour l'instant été retrouvé associé physiquement à Exo9 dans des expériences d'immunoprécipitation de l'exosome (Lange *et al.*, 2014, p.4) et à ce jour seul RRP6L2 a été impliqué dans des fonctions moléculaires de l'exosome, à savoir la maturation des ARNr (Lange *et al.*, 2008; Sikorski *et al.*, 2015). Exo9 s'associe aussi à un homologue de RRP44 chez Arabidopsis (Lange et al 2014) (Kumakura *et al.*, 2013; Lange *et al.*, 2014). A l'instar de la situation chez la levure et l'homme, l'activité catalytique de l'exosome d'Arabidopsis implique RRP44 et RRP6L2. Cependant, une caractéristique remarquable de l'exosome d'Arabidopsis est la présence d'une activité phosphorolytique de Exo9, conférée par la sous-unité RRP41 (Sikorska et al 2017). Les acides

aminés requis pour l'activité de RRP41 sont strictement conservés chez les plantes (Sikorska *et al.*, 2017).

In vivo l'exosome interagit avec des cofacteurs qui promeuvent son activité et/ou la reconnaissance de ses substrats. A ce titre, l'activité de l'exosome dans le cytoplasme dépend du complexe SKI, un assemblage d'un dimère de SKI8 avec l'hélicase à ARN SKI2, tous deux connectés par une interaction mutuelle avec SKI3 (Synowsky and Heck, n.d.; Halbach *et al.*, 2013). Chez *S. cerevisiae* le complexe SKI est également composé de Ski7p qui sert d'intermédiaire d'interaction entre SK2/SKI3/SKI8 et l'exosome (Araki *et al.*, 2001; Wang *et al.*, 2005; Halbach *et al.*, 2013; Kowalinski *et al.*, 2016). Aucune séquence homologue à Ski7p n'est retrouvée chez l'homme et la plante pour lesquels la fonction de ski7p semble être assurée par un variant d'épissage de HBS1, une protéine de la famille de eRF3 (Saito *et al.*, 2013; Zhang *et al.*, 2016; Kowalinski *et al.*, 2016; Kalisiak *et al.*, 2017; Marshall *et al.*, 2018).

Dans le cytoplasme l'exosome et le complexe SKI est notamment requis pour la dégradation des ARNm via les processus de surveillance NMD, NGD (**N**o-**G**o Decay) et NSD (**N**o-**S**top **D**ecay). Ils contribuent également à la dégradation générale des ARNm, des ARNm d'histones et à la défense antivirale (Łabno *et al.*, 2016; Houseley *et al.*, 2006; Chekanova *et al.*, 2007; Chlebowski *et al.*, 2013). En outre, SKI2, SKI3 et SKI8 sont impliquées dans la dégradation des fragments 5' après clivage des ARNm par RISC (Branscheid *et al.*, 2015).

3.3.2 L'exoribonucléases 3'5', DIS3L2/SOV

En plus de l'exosome, la dégradation de 3' en 5' dans le cytoplasme peut être effectuée par DISL2/SOV. DIS3L2 et ses homologues chez la plante (RRP44B/SOV) et chez *S. pombe* (Dis3L2) appartiennent à la même famille que Dis3, le co-facteur de l'exosome, mais ne possèdent pas le domaine PIN impliqué dans la liaison à l'exosome. Ces exoribonucléases 3'5' cytoplasmiques agissent indépendamment de l'exosome (Zhang *et al.*, 2010; Lubas *et al.*, 2013; Malecki *et al.*, 2013). DIS3L2 dégrade une large variété d'ARN codants et non codants dans le cytoplasme chez la levure à fission *S. pombe* (Malecki *et al.*, 2013), Arabidopsis (Zhang *et al.*, 2010; Sorenson *et al.*, 2018) et l'homme (Lubas *et al.*, 2013). Elle est, en revanche, absente chez *S. cerevisiae*. Nous verrons ultérieurement qu'elle dégrade préférentiellement les ARN avec une extrémité 3' uridylée (données résumées dans De Almeida et al., 2018; Scheer et al., 2016).

3.4 Mécanisme de surveillance des ARNm dans le cytoplasme

Les cellules eucaryotiques contiennent des mécanismes de surveillance des ARNm qui préviennent l'accumulation de transcrits non fonctionnels et dont la traduction conduirait à la production de protéines aberrantes et potentiellement toxiques pour la cellule. Il existe trois principales voies de surveillances des ARNm, la voie du **NMD** (**N**on**s**ense **M**ediated **D**ecay), du **NGD** (**N**o-**G**o **D**ecay) et du **NSD** (No-stop mediated decay). Chacune de ces trois voies a pour caractéristique commune d'assurer la reconnaissance, lors de la traduction, d'ARNm défectifs, et d'assurer leur élimination.

Le NMD a été initialement décrit comme la voie impliquée dans la dégradation spécifique des ARNm contenant un codon stop prématuré (PTC) et dont la traduction conduirait à la production de protéines tronquées. Le rôle du NMD s'est depuis lors considérablement élargi puisque des analyses à l'échelle du génome révèlent que le NMD cible également des transcrits « normaux », c.-à-d. qui ne présentent pas de défauts apparents et qui codent pour des protéines fonctionnelles (Karousis *et al.*, 2016). De manière générale, le NMD est impliqué, de manière directe ou indirecte, dans la régulation de l'expression de ~3 à 20% des transcrits chez la levure (Guan *et al.*, 2006; He *et al.*, 2003), *C. elegans* (Ramani *et al.*, 2009), la drosophile (Metzstein and Krasnow, 2006), l'homme (Mendell *et al.*, 2004; Tani *et al.*, 2012) et les plantes (Kurihara *et al.*, 2009; Rayson *et al.*, 2012). Certaines caractéristiques semblent favoriser le NMD, comme la présence d'une uORF (upstream ORF), d'une longue 3'UTR ou la présence d'intron dans la 3'UTR (He and Jacobson, 2015b).

La reconnaissance des substrats du NMD est principalement permise par l'association des trois protéines UPF1, UPF2 et UPF3 (Shaul, 2015; Nasif *et al.*, 2018). UPF1 joue un rôle central puisqu'elle est capable d'interagir avec les facteurs de terminaison de la traduction eRF1 et eRF3. L'activation du NMD implique la phosphorylation de UPF1 par SMG1 et résulte sur la dissociation de eRF1 et ERF3 et des ribosomes. *In fine*, la dégradation de l'ARNm est permise par l'activité endonucléolytique de **SMG6** chez l'homme et par le recrutement du complexe d'**enlèvement de la coiffe** par UPF1 phosphorylée. Deux facteurs additionnels SMG5 et SMG7 du NMD sont également capable de recruter le **complexe de déadénylation** CCR4-NOT (Shoemaker and Green, 2012; He and Jacobson, 2015b; Hug *et al.*, 2016; Karousis *et al.*, 2016).

Le **NSD** reconnait et dégrade les ARNm dépourvus de codon stop. L'absence de codon stop peut être la conséquence d'une polyadénylation prématurée du messager ou d'un clivage endonucléolytique. La traduction de la queue poly(A) en poly-lysines ou l'arrivée du ribosome à l'extrémité 3' du messager défectif entraine un blocage du ribosome sur l'ARNm. Le modèle actuel chez la levure implique une reconnaissance du ribosome par Ski7, ce qui induit le recrutement du complexe Ski, de l'exosome et la subséquente dégradation 3'5' du transcrit. Deux autres facteurs, Hbs1 et Dom34 sont impliqués dans la NSD chez la levure (Fourati and Graille, 2014). Chez les mammifères, la voie du NSD est permise par les homologues de Hbs1 et Dom34, HBS1 et PELOTA. Dom34-hbs1/PELOTA-HBS1 promeuvent la dissociation du complexe d'élongation de la traduction, le relargage de la chaine peptidique et de l'ARNm qui seront dégradés par le protéasome et l'exosome, respectivement (Graille and Séraphin, 2012; Inada, 2013; Inada, 2017; Simms *et al.*, 2017). Chez les plantes, des homologues de Pelota, Hbs1 and SKI2 sont aussi requis pour le NSD (Roy and Rajyaguru, 2018).

Enfin la voie du **NGD** est quant à elle activée lorsqu'un ribosome est bloqué sur l'ARNm du fait de la présence d'une structure secondaire, d'un codon rare ou d'une base endommagée. A l'instar du NSD, le NGD utilise Dom34 et Hbs1 pour permettre la dissociation du ribosome et de la chaine peptidique naissante. Le NGD résulte sur un clivage endonucléolytique par un facteur non encore identifié (Graille and Séraphin, 2012; Inada, 2013; Inada, 2013; Szádeczky-Kardoss *et al.*, 2018).

3.5 L'uridylation des ARNm : une nouvelle acteur majeur de la dégradation des ARNm :

L'uridylation des ARNm est un processus conservé chez les eucaryotes, mis à part la levure *S. cerevisiae*. L'uridylation qui joue un rôle fondamental dans le métabolisme des ARN, en favorisant notamment la dégradation 5'3' et 3'5' des ARNm. L'uridylation peut également avoir des fonctions alternatives chez certains organismes, à des stades particuliers de développement ou dans le cas d'ARNm spécifiques. L'ensemble des rôles de l'uridylation dans le métabolisme des ARN a fait l'objet de trois articles de revue dans Trends in Genetics en 2016, dans WIREs RNA en 2018 et enfin dans Philosophical Transactions of the Royal Society B, ce dernier article étant sous presse. Je signe ces trois articles en co-premier auteur, en second auteur et co-premier auteur, respectivement :

3.5.1 *Uridylation Earmarks mRNAs for Degradation …and More.*

Scheer H*, Zuber H*, De Almeida C, Gagliardi D.

Trends Genet. 2016 Oct;32(10):607-619. doi: 10.1016/j.tig.2016.08.003.

Dans ce premier article de revue, nous intégrons les connaissances relatives aux rôles de l'uridylation dans le métabolisme des ARNm. Nous décrivons la famille des TUTases impliquées dans l'uridylation des ARNm. Nous insistons bien sûr sur le rôle principal de l'uridylation en tant que signal de dégradation des ARNm chez les eucaryotes. Mais nous soulignons les différences qui semblent exister entre les conséquences de l'uridylation des ARNm chez les plantes et les mammifères. Nous rapportons également des rôles alternatifs possibles de l'uridylation, notamment dans le contrôle de la traduction dans les mitochondries de trypanososmes ou de la dissociation des polysomes des ARNm ciblés par la voie de dégradation des ARNm du Nonsense Mediated Decay (NMD) chez *Aspergillus nidulans*.

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

De Almeida C, Scheer H, Zuber, Gagliardi D.

WIREs RNA 2018, 9:e1440. doi:10.1002/wrna.1440

Cette seconde revue traite de l'uridylation des ARN de manière générale et inclus les connaissances actuelles relatives à l'uridylation des ARNm mais également de différents types d'ARN non codants. L'uridylation est en effet également très largement impliquée dans les processus de maturation et dégradation des ARN non codants chez les eucaryotes.

3.5.3 *RNA uridylation and decay in plants*

De Almeida* C, Scheer* H, Gobert A, Fileccia V, Martinelli F, Zuber H and Gagliardi D

Accepté pour publication dans Philosophical Transactions of the Royal Society B

Ce troisième article est une synthèse des travaux sur les conséquences de l'uridylation sur la dégradation des ARN chez les plantes. Ces données ont été principalement obtenues chez *Arabidopsis thaliana* et dans une moindre mesure chez l'algue verte *Chlamydomonas reinhardtii*. Deux TUTases, responsables de l'uridylation des petits ARN et des ARNm, HESO1 et URT1, respectivement ont été identifiées chez Arabidopsis. Dans le cadre de cet article, j'ai analysé l'histoire évolutive de ces deux TUTases en sélectionnant des espèces représentatives de l'ensemble des plantes. Cette analyse est aussi intégrée dans la partie résultat. Elle démontre que URT1 et HESO1 forment chacune un groupe monophylétique conservée chez la plupart des espèces. A la suite de cette histoire évolutive des TUTases, nous détaillons les rôles respectifs et communs de HESO1 et URT1 dans la dégradation des ARN chez les plantes.

Depuis la publication de ces articles de revues, trois nouvelles études portantes sur le rôle de l'uridylation des ARNm ont été publiées.

- L'uridylation par les homologues de TUT4/TUT7 est requise pour les stades de développement précoces de développement chez le poisson zèbre et *X. laevis* (Chang *et al.*, 2018)

Dans des travaux menés dans des embryons chez le poissons zèbre et *X. laevis*, Chang et al. montrent que l'uridylation par TUT4/TUT7 est induite au début de la transition maternelle-zygotique (4-6h après fertilisation) et permet l'élimination ciblée d'ARN maternels. L'inactivation de TUT4/TUT7 prévient la dégradation des ARNm et induit des défauts développementaux lors de la gastrulation.

- L'uridylation par TUT4/TUT7 empêche la rétrotransposition de LINE-1 chez l'homme (Warkocki *et al.*, 2018)

Warkocki et al. montrent que chez l'homme TUT4/TUT7 coopèrent avec MOV10, une hélicase à ARN et facteur de restriction des rétrovirus et rétrotransposons pour prévenir la rétrotransposition des éléments LINE-1 chez les humains. Les auteurs proposent que l'uridylation pourrait restreindre la rétrotransposition de LINE-1 en inhibant sa transcription inverse après réimportation dans le noyau et en déstabilisant les ARNm LINE-1 dans le cytoplasme.

- L'uridylation des ARN viraux constitue une nouvelle voie du système immunitaire innée chez *C. elegans* et l'homme (Pen *et al.*, 2017)

La TUTase CDE-1 a été identifiée dans un crible génétique visant à trouver des nouveaux gènes impliqués dans la défense antivirale chez *C. elegans*. Le Pen et al. montrent que CDE-1 uridyle les extrémités 3' de l'ARN génomique du virus Orsay (OrV) et proposent qu'en se faisant elle promeut leur dégradation par l'exosome et par XRN1 et XRN2, les deux protéines paralogues au

exoribonucléases de type XRN chez *C. elegans*. Les auteurs montrent également que les TUTases humaines, TUT4 et TUT7 uridylent l'ARNm du virus l'influenza A (IAV) et que leur délétion conduit à l'accumulation du messager viral. Cette étude révèle donc que les TUTases sont des facteurs de restriction virale. Les TUTases étant maintenant considérées comme des acteurs à part entière de la machinerie de dégradation, il ne s'agit peut-être pas d'une nouvelle voie *per se* de restriction virale chez les eucaryotes, mais cette étude rapporte la première implication de l'uridylation dans ce processus.

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).

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].

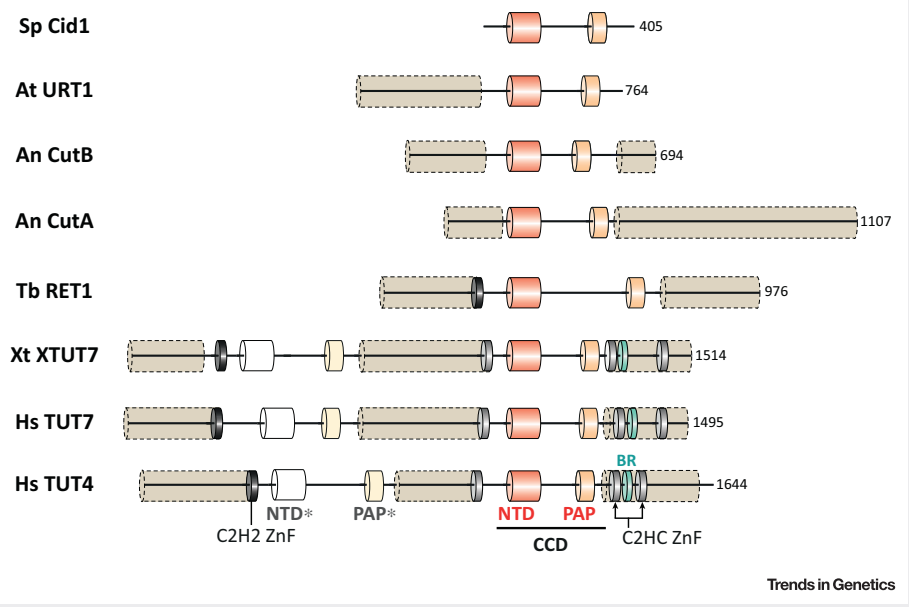


Figure I. Domain Organization of Terminal Uridylyltransferases (TUTases) Uridylating mRNAs. Nucleotidyltransferase domain (NTD) is shown in red, poly(A) polymerase-associated domain (PAP) in orange. Together, the NTD and PAP form the catalytic core domain (CCD). Nonfunctional NTD and PAP domains are 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.

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

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.

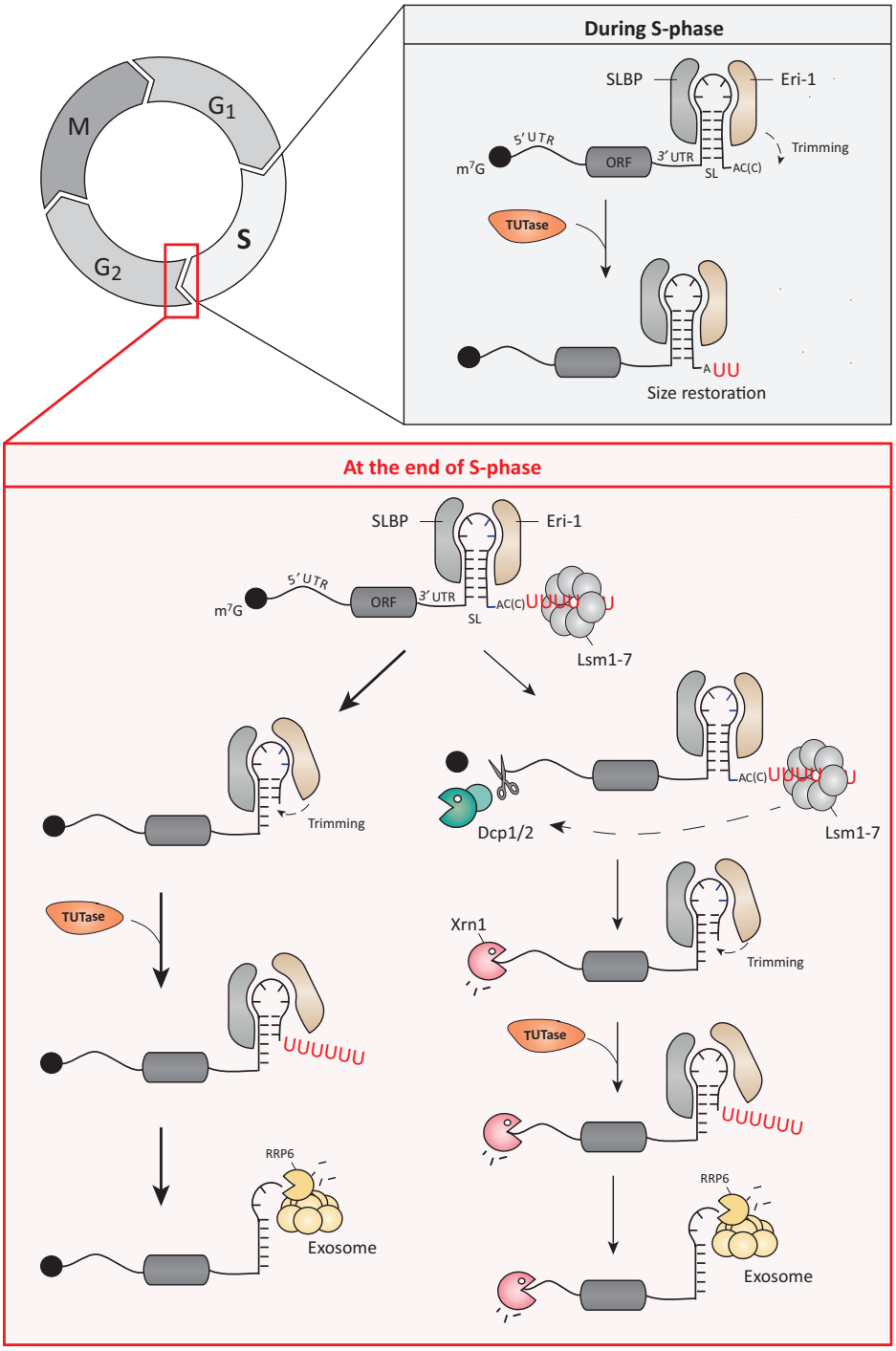
S-phase: S-phase (synthesis phase) is the phase of the cell cycle during which chromosome replication occurs. Histone mRNA levels increase considerably because the production of new histones is required for nucleosome assembly.

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.

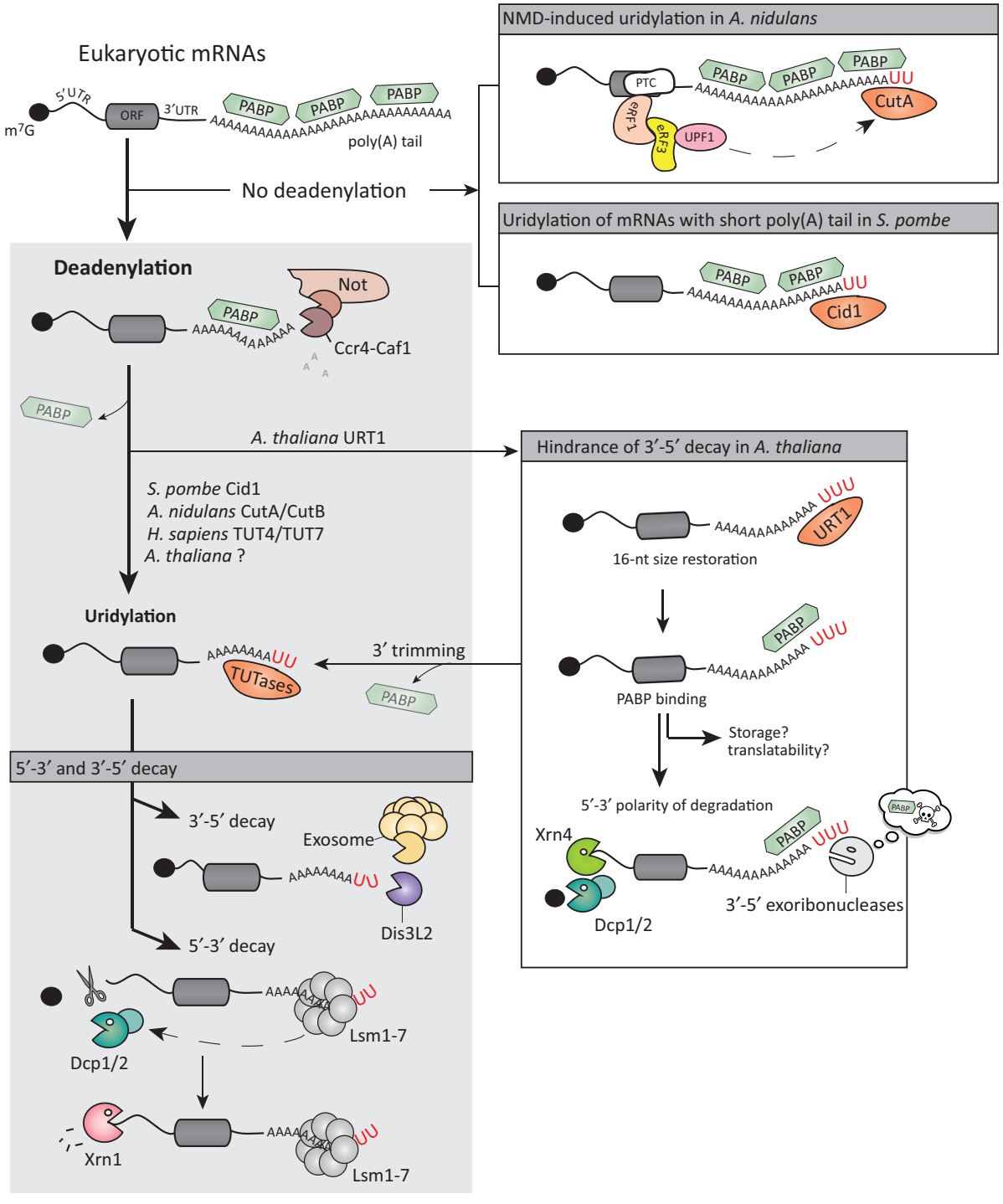
The next major breakthrough toward the realization that uridylation is an integral step of mRNA decay was the discovery that uridylation elicits the degradation of replication-dependent histone mRNAs in humans (Figure 1). Upon inhibition of DNA replication or at the end of the **S-phase**, histone mRNAs 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.

CellPress

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.

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.

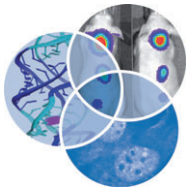
References

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/rna.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

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.

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.

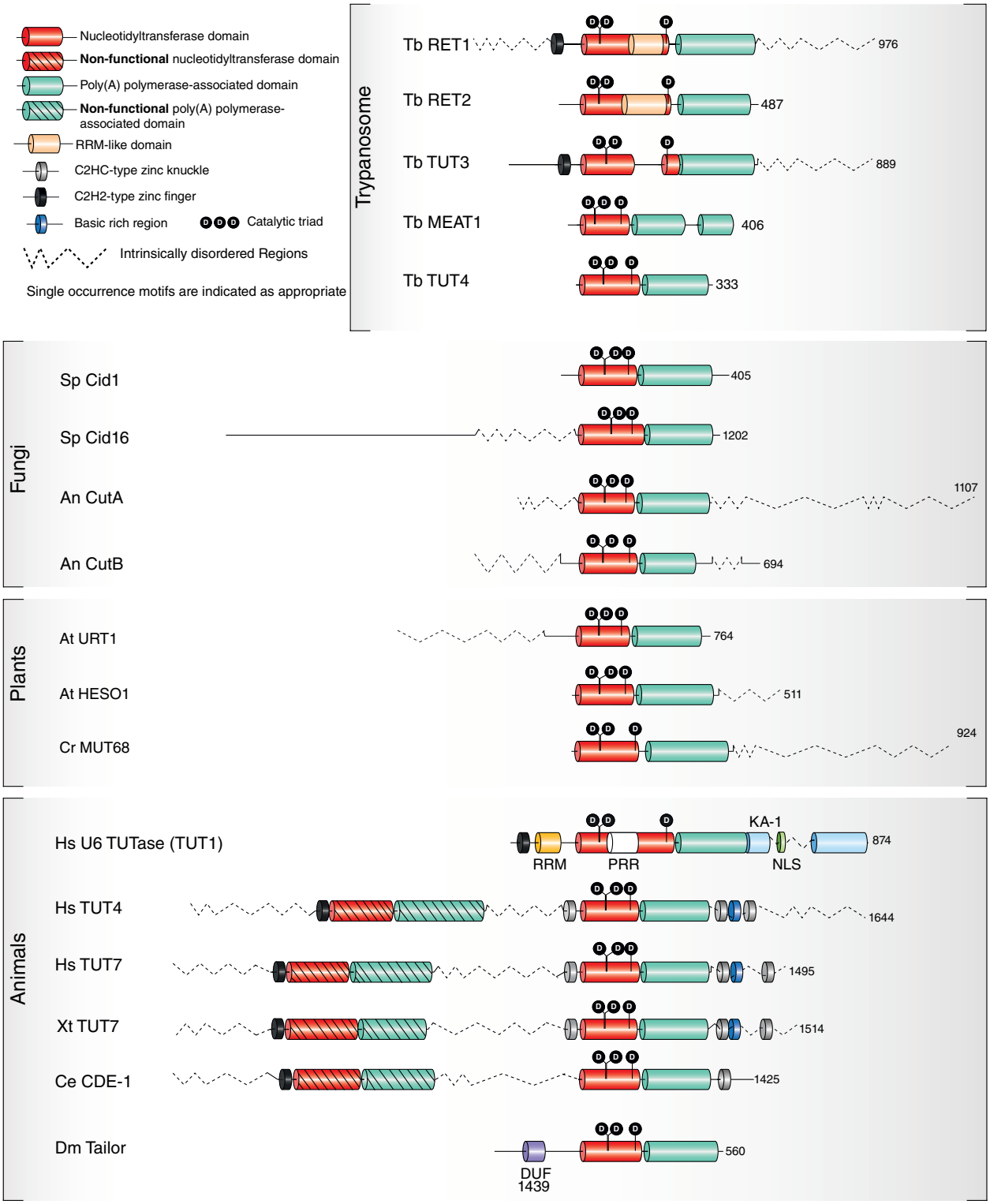


FIGURE 1 | Domain architectures of TUTases are diverse across organisms. Tb, *Trypanosoma brucei*; Sp, *Schizosaccharomyces pombe*; Hs, *Homo sapiens*; An, *Aspergillus nidulans*; At, *Arabidopsis thaliana*; Cr, *Chlamydomonas reinhardtii*; Xt, *Xenopus tropicalis*; Ce, *Caenorhabditis elegans*; Dm, *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.

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

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*)

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

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]

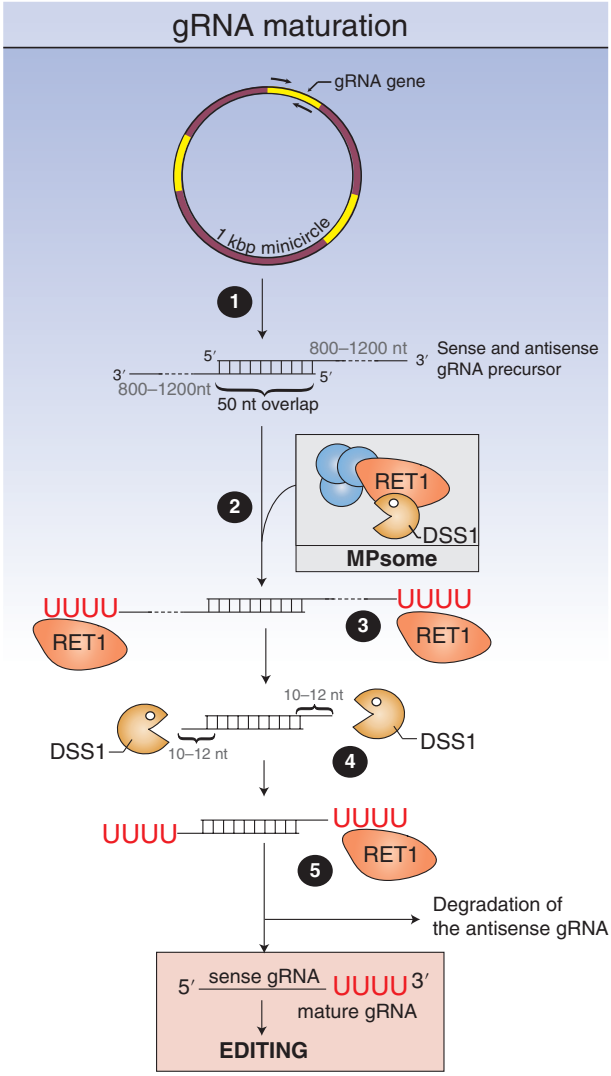


FIGURE 2 | Uridylation and guide RNA (gRNA) maturation in trypanosome mitochondria. gRNAs are processed through a sequential maturation process by the mitochondrial 3′ processome (MPsome), containing the TUTase RET1 in complex with the 3′→5′ exoribonuclease DSS1. (1) gRNAs are generated by bidirectional 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

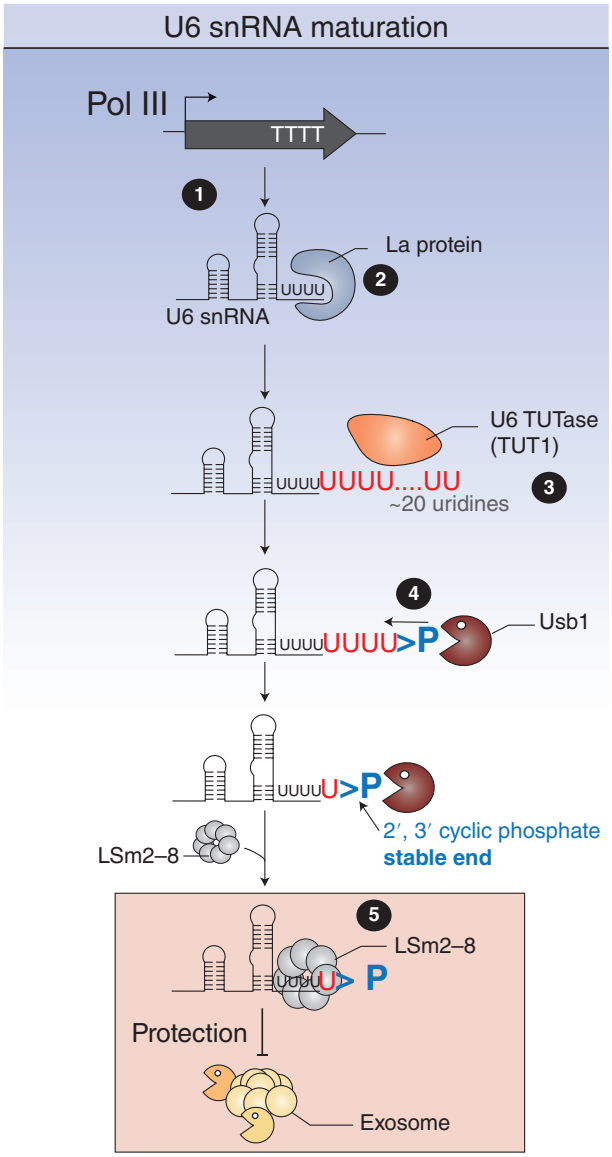


FIGURE 3 | Uridylation is critical for U6 snRNA maturation. U6 snRNAs are transcribed by polymerase III (Pol III) which terminates transcription by a stretch of four encoded uridines (1) that are immediately bound by the La protein (2). U6 snRNAs are then uridylated by U6 TUTase (TUT1) (3) which favors 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]

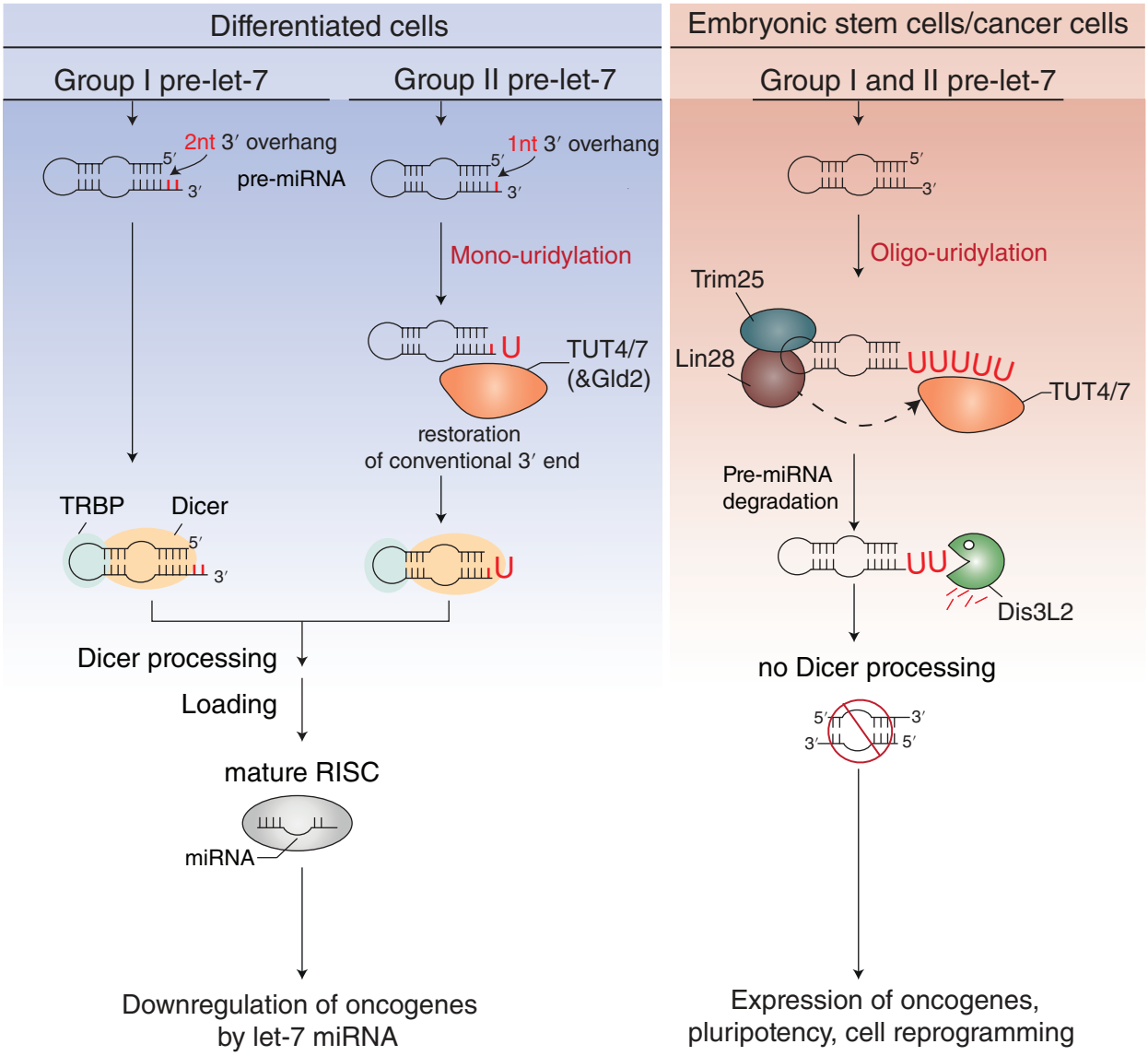


FIGURE 4 | Dual role of uridylation in let-7 miRNA maturation. Group I pre-let-7 miRNAs end with a 2 nt 3′ overhang and are further processed by Dicer to generate mature let-7 which downregulate oncogenes in differentiated cells. Group II pre-let-7 miRNAs end with a 1 nt 3′ overhang and are mono-uridylated 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

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

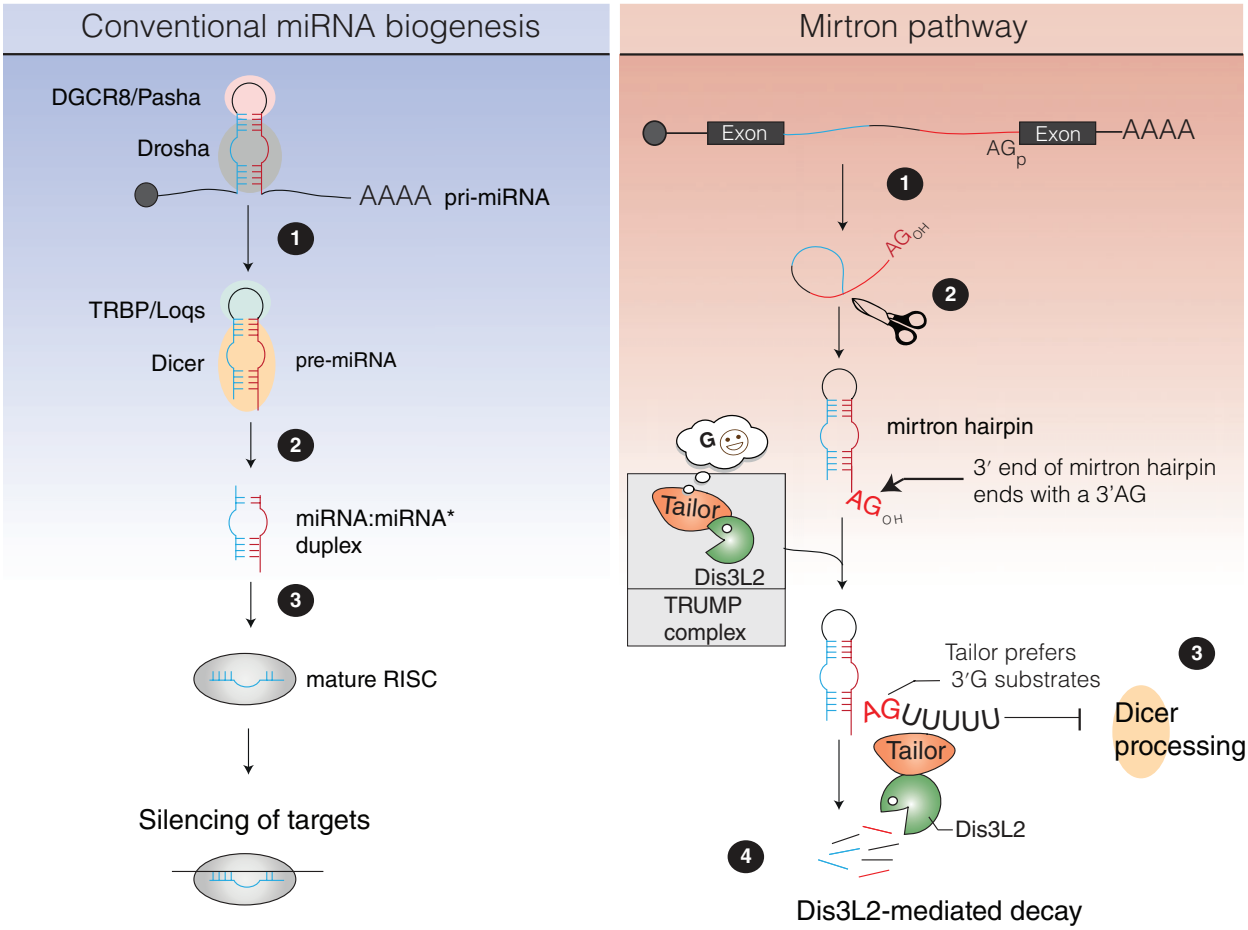


FIGURE 5 | Uridylation restricts mirtron accumulation. Left panel: (1) The conventional processing of animal miRNAs starts with the cleavage of the pri-miRNA by the nuclear Drosha/DGCR8 complex. (2) The generated pre-miRNA hairpin is then further processed by Dicer/TRBP in the cytoplasm. (3) The resulting 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]

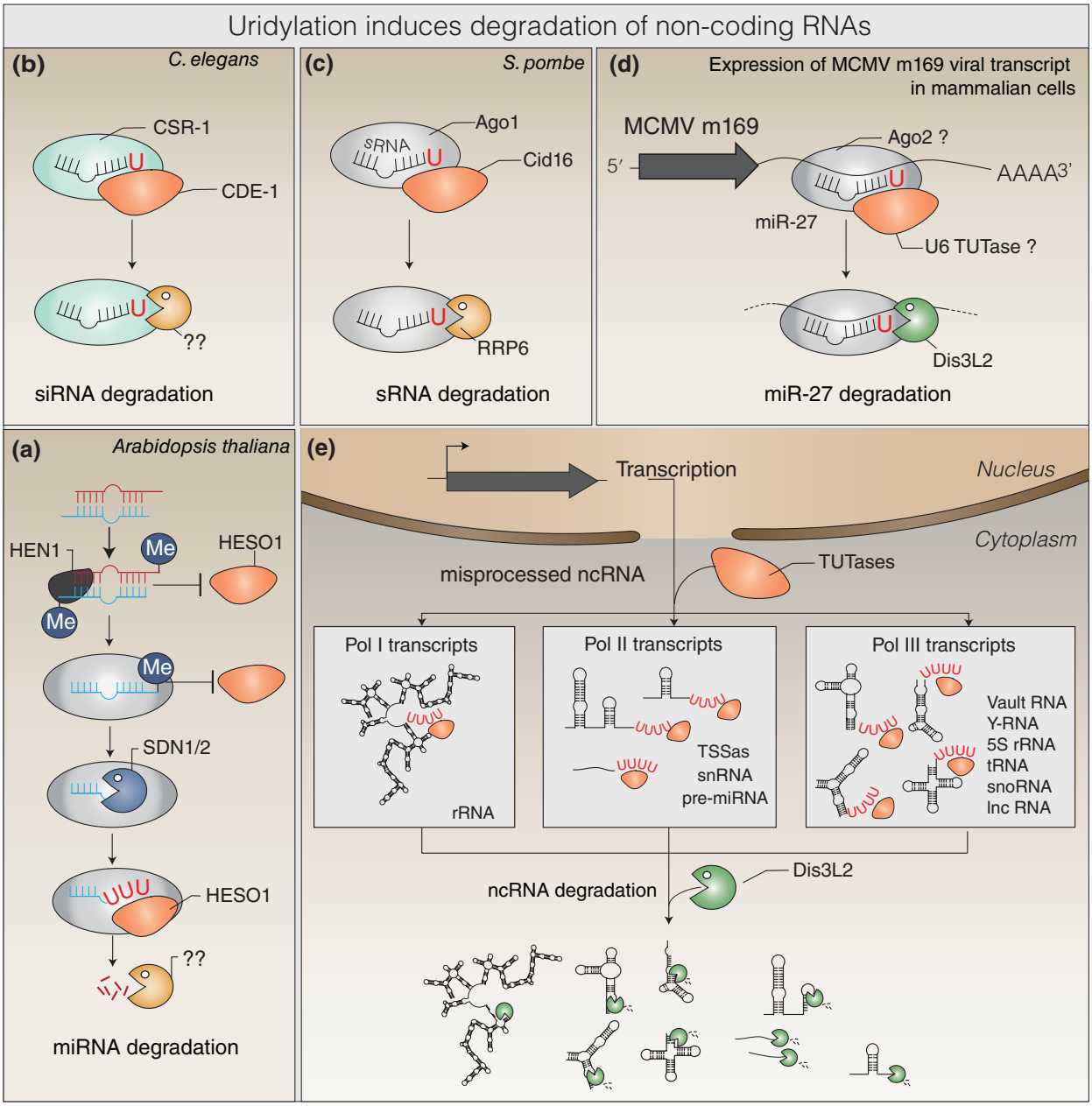


FIGURE 6 | Uridylation induces the degradation of various noncoding RNAs. A few examples were selected to illustrate the destabilizing role of uridylation on noncoding RNAs. (a) Methylation of small RNAs in *Arabidopsis* preventing uridylation by HESO1. SDN1/2 3′→5′ exoribonucleases trim small RNAs, 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

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

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

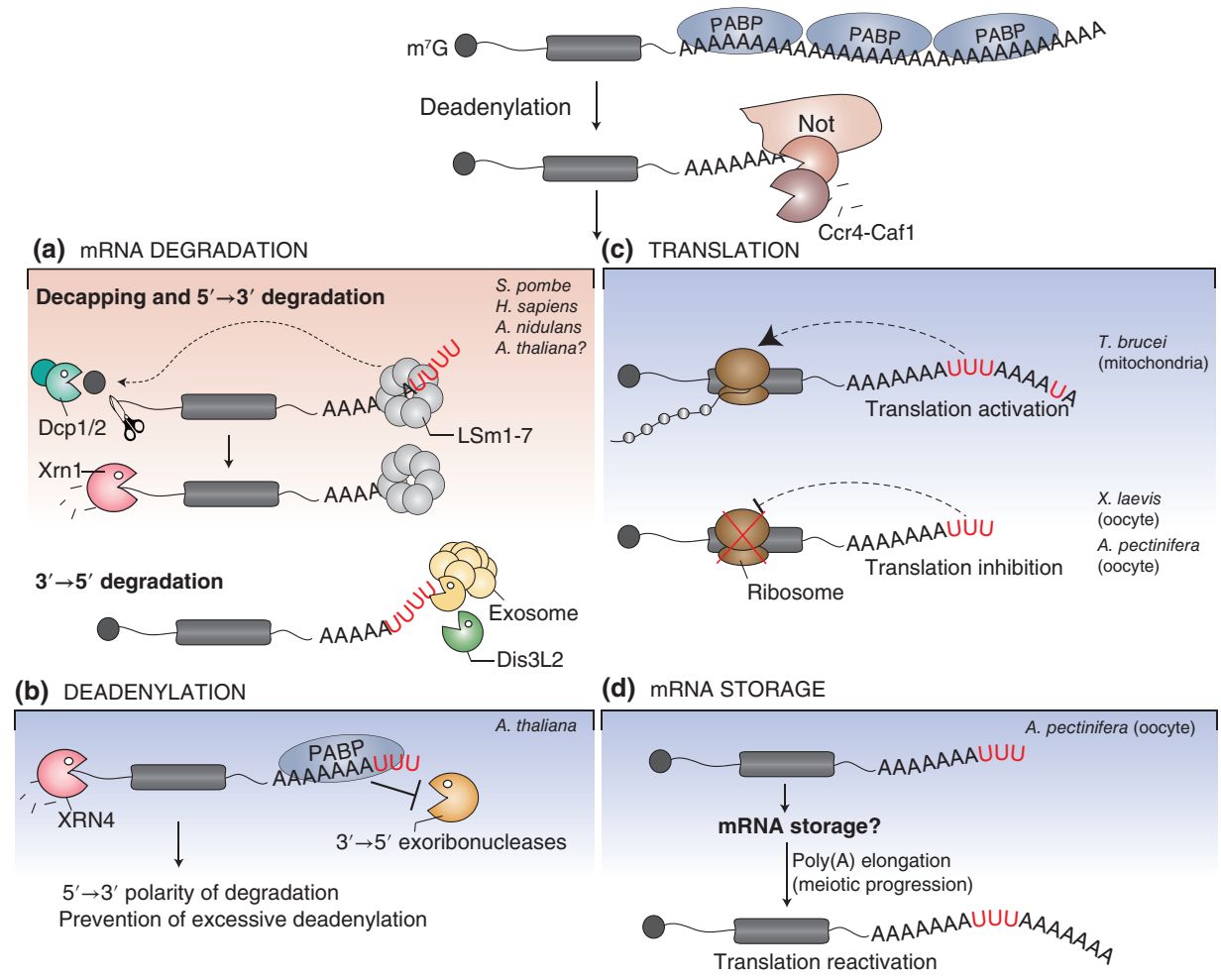


FIGURE 7 | Uridylation plays diverse roles in mRNA metabolism. Uridylation usually occurs after a deadenylation step. (a) The conserved effect of mRNA uridylation is to trigger degradation in eukaryotes. Recognition of uridylated oligoadenylated mRNAs by the LSm1-7 complex induces decapping and subsequent 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.

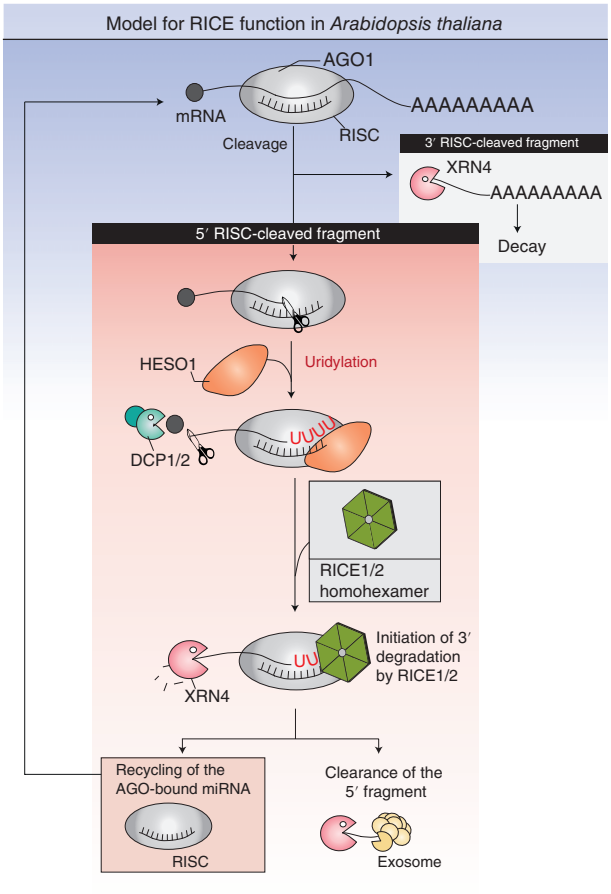


FIGURE 8 | Recycling of RNA-induced silencing complex (RISC) by RICE1/2. MicroRNA is incorporated into the RISC for target recognition. Perfect pairing of plant miRNA with its target supports slicing of the mRNA by the Argonaute protein. This cleavage results in two pieces, known as the 5′ and 3′ RISC-cleaved 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,

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

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

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. Multi-domain 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*

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.

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.

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.

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,

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.

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.

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.

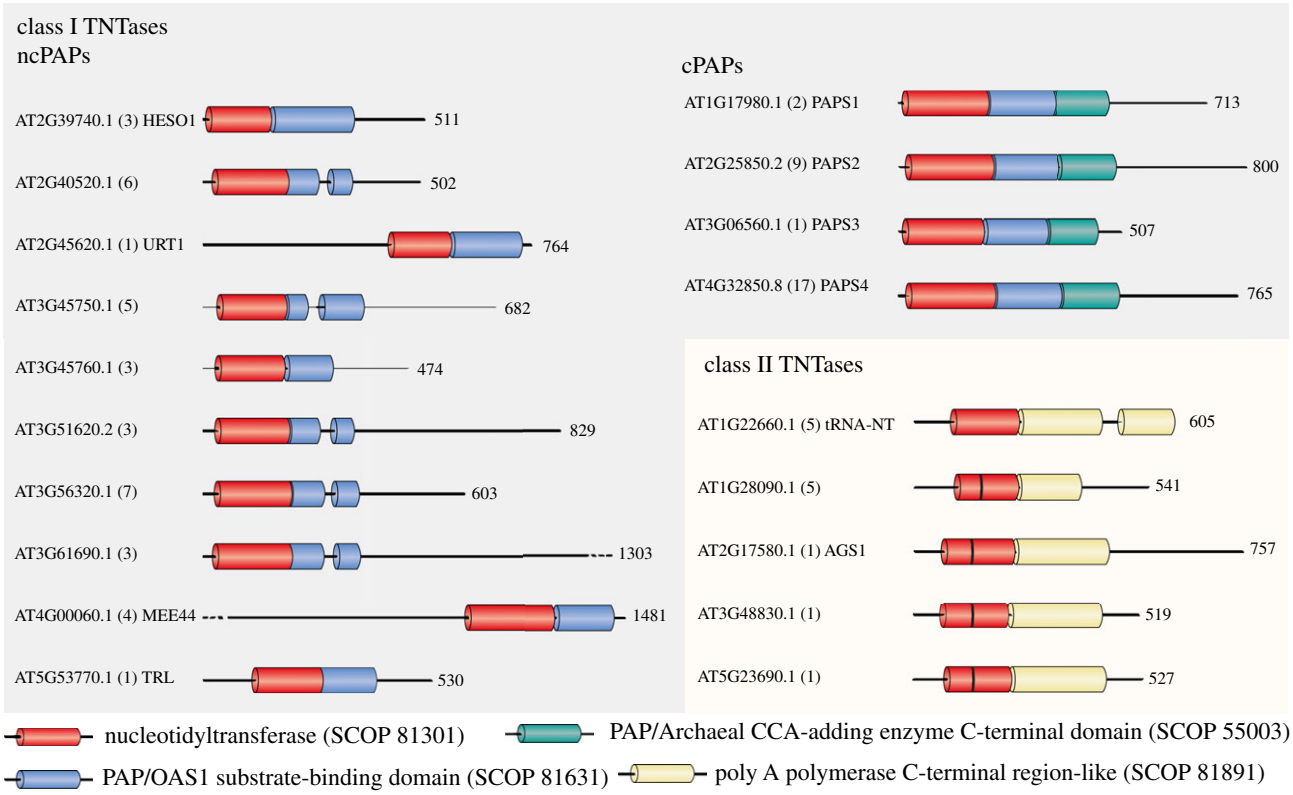


Figure 1. Domain organization of class I and class II TNTases of *A. thaliana*. The class I TNTase family is composed of 10 non-canonical poly(A) polymerases (ncPAPs) and four canonical poly(A) polymerases (cPAPs). The class II TNTase family contains the tRNA nucleotidyltransferase (tRNA-NT), also called 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

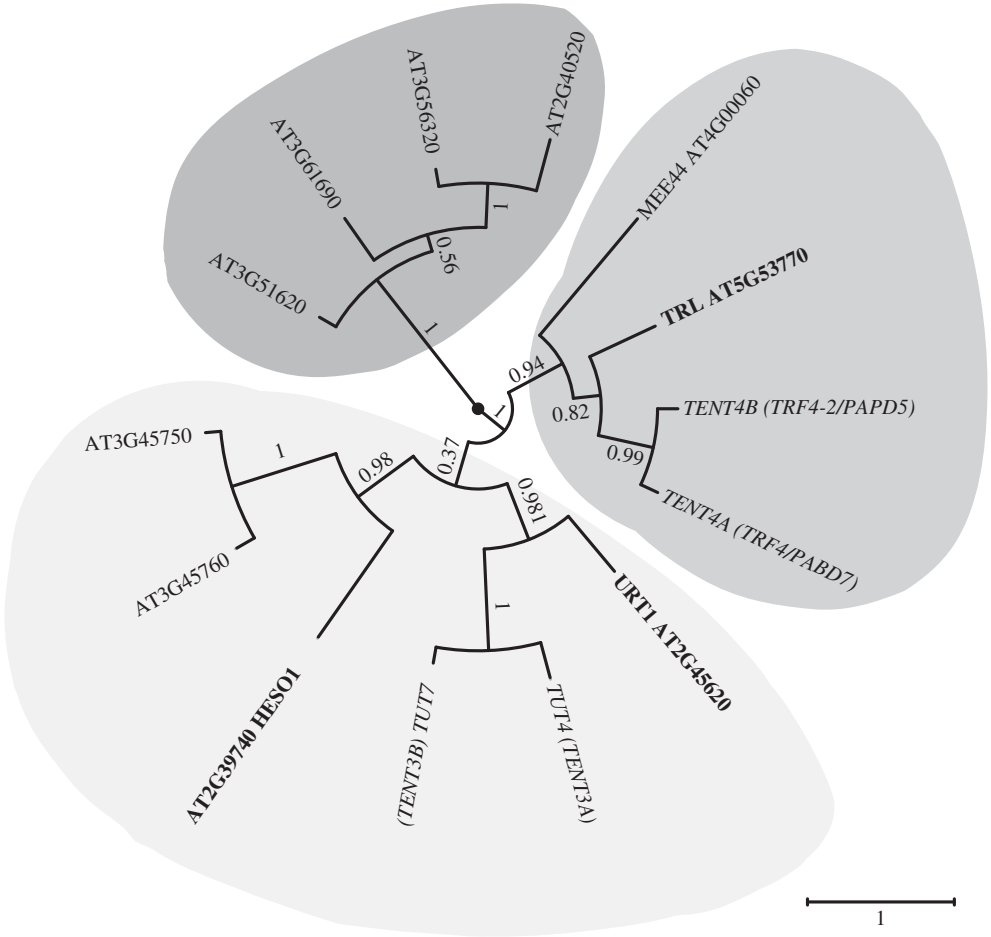


Figure 2. Phylogenetic relationship among *A. thaliana* class I ncPAPs and four human ncPAPs. The nucleotidyltransferase domains SCOP 81301 and the PAP/OAS1 substrate-binding domains SCOP 81631 of the 10 class I ncPAPs of *Arabidopsis* and four human ncPAPs were aligned with Muscle (v. 3.8.31) [37]. 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

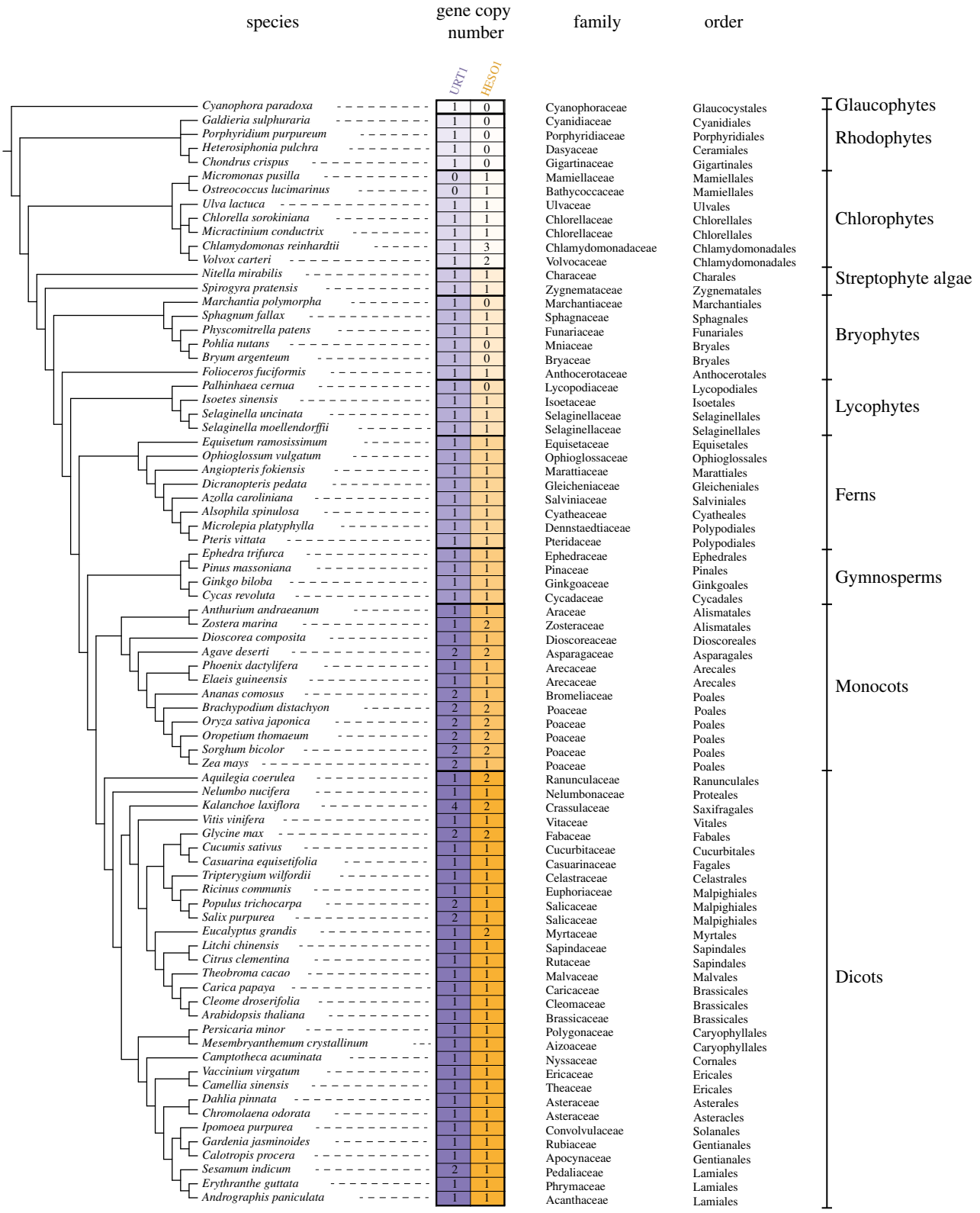


Figure 3. Copy number of HESO1 and URT1 in 79 representative species of Archaeplastida. The phylogenic relationship between the 79 species analysed in this study was visualized using Phylostatic [44]. Full species names are grouped by taxonomic clades indicated on the right. The colour code for URT1 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

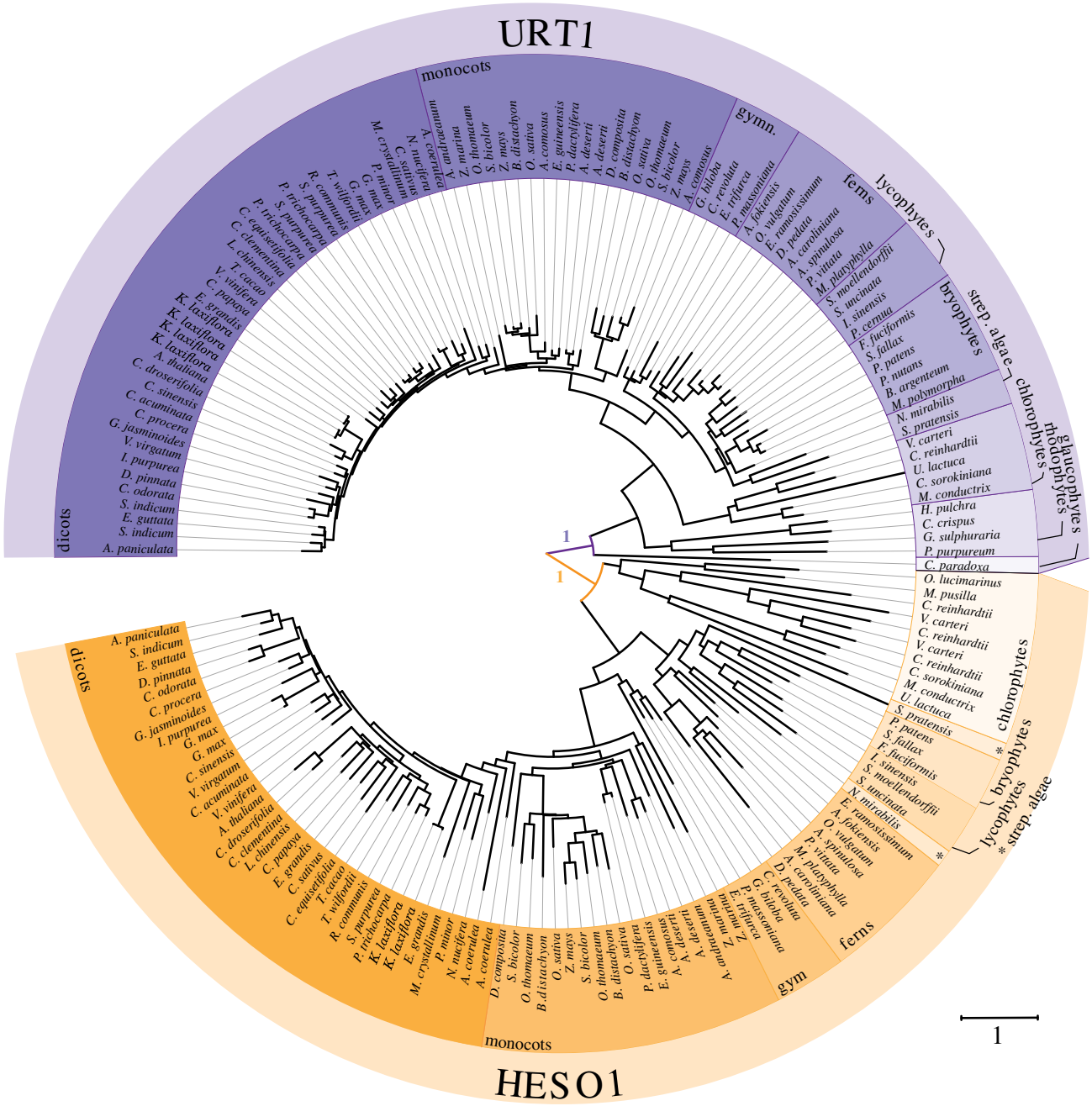


Figure 4. Phylogenetic relationship between URT1 and HESO1 sequences among 79 representative species of Archaeplastida. The phylogenetic tree was generated using the maximum-likelihood method and WAG substitution model implemented in PhyML (v. 3.1) [45,46]; see electronic supplementary material, Dataset 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

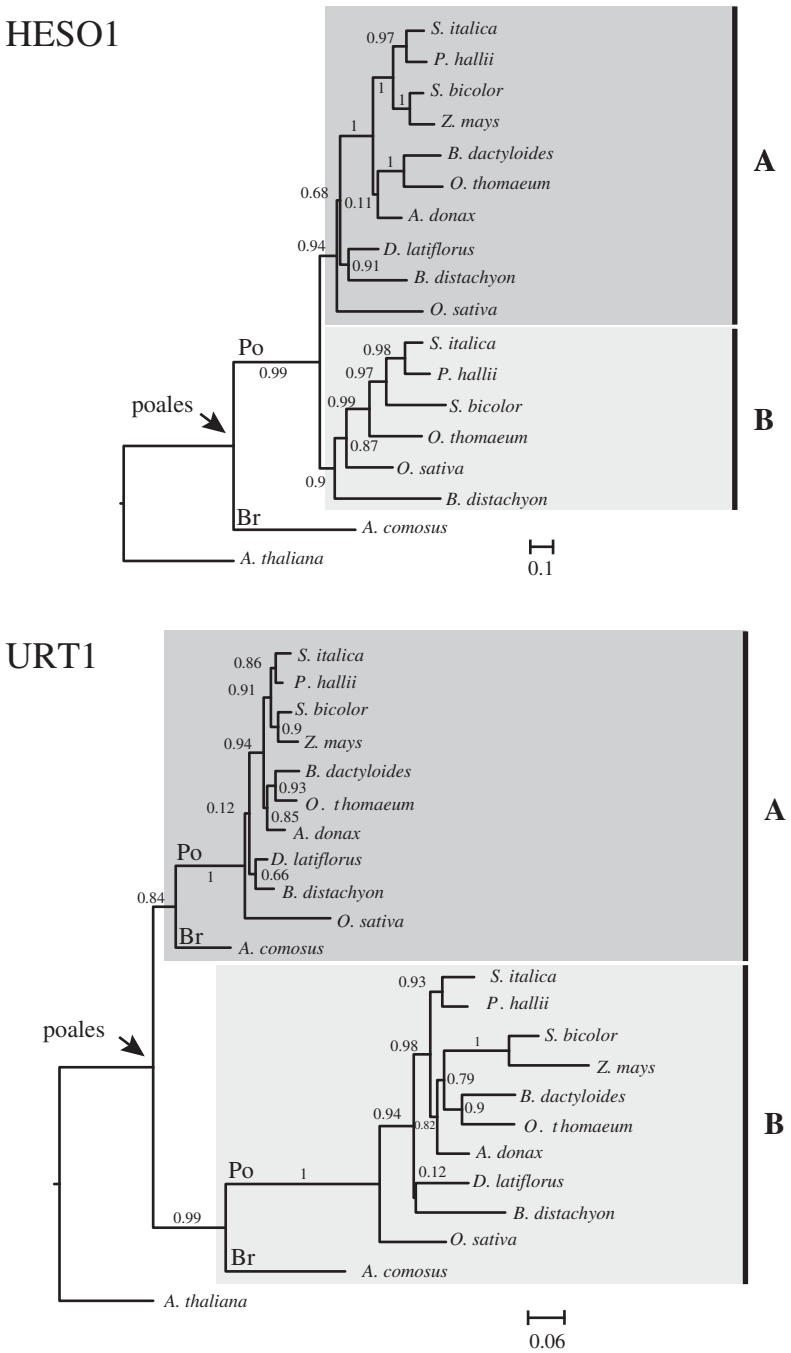


Figure 5. Phylogenetic relationship of URT1 and HESO1 isoforms among 11 species of Poales. Sequences of HESO1 and URT1 isoforms from 10 Poaceae (Po) and 1 Bromeliaceae (Br) were aligned separately. Full names of species are given in figure 3. The analysis was performed as described for figure 4 except 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

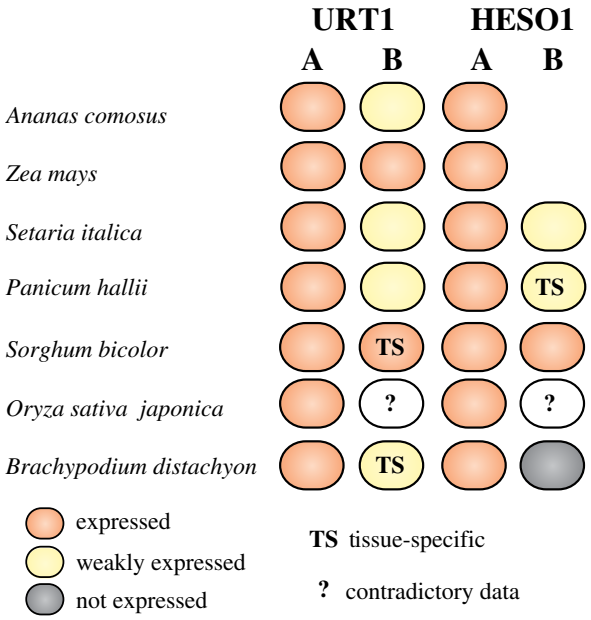


Figure 6. Expression of URT1 and HESO1 isoforms from selected species of the Poaceae family. The diagram was drawn based on the transcriptomic data deposited in Phytozome v. 12.1 (https://phytozome.jgi.doe.gov) [49], in Next-Gen Sequence DBs (https://mpss.danforthcenter.org/dbs) [50], in eFP browser (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

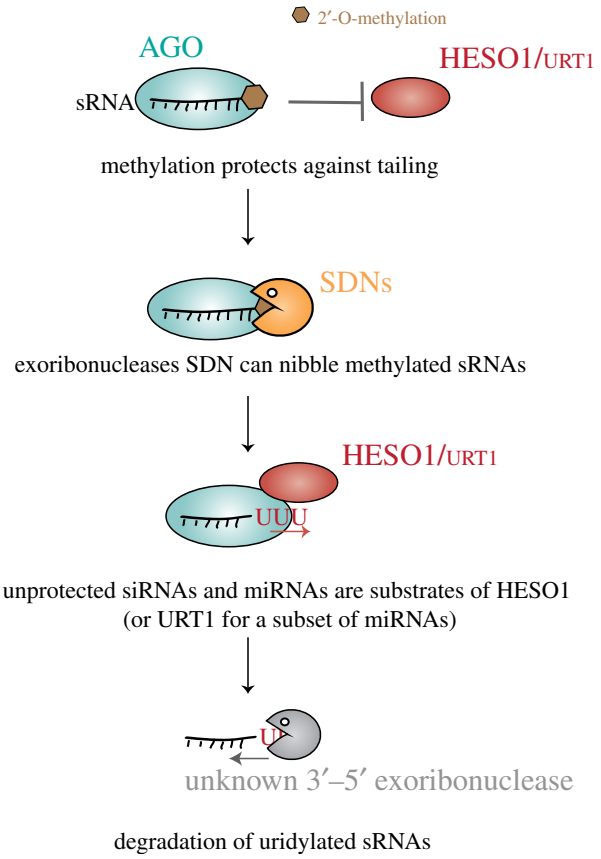


Figure 7. Small RNA uridylation and decay. 2′-O-methylation deposited by the methyltransferase HEN1 protects against uridylation by HESO1 or by URT1. The exoribonucleases' SDNs can nibble methylated sRNAs that are loaded in AGO, thereby generating nibbled, unprotected sRNAs. These unprotected siRNAs 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

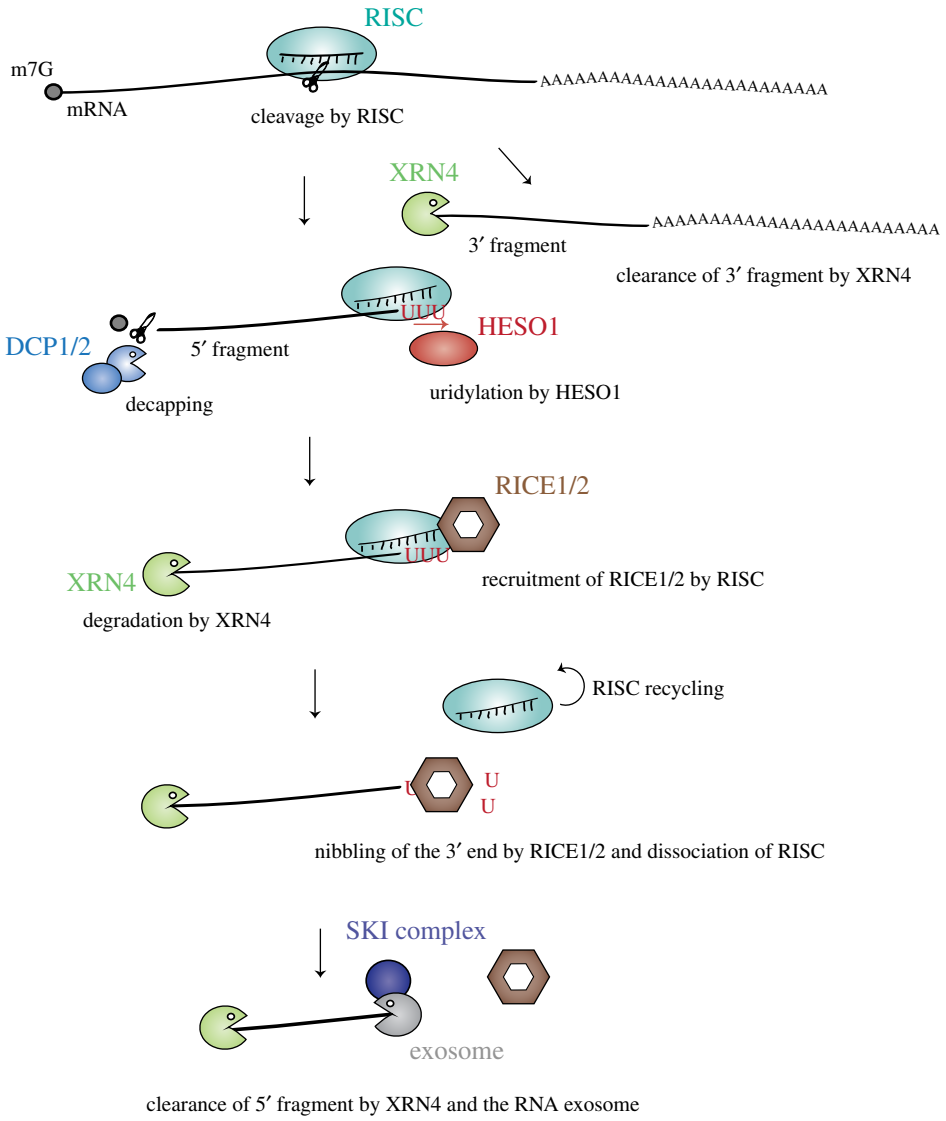


Figure 8. Uridylation by HESO1 promotes degradation of 5′ fragments of RISC-cleaved mRNAs. RISC cleavage of mRNAs generates a 3′ cleavage fragment that is degraded by XRN4, and a 5′ cleavage fragment. The 5′ cleavage fragment is uridylated by HESO1, which binds RISC, but can also be decapped by the DCP1/2 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)

4. Le rôle des processing bodies dans le métabolisme des ARN

Dans le cytosol, les ARNm sont traduits, dégradés ou agrégés dans des granules ribonucléoprotéiques (granules RNP). Parmi les granules RNP cytosoliques décrits figurent les P-bodies, les granules de stress (SGs) et les granules présents dans certaines cellules spécialisées tels que les cellules neuronales et germinales. Ces granules sont des structures dynamiques ayant des propriétés physiques similaires à ceux des liquides, ils s'assemblent et se désagrègent, se rencontrent et fusionnent. La formation de ces granules dépourvus de membranes est issue d'un phénomène de séparation de phase, un processus physique qui se produit lorsqu'une solution sursaturée (en protéines et ARN, par exemple) se sépare spontanément en deux phase distincts, l'une diluée et l'autre condensée (Shin and Brangwynne, 2017; Boeynaems *et al.*, 2018). Au niveau moléculaire cette séparation est promue par différent types d'interactions moléculaires (1) des interactions spécifiques entre protéines (2) des interactions de faibles affinités entre des domaines de faibles complexités ou des domaines intrinsèquement désorganisées (3) des interaction protéines-ARN (Banani *et al.*, 2017).

4.1 Composition des P-bodies

Bien que les P-bodies soient constitutivement présents dans les cellules, leur taille et leur nombre peuvent augmenter en réponse à des stress et dans des conditions qui conduisent à l'inhibition de la traduction (Kedersha *et al.*, 2005; Teixeira *et al.*, 2005). Les P-bodies sont composés d'ARNm en complexe avec des protéines impliquées dans la répression de la traduction et/ou la dégradation des ARNm. La mise au point récente d'une méthode de purification des P-bodies par FAPS (Fluorescence-Activated Particle Sorting) a notamment permis de définir le protéome des P-bodies dans des cellules humaines en culture (Hubstenberger *et al.*, 2017). Les composants principaux des P bodies sont conservés chez les eucaryotes. On trouve notamment la déadénylase CCR4, les facteurs impliqués dans la dégradation 5'3' des ARNm, tel que PAT1, le complexe LSM1-7, le complexe de decapping et ses cofacteurs associés (DCP1, DCP2, HEDLS/Ge-1/VCS, LSM14/DCP5, DDX6/RH), l'exoribonucléase XRN1/XRN4, les facteurs impliqués dans le NMD (UPF1, SMH7) et les protéines du silencing Argonautes (AGO) (Eulalio *et al.*, 2007; Parker and Sheth, 2007; Kulkarni *et al.*, 2010; Maldonado-Bonilla, 2014; Chantarachot and Bailey-Serres, 2018). Les uridylyltransférases TUT4 et URT1 sont également trouvées dans les P bodies chez l'homme et chez Arabidopsis, respectivement (Sement *et al.*, 2013; Hubstenberger *et al.*, 2017). A noter tout de même la présence chez les mammifères de eIF4E, de sa protéine de liaison 4E-T (Andrei *et al.*, 2005; Ferraiuolo *et al.*, 2005) et du facteur impliqué dans le silencing par les miARN, GW182 (Eystathioy, 2003). Chez la plante eIF4E n'est pas retrouvé accumulé dans les P bodies et il n'existe pas d'homologues connus à 4E-T et GW182. La fonction de GW182 chez Arabidopsis est présumablement assurée par la protéine SUO, un analogue fonctionnel de cette dernière qui contient des répétitions GW, promeut l'inhibition de la traduction par les miRNA et co-localise avec

DCP1 dans les P bodies (Yang *et al.*, 2012). Chez l'homme trois protéines, DDX6, 4E-T et LSM14A (Ayache *et al.*, 2015) sont essentielles pour la formation des P bodies.
Les P bodies sont bien sûr également composés d'ARNm (Texeira 2005, Pillai 2005, Liu 2005) et un traitement à la RNase A dissocie des P bodies semipurifiés.

4.2 Fonction des P-bodies

Les P-bodies ont été initialement découverts lors d'expériences de localisation de protéines associées avec la voie de dégradation 5'3' des ARNm et de ce fait ils ont été décrits comme des compartiments cellulaires dédiés à la dégradation des ARNm. Cette idée a été confortée par des expériences menées chez la levure montrant que l'inhibition de Dcp1 ou de Xrn1 conduit à l'augmentation de la taille et du nombre de P-bodies (Sheth, 2003). Elle a cependant été rapidement remise en question par le fait que la dégradation des ARNm, le NMD et le silencing pouvaient avoir lieu dans des cellules dépourvus de P-bodies visibles (Eulalio *et al.*, 2007). Ces mêmes travaux montrent que le blocage du silencing prévient la formation des P-bodies. Les auteurs suggèrent que formation des P-bodies serait une conséquence et non une cause de la dégradation des ARNm et du silencing. Le fait que les ARN puissent être dégradés co-traductionnellement s'oppose également à l'idée que les P bodies soient le site principal de la dégradation cytoplasmique (Hu *et al.*, 2009; Sweet *et al.*, 2012; Pelechano *et al.*, 2015; Yu *et al.*, 2016). L'étude récente de la composition en ARNm des P-bodies démontre que les ARNm stockés dans les P bodies sont traductionnellement inactifs et ne sont pas dégradés (Hubstenberger *et al.*, 2017). De plus, le suivi en temps réel de la dégradation d'un ARNm par marquage fluorescent montre que la dégradation a lieu dans le cytoplasme et ne révèle aucune accumulation de produits de dégradation dans les P bodies (Horvathova *et al.*, 2017; Tutucci *et al.*, 2018). En outre la reconstitution *in vitro*, par séparation de phase, de granules qui inclus Dcp2 ou une endonucléase résulte en une forte diminution de leur activité enzymatique (Schütz *et al.*, 2017). Ceci pourrait suggérer que les enzymes impliquées dans la dégradation des ARN soient en réalité inactives au sein des P bodies et que les ARNm associés aux P-bodies soient protégés de l'action des exoribonucléases et endonucléases cytoplasmiques.
Les P- bodies constituent à l'inverse un site de stockage pour les ARNm non traduits, site duquel ils peuvent être recyclés et renvoyés au niveau des polysomes selon les besoins cellulaires. L'inhibition traductionnelle causée par une carence en glucose chez la levure conduit à l'accumulation des ARNm dans le P bodies, une fois la carence levée ces ARNm sont relocalisés dans le cytoplasme et forment un pool d'ARNm actif pour la traduction (Arribere *et al.*, 2011; Brengues, 2005). Ce phénomène est cependant restreint un nombre limité d'ARNm et ne semble pas constituer un mécanisme général chez la levure (Arribere *et al.*, 2011). Chez Arabidopsis, les ARNm envoyés dans les P-bodies tendent également à être stabilisés plutôt que dégradés en dépit de la présence de facteurs de dégradation dans ces structures ce qui suggèrent qu'elles pourraient être un lieu de stockage (Chantarachot and Bailey-Serres, 2018; Li *et al.*, 2015; Merchante *et al.*, 2015; Scarpin *et al.*, 2017). Certaines études suggèrent un double rôle des P-bodies qui peuvent à la fois être un lieu de dégradation et de stockage selon la population d'ARNm considéré et selon le

contexte (Aizer *et al.*, 2014; Wang *et al.*, n.d.). Le résultat décisif pour trancher cette question a été l'identification récente de la population d'ARNm associée au P bodies (Hubstenberger *et al.*, 2017). Cette population est enrichie en transcrits qui codent pour des fonctions régulatrices alors que la population d'ARNm exclue est enrichie pour des transcrits qui codent pour fonctions cellulaires de base (Housekeeping) (Hubstenberger *et al.*, 2017; Standart and Weil, 2018). Les P bodies constituent donc vraisemblablement un réservoir d'ARNm qui peuvent être mobilisés et traduits rapidement en fonction des besoins cellulaires.

OBJECTIFS DE THESE

La dégradation des ARNm est une composante essentielle de la régulation de l'expression des gènes, cette régulation étant essentielle au développement et à la survie des organismes. Depuis peu, l'**uridylation** est établie comme faisant partie intégrante de la voie générale de dégradation des ARNm. Au début de ma thèse, l'uridylation des ARNm n'avait été rapportée que chez *S. pombe*, *A. nidulans* et *A. thaliana* (Rissland and Norbury, 2009; Morozov *et al.*, 2010; Morozov *et al.*, 2012)/ Depuis, nous savons que l'uridylation des ARNm existe également chez les trypanosomes, la souris, l'homme, et probablement chez la majeure partie des eucaryotes, excepté *S. cerevisiae* (et les levures apparentées) dont le génome ne code plus pour des TUTases (Munoz-Tello *et al.*, 2015; Scheer *et al.*, 2016; De Almeida *et al.*, 2018; Chung *et al.*, 2017). La fonction majeure de l'uridylation, conservée chez tous les autres eucaryotes étudiés à ce jour, est de promouvoir la dégradation des ARNm. Notre connaissance des mécanismes moléculaires sous-jacents à cette fonction principale de l'uridylation et à sa régulation reste toutefois fragmentaire. Par exemple, les réseaux d'interaction des TUTases dans le cytosol de divers eucaryotes sont pratiquement inconnus. De plus, l'étendue des fonctions diverses de l'uridylation dans le métabolisme des ARNm reste à être découverte.

Dans ce contexte, les objectifs de ma thèse ont été l'identification de nouvelles caractéristiques moléculaires liées à l'activité de la TUTase **URT1** chez Arabidopsis**.** URT1 a été découverte dans le laboratoire où j'ai réalisé mes travaux de thèse et est majoritairement responsable de l'uridylation des ARNm chez Arabidopsis. URT1 cible les ARNm engagés dans le processus de dégradation et, notamment, prévient leur déadénylation excessive. Les objectifs précis de ma thèse ont été :

- De contribuer à la démonstration que URT1 « répare » les extrémités 3' d'ARNm déadénylés en restaurant un site de fixation pour une PAPBP.
- De mettre au point un protocole d'analyse des extrémités 3' des ARNm compatible avec la technologie Illumina. Ce protocole de 3'RACE-seq a notamment été essentiel à une meilleure caractérisation moléculaire des extrémités 3' d'ARNm ciblés par URT1.
- De définir le réseau d'interaction proteine-protéine de URT1 et de valider ces interactions.
- De décrypter les conséquences moléculaires de l'uridylation en cas d'expression de URT1 sauvage ou mutées sur la déadénylation, l'accumulation et la traduction d'ARNm rapporteurs exprimés dans un système hétérologue (*Nicotiana benthamiana*).

Les résultats correspondant à chacun de ces objectifs sont décrits en quatre parties, et pour certains sous forme de publications. Parmi les résultats marquants de ma thèse, j'ai identifié une interaction entre URT1 et DCP5, un facteur conservé chez les eucaryotes (Scd6 chez la levure *S. cerevisiae* et LSM14 chez l'homme) et impliqué notamment dans l'inhibition de la traduction. Il s'agit de la

première interaction décrite entre une TUTase et ce type de facteur. Je montre qu'un motif peptidique chez URT1 est nécessaire à l'interaction *in vitro* avec DCP5. Ce motif est situé dans une région intrinsèquement désorganisée de URT1 et il est parfaitement conservé chez l'ensemble des plantes terrestres, suggérant que cette interaction pourrait aussi être conservée au cours de l'évolution de tous les embryophytes. J'ai également participé à la mise au point du protocole de 3'RACE-seq, qui s'est révélé une méthode essentielle à la caractérisation des rôles moléculaires de URT1 dans l'uridylation des ARNm. Grâce au 3'RACE-seq, j'ai pu démontrer que les profils de distributions de taille des queues poly(A) sont modulés par l'uridylation dans un système d'expression transitoire, et ceci en fonction de la surexpression de URT1 ou de différentes versions mutées ou tronquées. De manière très intéressante, la modification des profils de distributions de taille des queues poly(A) uridylées et non-uridylées est associée à la répression de l'expression d'un gène rapporteur. Cette répression ne s'accompagne pas d'une modification de l'accumulation des ARNm rapporteurs.

Mes résultats contribuent à une meilleure compréhension des rôles moléculaires de l'uridylation des ARNm chez Arabidopsis et je discute les perspectives nouvelles que ces travaux ouvrent quant aux rôles divers de l'uridylation dans le métabolisme des ARNm.

RESULTATS

Première partie :

L'uridylation des ARNm par URT1 et la réparation de leur extrémité 3'

Cette première partie décrit l'uridylation des ARNm comme un phénomène global *chez Arabidopsis thaliana* et rapporte l'identification d'un nouveau rôle de l'uridylation dans le métabolisme des ARNm. Nos travaux démontrent en effet que la TUTase URT1 uridyle les queues poly(A) des ARNm ayant subi une étape de déadénylation afin de restaurer un site de fixation pour une PABP.

Ces travaux ont fait l'objet d'un article publié dans la revue Cell Reports en 2016, article que je signe en second auteur :

Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in Arabidopsis.
Zuber H, Scheer H, Ferrier E, Sement FM, Mercier P, Stupfler, Gagliardi D.
Cell Rep. 2016 Mar 22;14(11):2707-17. doi: 10.1016/j.celrep.2016.02.060

1. Résumé général

Dans la première section de cet article nous montrons que URT1 uridyle les ARNm après une première étape de déadenylation et que la surexpression de la protéine ne suffit pas à court-circuiter cette étape *in vivo*. Le statut d'uridylation et la taille des queues poly(A) ont été analysés par 3' RACE-PCR dans des lignées de complémentation du mutant *urt*1 : une lignée d'expression inférieure au niveau d'expression endogène (lignée transgénique T1), et une lignée de surexpression (lignée transgénique T2). Ces analyses révèlent que le niveau d'accumulation de transcrits excessivement déadénylés (ARNm avec des queues poly(A) inférieures à 10 adénosines) est inversement proportionnel au niveau d'uridylation observé dans les deux lignées. Elles montrent également que la surexpression de URT1 conduit à une légère mais significative augmentation de la taille des queues poly(A) des transcrits uridylés et non-uridylés. Ces éléments vont dans le sens d'un modèle où l'uridylation par URT1 ralentit la dégradation des extrémités 3' et prévient, en ce faisant, l'accumulation de transcrits excessivement déadénylés. La surexpression de URT1 *in vivo* n'entraine cependant pas à une augmentation dramatique du niveau d'uridylation ce qui suggère que les processus de déadénylation précèdent et constituent une étape limitante à l'activité d'uridylation par URT1.
Notre étude révèle ensuite que l'uridylation représente une modification extrêmement fréquente de l'extrémité 3' des ARNm déadénylés chez Arabidopsis et que l'uridylation attribuée à l'activité de URT1 permet de restaurer une taille de queue poly(A) propice à la liaison d'une PABP. Les extrémités 3' des ARNm ont été analysées par TAIL-seq dans des plantes sauvages (WT) et mutantes *urt1*, *xrn4* et *urt1 xrn4*. Le TAIL-seq est une méthode développée par le laboratoire de Narry Kim qui permet une analyse globale et à haut débit des extrémités 3' des ARNm (Chang *et al.*, 2014). L'utilisation de la méthode TAIL-seq nous a permis de confirmer à une échelle globale que l'uridylation cible essentiellement des extensions oligo(A) inférieures à 20 A. Le résultat le plus marquant de l'analyse de ces librairies TAIL-seq a été la découverte que le nombre de U rajouté par URT1 est inversement corrélé à la taille des extensions oligo(A). La distribution de la taille des

extensions oligo(A) inférieures à 30A pour les transcrits non-uridylés est centrée autour d'un pic majoritaire à 16nt alors que celle des transcrits uridylés est centrée sur un pic à ~12 - 14nt. Étonnamment un profil de distribution centré à 16nt est retrouvé si l'on analyse la distribution des extensions A+U des transcrits uridylés dans WT et *xrn4* mais pas dans *urt1* et *urt1 xrn4*. Ceci indique que l'uridylation par URT1 permet de restaurer une extension d'environ 16nt. Nous proposons que la restauration d'une extension de cette taille permet de rétablir un site de fixation pour une PABP. Cette idée est soutenue par des expériences d'immunoprécipitations combinées à des analyses par 3' RACE-PCR montrant que les PABP se lient à des ARN uridylés *in vivo* et que PAB2, l'une des PABP exprimée de manière constitutive chez Arabidopsis, est capable de contrôler la taille des extensions U ajoutés par URT1 *in vitro*.

Ce rôle de l'uridylation par URT1 diffère du rôle de l'uridylation comme signal de dégradation tel qui l'est décrit chez d'autres organismes (voir le paragraphe « Uridylation as a New Integral Step of Polyadenylated mRNA Decay », introduction). Nos travaux n'excluent pas un rôle de l'uridylation dans la dégradation des ARNm chez Arabidopsis mais révèlent un rôle additionnel de cette modification. D'une part l'uridylation par URT1 qui cible des ARNm avec des queues poly(A) de ~14nt et restaure leur extrémité 3' et d'autre part l'uridylation qui conduirait à la dégradation des ARNm excessivement déadénylés. Cette seconde uridylation impliquerait URT1 et une seconde TUTase capable d'uridyler les ARNm. En effet nous observons une accumulation, dans le mutant *xrn4*, d'une population d'ARNm uridylés avec des queues poly(A) < 10 qui semble être en partie indépendante de l'activité de URT1 puisqu'elle ne disparaît pas complètement dans le double mutant *urt1 xrn4*.

Les travaux présentés dans cet article publié révèlent donc un rôle inédit de l'uridylation des ARNm par URT1 chez Arabidopsis et illustre le rôle complexe joué par cette modification dans le métabolisme des ARNm.

2. Contribution personnelle.

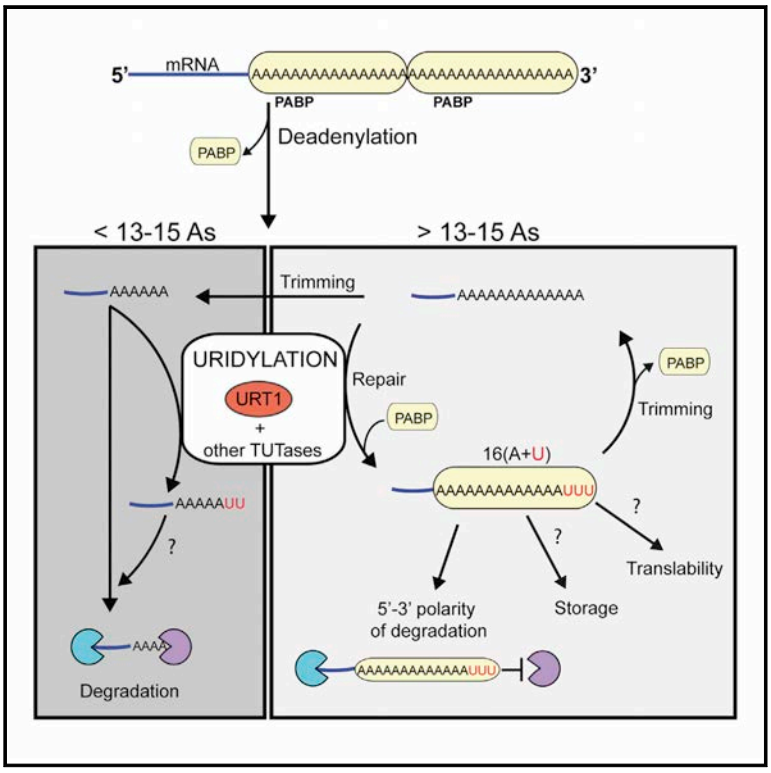
J'ai contribué à cette étude en sélectionnant des lignées stables qui expriment myc-URT1 à des niveaux d'expression différents (lignées transgéniques T1 et T2) et en analysant dans ces dernières les extrémités 3' des ARNm par 3' RACE PCR pour plusieurs réplicas biologiques. J'ai également réalisé les expériences d'immunoprécipitation des PABP et analysé par 3' RACE-PCR les queues poly(A) des ARNm associés. Ces expériences ont contribué à montrer que les ARNm uridylés par URT1 peuvent être reconnus par des PABP.

Article

Cell Reports

Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in *Arabidopsis*

Graphical Abstract



Highlights

- TAIL-seq analysis reveals widespread mRNA uridylation in *Arabidopsis*
- Uridylation repairs mRNA deadenylated 3′ ends
- PABP binds to uridylated oligo(A) tails and determines the size of U-tail extension
- Uridylation determines the extent of mRNA deadenylation

Authors

Hélène Zuber, Hélène Scheer, Emilie Ferrier, François Michaël Sement, Pierre Mercier, Benjamin Stupfler, Dominique Gagliardi

Correspondence

dominique.gagliardi@ibmp-cnrs.unistra.fr

In Brief

Zuber et al. report that uridylation can repair deadenylated mRNAs in *Arabidopsis*. Uridylation and poly(A) binding proteins (PABP) cooperate to restore a defined tail length and control the extent of mRNA deadenylation.

Accession Numbers

GSE72458

Zuber et al., 2016, Cell Reports *14*, 2707–2717
March 22, 2016
http://dx.doi.org/10.1016/j.celrep.2016.02.060

CellPress

Cell Reports
Article

Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in *Arabidopsis*

Hélène Zuber,[1] Hélène Scheer,[1] Emilie Ferrier,[1] François Michaël Sement,[1] Pierre Mercier,[1] Benjamin Stupfler,[1] and Dominique Gagliardi[1,*]
[1]Institut de Biologie Moléculaire des Plantes, Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg, 67000 Strasbourg, France
*Correspondence: dominique.gagliardi@ibmp-cnrs.unistra.fr
http://dx.doi.org/10.1016/j.celrep.2016.02.060

SUMMARY

Uridylation emerges as a key modification promoting mRNA degradation in eukaryotes. In addition, uridylation by URT1 prevents the accumulation of excessively deadenylated mRNAs in *Arabidopsis*. Here, we show that the extent of mRNA deadenylation is controlled by URT1. By using TAIL-seq analysis, we demonstrate the prevalence of mRNA uridylation and the existence, at lower frequencies, of mRNA cytidylation and guanylation in *Arabidopsis*. Both URT1-dependent and URT1-independent types of uridylation co-exist but only URT1-mediated uridylation prevents the accumulation of excessively deadenylated mRNAs. Importantly, uridylation repairs deadenylated extremities to restore the size distribution observed for non-uridylated oligo(A) tails. In vivo and in vitro data indicate that Poly(A) Binding Protein (PABP) binds to uridylated oligo(A) tails and determines the length of U-extensions added by URT1. Taken together, our results uncover a role for uridylation and PABP in repairing mRNA deadenylated ends and reveal that uridylation plays diverse roles in eukaryotic mRNA metabolism.

INTRODUCTION

The control of mRNA stability and translatability is crucial for regulating genome expression. Primary determinants of mRNA stability and translation are the 5′ m^7G cap and the 3′ poly(A) tail. These structures are bound by the eukaryotic initiation factor 4E (eIF4E) and poly(A) binding proteins (PABP), respectively. eIF4E contacts eIF4G, which also interacts with PABP, thereby circularizing mRNAs into a stable and translatable entity (Mangus et al., 2003). mRNA stability and translation can be modulated by RNA binding proteins and miRNAs. Besides these well-studied transfactors, mRNA fate is also regulated by RNA modifications such as the chemical modification of nucleotides and the post-transcriptional untemplated 3′ addition of ribonucleotides or tailing (Lee et al., 2014; Norbury, 2013; Munoz-Tello et al., 2015; Viegas et al., 2015). Tailing includes non-canonical adenylation, which is a widespread modification present in the three domains of life. It triggers the degradation of non-coding RNAs in almost all genetic systems but also the destruction of mRNAs in bacteria, in most Archaea, in chloroplasts and in plant and human mitochondria (Lange et al., 2009; Norbury, 2013). In addition, cytoplasmic adenylation is crucial for activating translation of target mRNAs at several developmental or physiological transitions including oocytes maturation or synapse function (Charlesworth et al., 2013). Besides tailing by adenylation, guanylation and cytidylation of mRNAs have also been recently described in humans, and their respective roles remain to be elucidated (Chang et al., 2014). By contrast, accumulating evidence points toward a role of uridylation in influencing mRNA stability. The uridylation of mRNAs was first demonstrated for human cell-cycle-dependent histone mRNAs, which are not polyadenylated (Mullen and Marzluff, 2008). With reports in *Schizosaccharomyces pombe*, *Aspergillus nidulans*, *Arabidopsis thaliana, Trypanosoma brucei*, and humans, it is now evident that uridylation of polyadenylated mRNAs also exists and is a conserved feature of mRNA metabolism in eukaryotes (Chang et al., 2014; Knüsel and Roditi, 2013; Morozov et al., 2010, 2012; Rissland and Norbury, 2009; Sement et al., 2013; Thomas et al., 2015). The TAIL sequencing (TAIL-seq) method, designed to detect transcriptome-wide nucleotide tailing, revealed the pervasiveness of uridylation for human mRNAs (Chang et al., 2014). Human mRNAs are uridylated by both TUT4 and TUT7. Their simultaneous downregulation increases global mRNA half-lives, demonstrating the impact of uridylation in influencing mRNA degradation (Lim et al., 2014). Importantly, uridylation triggers both 5′-3′ and 3′-5′ mRNA degradation in *S. pombe* and humans (Lim et al., 2014; Malecki et al., 2013; Mullen and Marzluff, 2008; Rissland and Norbury, 2009; Slevin et al., 2014; Su et al., 2013).

Triggering mRNA degradation by uridylation is likely conserved in plants, although this remains to be formally demonstrated. Yet, we recently proposed a role for uridylation in preventing the trimming of oligoadenylated mRNAs in *Arabidopsis* (Sement et al., 2013). We identified UTP:RNA URIDYLYTRANSFERASE1 (URT1) as the main Terminal UridylylTransferase (TUTase) responsible for mRNA uridylation in *Arabidopsis*. *urt1* mutants accumulate mRNAs that are excessively deadenylated, although their decay rate is not affected (Sement et al., 2013). In the present study, we show that URT1

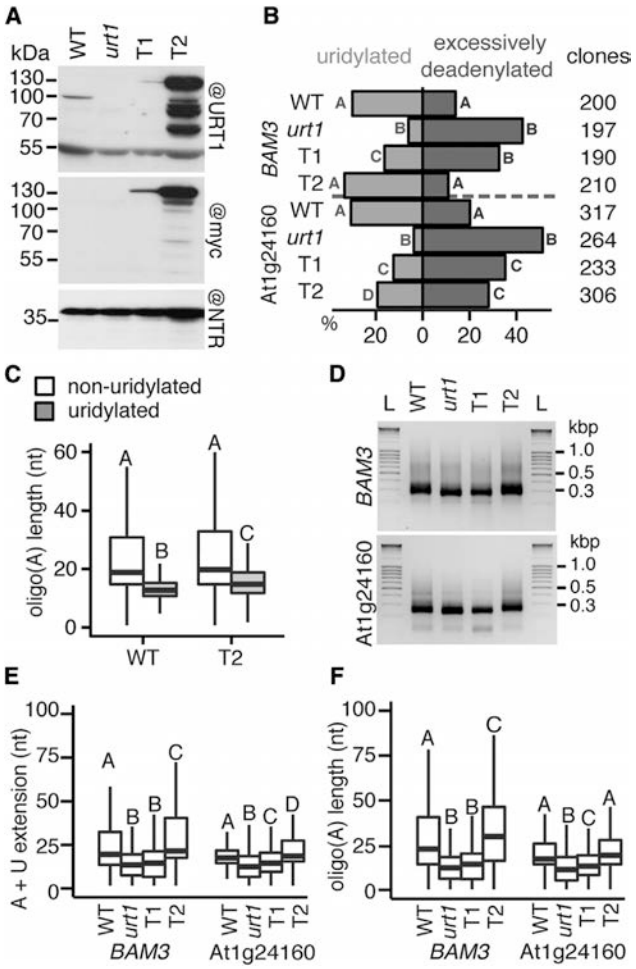


Figure 1. URT1-Mediated Uridylation Modulates the Population of Deadenylated mRNAs

(A) Transgenic lines T1 and T2 express different levels of myc-URT1. Western blot analysis of WT, *urt1*, T1, and T2 flowers using anti-URT1 and anti-myc antibodies. Antibodies against NADPH-dependent thioredoxin reductase AtNTRB (NTR) were used for controlling loading.

(B) Uridylation and accumulation of excessively deadenylated mRNAs are linked to URT1 expression level. Percentage of uridylated mRNAs (left, in light gray) or excessively deadenylated (from 0 to 10 As) mRNAs (right, in dark gray) determined by 3′ RACE-PCR for *BAM3* (At4g20270) and At1g24160. The number of analyzed clones for two to three biological replicates (see Supplemental Experimental Procedures) is indicated for each gene and genotype. Significant differences determined by chi-square contingency table tests are indicated by letters.

(C) Deadenylation precedes URT1-mediated uridylation in both WT and T2. Boxplot analysis of poly(A) length distribution for non-uridylated (white) or uridylated (gray) *BAM3* and At1g24160 mRNAs in WT and T2 lines. The upper and lower edges correspond to the first and third quartiles, respectively. The median is indicated by a horizontal bar and whiskers show data range except far outliers. Significant differences determined by Mann-Whitney tests are indicated by letters.

(D–F) Progressive recovery of tail sizes from *urt1* to T1 and T2. (D) 3′ RACE-PCR profiles for *BAM3* and At1g24160 from flower samples. Negative images of ethidium bromide stained agarose gels. L, DNA ladder. (E and F) Boxplot analysis showing extension sizes for uridylated and non-uridylated sequences (E, poly[A] tail + Us) and for non-uridylated sequences (F, only As) in WT, *urt1*, T1, and T2 lines for *BAM3* and At1g24160. Significant differences determined by Mann-Whitney tests are indicated by letters.

See also Figure S1.

controls the extent of mRNA deadenylation. By applying TAIL-seq analysis to *Arabidopsis*, we identified two types of uridylation. Only URT1-mediated uridylation prevents the accumulation of excessively deadenylated mRNAs. More importantly, we show that uridylation repairs deadenylated ends to restore a binding site for Poly(A) Binding Protein (PABP). Our results support a model in which uridylation and PABP cooperate to control the extent of mRNA deadenylation in *Arabidopsis*.

RESULTS

URT1-Mediated Uridylation Modulates the Population of Deadenylated mRNAs

To get further insight into the role of URT1-mediated uridylation, we complemented the *urt1-1* mutant with a myc-tagged version of URT1 and selected two lines with highly dissimilar expression levels of the transgene. Transgenic line T1 expresses myc-URT1 to a lower level than the endogenous URT1 in wild-type (WT), while line T2 overexpresses the transgenic protein (Figure 1A). We then determined the uridylation status of two mRNAs, *BAM3* (At4g20270) and At1g24160, previously shown to be targets of URT1 (Sement et al., 2013). The uridylation status of *BAM3* and At1g24160 in WT, *urt1*, T1, and T2 was determined from 797 and 1,120 clones obtained by using a modified 3′ rapid amplification of cDNA ends (RACE)-PCR protocol designed to detect nucleotides added 3′ to poly(A) tails (Sement and Gagliardi, 2014). In agreement with previous results, uridylation of both *BAM3* and At1g24160 is detected in WT and decreases sharply in *urt1* mutant, because URT1 is the main uridylytransferase responsible for the uridylation of these two mRNAs (Figure 1B). Interestingly, we observed a gradual increase in the level of uridylation from *urt1* to T1 and T2 lines (Figure 1B). In addition, the accumulation of excessively deadenylated mRNAs (defined as mRNAs with an oligo(A) tail from 0 to 10) (Sement et al., 2013) is inversely correlated to the uridylation status in all samples ($r = -0.98$; p value = 1.68E-05) (Figure 1B). Therefore, varying the expression level of URT1 modulates both the extent of uridylation and the accumulation of excessively deadenylated *BAM3* and At1g24160 mRNAs. Yet, uridylation in line T2 was not dramatically increased as compared to WT, albeit T2 overexpresses myc-URT1 well above WT endogenous level (Figures 1A and 1B). A possible explanation is that the deadenylation step, which precedes URT1-mediated uridylation (Sement et al., 2013), remains limiting. Confirming this hypothesis, a significant reduction in the poly(A) size for uridylated tails as compared with non-uridylated ones is still observed in T2 (Figure 1C). Therefore, deadenylation appears as a prerequisite for URT1-mediated uridylation.

The accumulation of excessively deadenylated mRNAs in *urt1* as compared to WT can be visualized on agarose gels as a slight shift downward due to the faster migration of shorter 3′ RACE-PCR products that correspond to oligoadenylated mRNAs (Figure 1D). A gradual delay in the migration of these 3′ RACE-PCR products is observed from *urt1* to lines T1 and T2 (Figure 1D).

This increase is solely due to the modification of extension sizes since the position of poly(A) sites are identical in the four genotypes (Figure S1A). Plotting the tail sizes (adenosines and uridines) shows that the 3′ extensions of *BAM3* and At1g24160 mRNAs are gradually increased from *urt1* to lines T1 and T2, mirroring myc-URT1 expression levels (Figure 1E). A gradual increase in extension sizes is also observed when considering only non-uridylated poly(A) tails (Figure 1F). This is consistent with the proposed role of URT1 in preventing trimming of oligoadenylated mRNAs. Alternatively, uridylation could also favor the decay of excessively deadenylated mRNAs, thereby shifting upward the population of deadenylated mRNAs. However, the size of uridylated oligo(A) tails is also slightly, but significantly, increased in T2 as compared with WT (Figures 1C and S1B). This observation strongly supports the idea that URT1 overexpression can antagonize the deadenylation step. In conclusion, URT1 expression level influences the sizes of both uridylated and non-uridylated oligo(A) tails. These data do not rule out the possibility that uridylation could trigger the decay of excessively deadenylated mRNAs. However, they also support the idea that URT1-mediated uridylation determines the extent of deadenylation and that an initial deadenylation step cannot be overcome, regardless of URT1 expression level.

Widespread Uridylation of mRNAs in *Arabidopsis*

To obtain a global view of mRNA uridylation in *Arabidopsis*, we generated TAIL-seq libraries from WT plants, *urt1* and *xrn4* single mutants, and *urt1 xrn4* double mutant. The TAIL-seq protocol was recently developed to deep sequence the 3′ ends of RNAs (Chang et al., 2014). Briefly, rRNA-depleted RNA samples are ligated to a biotinylated 3′ adaptor and fragmented, and the affinity-purified 3′ most fragments are ligated to a 5′ adaptor prior to cDNA synthesis and library amplification. Paired-end sequencing of TAIL-seq libraries allows the identification of the RNA (read 1) and the analysis of any nucleotides added at its 3′ extremity (read 2) (Chang et al., 2014). Because URT1-mediated uridylation occurs on mRNAs with short poly(A) tails (Sement et al., 2013), uridylation of oligoadenylated mRNAs was determined using a base-call analysis protocol, which is suitable for analyzing poly(A) tails up to 30 As (Figure S2) (Chang et al., 2014). As detailed in Supplemental Experimental Procedures, we obtained reads for 2,716, 5,501, 2,571, and 4,077 unique genes in WT, *urt1*, *xrn4*, and *urt1 xrn4* libraries, respectively, though most genes had a low number of reads preventing a reliable gene-to-gene comparison between the four libraries. However, when reads are considered globally, our TAIL-seq data provide an unbiased view of mRNA modification by nucleotide addition in *Arabidopsis*. As detailed below, uridylation emerges as the most common modification. We also observed that about 5% and 2% of the reads of all libraries correspond to guanylated and cytidylated mRNAs, respectively (Figure 2A). A similar observation was recently reported in humans (Chang et al., 2014), indicating that mRNA guanylation and cytidylation is conserved from plants to human.

About one-third of the reads (32%) for mRNAs with poly(A) tails <31 As corresponded to uridylated sequences in WT (Figure 2A). The proportion of uridylated mRNAs drops to 7% in *urt1*, confirming that URT1 is the main terminal uridylytransferase modifying mRNAs (Figure 2A). The residual uridylation detected in the *urt1-1*-null mutation indicates that *Arabidopsis* contains at least a second enzyme able to uridylate mRNAs. In *xrn4*, the ratio of uridylated mRNAs raises to 40%, suggesting that mRNA uridylation is favored when 5′-3′ RNA degradation is compromised, as recently reported in humans (Lim et al., 2014). To determine the relative contribution of URT1 in uridylating mRNAs in *xrn4*, the TAIL-seq analysis was also performed for *urt1 xrn4*. Uridylation drops from 40% in *xrn4* to 13% in *urt1 xrn4* revealing that URT1 is indeed partially responsible for the uridylation observed in *xrn4* and confirming the existence of at least a second TUTase involved in mRNA uridylation (Figure 2A).

Plotting the frequency of uridylation against the poly(A) length unambiguously shows at genomic scale that mRNA uridylation in WT occurs preferentially on oligoadenylated mRNAs (Figure 2B). Indeed, uridylation increases for tail lengths shorter than 20 As and peaks at 11 As, representing about 75% of the reads at this tail length. In *xrn4* as well, uridylation is detected mostly on oligoadenylated mRNAs (Figure 2B). The TAIL-seq data for *xrn4* also reveal that the absence of XRN4 increases the frequency of uridylated mRNAs with short tails. For instance, more than 50% of reads for mRNAs with an oligo(A) tail of six to seven As are uridylated (Figure 2B). In *urt1 xrn4*, the distribution of uridylated oligo(A) tails is also biased toward short tails, as observed in WT and *xrn4* (Figure 2B). The higher proportion of shorter tails that are uridylated in *xrn4* versus WT and in *urt1 xrn4* versus *urt1* can be explained by two alternative, but not mutually exclusive, possibilities: (1) uridylation favors 5′-3′ degradation and/or (2) the degradation of deadenylated mRNAs is compromised in the absence of XRN4 and those mRNAs are longer accessible for uridylation. The bias toward shorter tails could also be influenced by an intrinsic preference of the second TUTase for short tails, similar to what is observed for URT1. Interestingly, uridylation drops for oligo(A) tails of 8 nt both in *xrn4* and to a lesser extent in WT (Figure 2B). This drop in uridylation could reveal the binding of a factor that could mask 8-nt A-tails and hinder accessibility by TUTases.

Besides mRNAs, URT1 has recently been identified as uridylating miRNAs in the absence of the small RNA methyltransferase HEN1 and the uridylyltransferase HESO1 (Tu et al., 2015; Wang et al., 2015). To directly compare the impact of URT1-mediated uridylation on mRNAs and small RNAs, we deep sequenced small RNA libraries for WT and *urt1* duplicate samples at the same developmental stage that was analyzed by TAIL-seq, i.e., 2-week-old seedlings. The overall level of nucleotide tailing of miRNAs was not significantly affected by the lack of URT1 (Figure 2C). No significant changes were observed for ten miRNAs displaying the highest uridylation percentage (>10%) and for eight small interfering RNA (siRNA) loci with at least 40 reads corresponding to uridylated siRNAs (Figures 2D and 2E). Therefore, at the seedling stage investigated here, URT1 is dispensable for bulk small RNA uridylation in a WT context. This observation is in agreement with the recent studies reporting that HESO1 will outcompete URT1 for miRNA tailing (Tu et al., 2015; Wang et al., 2015). Taken together, these data show that URT1 appears dispensable for bulk small RNA uridylation,

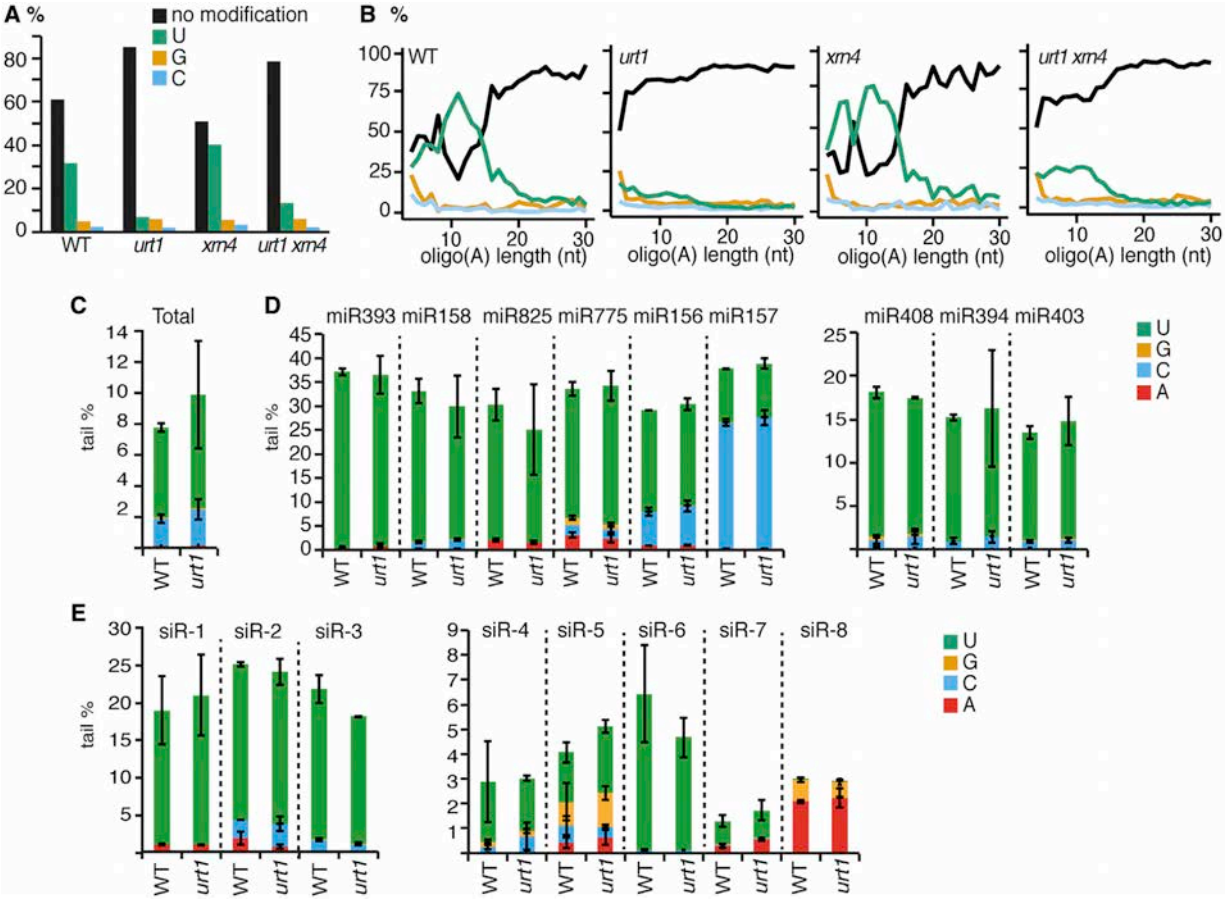


Figure 2. Widespread Uridylation of Oligoadenylated mRNAs in *Arabidopsis*

(A) Frequency of modifications at mRNA 3′ ends for WT, *urt1*, *xrn4*, and *urt1 xrn4* as determined by TAIL-seq analysis.

(B) Frequency of 3′ end modifications plotted against poly(A) tail sizes calculated from TAIL-seq analysis. Note that mRNAs with zero to three As are not considered during TAIL-seq data processing (see Supplemental Experimental Procedures).

(C–E) Frequency of the different modifications at the 3′ end of miRNAs and siRNAs for WT and *urt1*, two biological replicates each. For all panels, Mann-Whitney tests show no significant differences of tailing between WT and *urt1*. (C) Overall frequency of miRNA tailing. Modification frequencies were calculated as percentage of the total number of reads mapping to miRNAs. (D) Tailing frequency for the ten miRNAs that show the highest uridylation level (>10%). Modification frequencies were calculated for each individual miRNA as percentage of the total number of reads for the corresponding miRNA. (E) Tailing frequency for eight siRNA loci. Modification frequencies were calculated for each individual siRNA loci as percentage of the total number of reads for the corresponding siRNA. See also Figure S2.

whereas it plays a prominent role in mRNA uridylation as shown by TAIL-seq analysis.

Distinct Roles for URT1-Dependent and URT1-Independent Uridylation of mRNAs

The TAIL-seq analysis also revealed the existence of at least a second TUTase involved in mRNA uridylation, besides URT1. This second TUTase(s) cannot fully complement the absence of URT1 because excessively deadenylated mRNAs accumulate in *urt1* single mutants (Sement et al., 2013; Figure 1). This accumulation in *urt1* can be explained either because URT1-mediated and URT1-independent uridylation do not have fully redundant functions or because of a lower global uridylation of mRNAs observed in absence of URT1 (Figure 2A). To distinguish between these possibilities, we took advantage of the increased mRNA uridylation detected in absence of XRN4, and we looked for individual mRNAs with similar uridylation levels in WT and *urt1 xrn4*. For five model mRNAs investigated by 3′ RACE-PCR, uridylation is detected in WT, drops in *urt1*, and increases in *xrn4* and is increased in *urt1 xrn4* as compared to *urt1* (Figure 3A). Yet, uridylation level is not significantly different between WT and *urt1 xrn4* for four out of the five mRNAs (Figure 3A). By contrast, a significant increase in excessively deadenylated mRNAs is observed in *urt1 xrn4* as compared to WT (Figure 3A). In addition, both the migration of shorter 3′ RACE-PCR products (Figure 3B) and the analysis of poly(A) sizes (Figure 3C) show that the five mRNAs have shorter oligo(A) tails in *urt1 xrn4* as compared to WT. However, for both migration of PCR products and size of oligo(A) tails, we observed no significant differences between *urt1* and *urt1 xrn4*. These results indicate that URT1-mediated and URT1-independent types of uridylation play distinct roles, with only

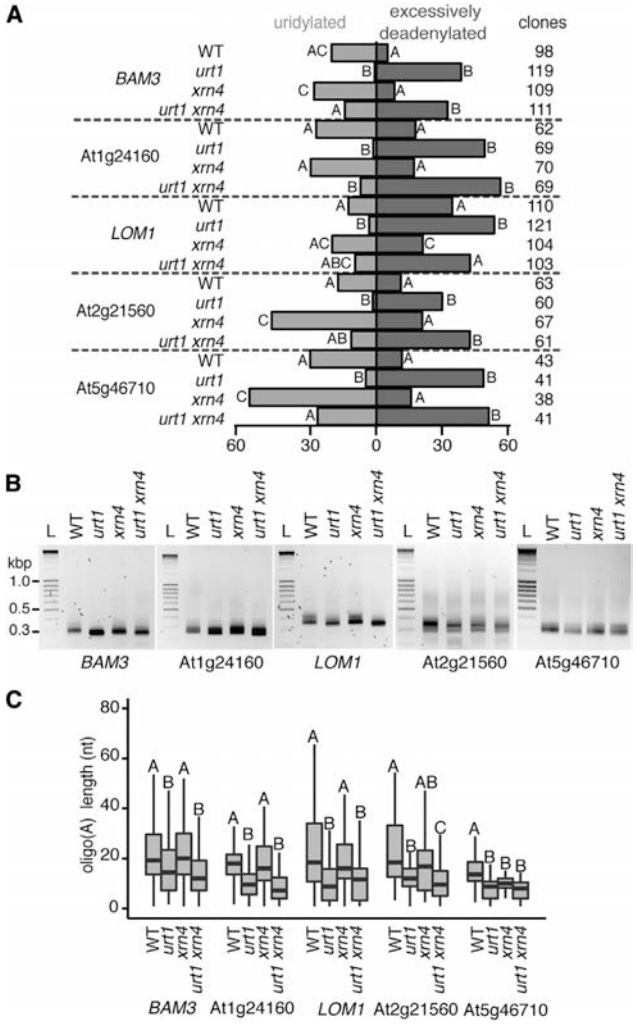


Figure 3. Distinct Roles for URT1-Dependent and URT1-Independent Uridylation of mRNAs

3′ RACE-PCR experiments were performed using leaf RNA from two biological replicates from WT, *urt1*, *xrn4*, and *urt1 xrn4* plants. Five model genes were analyzed: *BAM3*, At1g24160, *LOM1*, At2g21560, and At5g46710.

(A) Only URT1-mediated uridylation prevents accumulation of excessively deadenylated mRNAs. Percentage of uridylated mRNAs (light gray) or excessively deadenylated (from 0 to 10 As) mRNAs (dark gray) determined by 3′ RACE PCR. The number of analyzed clones is indicated for each gene and genotype. Significant differences determined by chi-square contingency table tests are indicated by letters.

(B) 3′ RACE-PCR profiles. Negative images of ethidium bromide stained agarose gels. L, DNA ladder.

(C) Boxplot analysis showing poly(A) sizes for non uridylated sequences. Significant differences determined by Mann-Whitney tests are indicated by letters.

URT1-mediated uridylation preventing the accumulation of excessively deadenylated mRNAs.

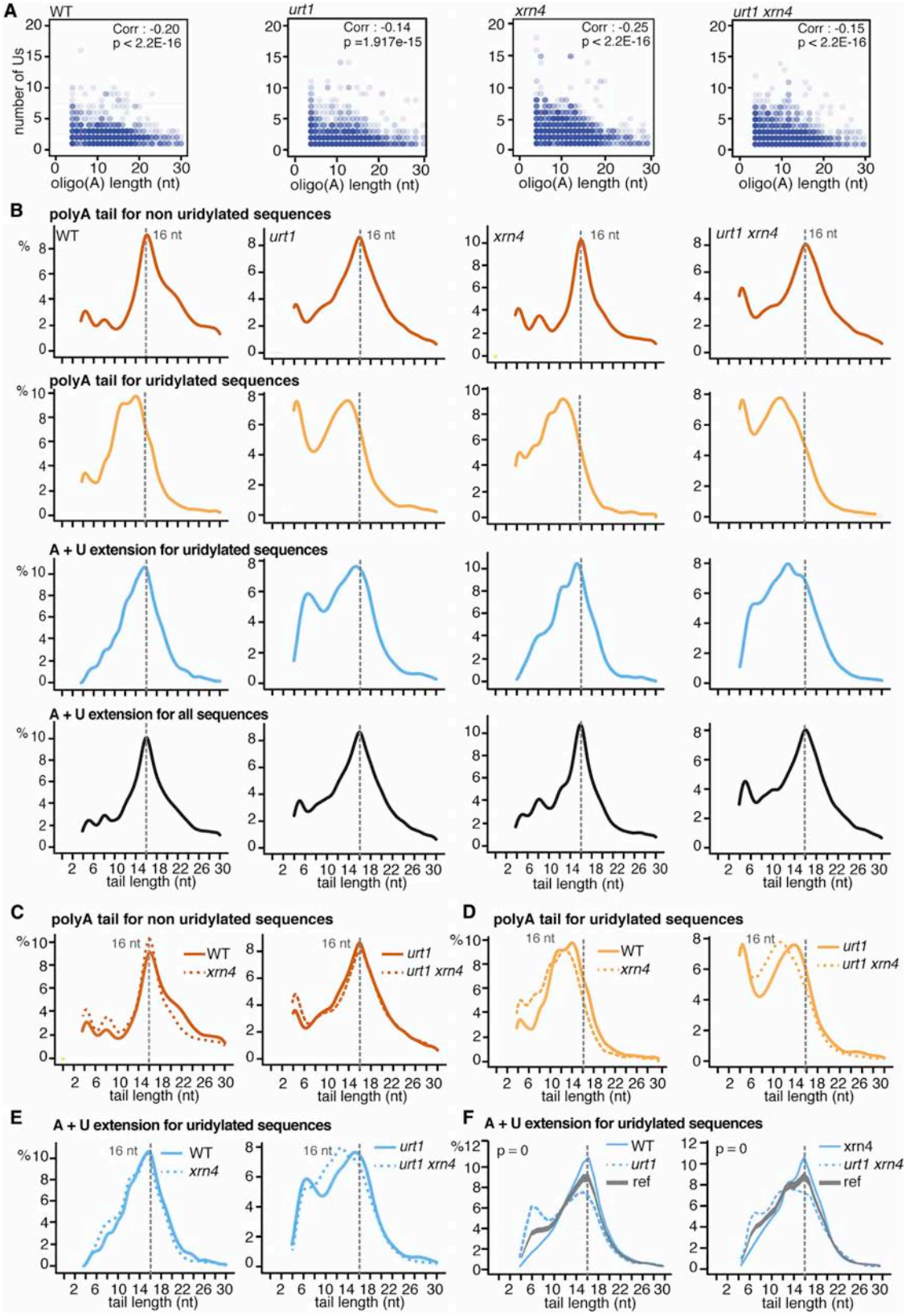
Uridylation Repairs Deadenylated mRNAs

To further investigate the link between uridylation and deadenylation, the TAIL-seq datasets were analyzed by plotting the number of Us added to mRNA 3′ ends against the size of the oligo(A) tails (Figure 4A). Despite relatively low Pearson correlation coefficients, p values inferior to 2E-15 indicate that the inverse correlation observed between the size of the oligo(A) tails and the numbers of U added is significantly different from 0. In other words, the shorter the oligo(A) size, the more Us are added. This observation prompted us to compare the size distribution of oligo(A) tails between non-uridylated and uridylated mRNAs (Figure 4B). For non-uridylated mRNAs, the tail distribution of oligo(A) tails (i.e., As only, from four to 30 As) peaks at 16 nt (Figure 4B). The presence of short oligo(A)-tailed mRNAs in *Arabidopsis* was validated with a method devoid of any potential PCR bias (Figures S3A and S3B). As expected, oligo(A) tail size distribution is influenced by the lack of URT1 since a higher proportion of short oligo(A) tailed mRNAs accumulates in *urt1* and *urt1 xrn4* (Figures 4B–4E). This accumulation is likely underestimated because oligo(A) tails less than four As are excluded during TAIL-seq data processing (see Experimental Procedures). This observation confirms at genomic scale that excessively deadenylated mRNAs accumulate in absence of URT1, as we previously showed for model mRNAs (Sement et al., 2013).

Interestingly, plotting the distribution size of oligo(A) tails (As only) for uridylated sequences revealed a clear shift of the main peak by a few nucleotides toward smaller sizes as compared to non-uridylated sequences (Figures 4B and 4D). This shift is observed for all four genotypes: WT, *urt1*, *xrn4*, and *urt1 xrn4*. Importantly, the overall size distribution centered on 16 nt is restored in both WT and *xrn4* when the whole extensions (As + Us) for uridylated sequences are considered (Figures 4B and 4E). This restoration indicates that uridylation repairs deadenylated ends.

URT1 is expected to play a prominent role in this repair process since it is the main activity uridylating mRNAs in WT and *xrn4*. Indeed, a WT-like distribution for uridylated mRNAs is not observed neither in *urt1* nor *urt1 xrn4*, with a significant higher proportion of shorter tails as compared with WT and *xrn4* (Figures 4B, 4E, and 4F). This shows that, despite the inverse correlation between the size of the oligo(A) tails and the numbers of U observed in the absence of URT1 (Figure 4A), uridylation by URT1 remains essential to fully restore the size distribution observed for non-uridylated sequences. URT1-independent uridylyltransferase activity(ies) could be either inefficient in repairing deadenylated mRNA ends or unable to cope with the increase of excessively deadenylated mRNAs due to the absence of URT1. Either way, a functional URT1 is required to observe a similar size distribution of extension size for non-uridylated and uridylated oligo(A) mRNAs. Similar results were obtained by compiling 1,516, 1,152, 1,022, and 385 sequences for WT, *urt1*, *xrn4*, and *urt1 xrn4*, respectively, and for seven model mRNAs analyzed by 3′ RACE PCR (Figure S3C), which validates the TAIL-seq data. Taken together, these results indicate that uridylation repairs deadenylated ends to restore the size distribution observed for non-uridylated sequences.

In addition, a significant increase in the density of reads representing short oligo(A) tails of uridylated mRNAs is detected by TAIL-seq and 3′ RACE-PCR analyses in *xrn4* as compared with WT (Figures 4D and S3C). If uridylation restores the normal size distribution of deadenylated tails, we would expect that



(legend on next page)

the shorter oligo(A) tails observed in *xrn4* have longer U-extensions as compared to the U-extensions detected in WT. Indeed, in WT, 25% of U-extensions are longer than two Us, and only 7% of U-tails are longer than three Us. In *xrn4*, 42% of U-extensions are longer than 2 Us, while 19% are longer than 3 Us. Both the average U-tail size and the distribution of U-extensions are significantly different between WT and *xrn4* with <2E-16 and 5.5E-9 p values for Mann-Whitney and chi-square tests, respectively. A similar, yet exacerbated, phenomenon was observed for the seven mRNAs analyzed by 3′ RACE-PCR. Only 6% of U-tails are larger than three Us in WT, whereas 34% of U-tails are more than three Us in *xrn4* (p values for Mann-Whitney and chi-square tests, 7.4E-15 and 2.851E-14, respectively). Therefore, both TAIL-seq and 3′ RACE-PCR data show that U-tails are significantly longer in *xrn4* as compared to WT, in line with a preliminary observation obtained with independent samples and for a limited number of mRNAs (Sement et al., 2013).

In conclusion, the most striking information gained through TAIL-seq analysis is that uridylation restores the oligo(A) size distribution observed for non-uridylated oligoadenylated mRNAs. U-extensions are larger in *xrn4*, compensating for the shorter oligo(A) tails observed when 5′-3′ degradation is compromised.

PABP Binds to Uridylated mRNAs and Determines the Size of U-Extensions

The predominance of oligo(A) tails (<31 As) peaking around 16 nt for oligoadenylated mRNAs likely reveals the footprint of a factor bound to oligo(A) tails. A candidate for this binding activity is the cytoplasmic PolyA Binding Protein (PABP). PABP requires at least 12 nt for binding and can bind other homopolymers than poly(A) (Baejen et al., 2014; Eliseeva et al., 2013; Kini et al., 2015; Tuck and Tollervey, 2013). The *Arabidopsis* genome contains eight *PAB* genes, *PAB2*, *PAB4*, and *PAB8* showing high and broad expression in vegetative tissues (Belostotsky, 2003). PABPs are multifunctional proteins with major roles in mRNA stabilization and translation. Therefore, we did not use a reverse genetic strategy to validate a potential involvement of PABPs in the metabolism of uridylated oligo(A) tails. Rather, we used two complementary biochemical approaches. First, we tested whether PABP binds uridylated oligo(A) tails in vivo. To this end, we performed RNA immunoprecipitation (RIP) experiments using anti-PABP antibodies. The uridylation status of At1g24160 mRNAs in RNA samples extracted from immunoprecipitated fractions was determined by 3′ RACE PCR in two biological replicates. Mock reactions, i.e., without anti-PABP antibodies, were used as negative controls (Figure 5A). 3′ RACE-PCR products were obtained only from RNA samples extracted from immunoprecipitated fractions (Figure 5B). Sequence analysis revealed that 27% of 51 clones correspond to uridylated At1g24160 mRNAs showing that PABP binds uridylated oligo(A) tails in vivo (Figure 5C). Interestingly, the sizes of A + U extensions for the immunoprecipitated mRNAs vary from 15 to 18 nt, with a median size of 17 nt. This size range fits well to the prevalent sizes observed for oligoadenylated mRNAs by TAIL-seq and 3′ RACE-PCR. Next, we determine the impact of one of the constitutively expressed PABP (PAB2) on U-tail synthesis by URT1 in vitro using recombinant proteins (Figure 5D). To prevent the foldback of U-tails on a oligo(A) sequence that could interfere with the assay, we first used a non-adenylated 21-nt RNA substrate. As previously shown (Sement et al., 2013), URT1 adds large U-tails (>100 Us) to this RNA substrate when incubated in excess of UTP (1 mM). These U-tails of undefined sizes appear as large smears on acrylamide gels. Strikingly, adding PAB2 does not prevent the initial extension by URT1 but limits the number of added uridines to generate a product of defined length (40 nt) rather than the smear typically observed in absence of PAB2 (Figure 5E). Therefore, PAB2 can efficiently limit the extent of U-addition by URT1 in vitro. We then used a more physiological RNA substrate corresponding to the last 336 nt of *AGO1* mRNA plus 15 As. URT1 is able to uridylate this template as judged by the size increase observed by PAGE following incubation of the substrate with URT1 and UTP (Figure 5F). PAB2 efficiently blocks extension by uridine addition, likely by binding to the oligo(A) tail and preventing URT1 access. Taken together, these in vitro assays show that PAB2 is intrinsically able to limit uridine extension by URT1 and that an initial oligo(A) tail of 15 nt is sufficient for this inhibitory effect. Altogether, these results show that PABP can determine the size of U extensions added by URT1 and that PABP binds uridylated oligo(A) tails in vivo.

DISCUSSION

Our data reveal that uridylation repairs deadenylated mRNAs to restore a defined tail length, which allows for PABP binding. The detection of PABP bound to uridylated mRNAs may profoundly extend our view on the roles played by uridylation in mRNA metabolism. Previous studies have demonstrated that uridylation is linked to cytosolic mRNA degradation in eukaryotes (Lim et al., 2014; Morozov et al., 2010, 2012; Mullen and Marzluff, 2008; Rissland and Norbury, 2009; Slevin et al., 2014; Su et al., 2013). Such a destabilizing role for uridylation certainly exists as well in *Arabidopsis*. In fact, our current data do definitely not exclude the possibility that uridylation by URT1 could also induce mRNA degradation, besides its role in controlling the

Figure 4. Uridylation Repairs Deadenylated mRNAs

(A) Number of Us plotted against poly(A) tail length for uridylated sequences determined by TAIL-seq analysis. Pearson correlation coefficients were estimated between the poly(A) tail length and the number of Us.

(B) Distribution of tail sizes determined by TAIL-seq analysis. Plots display a smooth density estimate of the extension sizes (1D Gaussian kernel).

(C) Overlay of oligo(A) tail size distribution for non-uridylated mRNAs in WT and *xrn4*, and *urt1* and *urt1 xrn4*.

(D) Overlay of oligo(A) tail size distribution for uridylated mRNAs in WT and *xrn4*, and *urt1* and *urt1 xrn4*.

(E) Overlay of size distribution for A+U extension for uridylated mRNAs in WT and *xrn4*, and *urt1* and *urt1 xrn4*.

(F) Comparison of univariate density estimates with the sm R package for A+U extension for uridylated mRNAs between WT and *urt1*, and *xrn4* and *urt1 xrn4*. Grey bands display upper and lower end points of the reference (ref) band for equality.

Note that mRNAs with zero to three As are not considered during TAIL-seq data processing (see Supplemental Experimental Procedures).

See also Figure S3.

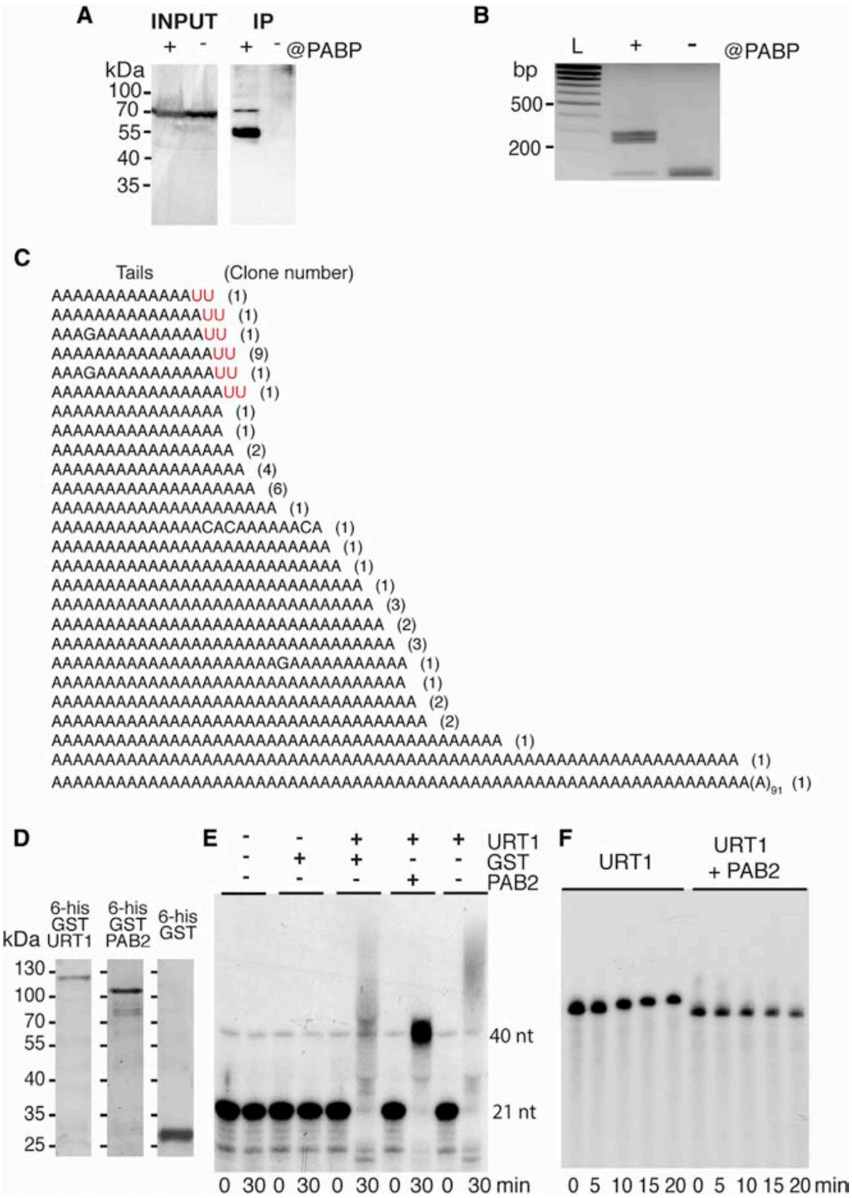


Figure 5. PABP Binds to Uridylated mRNAs In Vivo and Determines the Size of U-Extensions In Vitro

(A) Western blot analysis of PABP immunoprecipitation (IP) performed from WT plants.

(B) 3′ RACE-PCR profiles for At1g24160 mRNAs co-precipitated with PABP.

(C) Tails of 51 3′ RACE clones. Numbers of clone are indicated in brackets for each tail.

(D) Coomassie-blue-stained gels of 6-his-GST-URT1, 6-his-GST-PAB2, and GST expressed in *E. coli* and purified using Ni-NTA and glutathione Sepharose resins.

(E) Impact of PAB2 on U-tail synthesis by URT1 in vitro. GST-URT1 were incubated for 30 min with a non-adenylated 21-nt RNA 5′-labeled substrate and UTP (1 mM), and with or without GST-PAB2 or GST as indicated.

(F) PAB2 efficiently blocks uridine addition to a mRNA with a tail of 15 As. Time course of GST-URT1 incubation with UTP (1 mM), a 5′ [^{32}P]-labeled 336-nt fragment of *AGO1* mRNA tailed with 15 As, and with or without GST-PAB2.

extent of deadenylation. A plausible scenario based on our overall data would be that deadenylation would reach a stage determined by an oligo(A) size >13–15 As where uridylation competes with deadenylation (the "repair" step described here and involving PABP). Even if slowed down, deadenylation could still proceed and beyond a certain size, the repair is no longer effective to allow for PABP binding. These shorter uridylated tails could be recognized by decay factors to promote either by 5′-3′ or 3′-5′ degradation. However, this potential destabilizing role of URT1-mediated uridylation remains to be formally demonstrated. So far, we have observed that URT1-mediated uridylation on selected model mRNAs does not affect decay rates while preventing trimming of oligoadenylated mRNAs (Sement et al., 2013). By contrast, the second type of uridylation detected by the TAIL-seq analysis in *urt1 xrn4* mutants does not protect deadenylated 3′ ends and could be involved in mRNA destabilization. HESO1, a second uridylyltransferase identified in *Arabidopsis*, represents a good candidate for this alternative uridylation activity. Uridylation by HESO1 favors miRNAs and RISC-cleaved transcripts degradation (Ren et al., 2012, 2014; Zhao et al., 2012), and future experiments will reveal whether HESO1 also facilitates the degradation of mRNAs. However, URT1-mediated uridylation definitely plays an additional role in mRNA metabolism, and this is the primary focus of the present study.

The most straightforward interpretation of the data presented here is that a dynamic equilibrium between URT1 and the deadenylase activity(ies) could define the actual length of oligoadenylated mRNAs in *Arabidopsis*. This implies a two-step scenario. In a first step, URT1 needs to gain access to the 3′ extremity of deadenylated mRNAs, possibly when the deadenylase switches between a processive to a distributive activity when the oligo(A) tails get shorter (Viswanathan et al., 2003). Alternatively, URT1 may intrinsically prefer short oligo(A) tails as recently shown for the human TUT4 and TUT7 (Lim et al., 2014). In a second step, uridylation would impede deadenylation according to one of the following possibilities: (1) the presence of one or two Us could directly hinder the recognition by deadenylases, (2) both the deadenylase and URT1 could compete for the oligoadenylated mRNA 3′ extremity and URT1 would impede deadenylation by steric hindrance, (3) uridylation would favor the binding of a protective factor. These three possibilities are not mutually exclusive.

However, our present data support the latter mechanism since we observed that uridylation restores a size distribution of oligo(A) tails centered at 16 As.

Several lines of evidence indicate that a factor that binds uridylated mRNAs in vivo is PABP. First, PABPs have the capacity to bind other sequences than poly A tails (Baejen et al., 2014; Eliseeva et al., 2013; Kini et al., 2015; Tuck and Tollervey, 2013) and by combining PABP immunoprecipitation with 3′ RACE, we demonstrate that PABP can bind uridylated oligo(A) tails in *Arabidopsis*. Second, after an initial step of uridylation, *Arabidopsis* PAB2 very efficiently limits U-extension of a non-adenylated RNA substrate by URT1 in vitro. This experiment shows that PAB2 does not inhibit URT1 activity directly but rather suggests that URT1 starts synthesizing a U-tail, which is bound by PAB2 once the tail reaches a defined length, which then prevents further uridylation by URT1. Importantly, URT1 is a distributive enzyme for the first added nucleotides (Sement et al., 2013). Hence, URT1 releases its RNA substrate after each U-addition, which likely facilitates binding of PABP. Under the conditions used, 19 Us are added. Taking into consideration that two As are present at the 3′ end of the RNA substrate, a total of 21 U + As is present before URT1 ceases to elongate the tail. Considering that in vitro, tail length might be biased by several parameters such as the nucleotide composition of the tail, the absence of a competing deadenylase activity and the high concentration of URT1 and substrates (RNA and nucleotides), a tail length of 21 nt is in good adequation with the oligo(A) length determined in vivo. Finally, PAB2 prevents URT1-mediated uridylation of a reporter sequence with an already existing tail of 15 As. Interestingly, human PABPC1 also suppresses uridylation of tails of 25 or 50 As by TUT4 and TUT7 in vitro (Lim et al., 2014). Taken together, these data support the idea that PABP interacts with uridylated oligo(A) tails and that PABP will inhibit URT1 elongation once a sufficient length is attained for its binding. This also explains the observed inverse correlation between the size of the oligo(A) tails and the numbers of Us added, and the restoration of oligo(A) size distribution by uridylation. Interestingly, the median size of uridylated oligo(A) tails (A + U) immunoprecipitated with PAB2 is 17 nt, which fits remarkably well to the peak of oligo(A) tails detected by TAIL-seq analysis. A size of 17 As may seem modest given that PABP can occupy up to 27 As (reviewed in Eliseeva et al., 2013). However, the minimal oligo(A) size bound by PABP is 12 As (Sachs et al., 1987). Hence, the peak centered at 16 nt observed by TAIL-seq may reflect the equilibrium set by competing PABP, URT1, and deadenylases.

The recognition of uridylated oligo(A) tails by PABP raises interesting questions on the biological roles of uridylation. URT1-mediated uridylation could be both part of the general mRNA degradation pathway and play also other roles in mRNA metabolism. These additional roles may vary with the different subcellular localization of uridylated mRNAs and, therefore, with the distinct and particular protein environment. URT1 is not only diffusely distributed in the cytosol, but also is localized in both P-bodies and stress granules (Sement et al., 2013). In stress granules, URT1 and PABP could play an obvious role in mRNA storage through URT1-mediated protection of deadenylated mRNA 3′ ends by PABP binding. Whether and how URT1 influences mRNA fate in P-bodies is unknown at present. P-bodies components include the decapping machinery, the 5′-3′ exoribonuclease and deadenylases but no detectable PABP. Uridylation by URT1 could yet favor 5′ to 3′ polarity of degradation in P-bodies either by directly impeding deadenylase activities and/or promoting decapping. Another protein than PABP could also recognize uridylated oligo(A) tails in P-bodies. A 5′ to 3′ polarization of degradation could be important to prevent the formation of aberrant transcripts such as mRNAs excessively trimmed by deadenylases. These aberrant RNAs could be substrates of the potent siRNA pathway of plants (Zhang et al., 2015). However, the most intriguing question raised by the binding of PABP to uridylated mRNAs is a potential role of URT1-mediated uridylation on polysomes. We have previously shown that uridylated transcripts are indeed present on polysomes, and that excessive deadenylation occurs on polysomal mRNAs in *urt1* (Sement et al., 2013). The protection conferred by URT1 to 3′ ends of polysomal mRNAs could be important for the 5′-3′ polarity of degradation while mRNAs are still engaged on polysomes. Indeed, co-translational degradation of mRNAs emerges as a conserved feature in eukaryotes (Hu et al., 2009; Merret et al., 2015; Pelechano et al., 2015). Alternatively, binding of PABP to uridylated oligo(A) tails could affect their translatability. At present, it cannot be excluded that deadenylated uridylated mRNAs bound by PABP are actively translated. However, initial experiments in *Xenopus* oocytes and *A. nidulans* suggests that uridylation is linked with translation inhibition (Lapointe and Wickens, 2013; Morozov et al., 2012). Whether this inhibitory role is conserved in *Arabidopsis* and/or whether URT1-mediated uridylation restores translation of deadenylated mRNA constitutes an exciting area for future investigation. In fact, we are just beginning to appreciate the functions conferred by uridylation in mRNA metabolism and their diversity across eukaryotes. The finding that uridylation repairs deadenylated ends to restore a binding site for PABP constitutes a key step toward our general understanding of the overall function of uridylation in mRNA metabolism.

EXPERIMENTAL PROCEDURES

Material

Arabidopsis thaliana plants were of Col-0 ecotype and grown on soil with 16-hr-light/8-hr-darkness cycles. *urt1-1* (Salk_087647C) and *xrn4-3* (SALK_014209) have been previously described (Gazzani et al., 2004; Sement et al., 2013) and were crossed to produce *urt1 xrn4*. myc-URT1-expressing lines were produced by transforming the *urt1-1* mutant by the floral dip method with the genomic sequence of URT1 cloned in pGWB621. All primer sequences and vectors are detailed in Supplemental Experimental Procedures.

Western Blotting

Immunoblots were incubated with anti-URT1 antibodies raised in rabbits against the full-length recombinant URT1, or with anti-myc (Roche), anti-NTRB (kind gift from Géraldine Bonnard), or anti-PAB2 (kind gift from Cécile Bousquet-Antonelli) antibodies in Tris-buffered saline (TBS); 5% (w/v) milk; 0.02% (v/v) Tween 20. Following incubation with horseradish-peroxidase-coupled secondary antibodies and Lumi-Light Western Blotting Substrate (Roche), signals were recorded using the Fusion-FX system (Fisher Biotech).

3′ RACE-PCR

The 3′ RACE PCR protocol used to sequence mRNAs 3′ ends is detailed in Sement and Gagliardi (2014). See Supplemental Experimental Procedures for an outline of the protocol.

TAIL-Seq

TAIL-seq libraries were prepared from WT, *urt1*, *xrn4*, and *urt1 xrn4* 2-week-old seedlings according to Chang et al. (2014). After rRNA depletion, RNAs were ligated to a biotinylated 3′ adaptor and partially digested by RNase T1. RNA 3′ fragments were purified with streptavidin beads, phosphorylated, and gel purified (500–1,500 nt). The purified RNAs were ligated to a 5′ adaptor, reverse transcribed, and amplified by PCR. PCR products were purified and sequenced on the Illumina HiSeq 2500 (50 × 240 bp paired end run). Sequences were processed using the Base calls acquired from HiSeq 2500 after processing by Illumina CASAVA-1.8.2. A detailed protocol of both library preparation and data processing are provided in Supplemental Experimental Procedures.

Small RNA Library Preparation and Sequencing

Small RNA libraries were generated from 3 μg of total RNA extracted with TRIzol reagent (Invitrogen) from WT and *urt1* 2-week-old seedlings, two biological replicates each. Small RNA libraries were produced using the Illumina Small RNA TruSeq protocol and sequenced using a HiSeq 2000 sequencer. Sequences were processed using the base calls acquired from HiSeq 2000 after processing by Illumina CASAVA-1.8.2. A full protocol of the data processing is provided in Supplemental Experimental Procedures.

Expression and Purification of Recombinant Proteins

6his-GST-URT1, 6his-GST-PAB2, and GST were produced in BL21(DE3) cells grown at 17°C. Cells were disrupted by sonication in 20 mM MOPS (pH 7.5), 250 mM KCl, 15% (v/v) glycerol, 1 mM DTT, and 0.1% (v/v) Tween 20 in presence of protease inhibitors (Roche). Recombinant proteins were purified on Ni-NTA resin followed by glutathione affinity chromatography. Purified proteins were dialyzed against 20 mM MOPS (pH 7.5), 100 mM NaCl, 15% (v/v) glycerol, 0.1% (v/v) Tween 20. Aliquots were snap-frozen in liquid nitrogen and stored at −80°C.

Activity Assays

In vitro assays shown in Figure 5 contained 100 nM of GST-URT1, 20 mM MOPS (pH 7.5), 100 mM NaCl, 15% (v/v) glycerol, 0.1% (v/v) Tween 20, 1 mM $MgCl_2$, 0.5 μg/μl BSA, and 1 mM UTP. For Figure 5E, GST-URT1 was incubated for 30 min with or without GST-PAB2 (20 nM) or GST (20 nM) and with a non-adenylated 21-nt RNA substrate labeled by T4 Polynucleotide Kinase (NEB) and [γ-^{32}P]-ATP. Reaction products were separated by denaturing 5% (w/v) polyacrylamide gel electrophoresis before autoradiography. For Figure 5F, GST-URT1 was incubated for different time points with or without GST-PAB2 (20 nM) and with the last 336 nt of AGO1 mRNA plus 15 As labeled by T4 Polynucleotide Kinase (NEB) and [γ-^{32}P]-ATP. Reaction products were separated by denaturing 17% (w/v) polyacrylamide gel electrophoresis before autoradiography.

RNA Immunoprecipitation

For PABP immunoprecipitation, 300 mg of flowers were ground in 1 ml of 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% Triton X-100, 1 × Complete Protease Inhibitor EDTA free (Roche), and 10 mM ribonucleoside-vanadyl complex (NEB). Lysates were clarified by centrifugation at 4°C at 16,000 × *g* for 5 min, incubated with anti-PABP antibodies, and purified using magnetic protein A MicroBeads (Miltenyi Biotech) according to the manufacturer's instruction. Protein-RNA complexes were directly eluted from magnetic beads using 100 μl of 20 mM MOPS (pH 7.5), 10 mM NaCl, and 0.1% Triton X-100. 20 μl of the eluates was separated by SDS-PAGE electrophoresis and analyzed by western blotting. 5 μg of yeast total RNA was added to the remaining 80 μl eluates, and RNA was purified using RNeasy MinElute columns (QIAGEN). 3′ end extremities were analyzed by the modified 3′ RACE-PCR protocol previously described.

ACCESSION NUMBERS

Sequenced reads have been deposited in the NCBI Gene Expression Omnibus (GEO): GSE72458.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and three figures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.02.060.

AUTHOR CONTRIBUTIONS

H.Z. designed, conducted, and analyzed experiments and performed all bioinformatics analyses. H.S., E.F., F.M.S., P.M., and B.S. performed and analyzed experiments. D.G. designed and analyzed experiments and supervised the project. The manuscript was written by D.G. and H.Z.

ACKNOWLEDGMENTS

We thank V. Narry Kim, Jaechul Lim, and Hyeshik Chang (Seoul, Korea) for advice on TAIL-seq library construction. We also thank Cécile Bousquet-Antonelli and Jean-Marc Deragon (Perpignan, France) for providing the anti-PAB2 antibodies. This work was supported by the Center National de la Recherche Scientifique (CNRS) and a research grant from the French National Research Agency as part of the "Investments for the future" program in the framework of the LABEX ANR-10-LABX-0036_NETRNA to D.G.

Received: August 26, 2015
Revised: December 8, 2015
Accepted: February 11, 2016
Published: March 10, 2016

REFERENCES

Baejen, C., Torkler, P., Gressel, S., Essig, K., Söding, J., and Cramer, P. (2014). Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition. Mol. Cell *55*, 745–757.

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.

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.

Charlesworth, A., Meijer, H.A., and de Moor, C.H. (2013). Specificity factors in cytoplasmic polyadenylation. Wiley Interdiscip. Rev. RNA *4*, 437–461.

Eliseeva, I.A., Lyabin, D.N., and Ovchinnikov, L.P. (2013). Poly(A)-binding proteins: structure, domain organization, and activity regulation. Biochemistry (Mosc.) *78*, 1377–1391.

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

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.

Kini, H.K., Silverman, I.M., Ji, X., Gregory, B.D., and Liebhaber, S.A. (2015). Cytoplasmic poly(A) binding protein-1 binds to genomically encoded sequences within mammalian mRNAs. RNA *22*, 61–74.

Knüsel, S., and Roditi, I. (2013). Insights into the regulation of GPEET procyclin during differentiation from early to late procyclic forms of Trypanosoma brucei. Mol. Biochem. Parasitol. *191*, 66–74.

Lange, H., Sement, F.M., Canaday, J., and Gagliardi, D. (2009). Polyadenylation-assisted RNA degradation processes in plants. Trends Plant Sci. *14*, 497–504.

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.

Lee, M., Kim, B., and Kim, V.N. (2014). Emerging roles of RNA modification: m(6)A and U-tail. Cell *158*, 980–987.

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.

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. EMBO J. *32*, 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.

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., and Bousquet-Antonelli, C. (2015). Heat-induced ribosome pausing triggers mRNA co-translational decay in Arabidopsis thaliana. Nucleic Acids Res. *43*, 4121–4132.

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.

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.

Munoz-Tello, P., Rajappa, L., Coquille, S., and Thore, S. (2015). Polyuridylation in Eukaryotes: A 3′-End Modification Regulating RNA Life. BioMed Res. Int. *2015*, 968127.

Norbury, C.J. (2013). Cytoplasmic RNA: a case of the tail wagging the dog. Nat. Rev. Mol. Cell Biol. *14*, 643–653.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2015). Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. Cell *161*, 1400–1412.

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. USA *111*, 6365–6370.

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.

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.

Sement, F.M., and Gagliardi, D. (2014). Detection of uridylated mRNAs. Methods Mol. Biol. *1125*, 43–51.

Sement, F.M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J.-M., Bousquet-Antonelli, C., Lange, H., and Gagliardi, D. (2013). Uridylation prevents 3′ trimming of oligoadenylated mRNAs. Nucleic Acids Res. *41*, 7115–7127.

Slevin, M.K., Meaux, S., Welch, J.D., Bigler, R., Miliani de Marval, P.L., Su, W., Rhoads, R.E., Prins, J.F., and Marzluff, W.F. (2014). Deep sequencing shows multiple oligouridylations are required for 3′ to 5′ degradation of histone mRNAs on polyribosomes. Mol. Cell *53*, 1020–1030.

Su, W., Slepenkov, S.V., Slevin, M.K., Lyons, S.M., Ziemniak, M., Kowalska, J., Darzynkiewicz, E., Jemielity, J., Marzluff, W.F., and Rhoads, R.E. (2013). mRNAs containing the histone 3′ stem-loop are degraded primarily by decapping mediated by oligouridylation of the 3′ end. RNA *19*, 1–16.

Thomas, M.P., Liu, X., Whangbo, J., McCrossan, G., Sanborn, K.B., Basar, E., Walch, M., and Lieberman, J. (2015). Apoptosis triggers specific, rapid, and global mRNA decay with 3′ uridylated intermediates degraded by DIS3L2. Cell Rep. *11*, 1079–1089.

Tu, B., Liu, L., Xu, C., Zhai, J., Li, 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.

Tuck, A.C., and Tollervey, D. (2013). A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. Cell *154*, 996–1009.

Viegas, S.C., Silva, I.J., Apura, P., Matos, R.G., and Arraiano, C.M. (2015). Surprises in the 3′-end: 'U' can decide too!. FEBS J. *282*, 3489–3499.

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.

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.

Zhang, X., Zhu, Y., Liu, X., Hong, X., Xu, Y., Zhu, P., Shen, Y., Wu, H., Ji, Y., Wen, X., et al. (2015). Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. Science *348*, 120–123.

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B., and Chen, X. (2012). The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. Curr. Biol. *22*, 689–694.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

3’ RACE-PCR to detect uridylated mRNAs

The 3' RACE PCR protocol used to sequence mRNAs 3’ ends is detailed in (Sement and Gagliardi, 2014). Briefly, 5 µg of total RNA were purified with Nucleospin® RNA plant columns (Macherey Nagel), dephosphorylated with Calf Intestinal alkaline Phosphatase (CIP, Promega) before ligation to the RNA primer (5’P-CUAGAUGAGACCGUCGACAUGAAUUC-3’NH_2) with T4 RNA ligase (Fermentas). This sequence was used as an anchor to initiate cDNA synthesis by SuperScript® III reverse transcriptase (Invitrogen) and the anchor rv1 primer (see primer list below). Two nested PCR amplifications of 30 cycles were performed using anchor rv1 and rv2 primers in combination with gene-specific sense primers. 3' RACE products were cloned in pGEM®-T Easy (Promega) prior to sequence analysis. For data presented in Figure 1, two and three biological replicates were analyzed in two distinct tissues, flowers and rosette leaves, respectively. Identical results were observed for both developmental stages, indicating that uridylation of these two mRNAs is not differentially regulated in these tissues. Datasets were therefore pooled to increase sample size, the number of clones analyzed for one gene in a given phenotype ranging from 190 to 317. For data presented in Figure 3 and Figure S3, at least two biological replicates were analyzed in rosette leaves. For Figure S3, analysis also includes clones from Sement et al., 2013. Altogether, data of seven genes, i.e. At4g20270 (*BAM3)*, At1g24160, At2g45160 (*LOM1)*, At2g21560, At5g46710, At5g48250 and At2g40340 (*DREB2C*), were pooled to increase sample size, allowing the analysis of 385 to 1516 clones for one given genotype. For data presented in Figure 3, data were analysed for five genes: At4g20270 (*BAM3)*, At1g24160, At2g45160 (*LOM1)*, At2g21560 and At5g46710. The number of clones analyzed for one gene in a given phenotype ranges from 41 to 121.

TAIL-seq library preparation

TAIL-seq libraries were prepared from WT, *urt1*, *xrn4* and *urt1 xrn4* two week-old seedlings according to (Chang et al., 2014). Total RNA was extracted using Tri-reagent (Euromedex), treated by DNase I (Fermentas) and purified by RNeasy MinElute column (Qiagen) according to manufacturer’s instruction. 10 µg of RNA were depleted from rRNA using Ribo-Zero Magnetic Kit (Plant Leaf, Epicentre). Depleted RNAs were ligated to 10 pmol of the 3’ adaptor (see primer list) using 10 U of truncated T4 RNA ligase 2 (NEB). 3’ ligated RNA were then partially digested with 2 U of RNase T1 (Invitrogen) during 5 minutes at 22°C. 3’ RNA fragments were pull down using streptavidin beads (Dynabeads® M-280 Streptavidin, Invitrogen) and phosphorylated using 20 U of T4 Polynucleotide Kinase (ThermoScientific). RNAs were purified on denaturing 6% polyacrylamide gel and fragments ranging from 500nt to 1500nt were eluted from the gel. The purified RNAs were ligated to 5 pmol of the 5’ adaptor (see primer list) using 8 U of T4 RNA ligase 1 (NEB) and subjected to reverse-transcription using 200 U of Superscript III reverse-transcriptase (Invitrogen). cDNAs were amplified during 19 cycles using Phusion polymerase (Phusion High-Fidelity PCR Master Mix, ThermoScientific), 25 pmol of PCR primer fw and 25 pmol of PCR primer rev in 50 µl final volume. To discriminate samples, different barcodes were integrated during the PCR step. PCR products were purified both on 6% polyacrylamide gel and using AMPure XP beads (Beckman Coulter, Inc.). Libraries were sequenced on Illumina HiSeq 2500 (50 x 240 paired-end run, HiSeq Rapid Run (RR), HiSeq Rapid Flow Cell v) with 20% of the PhiX control library and 1% of the spike-ins mixture (see below).

Spike-in preparation

Spike-ins (IDT ultramer service), with different lengths of polyA stretches (from 0 A to 118 A), were added to the libraries. Each spike-in was amplified during 10 cycles using Phusion polymerase (Phusion High-Fidelity PCR Master Mix, ThermoScientific), 5 pmol of PCR primer fw and 5 pmol of PCR primer rev in 10µl final volume. Different barcodes for each spike-in were integrated during PCR. Spike-ins were then purified as for TAIL-seq library preparation.

TAIL-seq data processing

The base calls were acquired from HiSeq 2500 after processing by Illumina CASAVA-1.8.2. Sequencing quality was assessed with a PhiX reference spiked in the flow-cell that gave a sequencing error rate of 0.13% for read1 and 0.45% for read 2 with 97% and 89% of the reference nucleotides having a Q-score >=30 for read 1 and read 2, respectively (Phred calculation). In order to identify transcripts, read 1 sequences were mapped onto the *Arabidopsis thaliana* reference genome (TAIR 10) using Tophat (v2.0.9) (Langmead et al., 2009; Trapnell et al., 2009; Kim and Salzberg, 2011; Kim et al., 2013) and Bowtie2 (Langmead and Salzberg, 2012). Two mismatches were allowed. After removing multimapped reads, 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 onto cytosolic mRNAs and their corresponding reads 2 were extracted and used for further analyses. The sequences having completely identical nucleotides in read 1 (insert) and the 1st to 15th cycle in read 2 (degenerate bases in 3' adapter) were deduplicated. The degenerated sequence in 3' adapter was clipped out from read 2 by searching perfect match of delimiter sequence ('GTCAG' as in the direction of read 2) between 10th and 20th cycles in read 2. Reads 2 missing a delimiter sequence were removed from further analyses. In order to extract poly(A)-free 3' ends, the 40nt next to the 3' adapter of the reads 2 were mapped onto the Arabidopsis genome using Tophat with two allowed mismatches. A python script was then developed to extract polyA tails (>3nt) and their potential 3' end modification from the remaining read 2. Poly(A) were detected with a constraint that it must begin within the first 30 cycles, so the maximum detectable 3' end modification of poly(A) tails was limited to the last 29 nucleotides of insert. For poly(A) tails longer than 8nt, up to 3 non consecutive nucleotides other than A were tolerated in the polyA tail in order to include heteropolymeric poly(A) tails in the analysis. As the base call protocol is suitable to measure poly(A) tail up to 30 nt (Figure S2) (Chang et al., 2014), all plots and statistics were performed for reads with poly(A) tail <31nt. By doing so, we identified 10304, 43391, 9797 and 26300 poly(A) tail sequences <31nt that mapped to 2716, 5501, 2571 and 4077 unique genes for WT, *urt1*, *xrn4* and *urt1xrn4,* respectively. Of note, focusing on cytosolic mRNA analysis by TAIL-seq from plant samples present a major drawback: the overrepresentation of plastidial sequences, which represented about two thirds of the mapped reads. In fact, to analyze tailing of nuclear-encoded mRNAs, all reads corresponding to cytosolic organellar coding and non-coding RNAs, nuclear-encoded non-coding RNAs and a high number of reads that do not contain the delimiter sequence present at the end of the ligated RNA primer had to be removed. The future analysis of plant RNA samples by TAIL-seq will require technical optimization to allow gene-to-gene comparison. Yet, the numbers of reads obtained still allows an unbiased view of mRNA tailing.

Small RNA data processing

The base calls were acquired from HiSeq 2000 after processing by Illumina CASAVA-1.8.2. Sequencing quality was assessed with a PhiX reference spiked in the flow-cell that gave an average sequencing error rate of 0.05% with 98,71% of the reference nucleotides having a Q-score >=30 for read 1 and read 2, respectively (Phred calculation). 5' and 3' adapter were searched and removed using in-house Fasteris pipeline. Trimmed reads of 14nt to 30nt in length were selected to be mapped to TAIR10 genome using Bowtie (Langmead et al., 2009) by permitting up to 1 mismatch in the first 14 read nucleotides and no limit beyond. Reads that could map to more than 50 loci were discarded and only the best alignment(s) of each small RNA sequence was kept. The resulting alignment was annotated using the intersectBed tool from the BEDTools suite (v2.17.5) (Quinlan and Hall, 2010) and, either the Arabidopsis annotation file ath.gff with 427 mature miRNA sequences (http://www.mirbase.org/) either a custom annotation file with genome coordinates for 8 siRNA loci (siR-1: CACGGTCTTGGCTCAACCCGCC, siR-2: GAACGCTATGTTGGACTTAGGATG, siR-3: ATTATCCCCCGTGTTTTGTCC, siR-4: TGCCCGGACCCTGTCGCAC, siR-5: TTAGATTCACGCACAAACTCT, siR-6: CCTTTGCAGACGACTTAAATAC, siR-7: TCCGCTGTAGCACACAGGC, siR-8: CCTTTGTCGCTAAGATTCGA). Thus, reads mapping to miRNA, from 1925296 to 2783064, or to siRNA loci, from 310136 to 698499, were extracted and further used for tailing and trimming analysis. Tailing and trimming at the 3' end of small RNA was

identified using a homemade python script. Canonical variants were first isolated by looking for reads with perfect match for miRNA and siRNA sequences. Then, miRNA and siRNA sequences were successively trimmed from 1 nt to 7 nt in order to identify trimmed variants. At each matching step, 3' tails were extracted, *i.e.* untemplated nucleotides at the 3' end of sRNA reads. To minimize bias due to reads with sequencing errors, that would be considered as having 3' untemplated nucleotide, if the nucleotide after the first mismatch matched to the miRNA/siRNA sequence, the read was removed from the analysis. During the quantification process, multi-mapped reads were weighted by a factor 1/n, n being the number of identical loci where reads mapped. By doing so, we were able to analyse the amount and nature of trimming and tailing for 1925296 to 2783064 reads and 310136 to 698499 for miRNAs and siRNAs, respectively.

Poly(A) tail length assays

Total RNAs were extracted using Tri-reagent (Euromedex) from two week-old seedlings for two biological replicates of WT. One µg of RNA was labelled using 2U of Cid1 poly(U) polymerase (NEB) and 40µCi of [α-^{32}P]-UTP during one hour at 37°C. RNAs were then purified by RNeasy MinElute column (Qiagen) according to manufacturer's instruction and digested with 85U of RNAse T1 and 340ng of RNase A (Fermentas) during one hour at 37°C. RNAs were finally purified using 30µl (300µg) of oligod(T) beads (Ambion® Poly(A)Purist™ MAG Kit) according to manufacturer's instruction. Two controls were performed: 1) a non-adenylated 21-nt RNA substrate which was labelled using 10U of T4 Polynucleotide Kinase and 25µCi of [γ-^{32}P]-ATP and 2) a poly(A) mixture (Polyadenylic acid, Sigma) labelled with Cid1, as previously described, and purified or not on oligo(dT) beads. Purified poly(A) tails and control samples were separated on denaturing 10% (v/w) polyacrylamide gel electrophoresis before autoradiography. The relative signal density over each lane was retrieved using Image J software and used to plot the distribution of poly(A) tail between 8 and 30nt.

Statistical analyses

All plots and statistics were performed using the R software (R version 3.0.2). Plots were drawn using the ggplot2 R package (v1.0.0, http://cran.r-project.org/web/packages/ggplot2/). Mann-Whitney and Pearson's chi-squared tests were used to compare medians and frequencies, respectively, between two populations. Values were considered as significantly different when p-value <0.05. The linear correlation between two populations was estimated using the Pearson product-moment correlation coefficient. The sm.density.compare function of the SM R package (https://cran.r-project.org/web/packages/sm) was used to compare the distribution of extension sizes between two genotypes: this function allows a set of univariate density estimates to be compared in a permutation test of equality.

Supplemental references

Kim, D., and Salzberg, S.L. (2011). TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol. *12*, R72.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinforma. Oxf. Engl. *26*, 841–842.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinforma. Oxf. Engl. *25*, 1105–1111.

Sequences of primers used in this study. All adapters and primers were synthesized by IDT

Primers used for 3' RACE-PCR experiments		
Anchor 3' RACE	anchor rv1	GAATTCATGTCGACGGTCTCA
	anchor rv2	CATGTCGACGGTCTCATCTAG
At4g20270	BAM3 fw1	GGAAGAAGGGATAGACATTGTGC
	BAM3 fw2	TGTGGCAATGCTATGTGTGC
At1g24160	At1g24160 fw1	GGTTCTCTTCCAATGCGGA
	At1g24160 fw2	CCACCAATGTTGTTCCATCAG
At2g45160	LOM1 fw1	TGTTGATGAAACGACACCGT
	LOM1 fw2	AGTTTGAGTCAGATGGCGGA
AT2G21560	At2g21560 fw1	TCAGAGAATCACGAGGGCTC
	At2g21560 fw2	GAGGAGGAACCAAAAGGATGA
AT5G46710	AT5G46710 fw1	GCTTCCGCTTCTGCTCTATTG
	AT5G46710 fw2	TCAGTCCACCTACACCTCAAT
Primers used for genotyping		
SALK T-DNA	Salk LBb1	ATTTTGCCGATTTCGGAAC
At2g45160	URT1- fw1	CCGTCAATAATCCTTTTCCTCCG
	URT1- rv1	CCAAACTTTAGAATGCTCGCCC
At2g45161	XRN4-fw	TCCCATGAGAGCCATGCATTC
	XRN4-rv	ACCATCCTGCAGGTCCAAGAA
Primers used for cloning URT1 in pGWB621		
At2g45160	URT1-fw2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGGCGGACGGTGGGG
	URT1-rv2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCGGTGGAAGAGCTCGTCG
Primers used for preparation of TAIL-seq libraries		
5' ligated adaptor	RNA primer RA5	GUUCAGAGUUCUACAGUCCGACGAUC
3' ligated adaptor	DNA primer RA3	/5App/CTGACNNNNNNNNNNNNNNNTGGAATTCTCGGGTGCCAAGGC/iBiodT/ /iBiodT//3ddC/
Primer for RT	RT primer	GCCTTGGCACCCGAGAATTCCANNNNNNNNNNNNNNNGTCAG
Spike-ins	Poly A Spike_N	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNNNNNNNNNCTGAC GAGCTACTGTTGGAATTCTCGGGTGCCA
	Poly A Spike_8	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNBAAAAAAAACTGACG AGCTACTGTTGGAATTCTCGGGTGCCA
	Poly A Spike_16	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNNBAAAAAAAAAAAAAA AACTGACGAGCTACTGTTGGAATTCTCGGGTGCCA

	Poly A Spike_32	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNBAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAACTGACGAGCTACTGTTGGAATTCTCGGGTGCCA
	Poly A Spike_64	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNBAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTGA CGAGCTACTGTTGGAATTCTCGGGTGCCA
	Poly A Spike_118	TCAGAGTTCTACAGTCCGACGATCNNNNNNNNNNNNNNBAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACT GACGAGCTACTGTTGGAATTCTCGGGTGCCA
Fw primer for PCR	PCR Primer fw	AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGA
Rev primers for PCR	PCR Primer rev, Index 1	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCCTTGGCACCCG AGAATTCCA
	PCR Primer rev, Index 2	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCCTTGGCACCCG AGAATTCCA
	PCR Primer rev, Index 3	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCCTTGGCACCCG AGAATTCCA
	PCR Primer rev, Index 4	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCCTTGGCACCCG AGAATTCCA
	PCR Primer rev, Index 5	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCCTTGGCACCCGA GAATTCCA
	PCR Primer rev, Index 6	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCCTTGGCACCCG AGAATTCCA
	PCR Primer rev, Index 7	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCCTTGGCACCCG AGAATTCCA
	PCR Primer rev, Index 9	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCCTTGGCACCCGA GAATTCCA
	PCR Primer rev, Index 10	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCCTTGGCACCCG AGAATTCCA
	PCR Primer rev, Index 11	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCCTTGGCACCCG AGAATTCCA
DNA primers used for generating AGO1 polyadenylated RNA substrate for testing URT1 activity		
AT1G48410	AGO1-fw	CCAGAGACATCAGACAGTG
	AGO1-rv	TTTTTTTTTTTTTTTGAAAGGATCAAAGTCTGTT
RNA primer used for testing URT1 activity		
21nt RNA substrate		UCGCUUGGUGCAGGUCGGGAA

Vectors used in this study.

	AGI	Name	Cloned sequence	Vector
Arabidopsis transgenic lines	At2g45620	myc-URT1	URT1 gDNA	PGWB621
Production of recombinant proteins in *E coli*	At2g45620	6-his-GST URT1	URT1 cDNA	PHGGWA
	At4G34110	6-his-GST PAB2	PAB2 cDNA	PHGGWA
			GST cDNA	POGWA

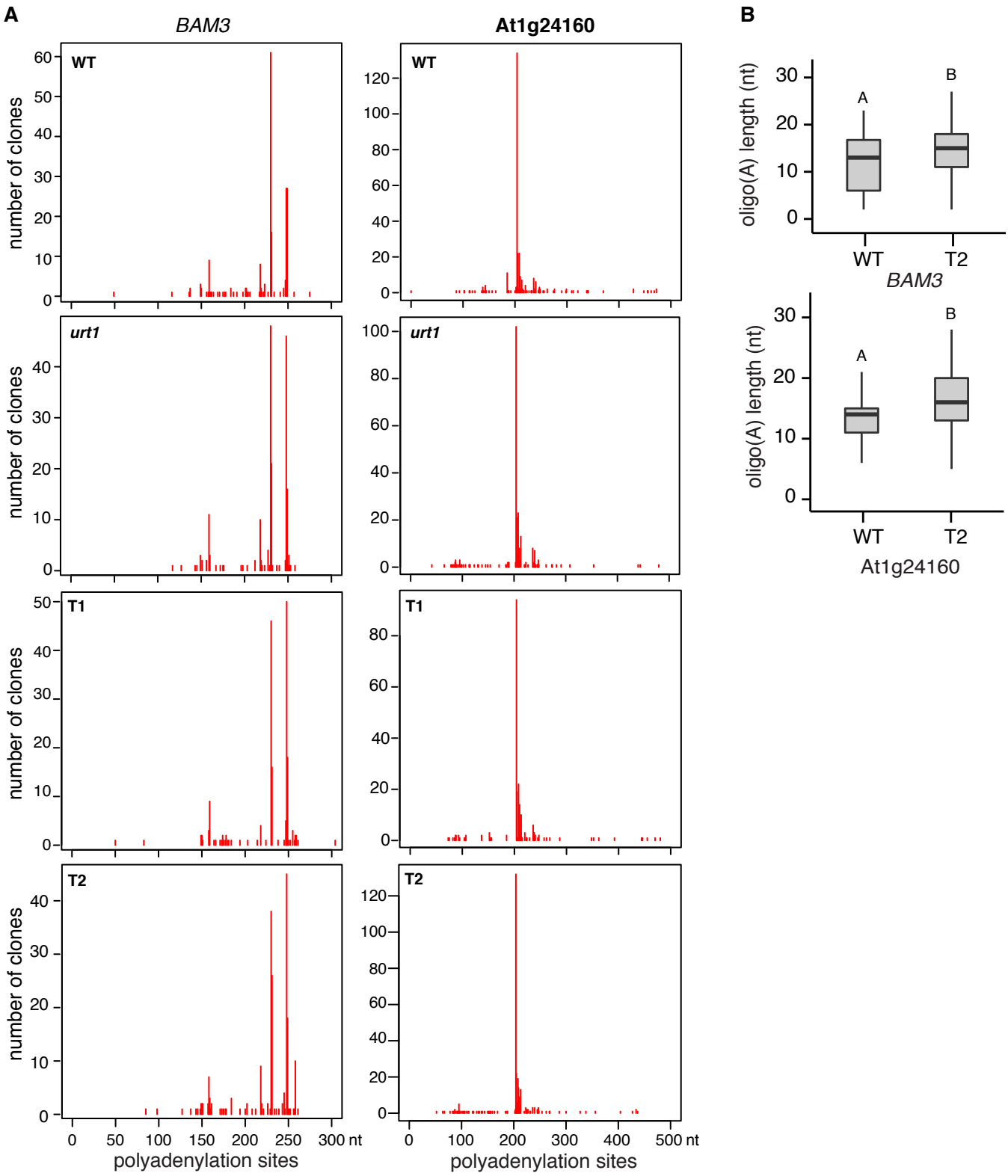


Figure S1, related to Figure 1
Identical poly(A) sites for *BAM3* and At1g24160 in WT, *urt1*, *T1* and *T2* lines and increase of uridylated oligo(A) tail size in T2 as compared with WT.

A) Numbers of clones (n=797 for *BAM3*, n=1120 for At1g24160) from at least two independent experiments were plotted against the 3' region of each respective mRNA as indicated.

B) Box plot analysis of poly(A) length distribution for uridylated *BAM3* and At1g24160 mRNAs in WT and T2 lines. The upper and lower edges correspond to the first and third quartiles, respectively. The median is indicated by a horizontal bar and whiskers show data range except far-outliers. Significant differences determined by Mann-Whitney tests are indicated by letters.

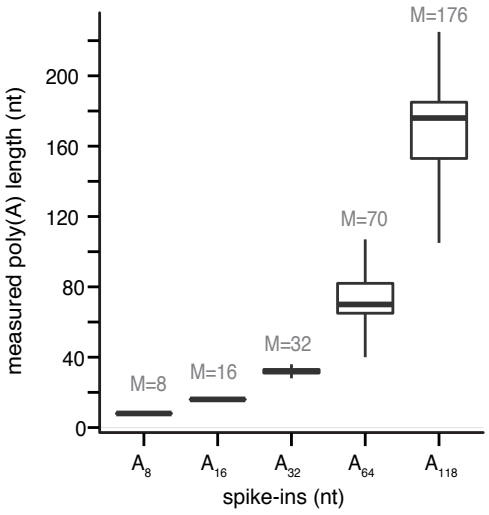
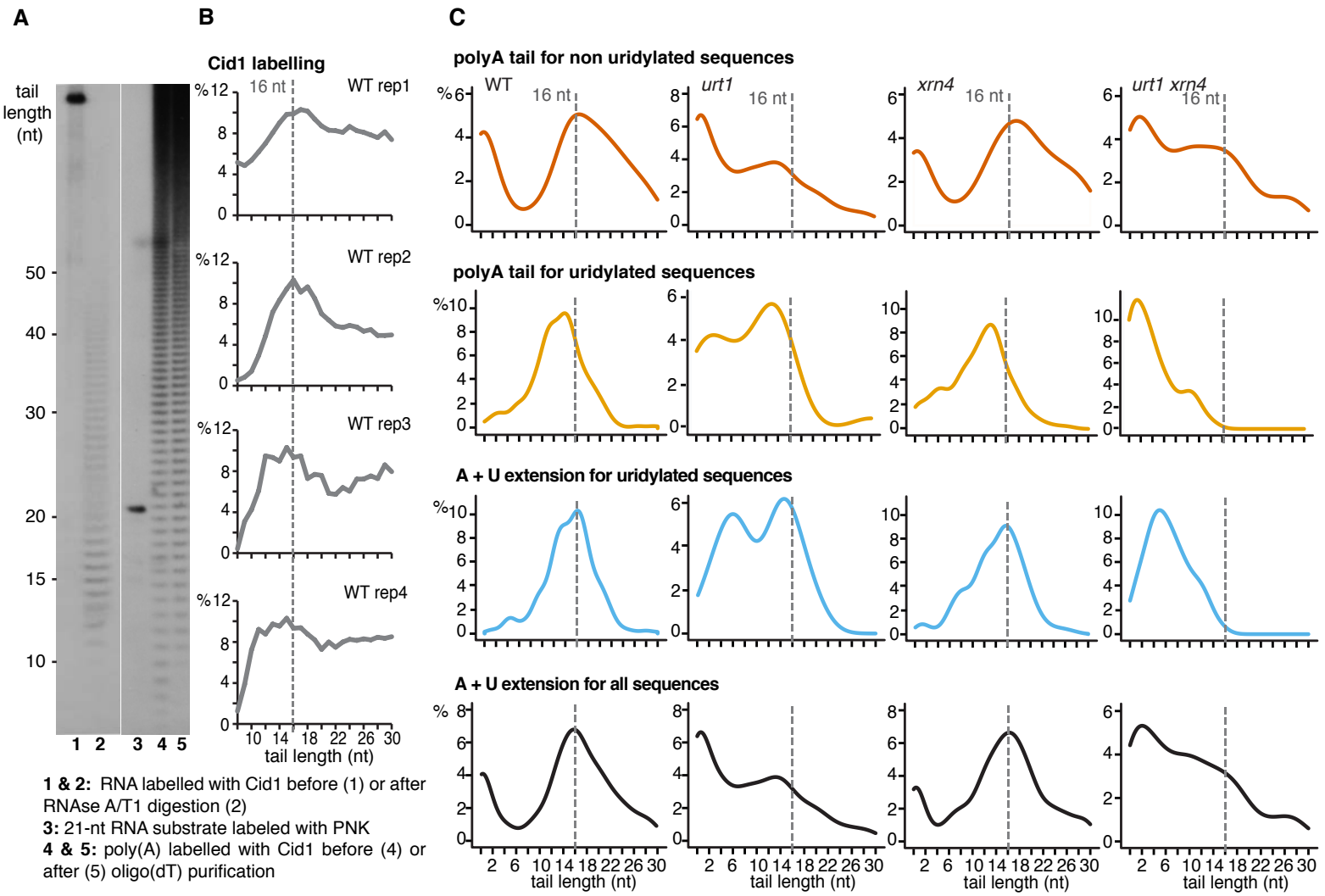


Figure S2, related to Figure 2

Base-call analysis protocol is suitable for analyzing poly(A) tails up to 30 As.

Box plot analysis of poly(A) length distribution for the synthetic spike-ins calculated using the base-call analysis protocol (see Experimental Procedures). The upper and lower edges correspond to the first and third quartiles, respectively. The median is indicated by a horizontal bar and whiskers show data range except far-outliers. M = median.



1 & 2: RNA labelled with Cid1 before (1) or after RNAse A/T1 digestion (2)
3: 21-nt RNA substrate labeled with PNK
4 & 5: poly(A) labelled with Cid1 before (4) or after (5) oligo(dT) purification

Figure S3, related to Figure 4

A population of oligo(A)-tailed mRNAs exists in Arabidopsis and URT1-mediated uridylation repairs deadenylated mRNAs.

(A,B) Distribution of oligo(A) tails for non uridylated sequences obtained by a method combining poly(U) polymerase labelling and RNase A/T1 treatment. To validate the presence of short oligo(A)-tailed mRNAs in Arabidopsis with a method devoid of any potential PCR bias, RNA from WT samples were 3' end-labeled with [α-32P]-UTP and digested with a mixture of RNases A and T1 (Supplemental Experimental Procedures). Digestion products were then separated on high-resolution polyacrylamide gels to visualize oligo(A) tail length. As control, a non-adenylated 21-nt RNA substrate, labelled by T4 Polynucleotide Kinase, and a poly(A) mixture, labelled with Cid1 and purified or not on oligo(dT) beads, were run along the samples. The experiment was performed for two different WT RNA, with two technical replicates. Note that oligo(A) tails modified by G, C and U addition will not be detected in these assays since the RNase A and T1 digestion step will remove the 3' label. In line with TAIL-seq results (Figure 4), oligo(A) tails centered at 16 As were also visualized on gels and integration of the four profiles confirmed the existence of this population. Although adenylation of RNAs other than mRNAs can potentially contribute to the observed profiles, these results fit well with the existence of mRNAs with short (A) tails as detected by TAIL-seq. The prevalence of oligo(A) tails centered on 16 As observed both by TAIL-seq and by 3' end labeling experiments likely reveals the footprint of a factor bound to these oligo(A) tails.

(B) Plots display relative signal density retrieved between 8 nt and 30 nt for each experiment using the ImageJ software.

(C) Distribution of tail sizes determined by compiling 1516, 1152, 1022 and 385 sequences for seven model mRNAs analysed by 3' RACE PCR. Plots display a smooth density estimate of the extension sizes (1D gaussian kernel).

Deuxième partie :

Le 3'RACE-seq , méthode d'analyse par séquençage à haut-débit des extrémités 3' d'ARN candidats

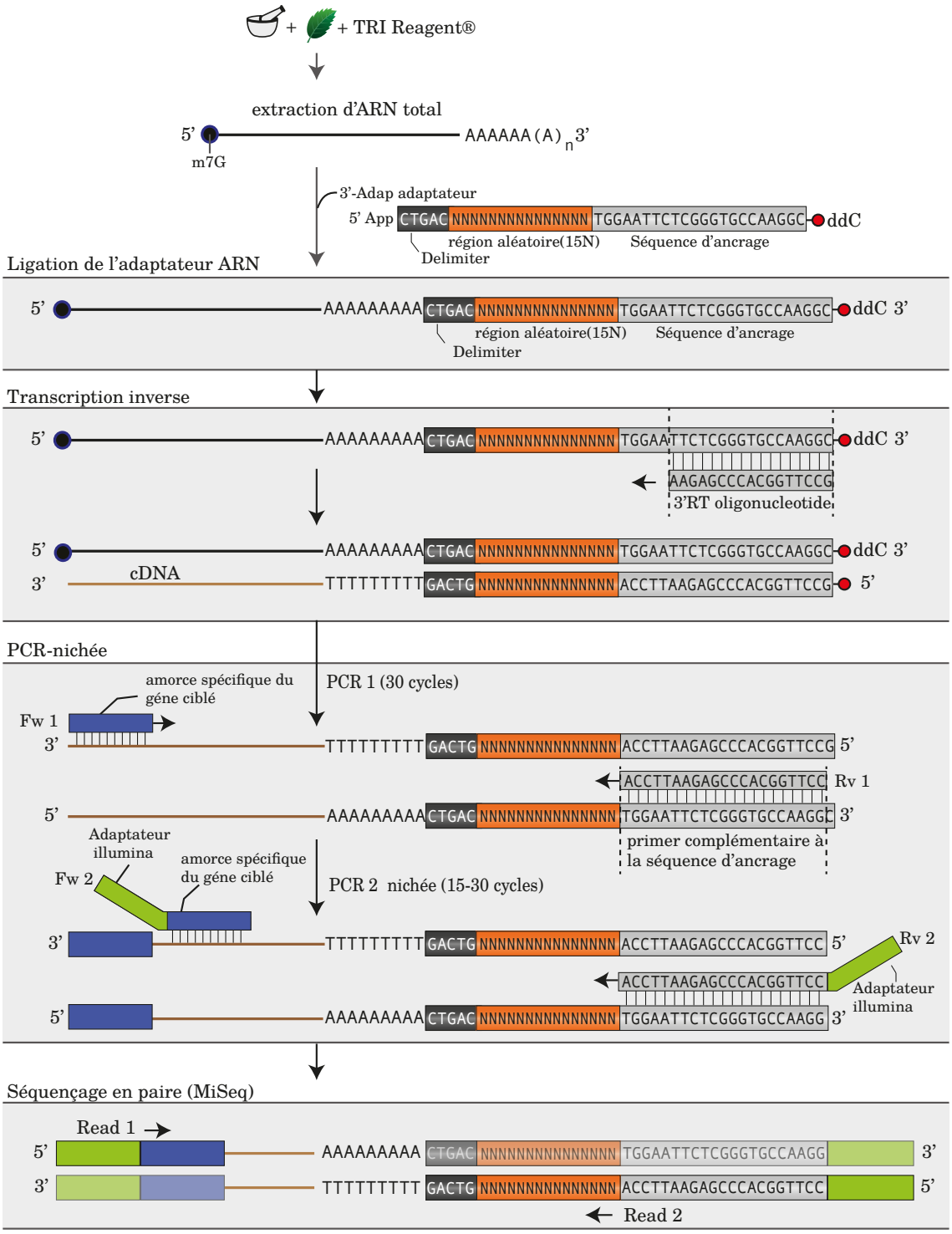


Figure 7 | Schéma des étapes principales du 3'RACE-seq
Le 3'RACE-seq débute par la ligation d'un adaptateur à l'extrémité 3' des ARN. La synthèse de cDNA est initiée en utilisant une amorce complémentaire à la séquence d'ancrage de l'adaptateur en 3' (3'RT oligonucléotide). Les extrémités 3' des cibles d'intérêts sont amplifiées de manière spécifique lors de la PCR nichée grâce à l'utilisation d'amorces sens complémentaires au transcrit d'intérêt (Fw 1, Fw 2) et antisens, complémentaire à l'adaptateur en 3' (Rv 1, Rv 2). Les amorces de la PCR2 sont supplémentées avec des séquences qui permettent la liaison à la « flow-cell » illumina et le séquençage de la Read 1 et de la Read 2. Une fois produite les librairies sont purifiées, quantifiées et leur qualité est déterminée avant de procéder au séquençage en paire (paired-end sequencing). La Read 1 permet d'identifier le transcrit et le Read2 permet d'analyser l'extrémité 3'.

L'étude de l'uridylation implique la comparaison du statut des extrémités 3' entre différents génotypes, différents tissus ou pour différentes conditions. Il est donc nécessaire de disposer d'une technique permettant l'analyse, avec une grande profondeur et pour un coût raisonnable, de plusieurs échantillons en parallèle (multiplexage). La méthode de TAIL-seq utilisée en partie 1 est difficilement utilisable en routine, en majeure partie du fait de sa complexité et de son coût. En outre, nous n'obtenons pas à l'heure actuelle une profondeur suffisante pour permettre une analyse gène-à-gène des extrémités 3' des ARNm chez *Arabidopsis thaliana*. Nous avons en conséquent mis au point au laboratoire, le 3' RACE-seq, une méthode d'analyse par séquençage à haut-débit des extrémités 3' d'ARN candidats.

Cette méthode permet l'analyse ciblée de l'uridylation et de la taille des queues poly(A) des ARNm. Elle a également été adaptée à l'analyse des extrémités 3' d'autres ARN tels que les fragments 5' issus du clivage d'ARNm par le complexe RISC ou de sous-produits de maturation des ARN ribosomiques.

1. Analyse des extrémités 3' des ARNm par 3'RACE-seq

1.1 Résumé de la technique

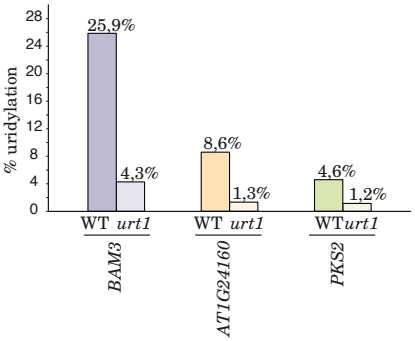
Le protocole de 3'RACE-seq mis au point au laboratoire est schématisé en Figure 7. Il débute par la ligation d'un adaptateur à l'extrémité 3'OH des molécules d'ARN. L'adaptateur choisi, est similaire à celui décrit pour le TAIL-seq (Partie 1 et (Chang *et al.*, 2014) excepté qu'il n'est pas biotinylé. Cet adaptateur contient 5 nucléotides à son extrémité 5' qui servent de délimiteur, suivis d'une région de 15 nucléotides de composition aléatoire (15N). La séquence du délimiteur sert à identifier, lors de l'analyse bioinformatique, les reads qui correspondent à de vraies extrémités 3' d'ARN et à éliminer ceux qui sont issus d'artefacts expérimentaux comme l'amorçage non-spécifique de la synthèse de l'ADN complémentaire (ADNc) pendant l'étape de transcription inverse. La séquence 15N constitue quant à elle un identifiant moléculaire unique qui permet de d'éliminer les reads qui correspondent à des dupliquas de PCR (étape de déduplication).

L'étape de ligation est suivie par la production d'ADNc par transcription inverse (RT) avec une amorce complémentaire à l'adaptateur en 3'. Lors des premiers essais de 3'RACE-seq, nous détections dans nos librairies une forte proportion de séquences qui ne contenaient pas la séquence du délimiteur et qui ne correspondaient pas aux extrémités 3' d'ARN. Ces artefacts étaient dus à la liaison aspécifique de l'amorce utilisée lors de la transcription inverse. Hélène Zuber, a dès lors conçu une amorce de RT plus courte que celle utilisée précédemment et complémentaire aux 17 derniers nucléotides de la séquence d'ancrage de l'adaptateur en 3'. La séquence de l'amorce utilisée lors de l'étape d'amplification PCR qui suit est *a contrario* complémentaires aux 22 nucléotides de la séquence d'ancrage de l'adaptateur en 3'. Cette astuce prévient l'amplification de produits de RT non-spécifiques.

Les ADNc sont sujets à deux étapes de PCR semi-nichées à l'aide d'amorces sens complémentaires aux transcrits d'intérêts et antisens complémentaires à l'adaptateur en 3'. Les amorces utilisées lors de la seconde étape de PCR sont supplémentées avec des séquences permettant le multiplexage

A

	BAM3	AT1G24160	*PKS2*
WT	24110	64362	15642
urt1	29666	77732	32109
TOTAL	51376	142094	47751



B

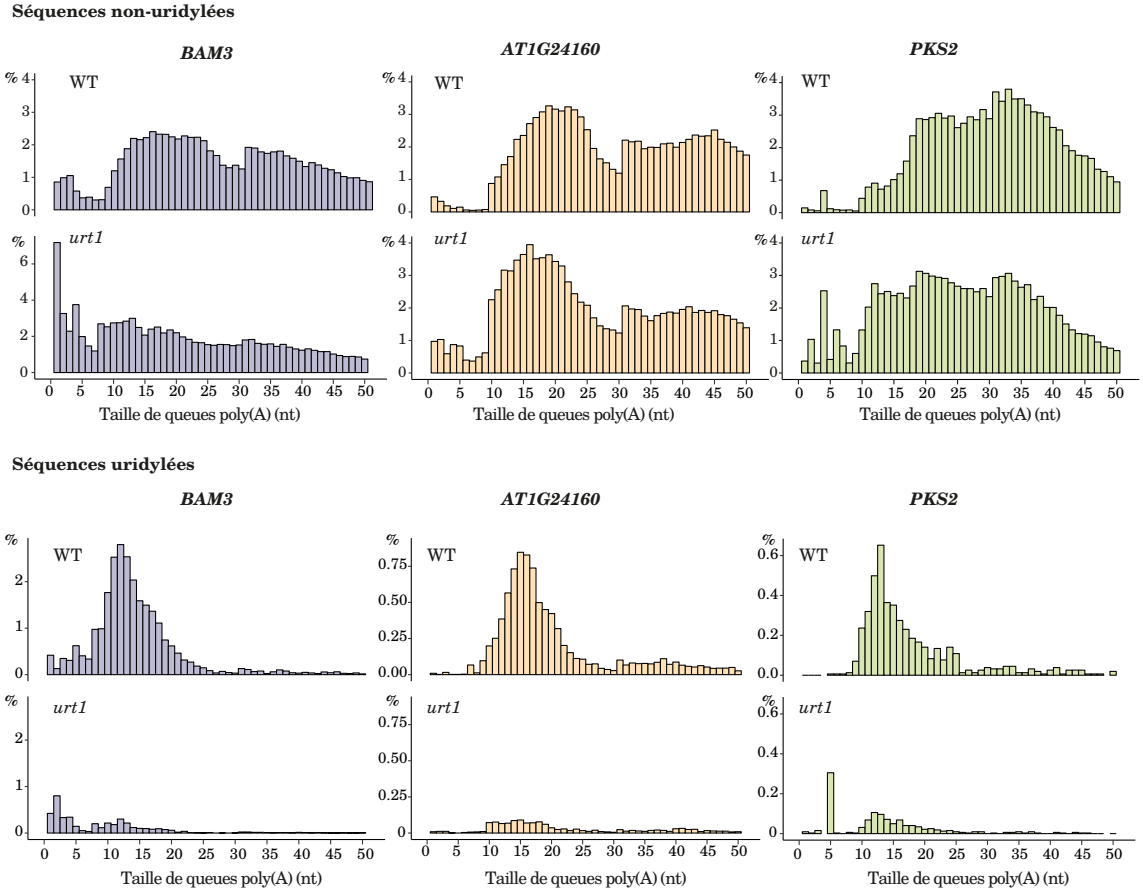


Figure 8 | Analyse des extrémités 3' de 3 transcrits par 3' RACE-seq
Des expériences de 3' RACE-seq ont été réalisées à partir d'ARN de fleurs de 4 réplicas biologiques de plantes WT et *urt1*. 3 gènes modèles ont été analysés: *BAM3* (pourpre), *AT1G24160* (orange) et *PKS2* (vert). **A|** Tableau récapitulant le nombre de reads obtenus pour chaque gène et chaque génotype. **B|** Diagramme en barres représentant le pourcentage de transcrits uridylés pour chaque transcrit et chaque génotype déterminé par 3' RACE-seq. **C|** Diagrammes en barres représentant la distribution de la taille des queues poly(A) déterminée par 3' RACE-seq pour les transcrits non-uridylés et uridylés. Le pourcentage de reads à chaque position a été normalisé par rapport au nombre total de reads correspondant à des tailles de queues poly(A) comprises entre 1 et 50A.

des échantillons, l'hybridation et le séquençage sur la cellule (« flow cell ») Illumina (amorce Fw2 et Rv2, Fig. 7). Les librairies obtenues sont purifiées avec des billes AMPure XP et quantifiées au fluorimètre Qubit (Invitrogen). *In fine*, la concentration de chaque banque est établie en tenant compte de la distribution de la taille estimée avec le système bioanalyzer 2100 (Agilent).
Le séquençage est réalisé en paire (paired end sequencing) grâce au système de séquençage Miseq, la read 1 est utilisée pour identifier le transcrit et la read 2 sert à identifier la taille et le statut des nucléotides ajoutés en 3'.
L'une des principales difficultés rencontrées lors de la mise au point du 3'RACEseq a été d'obtenir une bonne qualité de séquençage. La faible diversité de séquence des amplicons de PCR, et le biais en composition de base due à la présence de queues poly(A) affectent le score de qualité et le taux d'erreur obtenus en sortie de séquençage. Pour pallier ces problèmes il a fallu ajuster la concentration des librairies de 3'RACE-seq et de PhiX jusqu'à trouver un juste équilibre entre qualité et profondeur de séquençage. Le PhiX est une librairie fournie par Illumina, dérivée du génome du bactériophage PhiX et qui possède une composition équilibrée des 4 bases. L'ajout de cette librairie permet de compenser pour la faible diversité des librairies de 3'RACE-seq. Après séquençage, les reads ayant une trop faible qualité de séquençage (Q<10), correspondant à des séquences dupliquées ou ne contenant pas le délimiteur sont éliminés de l'analyse. En moyenne 85% des reads sont éliminés après déduplication et 15% des reads restants ne contiennent pas le délimiteur, ce qui illustre l'importance de la séquence 15N et de la séquence du délimiteur dans ce protocole de séquençage.

1.2 Preuve de concept

Afin de démontrer de l'efficacité et de la résolution de notre méthode de 3'RACE-seq nous avons analysé la taille de queues poly(A) de 3 transcrits issus des gènes *BAM3* (*AT4G20270*), *AT1G24160* et *PKS2* (*AT1G14280*) à partir d'ARN extraits de plantes WT et mutantes *urt1-1*. Le nombre de reads correspondant à des queues poly(A) comprises entre 1 et 50 As analysés pour chaque gène et chaque génotype est indiqué en Figure 8A.
Le taux d'uridylation des queues poly(A) comprises entre 1 et 50 As varie entre les gènes analysés (Fig. 8), avec *BAM3* étant le transcrit avec le taux d'uridylation le plus élevé (25,9%). Dans les trois cas et conformément à ce qui est attendu, le pourcentage de reads qui correspond à des transcrits uridylés est considérablement diminué dans le mutant *urt1* (Fig. 8B) confirmant que URT1 est la TUTase principalement responsable de l'uridylation des ARNm. A l'instar du TAIL-seq, nous détectons par cette méthode une uridylation résiduelle due à l'existence d'au moins une seconde TUTase capable d'uridyler les ARNm.
Dans nos mains, le TAIL-seq n'avait pas permis d'obtenir une profondeur suffisante pour permettre une analyse gène-à-gène de la distribution des queues poly(A). La distribution était calculée à partir des tailles de queues poly(A) d'un ensemble de gènes et aboutissait sur un pic centré à 16nt quand les extensions nucléotidiques de 4 à 30 étaient analysées (partie 1). L'analyse par 3'RACE-seq révèle quant à elle une distribution bimodale de la taille des queue poly(A) qui fluctue en fonction

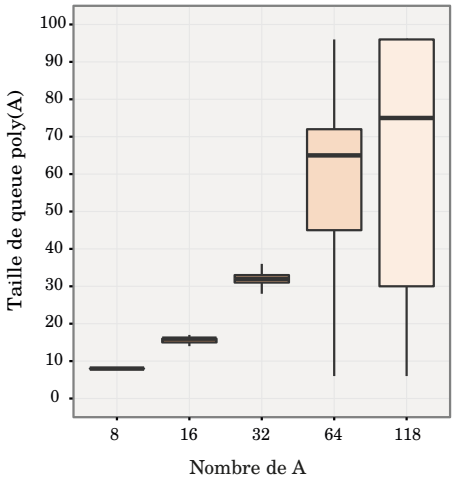


Figure 9 | Analyse de la distribution des tailles des queues poly(A) des spike-in synthétiques par 3'RACE-seq
Les bords supérieurs et inférieurs de la boîte à moustache représentent le 1er et le 3ème quartiles, respectivement. La droite horizontale au centre de la boîte à moustache indique la valeur médiane. Les moustaches indiquent l'étendue de la variable sans tenir compte des valeurs extrêmes.

du transcrit considéré (Fig. 8B). Dans les 3 cas nous observons un léger décalage de cette distribution dans le mutant *urt1* avec une accumulation de transcrits déadénylés par rapport à ce qui est observé dans WT. De manière intéressante, cet effet est le plus notable dans *BAM3*, transcrit pour lequel le taux d'uridylation est le plus élevé.
La distribution de taille des queues poly(A) des transcrits uridylés (Fig. 8B) est en revanche parfaitement similaire à ce qui est observé en partie 1 avec un unique pic centré entre 12 et 15 nucléotides. Cette population est considérablement diminuée en l'absence de *urt1* ce qui confirme qu'elle constitue le substrat préférentiel de URT1.

Ces analyses valident l'utilisation du 3'RACE-seq comme méthode d'analyse par séquençage à haut-débit des extrémités 3' d'ARN candidats. Le protocole est applicable en routine au laboratoire et offre la possibilité de multiplexer plusieurs réplicas, conditions ou génotypes en parallèle.

1.3 Limitation du 3' RACE-seq et de ses applications

1.3.1 Estimation de la taille des queues poly(A) et biais techniques

Une limitation de la technique de 3' RACE-seq est une mauvaise estimation de la taille des longues queues poly(A), à l'instar de ce qui est observé avec le TAIL-seq (Chang et al., Zuber et al.). Ce problème est dû à une mise en phase (phasing) ou pré-phase (pre-phasing), c'est à dire un décalage qui a lieu lorsque la lecture d'un brin est déphasé par rapport au cycle en cours. Une mauvaise mise en phase et en pré-phase entraine une sous-estimation ou une surestimation de la taille des homopolymères et dans notre cas de celle des queues poly(A). Le séquençage de spike-in[1] synthétiques révèlent en effet une forte imprécision dans l'estimation de la taille des queues poly(A) à partir de 64 adénosines (Fig. 9). Cette imprécision pourrait être également attribuable à une qualité moindre des spike-in de longues tailles qui sont relativement difficiles à produire. Une imprécision de la taille des longues poly(A) n'est cependant pas problématique pour l'analyse des queues poly(A) uridylées dans la mesure où l'uridylation affecte des ARNm ayant des populations de queues poly(A) relativement courtes et que le séquençage Illumina est très fiable pour des queues poly(A) allant au moins jusqu'à 32A. En revanche, ce biais technique pourrait s'avérer problématique pour détecter des changements de distribution de taille de queues poly(A) longues (>50-60). Il peut être résolu en analysant les données de fluorescences à la place des données de « base-calling» comme le propose (Chang *et al.*, 2014). Enfin, il faut également être conscient du fait que les distributions obtenues par 3'RACE-seq ou par TAIL-seq ne reflètent pas la distribution absolue de la taille des queues poly(A). D'une part, parce que la polymérase amplifie probablement plus efficacement des matrices courtes lors des étapes de PCR et d'autre part, parce que les

[1] Les spikes-in analysés dans le cadre de cette expérience sont des séquences d'ADN contenant un nombre défini d'adénosines et utilisé comme contrôle interne de taille de queue poly(A).

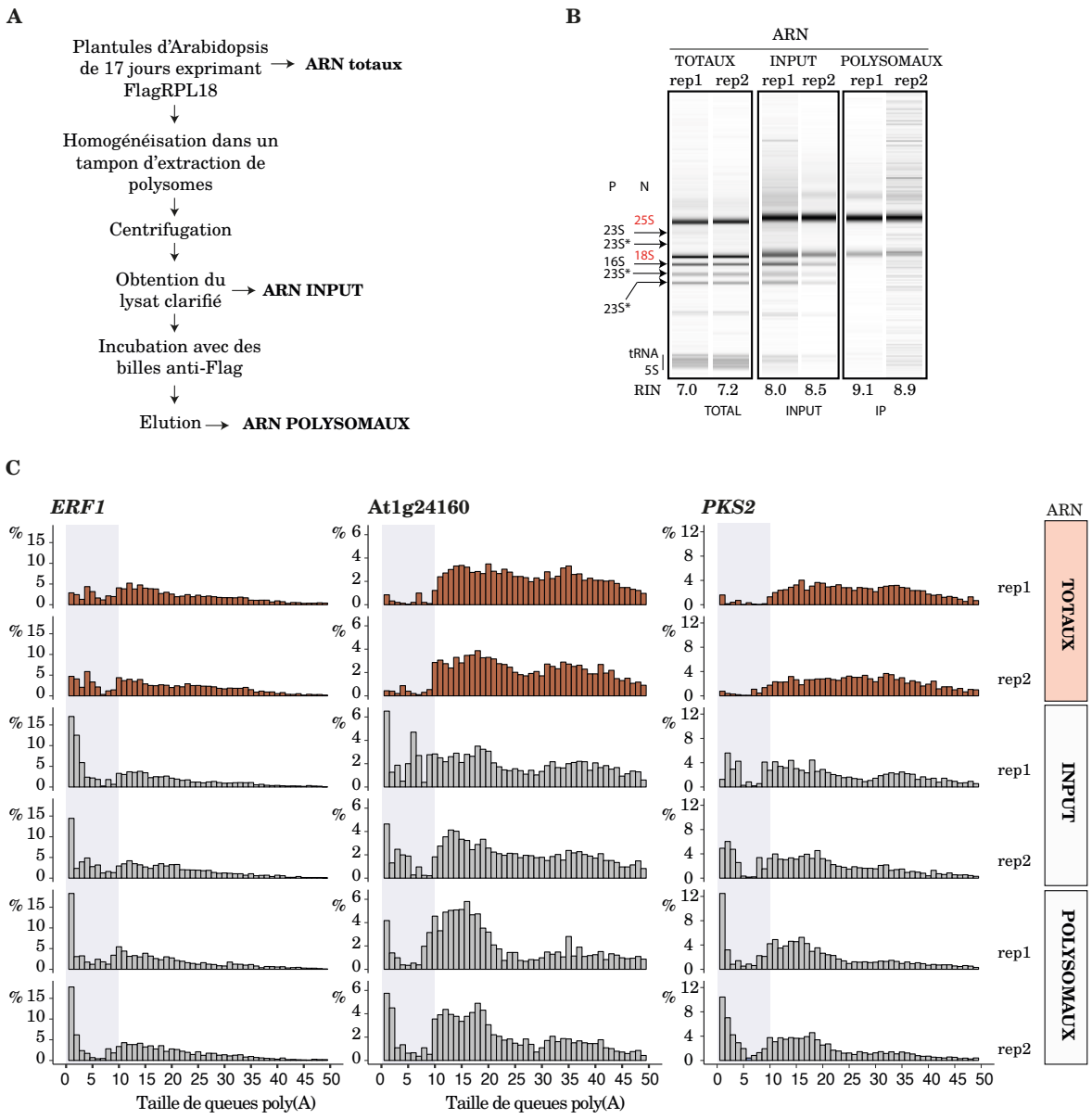


Figure 10 | Accumulation d'ARN fragmentés lors de la préparation des banques d'ARN polysomaux

A | Résumé des étapes principales du protocole de purification d'ARN polysomaux par TRAP (Translating Ribosomes Affiinity Purifiaation). Des plantules d'Arabidopsis de 17 jours sont homogénéisées dans un tampon adapté à l'extraction de polysomes. Le lysat obtenu est clarifié par centrifugation et incubé avec des billes anti-Flag. Les ARN polysomaux sont obtenus grâce à une élution par compétition avec des peptides Flag. La fraction des ARN totaux correspondent à des ARN directement extraits à partir des plantules exprimant FlagRPL18. Les ARN INPUT sont extraits à partir d'un aliquot de lysat clarifié et les ARN polysomaux correspondent aux ARN extraits après élution des billes anti-Flag. **B |** Profils bioanalyzer obtenus à partir des ARN totaux, INPUT et polysomaux pour deux réplicas biologiques indépendants. N, ARN ribosomiques nucléaires; P; ARN ribosomiques plastidiques (23S, 16S) et leur produits de fragmentation (23S*) **C |** Diagrammes en barres représentant la distribution de la taille des queues poly(A) déterminée par 3'RACE-seq et normalisée par rapport au nombre total de reads correspondants à des queues poly(A) comprises entre 0 et 50 A, pour les gènes *ERF1*, *AT1G24160* et *PKS2*. La distribution a été déterminée dans les fractions d'ARN totaux, INPUT et polysomaux.

amplicons de tailles inférieures se lient plus efficacement sur la cellule du séquenceur Illumina. Ces techniques restent néanmoins très utiles pour comparer des génotypes entre eux.

1.3.2 L'analyse par 3'RACE-seq révèle une fragmentation des ARN lors de la préparation d'ARN polysomaux

Les méthodes TAIL-seq et 3'RACE-seq sont habituellement utilisées pour analyser des ARN totaux extraits en conditions dénaturantes afin de préserver l'intégrité des ARNm et donc de leur queue poly(A). Cependant, diverses questions biologiques pourraient être abordées par l'analyse de sous-populations d'ARN obtenues suite à un fractionnement cellulaire ou après purification par affinité de protéines de liaison aux ARN. Ces étapes impliquent l'incubation des ARNm dans des lysats cellulaires complexes pendant la durée de l'expérience de fractionnement ou de purification. Cette incubation pourrait être préjudiciable à la qualité des ARNm et de leur queue poly(A). Pour tester ce type d'approche et dans l'optique d'analyser un lien éventuel entre uridylation des ARNm et leur traduction, nous avons voulu analyser par 3'RACE-seq le taux d'uridylation et la taille des queues poly(A) des ARNm engagés dans les polysomes. Pour ce faire nous avons purifié les ARN polysomaux en utilisant la méthode TRAP (**T**ranslating **R**ibosomes **A**ffinity **P**urification) (Zanetti *et al.*, 2005). Cette méthode est basée sur l'utilisation de lignées transgéniques qui expriment la protéine ribosomale RPL18 marquée avec un épitope Flag dans sa région N-terminale (FlagRPL18). Brièvement, un mélange de plantules FlagRPL18 est homogénéisé dans un tampon d'extraction de polysomes. Après homogénéisation le lysat est clarifié grâce à une étape de centrifugation et le lysat clarifié est mis à incuber en présence de billes couplées à des anticorps anti-Flag. Après incubation, et lavages des billes anti-Flag désormais interagissant avec FlagRPL18 les ARN polysomaux sont obtenus via une élution par compétition avec des peptides Flag. Nous avons réalisé des librairies de 3'RACE-seq à partir d'échantillons d'ARN totaux (c-à-d, directement extraits de plantules Flag-RPL18), d'ARN extraits à partir du lysat clarifié (INPUT) et après élution avec des billes anti-Flag (IP RPL18) (Fig. 10A). A noter que l'échantillon clarifié a été laissé sur glace pendant le temps de l'IP et traité en parallèle de l'échantillon d'élution IP RPL18.

Les profils obtenus au bioanalyzer après extraction des différents ARN ne révèlent aucun signe de dégradation dans les échantillons correspondant d'ARN INPUT et Polysomaux par rapport aux ARN totaux (Fig. 10B). Pour l'ensemble des profils d'ARN, la valeur de RIN (RNA integrity Number) est proche ou supérieure à celle recommandée pour des ARN issus de plantes d'Arabidopsis[2]. Cependant, l'analyse par 3'RACE-seq révèle que les ARNm analysés subissent une fragmentation variable au cours de l'expérience (Fig. 10C). En effet nous observons une forte accumulation de transcrits ayant des queues poly(A) courtes (<10 A) dans les fractions d'INPUT et d'ARN Polysomaux. Cette fragmentation est observée pour les 3 transcrits candidats analysés et résulte d'un artefact expérimental puisqu'elle n'est pas observée dans le cas d'ARN provenant de

[2] Agilent calcule un RIN de ~7,29 pour des ARN extraits de tissus d'Arabidopsis *https://www.agilent.com/cs/library/applications/5990-8850EN.pdf*

l'extraction directe à partir de plantules FlagRPL18 (ARN totaux). A noter que ce phénomène de fragmentation a également été observé dans le cas d'ARN polysomaux issus de purifications par gradient de saccharose (données non montrées).
L'analyse par 3'RACE-seq du statut d'uridylation et de la taille des queues poly(A) des ARN polysomaux serait extrêmement intéressante à mettre au point. Elle permettrait d'identifier la population de transcrits engagée dans les polysomes et d'observer d'éventuelles variations en cas de stress. Cette analyse requiert cependant une optimisation afin de réduire le phénomène de fragmentation des ARNm qui survient au cours du fractionnement cellulaire.

2. Utilisation du 3'RACE-seq pour l'analyse des extrémités 3' d'autres ARN

J'ai activement participé à la mise au point, au laboratoire, du protocole de 3'RACE-seq. Mes premiers essais sur les ARNm ont également ouvert la porte à l'analyse des extrémités 3' d'autres ARNm. Nous avons ainsi adapté ce protocole pour examiner les extrémités 3' des fragments 5' issus du clivage d'ARNm par le complexe RISC et des sous-produits de maturation des ARN ribosomiques.

2.1.1 Analyse par 3'RACE-seq de l'extrémité 3' des fragments 5' issus du clivage par le complexe RISC

Le 3'RACE-seq a été utilisé pour étudier la contribution respective des TUTases URT1 et HESO1 dans l'uridylation des fragments 5' issus du clivage par RISC des transcrits *MYB33* et *SPL13*. Nos résultats confirment que HESO1 est la TUTase principale de ces fragments comme précédemment montré (Ren *et al.*, 2014). De plus, nous démontrons que la seconde TUTase impliquée est URT1, et que l'absence d'uridylation du fragment 5' généré par le clivage de RISC provoque l'élimination par une activité inconnue de quelques nucléotides au site de clivage. Cet article fournit un protocole détaillé qui pourra être utilisé pour analyser de manière quantitative et qualitative les extrémités 3' de ces fragments. Cette technique est requise pour caractériser de manière fiable l'impact de différents facteurs impliqués dans la dégradation des fragments 5' produits par le clivage de RISC. Ces travaux font l'œuvre d'un article en révision d'une issue spéciale focalisée sur le métabolisme des ARN chez les plantes dans Frontiers in Plant Science, article que je signe en second auteur :

Respective contributions of URT1 and HESO1 to the uridylation of 5' fragments produced from RISC-cleaved mRNAs
Zuber H, Scheer H, Anne-Caroline Joly, Gagliardi D.
Manuscrit en révision à Front. Plant Sci.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 October 2018
doi: 10.3389/fpls.2018.01438

Respective Contributions of URT1 and HESO1 to the Uridylation of 5′ Fragments Produced From RISC-Cleaved mRNAs

Hélène Zuber, Hélène Scheer, Anne-Caroline Joly and Dominique Gagliardi**

Institut de Biologie Moléculaire des Plantes (IBMP), Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg, Strasbourg, France

In plants, post-transcriptional gene silencing (PTGS) represses gene expression by translation inhibition and cleavage of target mRNAs. The slicing activity is provided by argonaute 1 (AGO1), and the cleavage site is determined by sequence complementarity between the target mRNA and the microRNA (miRNA) or short interfering RNA (siRNA) loaded onto AGO1, to form the core of the RNA induced silencing complex (RISC). Following cleavage, the resulting 5′ fragment is modified at its 3′ end by the untemplated addition of uridines. Uridylation is proposed to facilitate RISC recycling and the degradation of the RISC 5′-cleavage fragment. Here, we detail a 3′ RACE-seq method to analyze the 3′ ends of 5′ fragments produced from RISC-cleaved transcripts. The protocol is based on the ligation of a primer at the 3′ end of RNA, followed by cDNA synthesis and the subsequent targeted amplification by PCR to generate amplicon libraries suitable for Illumina sequencing. A detailed data processing pipeline is provided to analyze nibbling and tailing at high resolution. Using this method, we compared the tailing and nibbling patterns of RISC-cleaved *MYB33* and *SPL13* transcripts between wild-type plants and mutant plants depleted for the terminal uridylyltransferases (TUTases) HESO1 and URT1. Our data reveal the respective contributions of HESO and URT1 in the uridylation of RISC-cleaved *MYB33* and SPL13 transcripts, with HESO1 being the major TUTase involved in uridylating these fragments. Because of its depth, the 3′ RACE-seq method shows at high resolution that these RISC-generated 5′ RNA fragments are nibbled by a few nucleotides close to the cleavage site in the absence of uridylation. 3′ RACE-seq is a suitable approach for a reliable comparison of uridylation and nibbling patterns between mutants, a prerequisite to the identification of all factors involved in the clearance of RISC-generated 5′ mRNA fragments.

Keywords: uridylation, TUTase, RISC, RNA silencing, Arabidopsis, RNA degradation, miRNA, Illumina

OPEN ACCESS

Edited by:
Ana Confraria, Instituto Gulbenkian de Ciência (IGC), Portugal

Reviewed by:
Tony Millar, Australian National University, Australia
Laura Arribas-Hernández, University of Copenhagen, Denmark

****Correspondence:***
Hélène Zuber
helene.zuber@ibmp-cnrs.unistra.fr
Dominique Gagliardi
dominique.gagliardi@ibmp-cnrs.unistra.fr

Specialty section:
This article was submitted to Plant Cell Biology, a section of the journal Frontiers in Plant Science

Received: *31 May 2018*
Accepted: *10 September 2018*
Published: *09 October 2018*

Citation:
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. doi: 10.3389/fpls.2018.01438

INTRODUCTION

Small RNAs are key regulators of gene expression (Borges and Martienssen, 2015; Bartel, 2018). They are classified as two main types, microRNAs (miRNAs) and short interfering RNAs (siRNAs), because of key distinctions in their respective mode of biogenesis (Martínez de Alba et al., 2013; Borges and Martienssen, 2015). miRNAs are processed from primary transcripts that fold as a

hairpin with an imperfectly paired stem. By contrast, siRNAs are generated from near-perfect double stranded RNAs (dsRNAs) or fully paired dsRNAs when the complementary strand is synthesized by a RNA-dependent RNA polymerase (RDR), which uses the sense strand as template. miRNAs and siRNAs are loaded onto members of the argonaute (AGO) protein family to form the core of RNA induced silencing complexes (RISCs) (Vaucheret, 2008; Zhang et al., 2015). RISCs are then guided to their targets by sequence complementarity with the loaded small RNA. In plants, the base pairing of miRNAs with their targets is rather extensive, and mRNAs regulated by RISCs are repressed by AGO1-mediated cleavage, but also by translation repression (Chen, 2004; Brodersen et al., 2008; Yang et al., 2012; Li et al., 2013; Iwakawa and Tomari, 2015; Reis et al., 2015; Arribas-Hernández et al., 2016). Cleavage of mRNAs by RISC produces a 5′ fragment and a 3′ fragment. As detailed below, both the 5′-3′ and 3′-5′ RNA degradation pathways contribute to the elimination of these fragments.

In Arabidopsis, the cytosolic 5′-3′ exoribonuclease XRN4 participates in the degradation of RISC 3′-cleavage fragments, as indicated by their accumulation in *xrn4* mutants (Souret et al., 2004). XRN4 was also proposed to be involved in the degradation of RISC 5′-cleavage fragments because the 5′ fragment resulting from the cleavage of *MYB domain protein 33* (*MYB33*) mRNA by miR159-loaded RISC accumulates in a *xrn4* mutant (Ren et al., 2014). RISC 5′-cleavage fragments are also definitely degraded by the 3′-5′ RNA degradation pathway because they accumulate in *ski2*, *ski3*, and *ski8* mutants (Branscheid et al., 2015). Together, SKI2-3-8 form the Ski complex, which is the major activator of the RNA exosome in the cytosol. Therefore, the involvement of the RNA exosome in the degradation of RISC 5′-cleavage fragments is likely in Arabidopsis, although it remains to be demonstrated using appropriate mutants. This implication of the RNA exosome would be consistent with previous findings in other organisms, such as *Drosophila melanogaster* (Orban and Izaurralde, 2005). In addition, two ribonucleases were recently described in Arabidopsis as taking part in the metabolism of RISC 5′-cleavage fragments: RISC-interacting clearing 3′-5′ exoribonucleases 1 and 2 (RICE-1 and -2) (Zhang et al., 2017). RICEs are homohexamers with a DnaQ-like exonuclease fold, and they interact with AGO1 and AGO10 (Zhang et al., 2017). RICEs are proposed to initiate the destabilization of RISC 5′-cleavage fragments thereby facilitating RISC dissociation. This would grant access of the 3′ extremity of RISC 5′-cleavage fragments to the RNA exosome and importantly, recycle RISC, which is essential to maintain functional RISC and miRNA abundance (Zhang et al., 2017). The access of the 3′ extremity of RISC 5′-cleavage fragments to the RNA exosome may also be facilitated by components of the non-stop decay (NSD) pathway when the RISC 5′-cleavage fragment is engaged in polysomes (Szádeczky-Kardoss et al., 2018). The prime function of NSD is to eliminate mRNAs lacking stop codons. Recently, orthologs of Pelota and Hbs1, two core components of NSD, were shown to participate in the elimination of RISC 5′-cleavage fragments in *Nicotiana benthamiana* and *A. thaliana*, provided that the cleavage site is within the coding region (Szádeczky-Kardoss et al., 2018). Likely, the NSD machinery promotes the dissociation of ribosomes stalled at the extremity of a RISC 5′-cleavage fragment to promote access to the RNA exosome (Szádeczky-Kardoss et al., 2018).

Besides exoribonucleases and RNA helicases, terminal uridylyltransferases (TUTases) constitute another type of enzymatic activities involved in the clearance of RISC 5′-cleavage fragments. Indeed a striking molecular event in this process is the untemplated addition of uridines at the 3′ extremity of RISC 5′-cleavage fragments (Shen and Goodman, 2004; Ren et al., 2014; Zhang et al., 2017). The uridylation of several of such fragments was originally reported in both Arabidopsis and mice (Shen and Goodman, 2004). Since then, uridylation has emerged as a conserved post-transcriptional process that shapes the coding and non-coding transcriptomes in eukaryotes (Munoz-Tello et al., 2015; Scheer et al., 2016; De Almeida et al., 2018). In Arabidopsis, two TUTases have been characterized: HEN1 SUPPRESSOR 1 (HESO1) and URIDYLYLTRANSFERASE 1 (URT1) (Kwak and Wickens, 2007; Ren et al., 2012; Zhao et al., 2012; Sement et al., 2013). Both HESO1 and URT1 contain the core catalytic domain (CCD) that defines proteins belonging to the terminal nucleotidyltransferase family. In addition, URT1 contains a large intrinsically disordered region (IDR) in its N-terminal region, while a shorter IDR is present in the C-terminal region of HESO1 (De Almeida et al., 2018). Those IDRs may mediate the recognition of protein partners by URT1 and HESO1, or be a key to their localization in P-bodies and/or stress granules (Sement et al., 2013; Ren et al., 2014; Wang et al., 2015). HESO1 was identified as the main TUTase uridylating miRNAs and siRNAs to trigger their degradation (Ren et al., 2012; Zhao et al., 2012). In addition, HESO1 was shown to uridylate three RISC 5′-cleavage fragments (Ren et al., 2014). Those fragments are generated from *MYB33*, *Auxin Response Factor 10* (*ARF10*), and *Lost Meristems 1* (*LOM1*) mRNAs, which are targets of miR159, miR160, and miR171, respectively. A residual uridylation of these RISC 5′-cleavage fragments is observed in *heso1* mutants (Ren et al., 2014) and, this secondary activity may be due to URT1, although experimental evidence supporting this hypothesis is lacking to date. URT1 is the main TUTase uridylating mRNAs in Arabidopsis, because mRNA uridylation is decreased by 70–80% in *urt1-1* mutants (Sement et al., 2013; Zuber et al., 2016). URT1 can also uridylate miRNAs, mostly when HESO1, the primary TUTase involved in small RNA uridylation, and HUA ENHANCER 1 (HEN1), a methyltransferase that methylate small RNA duplexes, are absent (Yu et al., 2005; Yang et al., 2006; Ren et al., 2012; Zhao et al., 2012; Tu et al., 2015; Wang et al., 2015). miRNAs are therefore the first documented example of shared RNA substrates between HESO1 and URT1 (Tu et al., 2015; Wang et al., 2015). Yet, both overlapping and distinctive roles in miRNA uridylation were attributed to each TUTase (Tu et al., 2015; Wang et al., 2015). mRNAs and RISC-cleaved transcripts could constitute other cases of shared RNA substrates between HESO1 and URT1. These possibilities remain to be experimentally addressed.

To date, the characterization of uridylated RISC 5′-cleavage fragments has relied on the use of 3′ RACE PCR followed by cloning and subsequent analysis based on Sanger sequencing. Although this experimental strategy has proven useful, it has

some inherent limitations. The first one is that this method is low-throughput. It is time-consuming and the depth is usually quite limited, with often less than 20–30 clones analyzed per genotype. The second major drawback is the lack of discrimination between amplicons and independent molecules. This turns out to be a real issue when analyzing low complexity samples by PCR amplification, with the majority (up to 90% as determined here during the analysis of RISC 5′-cleavage fragments) of the final PCR products that correspond to very few independent templates. Taken together, these limitations hinder the qualitative and quantitative analysis of the uridylation of RISC-cleaved transcripts. Such an analysis is crucial to reliably compare uridylation between wild-type (WT) and mutant genetic backgrounds, and this comparison is required to identify all factors involved in the metabolism of 5′ RNA fragments produced by RISC cleavage.

Here we detail a 3′ RACE-seq method that has been optimized for analyzing the uridylation of 5′ fragments from RISC-cleaved transcripts. Those molecules are usually low abundant within the complex mixture of all cellular RNAs, and they exhibit a rather poor diversity, with a few untemplated nucleotides usually added at a precise RISC-mediated cleavage site. We illustrate the use of 3′ RACE-seq to analyze the tailing and trimming patterns of *MYB33* and *SPL13* RISC 5′-cleavage fragments by comparing WT plants and mutants lacking HESO1 and URT1. This analysis revealed the respective contributions of both TUTases, and that the absence of uridylation results in the accumulation of 5′-cleavage fragments nibbled by a few nucleotides close to the site cleaved by RISC.

MATERIALS AND METHODS

Gene IDs and Primers

The Arabidopsis Genome Initiative (AGI) locus identifiers for the genes studied in this study are: AT2G39740 (*HESO1*), AT2G45620 (*URT1*), AT5G06100 (*MYB33*), and AT5G50570 (*SPL13A*). Please note that AT5G50570 (*SPL13A*) and AT5G50670 (*SPL13B*) have identical coding sequences and therefore cannot be discriminated in this study. For simplicity, the name *SPL13* is used thereafter. The sequence of all primers used in this study is shown in **Supplementary Table S1**.

Plant Material

The plant material used for analyzing RISC 5′-cleavage fragments by 3′ RACE-seq corresponds to Arabidopsis plantlets of Col-0 accession grown for 24 days *in vitro* on Murashige and Skoog media with 0.8% agar and 12 h light (22°C)/12 h darkness cycles (18°C). For other analyses, flowers were harvested from Arabidopsis of Col-0 accession and grown on soil with 16 h light/8 h darkness cycles. *urt1-1* (Salk_087647C) and *heso1-1* (GK-367H02-017041) T-DNA mutant lines have been previously described (Zhao et al., 2012; Sement et al., 2013). Double mutants were obtained by down regulating *URT1* by co-suppression in *heso1-1*. For this purpose, *heso1-1* plants were transformed with a construct expressing an inactive version of URT1 fused to YFP, which was fortuitously found to efficiently trigger co-suppression of the endogenous *URT1* gene. The sequence encoding the inactive version of URT1 ($URT1^{D491/3A}$) (Sement et al., 2013) was cloned in the pEarleyGate 104 Gateway plasmid under the control of the cauliflower mosaic virus (CaMV) 35S promoter. Analyses were performed on two biological replicates of three independent *heso1-1* $urt1^{SIL}$ lines. As a control, we also analyzed two biological replicates of a $urt1^{SIL}$ line obtained by co-suppressing *URT1* with a *YFP-URT1* sequence cloned in the pEarleyGate 104 Gateway plasmid.

Protein Extraction and Western Blot Analysis

Proteins were extracted from flowers of WT, *urt1-1*, $urt1^{SIL}$ and three independent *heso1-1* $urt1^{SIL}$ lines under denaturing conditions. Proteins were resolved on a 8% SDS-PAGE gel and transferred to an Immobilon-P membrane. Immunoblots were incubated with anti-URT1 antibodies raised in rabbits against the full-length recombinant URT1. Following incubation with horseradish peroxidase-coupled secondary antibodies and Lumi-Light Western Blotting Substrate (Roche), signals were recorded using the Fusion-FX system (Vilber Lourmat).

RNA Extraction and Northern Blot Analysis

Total RNA was extracted from 24-day-old plantlets and flowers for 3′ RACE-seq and northern blot analyses, respectively, with TRI Reagent® (Molecular Research Center) according to manufacturer's instructions. RNA was further purified by acid phenol: chloroform: isoamyl alcohol extraction followed by ethanol precipitation. For northern blot analysis of *MYB33* RISC 5′-cleavage fragments, 30 μg total RNAs from WT, *urt1-1* and *heso1-1* mutants were separated on a denaturing formaldehyde 1.5% (w/v) agarose gel and transferred onto a nylon membrane (Hybond^{TM}-N+, GE Healthcare Life Sciences^{TM}). Following UV-cross-link at 120 mJ/cm^2 for two times 30 s and incubation for 30 min in PerfectHyb Plus Hybridization buffer (Sigma), the membrane was hybridized overnight at 65°C with a probe that detects the 5′ fragment of *MYB33* RISC-cleaved transcripts. The probe was prepared by PCR amplifying a 219 bp sequence upstream of the RISC cleavage site (**Supplementary Table S1**) and by random primed labeling the PCR product using [α-^{32}P]-dCTP and DecaLabel DNA labeling kit (Thermo Scientific). For northern blot analysis of miR159, 10 μg total RNA from WT, *urt1-1*, and *heso1-1* mutants were separated on 17.5% polyacrylamide/7 M urea gels and transferred onto nylon membranes (Hybond^{TM}-NX, GE Healthcare Life Sciences^{TM}). Following UV-cross-link at 120 mJ/cm^2 for two times 30 s and incubation for 30 min in PerfectHyb Plus Hybridization buffer (Sigma), membranes were hybridized overnight at 50°C with a 5′ [^{32}P]-labeled oligonucleotide to detect miR159 (**Supplementary Table S1**). The probe was labeled using [γ-^{32}P] ATP and T4 PNK (NEB) according to manufacturer's instruction. Radioactive signals were detected by autoradiography and quantified using a Typhoon scanner (GE Healthcare Life Sciences) and Image Gauge software. Plant material used for

biological replicates 1 and 2 were common for both northern analyses.

3′ RACE-Seq Protocol

A 3′ RACE-seq protocol was adapted for analyzing RISC 5′-cleavage fragments. Total RNA was extracted from 24-day-old seedlings using TRI Reagent® as described above. Twenty pmoles of a 5′-riboadenylated DNA oligonucleotide (3′-Adap, **Figure 1** and **Supplementary Table S1**) were ligated to 10 μg of total RNA using 20 U of T4 ssRNA Ligase 1 (NEB) in a final volume of 100 μl for 1 h at 37°C and 1X T4 of RNA Ligase Reaction Buffer (NEB, 50 mM Tris–HCl, 10 mM MgCl2, 1 mM DTT, pH 7.5). The ligation products were purified from reagents and non-ligated adapter molecules with Nucleospin® RNA Clean-up columns (Macherey Nagel). RNA was then precipitated with ethanol, solubilized in water and quantified. cDNA synthesis was performed in two 20 μl-reactions for each sample. Each 20 μl-reaction contained 2 μg of purified ligated RNA, 50 pmol of the 3′-RT oligonucleotide (**Supplementary Table S1**), 10 nmol of dNTP, 0.1 μmol of DTT, 40 U of RNaseOUT (Invitrogen^{TM}), 200 U of SuperScript IV reverse transcriptase (Invitrogen^{TM}) and 1X of SuperScript IV RT buffer (Invitrogen^{TM}). Reactions were incubated at 50°C for 10 min, and then at 80°C for 10 min to inactivate the reverse transcriptase. The two 20 μl-reactions for each sample were pooled, the cDNAs were extracted with phenol–chloroform, precipitated with ethanol and dissolved in 8 μl Milli-Q water. Two nested PCR amplification rounds of 20 and 25 cycles, respectively, were then performed. PCR1 was run using cDNA synthesized from 1 μg of total RNA, i.e., 2 μl of concentrated cDNA, 10 pmol of *MYB33* or *SPL13* gene-specific sense primer 1, 10 pmol of RACEseq_rev1 primer (**Supplementary Table S1**), 10 nmol of dNTP, 1 U of GoTaq® DNA Polymerase (Promega) and 1X of Green GoTaq® Reaction Buffer (Promega) in a 20 μl final volume. The conditions for PCR1 were as follows: a step at 94°C for 30 s; 20 cycles at 94°C for 20 s, 50°C for 20 s and 72°C for 30 s; a final step at 72°C for 30 s. PCR2 was performed using 1 μl of PCR1 product, 10 pmol of *MYB33* or *SPL13* gene-specific sense primer 2, 10 pmol of a TruSeq RNA PCR index (RPI, **Supplementary Table S1**) 10 nmol of dNTP, 1 U of GoTaq® DNA Polymerase (Promega) and 1X of Green GoTaq® Reaction Buffer (Promega) in a 20 μl final volume. The conditions for PCR2 were as follows: a step at 94°C for 1 min; 25 cycles at 94°C for 30 s, 56°C for 20 s and 72°C for 30 s; a final step at 72°C for 30 s. For each sample, three to four 20 μl-reactions were run and pooled. All PCR2 products were purified using 1 volume of AMPure XP beads (Agencourt). Library concentrations were determined using a Qubit fluorometer (Invitrogen^{TM}). Libraries were analyzed on a 2100 Bioanalyzer system (Agilent) to assess quality and estimate size distribution. Library were paired-end sequenced with MiSeq (v3 chemistry) with 41 × 111 bp cycle settings. The respective numbers of sequencing cycles for read 1 and read 2 can be adjusted according to other samples that are co-analyzed. Read 1 is used to identify target transcript whereas read 2 is used to map 3′ extremities and analyze 3′ potential untemplated nucleotides. To compensate for the poor diversity of the amplicon libraries, 25–33% of phiX control library (Illumina) were included. Two rounds of RACE-seq experiments were performed. For the first round, four independent biological replicates were analyzed for WT and *heso1-1* genotypes. For the second round, two independent biological replicates were analyzed for WT, *urt1-1*, *heso1-1*, *urt1*^{SIL} line, and each of the three *heso1-1 urt1*^{SIL} lines (i.e., six *heso1-1 urt1*^{SIL} samples). Plant material used for biological replicates 1 and 2 was common for both rounds.

Bioinformatic Analysis of 3′ RACE-Seq Data

Sequencing run quality fit Illumina specification with more than 90% bases higher than Q30. After initial data processing by the MiSeq Control Software v 2.5 (Illumina), base calls were retrieved and further analyzed by a suite of homemade scripts (**Supplementary Data Sheets S1**, **S2**) using python (v2.7), biopython (v1.63), and regex (v2.4) libraries. Data processing pipeline was adapted from Sikorska et al. (2017). Reads with low quality bases (= < Q10) within the 15-base random sequence of the read 2 or within the 20 bases downstream the delimiter sequence, were filtered out. Sequences with identical nucleotides in 15-base random sequence were deduplicated. Next, the sequences AAGAATTCTCGTCGCCTGAA and GCCAGAGCTATGTTGTTGGT were searched into reads 1 to identify *MYB33* and *SPL13* corresponding reads, respectively. 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. In order to map 3′ extremities of target 5′ RISC generated fragments, the 20 nucleotide sequences downstream 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 PCR primer MYB33_RISC_fw2 or SPL13_RISC_fw2 to the last nucleotide of the miRNA binding site. 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 nt trimming step, until successfully mapped to the reference sequence or until a maximum of 30 nt has been removed. For each successfully mapped read 2, untemplated nucleotides at the 3′ end were extracted and analyzed for their size and composition. 3′ modifications longer than 1 were considered only if composed of at least 50 % of the same base (i.e., 50% U, 50% A, 50% C, or 50% G). As explained in the Results section and as illustrated in **Figures 2**, **7** for *MYB33* and *SPL13*, respectively, the sites cleaved by RISC were defined by using PARE-seq datasets to map the 5′ most nucleotide of the 3′ fragment. A single cleavage site was determined for *MYB33*, i.e., between nucleotides at position 0 and position +1, in contrast to two cleavage sites for *SPL13*, i.e., a major site between 0 and +1 and a minor site between +1 and +2. Because position +1 in *SPL13* is a uridine, we could not determine whether this U is encoded or added post-transcriptionally and SPL13 graphs were generated by considering only U-tails > 2. Finally, a supplemental deduplication was performed to increase stringency: sequences with 13 or more identical nucleotides in

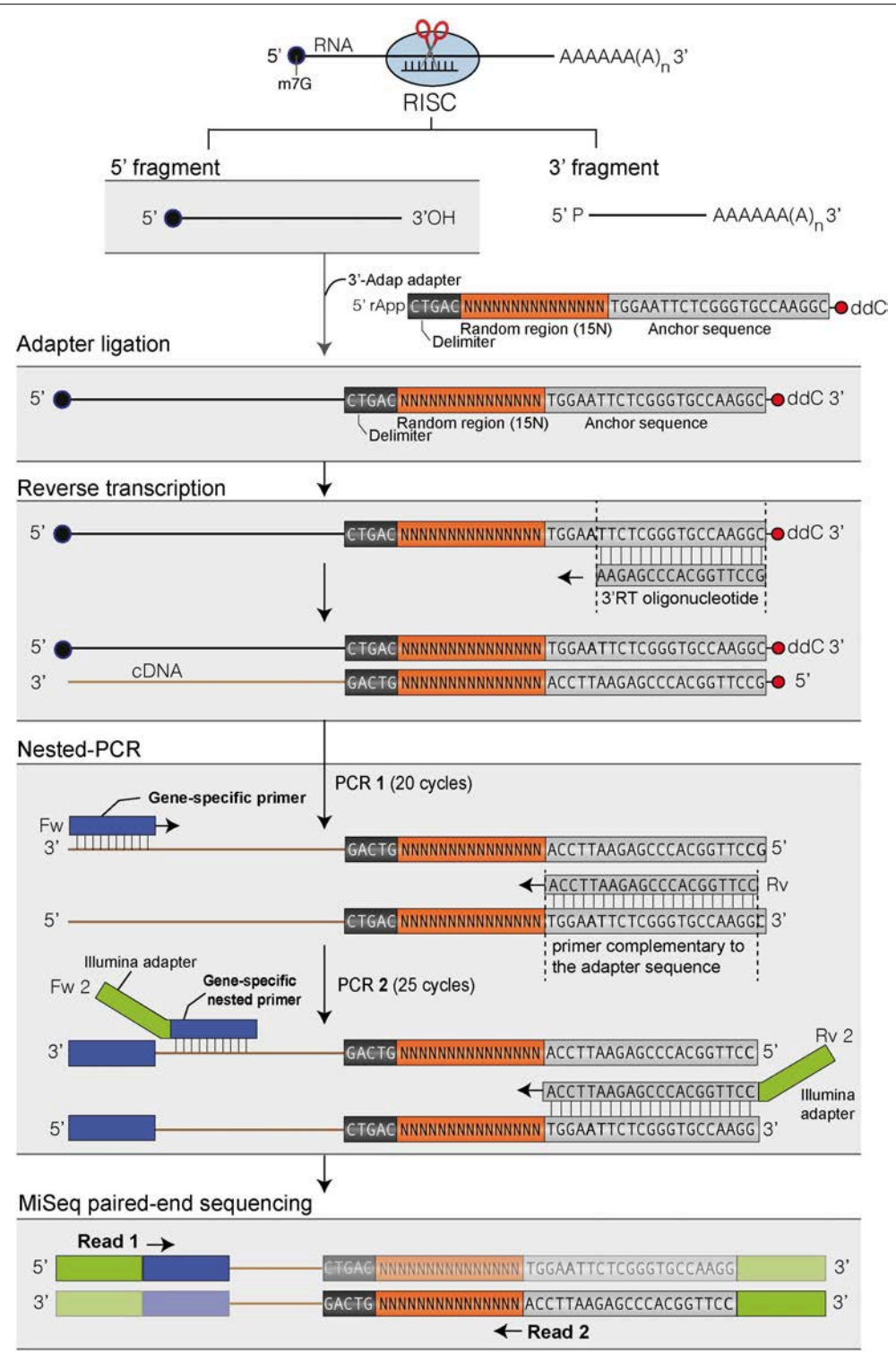


FIGURE 1 | Flowchart of the main steps to map the 3′ ends of RISC 5′-cleavage fragments by 3′ RACE-seq. Features of the 3′ adapter and the principle of the main steps are indicated. The experimental workflow begins by ligating the 3′ adapter to the RISC 5′-cleavage fragment. Please note that any RNA molecule with a 3′ hydroxyl end in the total RNA sample is ligated to the 3′ adapter. The target of interest is specifically amplified during PCR-1 and -2 due to the gene-specific sequences of the forward primers. The protocol is detailed in Methods and the scripts used to analyze data are given in **Supplementary Data Sheets S1**, **S2**.

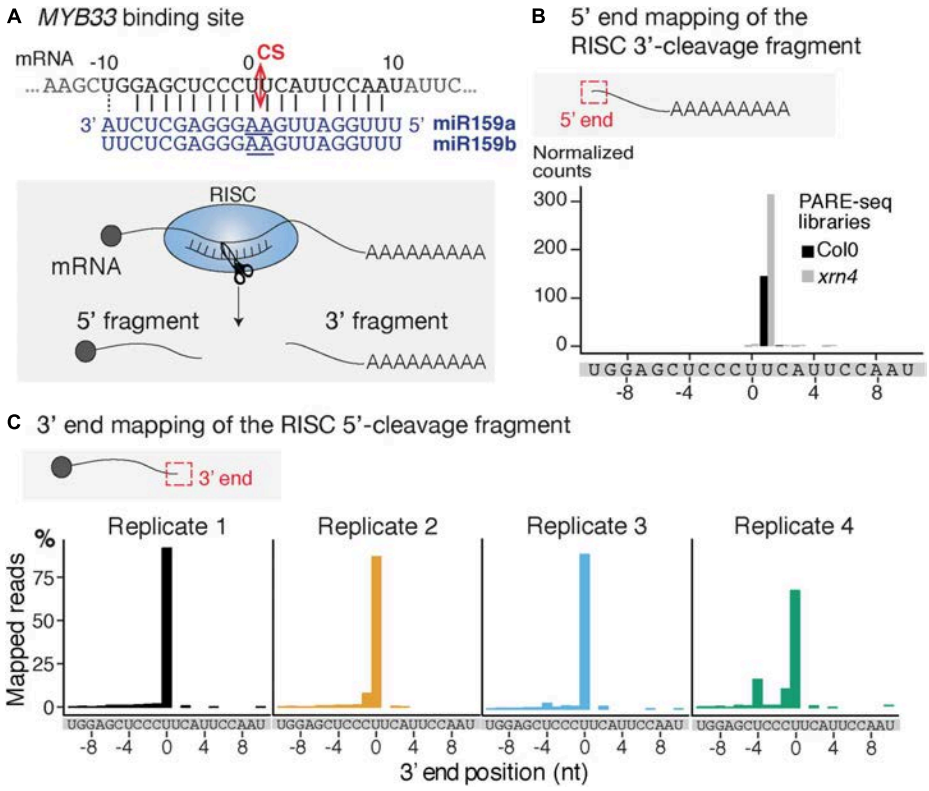


FIGURE 2 | Mapping of the 3′ extremity of *MYB33* RISC 5′-cleavage fragments. **(A)** Schematic representation of the cleavage of *MYB33* mRNA by AGO1 loaded with miR159. The cleavage site (CS) is predicted between the positions 0 and +1, and is defined by the 10th and 11th nucleotides (underlined) of miR159. **(B)** 5′ end mapping using PARE-seq data (German et al., 2008) of RISC 3′-cleavage fragments in WT and *xrn4*. **(C)** 3′ end mapping by 3′ RACE-seq of RISC 5′-cleavage fragments in WT within the –10/ + 10 window corresponding to the *MYB33* sequence recognized by miR159. For 3′ RACE-seq, most reads map to a single position in each of the four WT biological replicates.

the 15-base random sequence were deduplicated. Plotting and quantitative data analysis was performed with R software (v3.3.1) and ggplot2 R (v2.2.1). Percentages of uridylated fragments were plotted for reads with 3′ extremities that map at the cleavage site, with U-tails being defined as tail composed of more than 50% U. Data obtained from the two rounds of RACE-seq experiments, referred to as dataset #1 and dataset #2, have been deposited to the NCBI Gene Expression Omnibus (GEO) database with the accession code GSE115470.

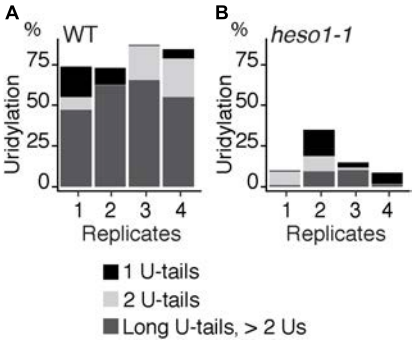


FIGURE 3 | HESO1 is the main TUTase uridylating *MYB33* RISC 5′-cleavage fragments. Percentages of uridylated *MYB33* RISC 5′-cleavage fragments in four biological replicates for **(A)** WT and **(B)** *heso1-1*. Percentages of long (>2 Us), 2 U- and 1 U-tails are indicated by dark gray, light gray, and black, respectively.

RESULTS AND DISCUSSION

Workflow for Mapping 3′ Ends of RISC 5′-Cleavage Fragments by 3′ RACE-Seq

The principle of 3′ RACE-seq to analyze the 3′ ends of RISC 5′-cleavage fragments is shown in **Figure 1**. Briefly, a 5′ pre-riboadenylated oligodeoxynucleotide adapter is ligated to the 3′ hydroxyl end of RNA molecules using T4 RNA ligase 1 and total RNA. The sequence of the 3′ adapter is identical to the one previously described for the TAIL-seq procedure (Chang et al., 2014). However, unlike for TAIL-seq, the 3′ adapter does not need to be biotinylated. The sequence features of the adapter are detailed in **Figure 1**. Five nucleotides at the 5′ end of the adapter form a delimiter sequence. All reads that do not contain

this delimiter sequence are discarded during the analysis. This ensures that we accurately map the 3′ extremity of a transcript that has been ligated to the 3′ adapter. Untemplated nucleotides are defined as any nucleotides present between the genome-encoded sequence and the delimiter. The delimiter is followed by a random sequence of 15 bases. This random sequence is essential to remove PCR duplicates during the bioinformatic analysis. This deduplication step is crucial when using 3′ RACE-seq to analyze low abundant RNA species with limited complexity, which is typically the case for RISC 5′-cleavage fragments. To further prevent amplicon biases due to the misincorporation of nucleotides in the random sequence during PCR amplification or due to sequencing errors of the random sequence, we enhance the stringency of the deduplication step by not tolerating up to two mismatches within the 15-base random sequence of deduplicated sequences.

The 3′ adapter then contains 22 additional bases, which provide an anchor sequence for cDNA synthesis and subsequent PCR amplification. Importantly, the primer used for cDNA synthesis is complementary to the sequence of the 3′ adapter but stops five bases downstream of the random sequence (**Figure 1**). By using a reverse primer for PCR amplification that extends up to the random sequence, we eliminate the vast majority of cDNAs that are due to priming artifacts and specifically analyze transcripts that have the 3′ adapter ligated at their 3′ ends. This trick greatly enhanced the quality and depth of our libraries. Finally, the 3′ adapter is terminated by a dideoxy-C to prevent self-ligation (**Figure 1**).

cDNAs are then subjected to two successive rounds of PCR amplification. For the first round, the forward primer is a gene-specific primer matching the sequence of a selected mRNA and located ideally about 200–400 nucleotides upstream of the predicted RISC-mediated cleavage site. The reverse primer matches the sequence of the 3′ adapter up to the random sequence (**Figure 1**). As mentioned above, this prevents the amplification of most cDNA priming artifacts. The second round of PCR is performed with a nested forward primer and a barcoded reverse primer complementary to the anchor sequence (**Figure 1**). Typically, thirty different barcodes can be used to simultaneously analyze different genotypes or replicates. Both forward and reverse primers contain 5′ extensions corresponding to the Illumina sequences that are used for flow cell hybridization and sequencing. The number of PCR cycles must be kept as low as possible for both PCRs and ideally should not exceed 20–25 per PCR. Amplicon libraries are purified using AMPure XP beads, quantified with an Invitrogen Qubit fluorometer and their size distribution is determined with a 2100 Bioanalyzer system (Agilent). Amplicon libraries are then sequenced using MiSeq paired-end sequencing for an average yield per run of 38 million of reads: 19 millions of read 1 and 19 millions of read 2.

Mapping of the RISC-Cleavage Site in *MYB33* mRNAs by 3′ RACE-Seq

We selected the *MYB33* mRNAs targeted by miR159 as a model substrate to set up the mapping of the 3′ ends of RISC 5′-cleavage fragments by 3′ RACE-seq. *MYB33* has been chosen in several studies to investigate uridylation of RISC 5′-cleavage fragments for two main reasons (Shen and Goodman, 2004; Ren et al., 2014; Zhang et al., 2017). First, *MYB33* RISC 5′-cleavage fragments are detectable by northern blot analysis, and therefore their accumulation can be compared between WT plants and

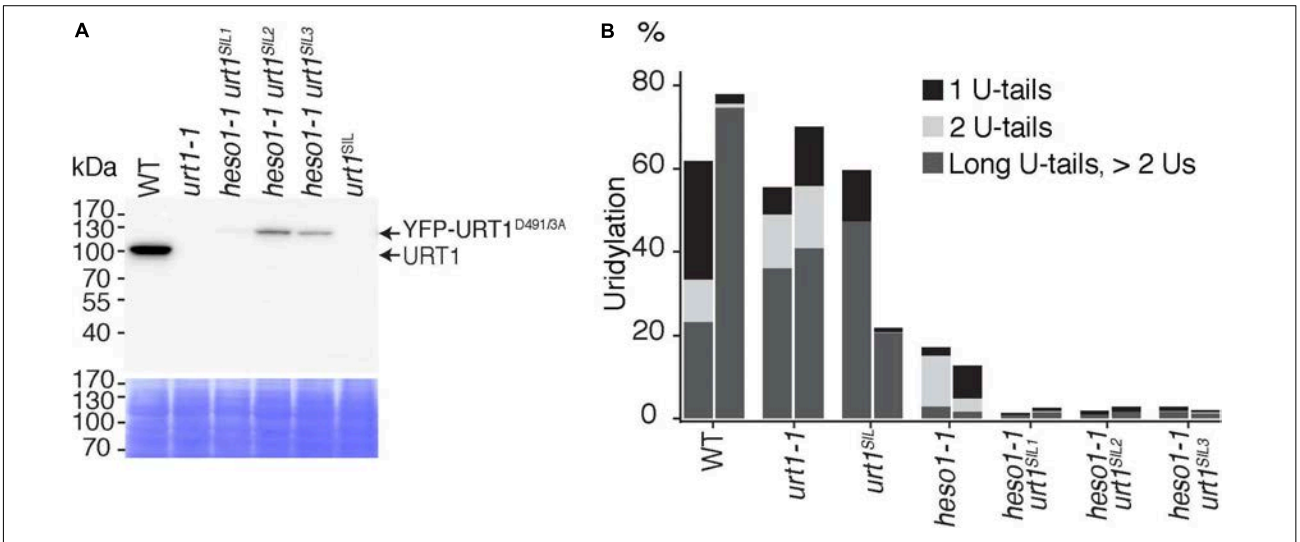


FIGURE 4 | Respective contributions of HESO1 and URT1 in the uridylation of *MYB33* RISC 5′-cleavage fragments. **(A)** Western blot analysis of WT, *urt1-1* T-DNA mutant, co-suppressed *urt1* mutant (*urt1*^{SIL}), and three double mutant lines *heso1-1 urt1*^{SIL}. Molecular weight markers are indicated in kDa. A portion of the membrane stained with Coomassie blue is shown as loading control. Please note that the construct used to co-suppress *URT1* in the *heso1-1* background expresses an inactive version of URT1 fused to YFP. Uncropped images are shown in **Supplementary Figure S1**. **(B)** Percentages of uridylated *MYB33* RISC 5′-cleavage fragments in two biological replicates for WT, *urt1-1*, the *urt1*^{SIL} line, *heso1-1*, and the three *heso1-1 urt1*^{SIL} lines (i.e., six *heso1-1 urt1*^{SIL} samples). Percentages of long (>2 Us), 2 U- and 1 U-tails are indicated by dark gray, light gray, and black, respectively.

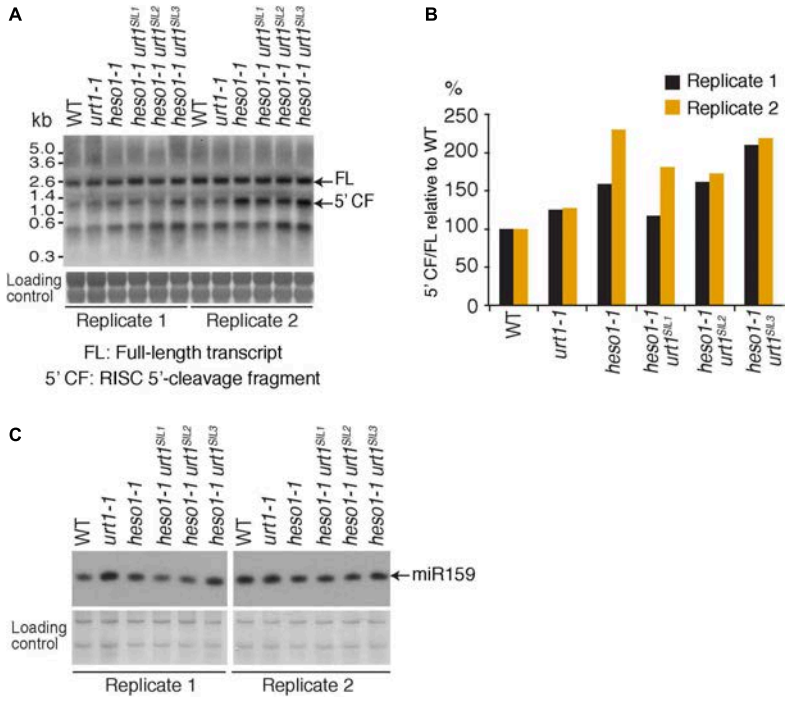


FIGURE 5 | *MYB33* 5′-cleavage fragments accumulate in *heso1-1* **(A)** Northern blot analysis of WT, *urt1-1*, and *heso1-1* mutant lines using a probe 5′ of the cleavage site to detect *MYB33* full-length mRNA and the RISC 5′-cleavage fragment. Molecular weight markers are indicated in kb. A portion of the membrane stained with methylene blue is shown as loading control. Uncropped images are shown in **Supplementary Figure S2**. **(B)** *MYB33* RISC 5′-cleavage fragments accumulate mostly in the absence of HESO1. The accumulation of *MYB33* RISC 5′-cleavage fragments (5′ CF) relative to its full-length (FL) mRNA was determined by integrating the signals in **B** with a Phosphorimager. 5′ CF/FL values for each lane were normalized to the 5′ CF/FL ratio obtained for the WT control of each replicate. Two replicates are shown in black and orange, respectively, for WT, *urt1-1*, the *urt1*^{SIL} line, *heso1-1*, and the three *heso1-1 urt1*^{SIL} lines (i.e., six *heso1-1 urt1*^{SIL} samples). **(C)** Northern blot analysis of WT, *urt1-1*, and *heso1-1* mutant lines using a probe to detect *miR159*. A portion of the membrane stained with methylene blue is shown as loading control. Uncropped images are shown in **Supplementary Figure S2**.

relevant mutants. Second, a high proportion of *MYB33* RISC 5′-cleavage fragments is uridylated in WT plants. This proportion was in fact high enough to allow detection of uridylated *MYB33* RISC 5′-cleavage fragments by sequencing of a limited number of clones (Shen and Goodman, 2004; Ren et al., 2014; Zhang et al., 2017). The high level of uridylation in WT plants is useful to monitor decrease of uridylation in mutants to identify factors that are involved in the metabolism of this RISC 5′-cleavage fragment. However, there is one drawback in choosing *MYB33* to study uridylation: the predicted cleavage site, which is specified by the tenth and eleventh nucleotides of miR159, is situated between two uridines (**Figure 2A**). This can lead to uncertainties as to whether some 3′ terminal uridines are genome-encoded or added post-transcriptionally by TUTases. To solve this issue, we took advantage of previous data generated using parallel analysis of RNA ends (PARE)-seq. PARE-seq is one of the sequencing methods designed to map 5′ hydroxylated end of RNAs and used to map small RNA cleavage sites (German et al., 2008). PARE-seq unambiguously identifies the position defined here as +1 as the 5′ nucleotide of the RISC-generated 3′ fragment of *MYB33* (**Figure 2B**). Therefore, cleavage of *MYB33* by miR159-loaded AGO1 does occur at the canonical site, which we defined here between positions 0 and +1 (**Figure 2A**). This was further experimentally validated in the present study because *MYB33* RISC 5′-cleavage fragments ending at position 0 accumulate in a genetic background abolishing uridylation (detailed later in **Figure 6B**).

To study *MYB33* RISC 5′-cleavage fragments by 3′ RACE-seq, we first analyzed the aerial part of 24-day-old plants grown *in vitro* corresponding to four biological replicates for WT and four biological replicates for the *heso1-1* mutant. We obtained a total of 29,689 reads for WT and 34,096 reads for *heso1-1* (**Supplementary Table S2** and **Supplementary Data Sheet S1**). The WT data were first used to monitor the distribution of 3′ extremities mapped in the sequence to which miR159 binds. The majority of reads (up to 85%) mapped at position 0 (**Figure 2C**). Therefore, we conclude that the 3′ extremities of RISC-cleaved *MYB33* are accurately mapped by 3′ RACE-seq.

Respective Contributions of HESO1 and URT1 in the Uridylation of *MYB33* 5′-Cleavage Fragments

To analyze untemplated nucleotides added after RISC-mediated cleavage of *MYB33* mRNAs, the nucleotide extensions for reads that map to position 0 were analyzed first for the WT samples. Up

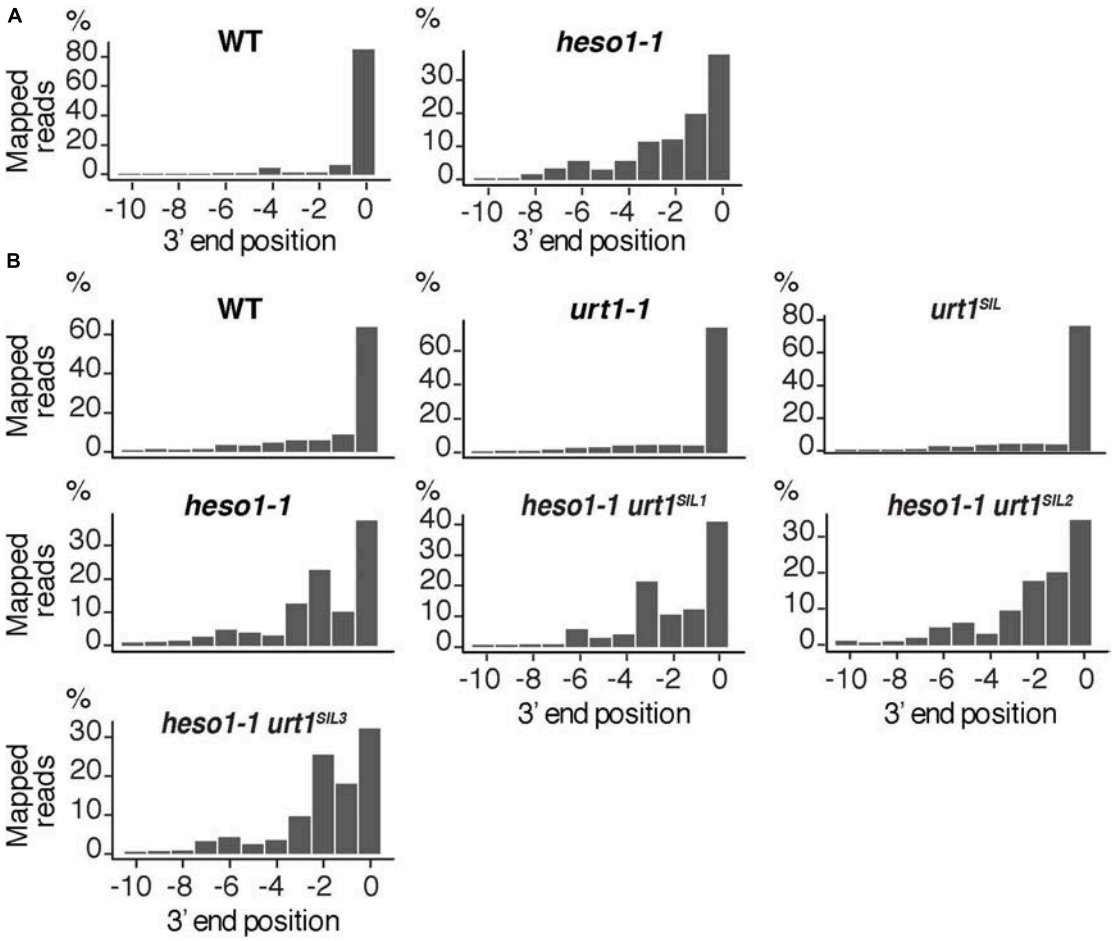


FIGURE 6 | Nibbled *MYB33* RISC 5′-cleavage fragments accumulate in the absence of HESO1. Positions of 3′ extremities of *MYB33* RISC 5′-cleavage fragments mapped in a –10/0 window for **(A)** four biological replicates, WT and *heso1-1*, or **(B)** two biological replicates for WT, *urt1-1*, the *urt1*^{SIL} line, *heso1-1*, and the three *heso1-1 urt1*^{SIL} lines. Graphs are shown separately for each of the replicates in **Supplementary Figures S3**, **S4**.

to 98 % of *MYB33* RISC 5′-cleavage fragments in WT are tailed by nucleotide extensions, which are predominantly composed of uridines (**Supplementary Table S3**). This result is in agreement with previous observations (Shen and Goodman, 2004; Ren et al., 2014; Zhang et al., 2017). Most U-rich tails were longer than 2 Us in the four WT biological replicates (**Figure 3A** and **Supplementary Table S3**). We then compared the impact of HESO1 on the uridylation *MYB33* RISC 5′-cleavage fragments. A major decrease in uridylation was observed in *heso1-1* as compared with WT samples (**Figure 3B**). This observation confirmed that HESO1 is the main TUTase uridylating *MYB33* 5′-cleavage fragments, as shown here using four independent biological replicates in the Col-0 genetic background. In addition, and as previously observed (Ren et al., 2014), the size of U-tails detected in *heso1-1* was reduced as compared to WT, with mainly short U-tails (<2 Us) detected in *heso1-1* (**Figure 3**).

The residual uridylation in *heso1-1* indicates the involvement of an alternative TUTase. A good candidate for this activity is URT1, the second TUTase that has been identified in Arabidopsis (Sement et al., 2013). To date, the possible involvement of URT1 in the uridylation of 5′ RISC-cleaved mRNAs, including *MYB33*, has been proposed but not tested experimentally. Testing this hypothesis requires the production of a *heso1 urt1* double mutant. To this end, we crossed the *heso1-1* and *urt1-1* single mutants. However, we failed to recover the expected double mutant in the F2 progeny. This failure is yet unexplained but we could obtain lines that were originally designed to overexpress an inactive version of URT1, but that in fact co-suppress the endogenous *URT1* gene in the *heso1-1* background. We selected three *heso1-1* lines for which the endogenous URT1 was not detected anymore by western blot analysis, revealing a drastic downregulation of URT1 (**Figure 4A**). These lines, which have no particular phenotype when grown under optimal conditions, are called *heso1-1 urt1*^{SIL1}, *heso1-1 urt1*^{SIL2}, and *heso1-1 urt1*^{SIL3} thereafter. The uridylation of *MYB33* 5′-cleavage fragments was down to background levels in both biological replicates for three *heso1-1 urt1*^{SIL} lines as compared with the single *heso1-1* mutant (**Figure 4B**). Therefore, both HESO1 and URT1 participate in

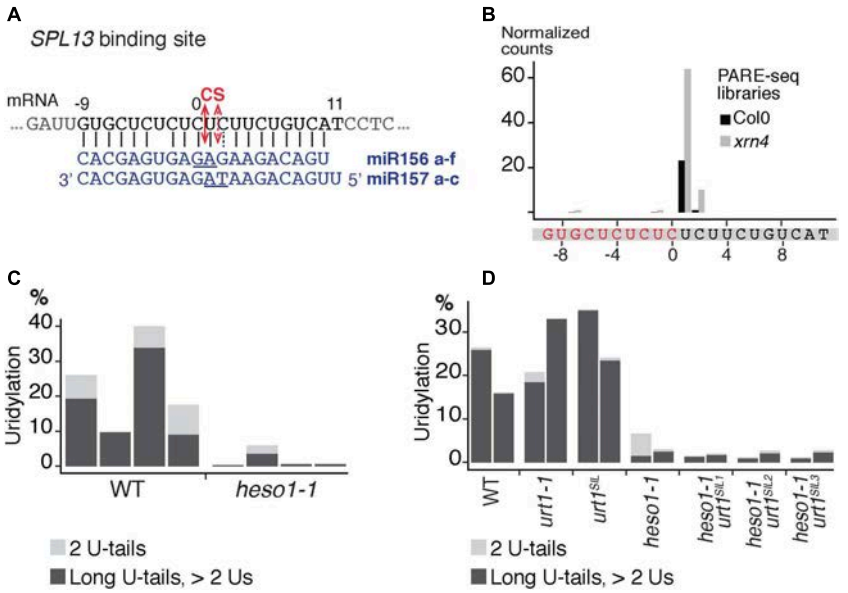


FIGURE 7 | Respective contributions of HESO1 and URT1 in the uridylation of *SPL13* RISC 5′-cleavage fragments. **(A)** Schematic representation of the cleavage of *SPL13* mRNA by AGO1 loaded with miR156 or miR157. The cleavage sites (CS) are predicted between the positions 0 and +1 or between position +1 and +2 for miR156 and miR157, respectively, and are defined by the 10th and 11th nucleotides (underlined) of the miRNAs. **(B)** 5′ end mapping using PARE-seq data (German et al., 2008) of RISC 3′-cleavage fragments in WT and *xrn4*. **(C,D)** Percentages of uridylated *SPL13* RISC 5′-cleavage fragments in **(C)** four biological replicates for WT and *heso1-1* or **(D)** two biological replicates for WT, *urt1-1*, the *urt1*^{SIL} line, *heso1-1*, and the three *heso1-1 urt1*^{SIL} lines (i.e., six *heso1-1 urt1*^{SIL} samples). Percentages of long (>2 Us) and 2 U-tails are indicated by dark gray and light gray, respectively.

uridylating *MYB33* 5′-cleavage fragments, albeit HESO1 is clearly the main TUTase involved in uridylating these fragments.

Of note, HESO1 and URT1 might have a distinct contribution in the uridylation of *MYB33* 5′-cleavage fragments. HESO1 can synthesize short and long U-extensions, but URT1 seems to add only one or two uridines (**Figure 4B**). Interestingly, a similar distinction was proposed for HESO1 and URT1 in uridylating small RNAs. URT1 was proposed to add a single uridine to small RNAs to favor the subsequent action of HESO1, which prefers 3′ extremities ending with uridines (Tu et al., 2015; Yu et al., 2017). A comparable scenario could exist for RISC 5′-cleavage fragments although additional investigation is required to confirm this hypothesis. In any case, and as previously observed for small RNAs, uridine addition by URT1 to RISC 5′-cleavage fragments does not seem to be a prerequisite to the action of HESO1, at least for *MYB33* 5′-cleavage fragments.

Respective Contribution of HESO1 and URT1 in the Accumulation of *MYB33* 5′-Cleavage Fragments

To further check the predominant role of HESO1 in the metabolism of *MYB33* 5′-cleavage fragments, we analyzed their accumulation by northern blot analysis and phosphorimager quantification (**Figure 5A**). The accumulation of *MYB33* RISC 5′-cleavage fragments in each sample was calculated relative to its full-length mRNA and each ratio was normalized to the ratio obtained for the WT control for each of the two replicates. As previously observed (Ren et al., 2014), *MYB33* 5′-cleavage fragments accumulated to higher levels in *heso1-1* with respect to WT (**Figure 5B**), although for unknown reasons the accumulation seemed variable in both replicates. Yet, our northern analysis confirmed that uridylation by HESO1 likely destabilizes *MYB33* 5′-cleavage fragments. The single *urt1* mutation seemed to have no major effect on this accumulation. Furthermore, *MYB33* 5′-cleavage fragments accumulated to similar levels in the *heso1-1 urt1*^{SIL} lines as compared to the single *heso1-1* mutant (**Figure 5**). In other words, there was no additive effect of the lack of URT1 and HESO1, and this observation points to HESO1 as the main TUTase controlling the accumulation of *MYB33* 5′-cleavage fragments. Of note, miR159 accumulated to similar levels when HESO1 is absent, ruling out a higher rate of production of *MYB33* 5′-cleavage fragments in *heso1-1* mutants (**Figure 5C**). Altogether, the 3′ RACE-seq and northern analyses indicate that HESO1 is the main TUTase modifying *MYB33* 5′-cleavage fragments. Although URT1 could add short uridine extensions to *MYB33* 5′-cleavage fragments, it does not appear to be a limiting factor neither in the uridylation nor in the destabilization of this fragment produced by RISC cleavage.

mRNA 5′ Fragments Are Nibbled at RISC Cleavage Site in the Absence of Uridylation

The 3′ truncation up to several hundreds of nucleotides upstream of the RISC cleavage site was previously observed for *MYB33*

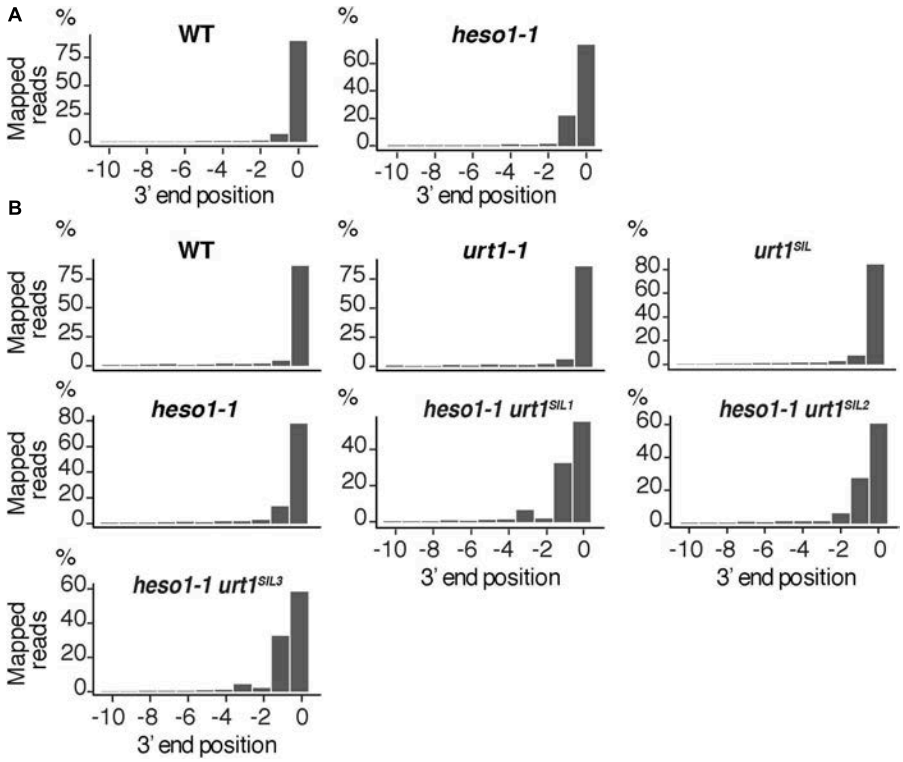


FIGURE 8 | Nibbled *SPL13* RISC 5′-cleavage fragments accumulate in the absence of uridylation. Positions of 3′ extremities of *SPL13* RISC 5′-cleavage fragments mapped in a –10/0 window for **(A)** four biological replicates, WT and *heso1-1*, or **(B)** two biological replicates for WT, *urt1-1*, the *urt1*^{SIL} line, *heso1-1*, and the three *heso1-1 urt1*^{SIL} lines. Graphs are shown separately for each of the replicates in **Supplementary Figures S5**, **S6**.

5′-cleavage fragments in the *heso1-2* mutant (Ren et al., 2014). We took advantage of the depth of the 3′ RACE-seq procedure to analyze at high resolution the 3′ extremities of *MYB33* 5′-cleavage fragments in the vicinity of the cleavage site. Although the vast majority of extremities in the four WT biological replicates mapped at position 0, different patterns were observed for *heso1-1*. The patterns were not completely identical in the four biological replicates, but they all revealed the same trend: the 3′ extremities were spread over positions from −10 to 0 (**Figure 6** and **Supplementary Figure S3**). This observation reveals that the *MYB33* 5′-cleavage fragments that accumulate in the absence of HESO1 are nibbled at close proximity to the cleavage site. This nibbling shortens *MYB33* 5′-cleavage fragments by up to 8–9 nucleotides (**Figure 6A** and **Supplementary Figure S3**). Such a nibbling was not observed in the single *urt1-1* mutant (**Figure 6B** and **Supplementary Figure S4**) but it was consistently observed in *heso1-1* and not aggravated in *heso1-1 urt1*^{SIL} mutants (**Figure 6B** and **Supplementary Figure S4**). Therefore, the nibbling is solely attributed to the absence of HESO1, but not of URT1, in the case of *MYB33* 5′-cleavage fragments.

We then analyzed the respective contribution of HESO1 and URT1 in uridylating the 5′ fragments produced by RISC cleavage of *Squamosa promoter-binding-like protein 13* (*SPL13*) mRNAs that are targets of miR156 and miR157 (**Figure 7A**). PARE-seq data identify a major and a minor 5′ extremity for the 3′ fragments produced by RISC cleavage (**Figure 7B**). Therefore, it is possible that in addition to the major cleavage site denoted 0 in **Figure 7A**, a minor site exists at position +1. This minor site at +1 presumably results from the action of miR157 (**Figure 7A**; He et al., 2018). Because nucleotide +1 is a U, it is not possible to determine in the 3′ RACE-seq data whether this U is encoded or added post-transcriptionally. To eliminate this uncertainty that could affect the proportion of uridylated versus non-uridylated fragments, we considered only tails of at least two nucleotides. Of note, not considering the 1 U extensions may lead to the underestimation of the action of URT1 and/or HESO1 in adding 1 U. The overall level of uridylation of *SPL13* 5′-cleavage fragments was lower than for *MYB33*, with percentage of uridylation below 40% and an increased variability between replicates (**Figures 7C,D** and **Supplementary Table S4**). Yet, a similar pattern was observed for both targets: uridylation of RISC 5′-cleavage fragments is mostly reduced in the absence of HESO1 and close to background levels in *heso1-1 urt1*^{SIL} lines (**Figures 7C,D**). Interestingly, the nibbling of RISC 5′-cleavage fragments was increased in the six replicates of *heso1-1 urt1*^{SIL} although to a lesser extent than the one observed for *MYB33* (**Figure 8**). This observation confirms the accumulation of RISC 5′-cleavage fragments that are nibbled close to the cleavage site in case of defective uridylation. The greater accumulation of nibbled fragments in absence of HESO1 and URT1 suggests

that in the case of *SPL13* 5′-cleavage fragments, the absence of uridylation *per se* is responsible for this accumulation.

Two, non-mutually exclusive, interpretations can explain the accumulation of nibbled RISC 5′-cleavage fragments in the absence of uridylation. First, uridylation of the nibbled fragments could trigger their degradation. Their fast turn-over would explain that they are not detected in WT plants. However, those fragments would accumulate in the absence of the TUTases. The second alternative possibility would be that in the presence of HESO1 and/or URT1, the 3′ extremities are not accessible to the activity, presumably a 3′-5′ exoribonucleolytic activity, that generates the nibbled RNA species. Such a possibility was previously evoked to explain the accumulation of truncated 5′-cleavage fragments in the *heso1-2* mutant (Ren et al., 2014). Solving this question entails the identification of all ribonucleases involved in the metabolism of 5′ RISC-cleaved transcripts.

CONCLUSION

Here, we report the respective contribution of HESO1 and URT1 in the metabolism of two 5′ RISC-cleaved mRNAs. In addition, we show the applicability of 3′ RACE-seq to map the 3′ ends of 5′ RISC-cleaved transcripts and to identify untemplated nucleotides added at these 3′ ends. The depth of 3′ RACE-seq will be useful for both qualitative and quantitative comparisons across different targets, tissues, conditions or genotypes. For instance, different RISC 5′-cleavage fragments could be investigated to identify both common and specific behaviors of these RNA fragments produced by post-transcriptional gene silencing. Also, the full machinery involved in the degradation of RISC 5′-cleavage fragments needs to be characterized. This is an on-going process with the recent identification of RICE exoribonucleases (Zhang et al., 2017) or the recent description that components of the NSD pathway and the Ski complex, a major co-factor of the cytosolic RNA exosome, are involved in the degradation of RISC 5′-cleavage fragments (Branscheid et al., 2015; Szádeczky-Kardoss et al., 2018). Yet the direct involvement of the RNA exosome in the clearance of RISC 5′-cleavage fragments remains to be demonstrated in Arabidopsis. The impact of SUPPRESSOR OF VARICOSE (SOV), whose ortholog is called Dis3L2 in non-plant eukaryotes, on the degradation of RISC 5′-cleavage fragments could also be investigated. Dis3L2 is a 3′-5′ exoribonuclease belonging to the RNase II family and whose activity is stimulated by uridylation in fission yeast, fruit fly or human cells (De Almeida et al., 2018). Whether SOV participates in the clearance of uridylated 5′ fragments of RISC-cleaved transcripts could be reliably addressed using 3′ RACE-seq and by comparing Col-0 and *Ler* accessions, because a point mutation affects SOV activity in Col-0 (Zhang et al., 2010). All these examples illustrate that a large number of samples must be analyzed with sufficient depth and replicates to draw reliable conclusions. The 3′ RACE-seq method adapted to the analysis of RISC 5′-cleavage fragments will contribute to fully characterize the tailing and nibbling events linked to the metabolism of these fragments and to address the respective roles of distinct factors of the RNA degradation machinery in this process.

AUTHOR CONTRIBUTIONS

DG and HZ conceived and designed the study, wrote the paper, and acquired funding. HZ, A-CJ, and HS performed the experiments. HZ performed the bioinformatics analysis. A-CJ and HS edited the manuscript. HZ and HS prepared the illustrations.

FUNDING

This work was supported by the Centre National de la Recherche Scientifique (CNRS, France) and the Agence Nationale de la Recherche (ANR, France) as part of the Programme d'Investissements d'Avenir in the frame of the LabEx NetRNA (ANR-2010-LABX-36) to DG and in the frame of the IdEx Unistra to HZ.

ACKNOWLEDGMENTS

The authors are grateful to Camille Noblet for technical help.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01438/full#supplementary-material

FIGURE S1 related to Figure 4 | Uncropped images of the western blot analysis and the membrane stained with Coomassie blue shown in **Figure 4**.

FIGURE S2 related to Figure 5 | Uncropped images of the northern blot analysis and the membrane stained with methylene blue for **(A)** *MYB33* 5′ fragment analysis and for **(B)** miR159 analysis.

FIGURE S3 related to Figure 6 | Nibbled *MYB33* RISC 5′-cleavage fragments accumulate in the absence of HESO1. Positions of 3′ extremities of *MYB33* RISC 5′-cleavage fragments mapped in a −10/0 window for four biological replicates in WT and *heso1-1*. Graphs are shown separately for each of the four replicates.

FIGURE S4 related to Figure 6 | Nibbled *MYB33* RISC 5′-cleavage fragments accumulate in the absence of HESO1. Positions of 3′ extremities of *MYB33* RISC 5′-cleavage fragments mapped in a −10/0 window for two biological replicates for WT, *urt1-1*, the *urt1*^{SIL} line, *heso1-1*, and the three *heso1-1 urt1*^{SIL} lines. Graphs are shown separately for each of the two replicates.

FIGURE S5 related to Figure 8 | Positions of 3′ extremities of *SPL13* 5′-cleavage fragments mapped in a −10/0 window for four biological replicates in WT and *heso1-1*. Graphs are shown separately for each of the four replicates.

FIGURE S6 related to Figure 8 | Nibbled *SPL13* RISC 5′-cleavage fragments accumulate in the absence of HESO1 and URT1. Positions of 3′ extremities of *SPL13* RISC 5′-cleavage fragments mapped in a −10/0 window for two biological replicates for WT, *urt1-1*, the *urt1*^{SIL} line, *heso1-1*, and the three *heso1-1 urt1*^{SIL} lines. Graphs are shown separately for each of the two replicates.

TABLE S1 | List of primers used in this study.

TABLE S2 | Summary of the number of reads analyzed at each step of the data processing for each 3′ RACE-seq library.

TABLE S3 | Exhaustive list of extensions found by 3′ RACE-seq for *MYB33* 5′-cleavage fragments in WT for four biological replicates from dataset #1.

TABLE S4 | Exhaustive list of extensions found by 3′ RACE-seq for *SPL13* 5′-cleavage fragments in WT for four biological replicates from dataset #1.

DATA SHEET S1 | Scripts for 3′ RACE-seq data processing, related to the analysis of 5′ mRNA fragments generated by RISC cleavage of *MYB33* mRNAs.

DATA SHEET S2 | Scripts for 3′ RACE-seq data processing, related to the analysis of 5′ mRNA fragments generated by RISC cleavage of *SPL13* mRNAs.

REFERENCES

Arribas-Hernández, L., Kielpinski, L. J., and Brodersen, P. (2016). mRNA decay of most Arabidopsis miRNA targets requires slicer activity of AGO1. *Plant Physiol.* 171, 2620–2632. doi: 10.1104/pp.16.00231

Bartel, D. P. (2018). Metazoan microRNAs. *Cell* 173, 20–51. doi: 10.1016/j.cell.2018.03.006

Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.* 16, 727–741. doi: 10.1038/nrm4085

Branscheid, A., Marchais, A., Schott, G., Lange, H., Gagliardi, D., Andersen, S. U., et al. (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. doi: 10.1093/nar/gkv1014

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., et al. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science* 320, 1185–1190. doi: 10.1126/science.1159151

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. doi: 10.1016/j.molcel.2014.02.007

Chen, X. (2004). A MicroRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* 303, 2022–2025. doi: 10.1126/science.1088060

De Almeida, C., Scheer, H., Zuber, H., and 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

German, M. A., Pillay, M., Jeong, D. -H., Hetawal, A., Luo, S., Janardhanan, P., et al. (2008). Global identification of microRNA–target RNA pairs by parallel analysis of RNA ends. *Nat. Biotechnol.* 26, 941–946. doi: 10.1038/nbt1417

He, J., Xu, M., Willmann, M. R., McCormick, K., Hu, T., Yang, L., et al. (2018). Threshold-dependent repression of SPL gene expression by miR156/miR157 controls vegetative phase change in *Arabidopsis thaliana*. *PLoS Genet.* 14:e1007337. doi: 10.1371/journal.pgen.1007337

Iwakawa, H., and 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

Kwak, J. E., and Wickens, M. (2007). A family of poly(U) polymerases. *RNA* 13, 860–867. doi: 10.1261/rna.514007.

Li, S., Liu, L., Zhuang, X., Yu, Y., Liu, X., Cui, X., et al. (2013). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. *Cell* 153, 562–574. doi: 10.1016/j.cell.2013.04.005

Martínez de Alba, A. E., Elvira-Matelot, E., and Vaucheret, H. (2013). Gene silencing in plants: a diversity of pathways. *Biochim. Biophys. Acta* 1829, 1300–1308. doi: 10.1016/j.bbagrm.2013.10.005

Munoz-Tello, P., Rajappa, L., Coquille, S., and Thore, S. (2015). Polyuridylation in eukaryotes: a 3′-end modification regulating RNA Life. *Biomed. Res. Int.* 2015:968127. doi: 10.1155/2015/968127

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. doi: 10.1261/rna.7231505

Reis, R. S., Hart-Smith, G., Eamens, A. L., Wilkins, M. R., and Waterhouse, P. M. (2015). Gene regulation by translational inhibition is determined by dicer partnering proteins. *Nat. Plants* 1:14027. doi: 10.1038/nplants.2014.27

Ren, G., Chen, X., and Yu, B. (2012). Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. *Curr. Biol.* 22, 695–700. doi: 10.1016/j.cub.2012.02.052

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. doi: 10.1073/pnas.1405083111

Scheer, H., Zuber, H., De Almeida, C., and Gagliardi, D. (2016). Uridylation earmarks mRNAs for degradation and more. *Trends Genet.* 32, 607–619. doi: 10.1016/j.tig.2016.08.003

Sement, F. M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J. -M., et al. (2013). Uridylation prevents 3′ trimming of oligoadenylated mRNAs. *Nucleic Acids Res.* 41, 7115–7127. doi: 10.1093/nar/gkt465

Shen, B., and Goodman, H. M. (2004). Uridine addition after microRNA-directed cleavage. *Science* 306, 997–997. doi: 10.1126/science.1103521

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. doi: 10.1038/s41467-017-02066-2

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. doi: 10.1016/j.molcel.2004.06.006

Szádeczky-Kardoss, I., Csorba, T., Auber, A., Schamberger, A., Nyikó, T., Taller, J., et al. (2018). The nonstop decay and the RNA silencing systems operate cooperatively in plants. *Nucleic Acids Res.* 46, 4632–4648. doi: 10.1093/nar/gky279

Tu, B., Liu, L., Xu, C., Zhai, J., Li, S., Lopez, M. A., 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

Vaucheret, H. (2008). Plant ARGONAUTES. *Trends Plant Sci.* 13, 350–358. doi: 10.1016/j.tplants.2008.04.007

Wang, X., Zhang, S., Dou, Y., Zhang, C., Chen, X., Yu, B., 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. doi: 10.1371/journal.pgen.1005091

Yang, L., Wu, G., and Poethig, R. S. (2012). Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. *Proc. Natl. Acad. Sci.* 109, 315–320. doi: 10.1073/pnas.1114673109

Yang, Z., Ebright, Y. W., Yu, B., and 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

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., et al. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science* 307, 932–935. doi: 10.1126/science.1107130

Yu, Y., Jia, T., and Chen, X. (2017). The "how" and "where" of plant microRNAs. *New Phytol.* 216, 1002–1017 doi: 10.1111/nph.14834

Zhang, H., Xia, R., Meyers, B. C., and Walbot, V. (2015). Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Curr. Opin. Plant Biol.* 27, 84–90. doi: 10.1016/j.pbi.2015.06.011

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. U.S.A.* 107, 15981–15985. doi: 10.1073/pnas.1007060107

Zhang, Z., Hu, F., Sung, M. W., Shu, C., Castillo-González, C., Koiwa, H., 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

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T. T., Meyers, B. C., et al. (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

Zuber, H., Scheer, H., Ferrier, E., Sement, F. M., Mercier, P., Stupfler, B., et al. (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

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

2.1.2 Analyse par 3'RACE-seq des sous-produits de maturation des ARN ribosomiques

La méthode de 3'RACE-seq a également été adaptée à l'analyse des extrémités 3' de 2 substrats types de l'exosome durant la maturation des ARNr : le précurseur de l'ARNr 5,8S et l'espaceur externe en 5' (5'ETS) issus de la maturation du précurseur polycistronique des ARNr 18S-5,8S-25/28S. Cette adaptation du protocole de 3'RACE-seq est publiée sous la forme d'un article accepté pour publication dans la revue Methods in Molecular Biology, article que je signe en tant que premier auteur.

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

Scheer H, De Almeida C, Sikorska N, Koechler S, Gagliardi D, Zuber H.

Accepté pour publication dans Methods in Molecular Biology

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,

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.

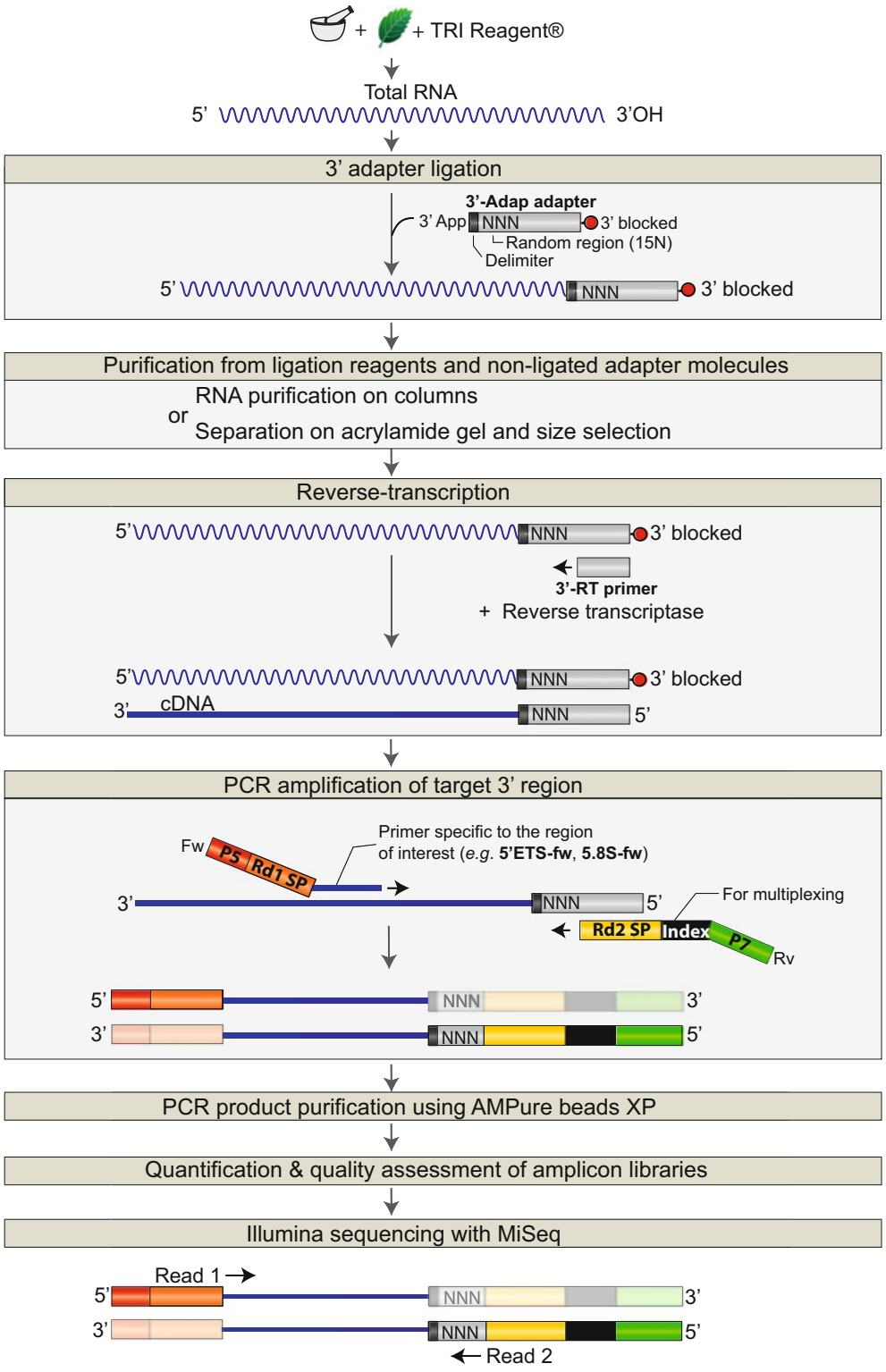


Fig. 1 Flowchart steps for 3′ RACE-seq procedure. After total RNA purification, the 3′ hydroxyl (3′ OH) end of each RNA molecule is ligated to the 3′-Adap adapter (*see* Fig. 2, Subheading 3.2 in the text). To remove nonligated 3′ adapters, ligated RNAs are purified using RNA purification columns or 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/CTGACNNNNNNNNNNNNNNNTGGAATTCTCGGGTGCCAAGGC/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.

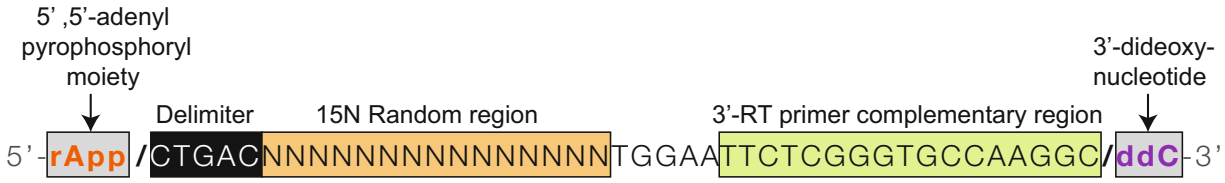


Fig. 2 Sequence details for the 3′-Adap adapter. 3′-Adap is a preadenylated oligonucleotide containing a 5′,5′-adenyl pyrophosphoryl moiety (*see* **Note 4**). The delimiter sequence allows for the demarcation between the sequence corresponding to the 3′ end of the RNA and the adapter sequence. This 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.

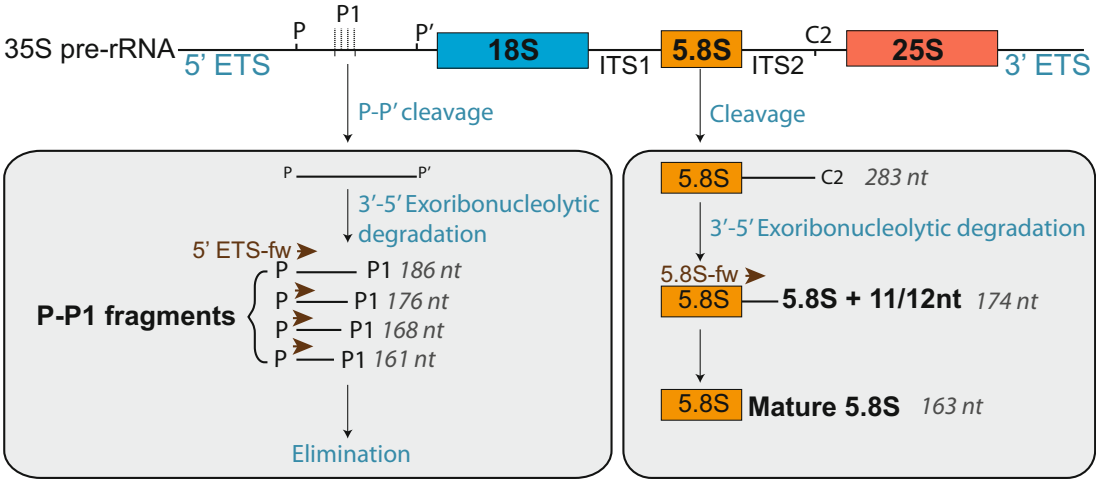


Fig. 3 Scheme of Arabidopsis 5′ ETS maturation by-products and 5.8S precursors analyzed by 3′ RACE-seq. 3′ RACE-seq procedure described in this chapter was used to map at high density the 3′ extremities of P-P1 fragments and of 5.8S precursors. P-P1 fragments result from endonucleolytic cleavage of the 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

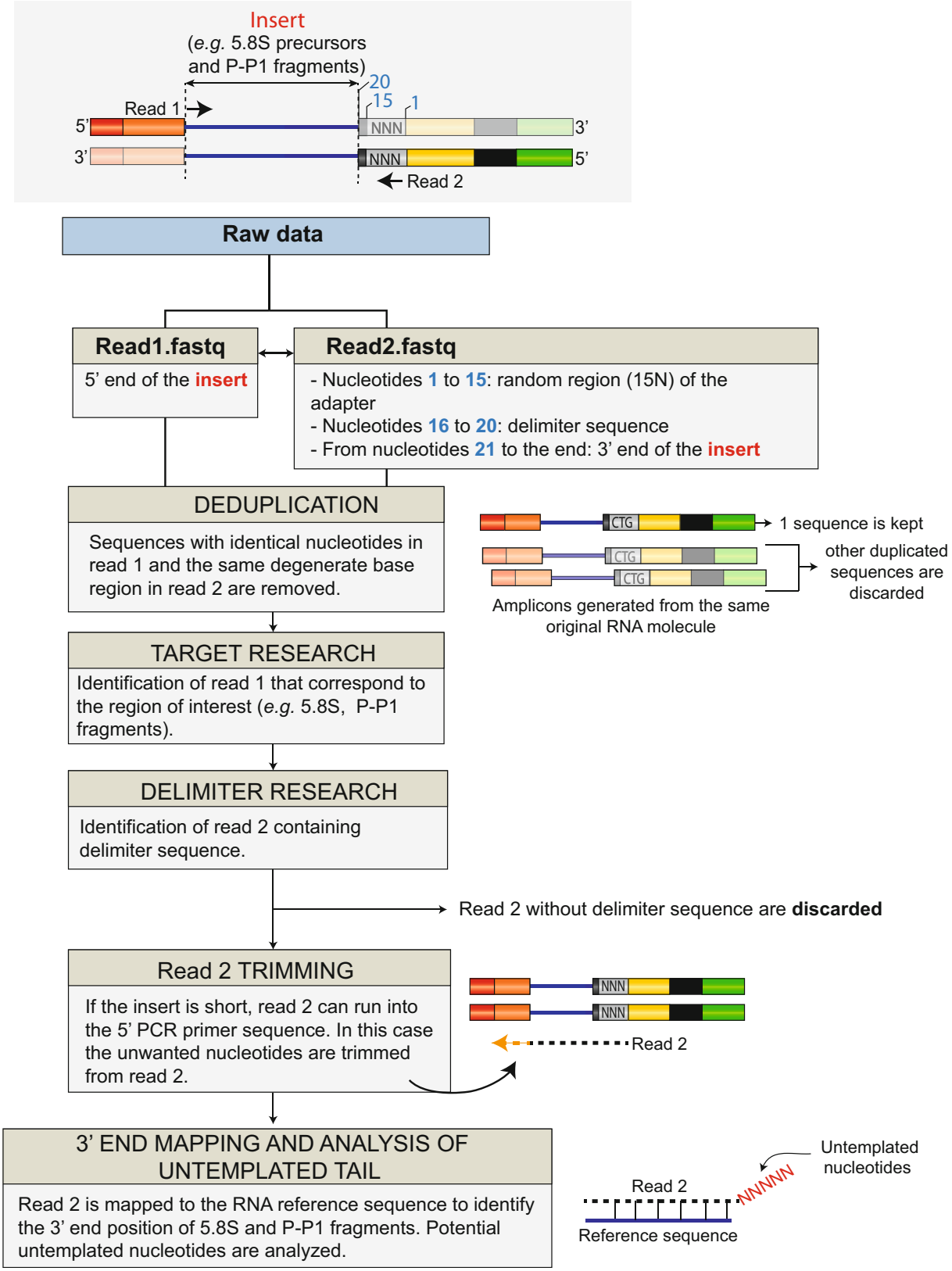


Fig. 4 Schematic overview of the bioinformatic pipeline for 3′ RACE-seq analysis of 5.8S and P-P1 fragments. Color code as in Fig. 1. Scripts are available in [6]

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.] × 10^6)/([bp length] × 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

Troisième partie :

URT1 interagit avec l'inhibiteur de traduction et cofacteur de décapping DCP5

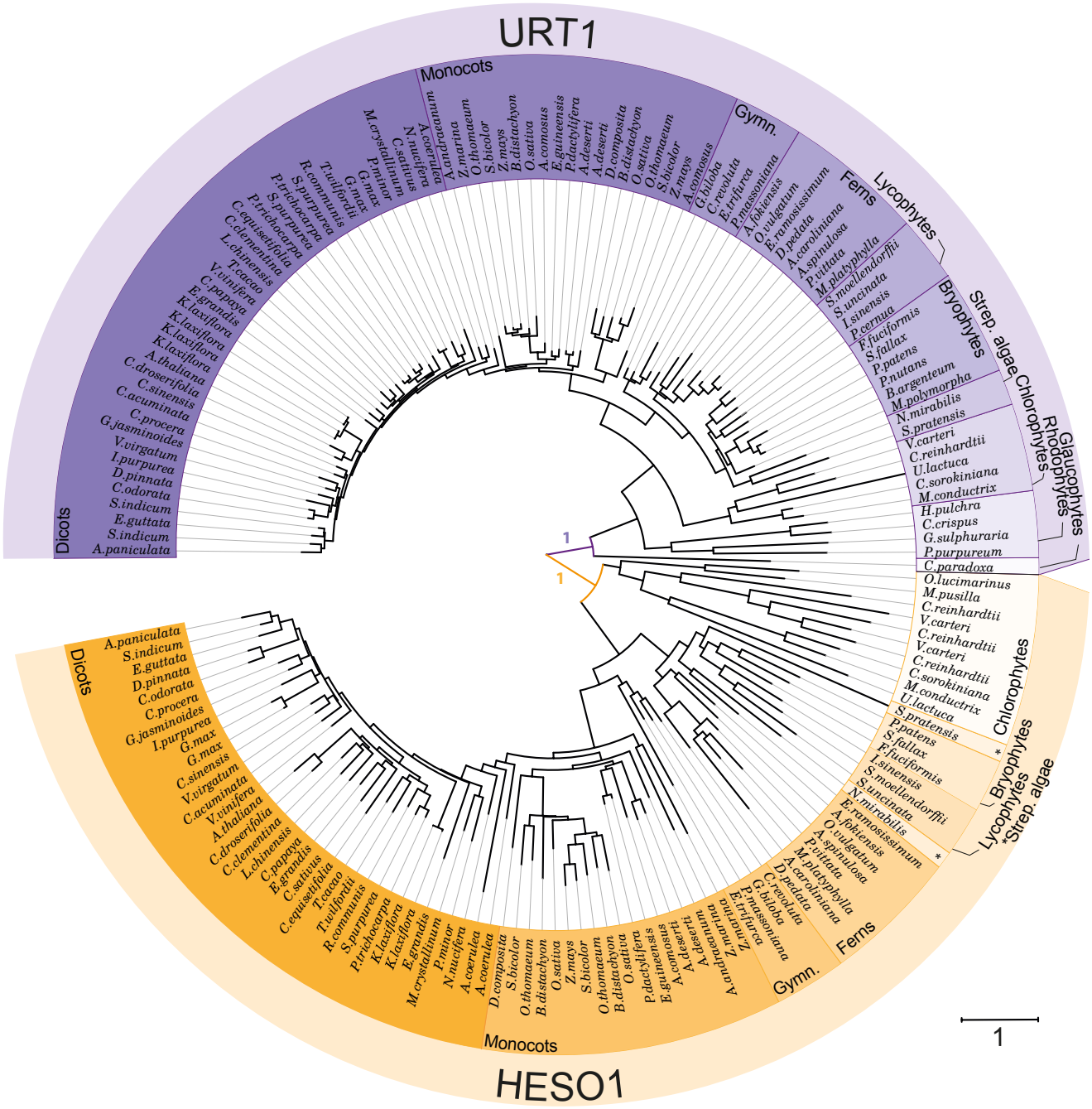


Figure 11| Analyse phylogénétique à partir des séquences homologues à URT1 et HESO1 chez 79 espèces représentatives d'Archaeplastida

L'arbre phylogénétique a été généré en utilisant la méthode de maximum de vraisemblance (aLRT) et le modèle de substitution WAG implémenté dans PhyML (v3.1). L'arbre obtenu a été édité en utilisant iTOL. Le code couleur utilisé pour les différents groupes taxonomiques est défini en Annexe 1. La valeur statistique obtenue pour les premières branches supporte le fait que HESO1 et URT1 forment deux clades distincts. L'échelle de mesure représente le nombre de substitutions par site.

URT1 possède une activité enzymatique similaire à la seconde TUTase présente chez Arabidopsis, HESO1, leurs rôles moléculaires semblent en revanche au moins partiellement distincts. Bien qu'il existe *in vivo* un certain recouvrement entre les types d'ARN reconnus par les deux protéines, URT1 est majoritairement responsable de l'uridylation des ARNm oligoadenylés dont l'une des conséquences est la protection face à déadenylation tandis qu'HES01 est principalement impliquée dans l'uridylation des petits ARNs (sRNA), favorisant leur dégradation. Comment expliquer une telle disparité de fonction pour deux protéines ayant une activité enzymatique similaire ?

L'analyse architecturale de URT1 apporte un élément de réponse. J'ai approfondi et actualisé cette analyse, initiée par Emilie Ferrier lors de sa thèse, et notamment pour l'analyse évolutive effectuée pour des espèces représentatives de l'ensemble des *Archaeplastida*. URT1 présente une architecture particulière avec une région N-terminale prédite comme intrinsèquement désorganisée, longue de ~400 acides aminés et dépourvue de domaine connu. Cette région possède en outre des motifs très spécifiques et extrêmement bien conservés. Nous pensons que cet appendice et ses caractéristiques particulières pourraient conférer à URT1 un rôle singulier et distinct des autres TUTases. Nous montrons notamment que des motifs conservés dans la région N-terminale sont impliqués dans l'interaction de URT1 avec l'inhibiteur de traduction et co-facteur de décoiffage DCP5, l'homologue à LSM14 et Scd6 chez l'homme et la levure, respectivement.

1. Analyse évolutive et architecture de URT1

1.1 URT1 forme un groupe monophylétique distinct de celui de HESO1

Bien que la conservation des protéines URT1 et HESO1 aient déjà été partiellement étudiée dans le laboratoire (travaux de thèse de Emilie Ferrier) ou dans d'autres études (Yang *et al.*, 2017; You *et al.*, 2017), il n'existe jusqu'à alors aucune analyse évolutive de URT1 et HESO1 complète et homogène chez l'ensemble des *Archaeplastida* (lignée verte). Nous avons ainsi réalisé une analyse *in silico* nous permettant d'identifier les protéines homologues de URT1 et HESO1 dans 79 espèces représentatives d'*Archaeplastida* (Annexe 1). Les séquences homologues de URT1 et HESO1 chez les bryophytes, les lycophytes et les fougères ont été obtenues grâce aux travaux du laboratoire de Xuemei Chen (You *et al.*, 2017). L'ensemble des séquences analysées nous ont permis d'obtenir l'arbre phylogénétique de URT1 et HESO1 chez les *Archaeplastida* (Fig. 11). Notre analyse phylogénétique montre que URT1 et HESO1 forment deux groupes monophylétiques distincts (représentés en orange et en violet sur la figure) et qu'au moins une copie de URT1 et HESO1 est systématiquement retrouvée chez les bryophytes (sens large[3]), lycophytes, gymnospermes, monocotylédones et dicotylédones. Certaines espèces plus primitives ne semblent en revanche posséder qu'une des deux terminal uridylyltransférases. Aucun homologue à HESO1 n'a ainsi pu être identifié chez les glaucophytes ou les rhodophytes alors qu'une séquence homologue à URT1 a pu être identifiée dans ces mêmes espèces. Ces résultats pourraient suggérer que l'apparition de la

[3] **Les bryophytes, au sens large**, comprennent les bryophyta (les mousses), les hépatiques et les anthocérotes. Les bryophytes au sens strict ne comprennent que les mousses.

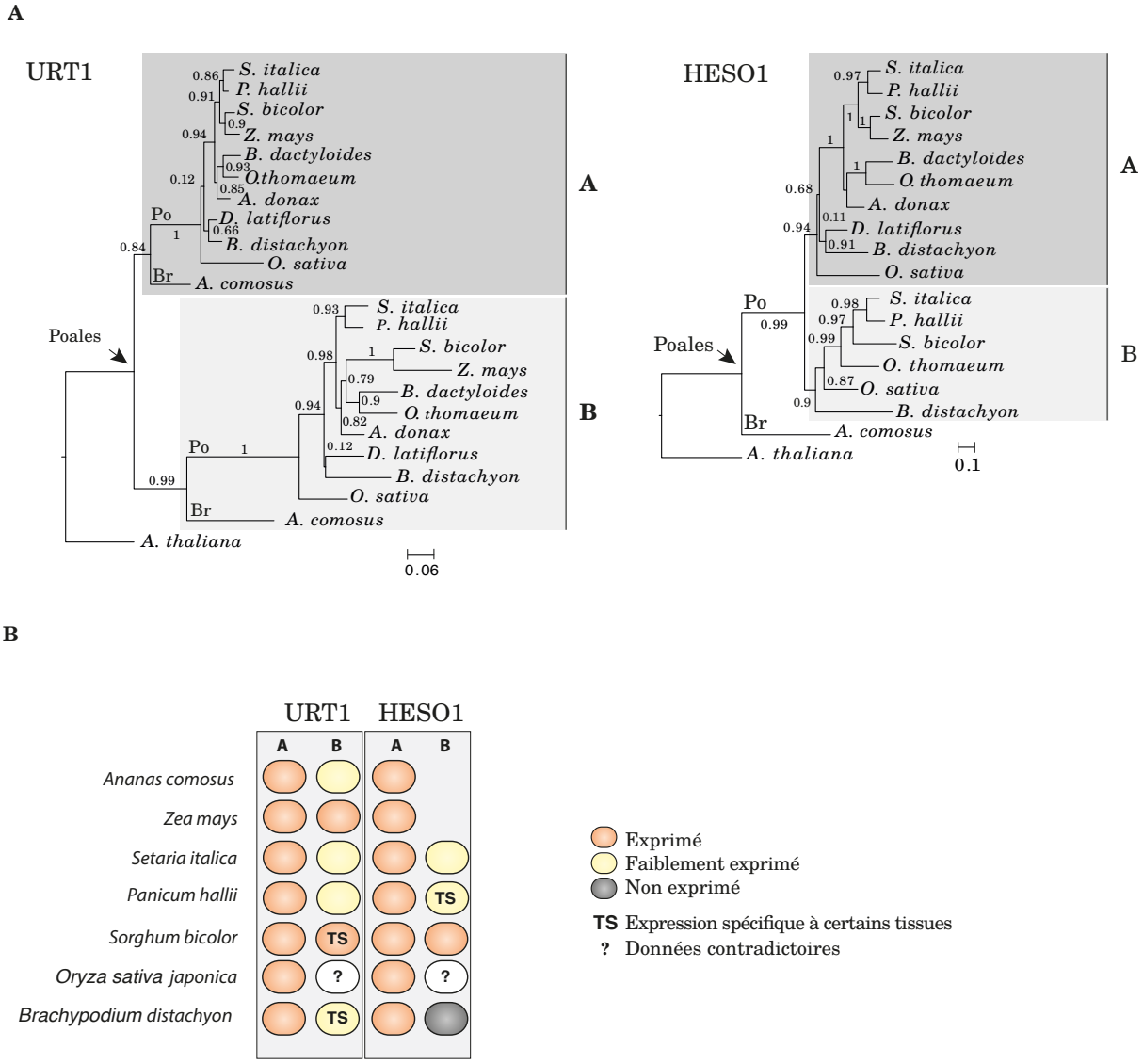


Figure 12 | Analyse phylogénétique des isoformes de URT1 et HESO1 dans 11 espèces de Poales
A | Les séquences des isoformes de HESO1 et URT1 de 10 espèces de Poacées (Po) et une espèce de Bromeliacée (Br) ont été alignées indépendamment. Les noms complets des espèces sont donnés en Annexe 1. L'analyse phylogénétique a été réalisée de la même façon qu'en Figure 4, excepté que les deux arbres obtenus ont été édités avec FigTree. Les valeurs de supports statistiques (aLRT) sont indiqués sur les branches. L'échelle de mesure représente le nombre de substitutions par site **B |** Niveau d'expression des isoformes de URT1 et HESO1 pour des espèces de Poales du clade BOP (*B. distachyon, O. sativa*) et du clade PACMAN (*P. hallii, S. italica, S. bicolor, Z. mays*) et pour une espèce de Broméliacée (*A. comosus*). Le diagramme a été construit à partir de données issues de Phytozome v12.1 (https://phytozome.jgi.doe.gov), de Next-Gen Sequence DBs (https://mpss.danforthcenter.org/dbs), de eFP browser (http://bar.utoronto.ca) et des données de RNAseq (Davidson et al. 2016).

première protéine homologue à URT1 et des fonctions moléculaires qui en découlent ont précédé celles de HESO1. URT1 est très largement conservée et n'est retrouvée absente que chez les Mamiellophyceae, des organismes unicellulaires, appartenant à la classe des chlorophytes et dont la taille du génome nucléaire est très réduite. De manière intéressante des séquences homologues à HESO1 ont été identifiées chez ces même organismes alors que les principaux composants de la machinerie de silencing y sont absents (Cerutti *et al.*, 2011).
Certaines espèces possèdent plusieurs copies de URT1 et HESO1, dans certains cas leurs séquences protéiques sont très similaires et l'apparition de deux copies est la conséquence d'une duplication récente du génome (p. ex. *Glycine max* et *Populus trichocarpa* (Panchy *et al.*, 2016). Dans d'autres cas, chez les Poales notamment, cette duplication est plus ancienne dans l'évolution et découle sur l'apparition de deux versions de URT1 (URT1A et URT1B) et HESO1 (HESO1A et HESO1B) qui forment deux sous-groupes distincts (Fig. 12A). Dans le cas de URT1, deux versions de la protéine sont systématiquement présentes dans les 10 espèces analysées et les protéines du groupe A semblent plus proches des dicotylédones que ne le sont celles du groupe B. Il est possible que les protéines du groupe A aient conservé les mêmes propriétés que la version de URT1 présente chez les dicotylédones et que les protéines du groupes B aient subi des modifications qui pourraient conduire à une néo-fonctionnalisation ou une pseudogénisation. Il serait dès lors très intéressant de comparer les fonctions cellulaires des protéines provenant des deux groupes afin de conclure.
Des données de transcriptomique provenant de certaines espèces de Poales dont le maïs, le riz, *Setaria italica* et *Brachypodium distachion* semblent indiquer que URT1A et HESO1A sont exprimées de manière constitutive contrairement à URT1B et HESO1B dont l'expression est faible ou restreinte à certains tissus (Fig. 12B).

Les analyses évolutives révèlent que URT1 forme un groupe monophylétique distinct de celui de HESO1 et que la protéine est conservée dans l'ensemble des *Archaeplastida*. Ces analyses ont été incluses dans un article accepté dans Philosophical Transactions of the Royal Society B, article que je co-signe en premier auteur et présenté en introduction. Nous nous sommes par la suite intéressés plus particulièrement à l'organisation de la protéine URT1, à ses caractéristiques particulières et à leur conservation à travers l'évolution.

1.2 La région N-terminale de URT1 possède des caractéristiques uniques qui sont conservées chez les *Archaeplastida*

URT1 est une protéine longue de 764 acides aminés. En plus de sa région C-terminale qui contient le cœur catalytique, cette protéine possède une région N-terminale ne comportant aucun domaine ou motif consensus reconnaissable (Fig. 13).
Le modèle structure-fonction qui prévalait aux débuts de la biologie structurale impliquait que la fonction moléculaire d'une protéine découlait d'une structure tridimensionnelle stable, rigide et

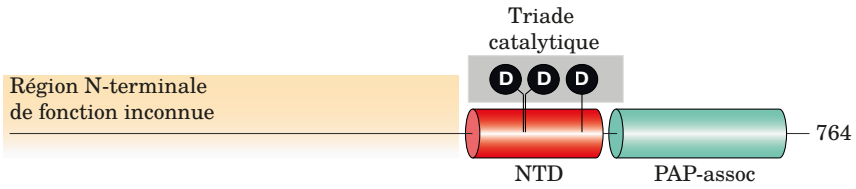


Figure 13 | Organisation structurelle de URT1
Schéma à l'échelle représentant l'organisation en domaine de URT1. URT1 possède deux domaines extrêmement bien conservés dans sa région C-terminale : le domaine nucléotidyltransférase (NTD) et le domaine Poly(A) polymérase associé (PAP-assoc). La région N-terminale ne contient aucun domaine ou motif connu.

bien définie. Il s'est très vite avéré, suite à la détermination massive de structures protéiques (environ ~140 000 disponibles aujourd'hui), qu'un très grand nombre de protéines possédaient en réalité un certain degré d'instabilité conformationnelle (Deller *et al.*, 2016), étaient partiellement ou complètement intrinsèquement désordonnées, et étaient difficilement cristallisables à l'état natif. Ces protéines intrinsèquement désorganisées (« intrinsically disordered proteins », **IDPs**) représentent ~30% des protéines cellulaires eucaryotes et sont donc très abondantes. Elles suscitent, et ce à juste titre, un intérêt croissant et le nombre d'outils bio-informatiques dédiés à leur analyse s'est considérablement accru. En réalité, le manque de contraintes structurelles confère aux IDPs des atouts moléculaires indubitables. Les régions intrinsèquement désorganisées (« intrinsically disordered region », **IDRs**) qui les composent sont **(i)** moins sensibles aux mutations aléatoires qui, bien qu'altérant la structures primaire, ne changeront pas la fonction moléculaire de l'IDR, **(ii)** apportent un site flexible pour la liaison à un ou plusieurs partenaires protéiques et **(iii)** contiennent bien souvent des motifs linéaires courts (« short linear motifs », SLiMS) également impliqués dans l'interaction avec d'autres protéines ou dans la reconnaissance aux substrats (Atkins *et al.*, 2015; Mészáros *et al.*, 2014; Oldfield and Dunker, 2014). La présence d'IDR(s) pourrait ainsi conférer à URT1 des propriétés intéressantes d'un point de vue moléculaire.

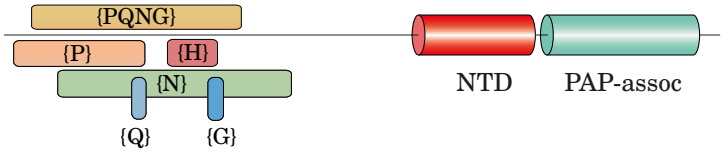
1.2.1 La région N-terminale de URT1 et de ses homologues chez les Archaeplastida est intrinsèquement désorganisée

Le terme « intrinsèquement » implicite que la désorganisation est une caractéristique dépendante de la séquence. En effet, les régions intrinsèquement désorganisées forment des segments protéiques de **faibles complexités** et **possèdent un biais très marqué de composition en acides aminés**, l'ensemble étant non propice à la formation de structures secondaires (Wright and Dyson, 1999).

La région N-terminale de URT1 est proposée comme étant une IDR (travaux de thèse d'Emilie Ferrier). Nous avons nous avons recherché *in silico* un biais de composition au sein de la protéine URT1. Nous avons tout d'abord utilisé le programme fLPS qui analyse la composition en acides aminés d'une séquence requête et définit des régions pour lesquelles un enrichissement en acide aminés est détecté. L'avantage de fLPS est qu'il détecte et localise aisément et extrêmement rapidement des biais pour un acide-aminé (unique) ou pour plusieurs acides aminés (multiple) et que les résidus qui introduisent ce biais peuvent être répartis de manière discontinue sur la séquence (Harrison, 2017). Le tableau représenté en Figure 14A récapitule les résultats de sortie d'une exécution en ligne de commande du fLPS. Pour une meilleure visualisation, les résultats sont également représentés sous forme de schémas. Cette analyse révèle un **biais de composition** en acides aminés Histidine, Proline, Glutamine, Asparagine et Glycine (H, P, G, N et Q, respectivement), un biais qui, nous le soulignons, est restreint à la région N-terminale de URT1. L'analyse de cette même région N-terminale de URT1 via Composition Profiler (Vacic *et al.*, 2007) corrobore ces résultats. Le logiciel Composition profiler réalise une analyse globale sur la séquence

A

Type de biais	LPS#	Début	Fin	Signature	P value
Unique	1	8	148	{P}	2.904e-11
Unique	2	134	149	{Q}	7.760e-10
Unique	3	172	226	{H}	3.539e-06
Unique	4	56	305	{N}	4.169e-06
Unique	5	216	233	{G}	6.503e-06
Multiple	1	1	27	{PQNG}	3.663e-18



B

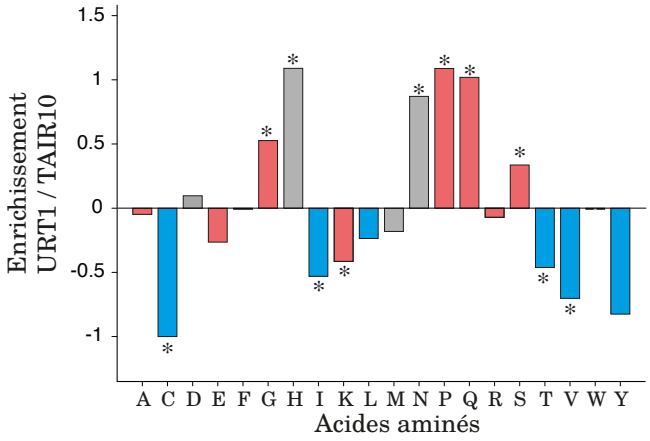


Figure 14 | Biais de composition en acides aminés dans la région N-terminale de URT1
A| Résultats de sortie après analyse du biais de composition en acides aminés de URT1 par le logiciel fLPS (Harrison et al. 2017). fLPS analyse des biais de composition pour un acide aminé unique ou pour plusieurs acides aminés (multiple). Les coordonnées de début et de fin indiquent la région où se situe le biais. Les résultats sont classés par ordre croissant de p-value et seuls ceux avec une p-value inférieure à 10^{-5} sont conservés. Les résultats sont également représentés sous la forme d'un schéma au-dessous du tableau. **B|** Analyse du biais de composition de la région N-terminale de URT1 (acides aminés 1 à 400) par le logiciel Composition Profiler (http://www.cprofiler.org). Le profil de composition de URT1 montre un enrichissement général pour les résidus qui promeuvent la désorganisation (en rouge) et une déplétion en ceux qui promeuvent l'ordre (en bleu) par rapport au protéome d'Arabidopsis (TAIR10). L'astérisque indique un enrichissement ou une déplétion significative (p-value >0,05).

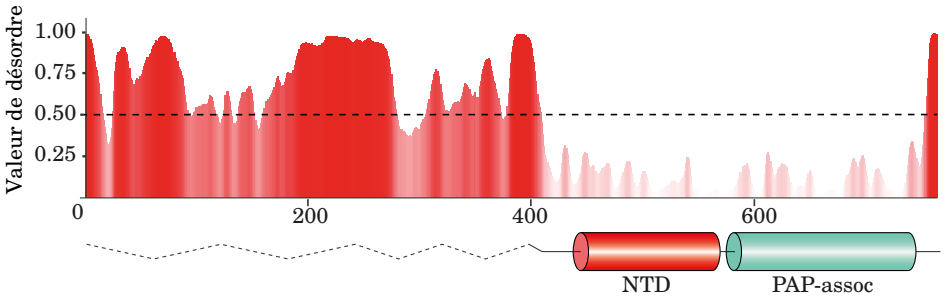


Figure 15 | La région N-terminale de URT1 est prédite comme étant intrinsèquement désorganisée
Détermination informatique des régions intrinsèquement désorganisées de URT1 en utilisant le prédicteur ESpritz-NMR du serveur FELLS (http://protein.bio.unipd.it/fells).

requête et détermine les acides aminés qui sont surreprésentés ou sous-représentés en prenant comme référence le protéome d'Arabidopsis (TAIR10). Le graphique représenté en Figure 14B confirme le biais observé avec fLPS pour les acides aminés H, P, G, N et Q et de manière intéressante révèle une déplétion, statistiquement significative, en acides aminés Cystéines, Isoleucine, Lysine, Tyrosine et Valine (C, I, K, Y et V, respectivement). Nous avons ajouté un niveau d'information supplémentaire aux résultats obtenus en marquant les acides aminés en fonction de leur propriété à promouvoir ou non la désorganisation. Cette analyse met en exergue un enrichissement, dans la région N-terminale de URT1, en acide-aminés promouvant la désorganisation (en rouge) et, à l'inverse, un appauvrissement en acides aminés promouvant « l'ordre » (en bleu). Ceci nous laisse présupposer que le biais de composition de la région N-terminale de URT1 est plus proche de celui d'une protéine intrinsèquement désorganisée qu'à celui d'une protéine structurellement organisée.

Afin de confirmer ces observations nous avons voulu analyser la séquence de URT1 avec des outils de **prédictions bio-informatiques de régions intrinsèquement désorganisées.** Il existe plus d'une soixantaine de serveurs de prédictions de régions intrinsèquement non structurées et ces derniers adoptent des méthodologies divergentes pour leurs prédictions (Atkins *et al.*, 2015). On distingue d'une part **les prédicteurs *ab initio*** qui ne se servent que de la structure primaire des protéines et calculent un degré de désordre en fonction de la propension des acides aminés à former des interactions inter-résidus favorables à la formation de régions structurées. Et d'autre part les prédicteurs utilisant **l'apprentissage automatique** (« machine learning ») pour la prédiction des IDPs. D'autres méthodes compilent les résultats de différents prédicteurs sous forme de méta-analyses. Nous avons fait le choix de ESpritz-NMR (Walsh *et al.*, 2012), un prédicteur de type « machine learning » dont l'apprentissage machine a été réalisé sur des données de résonance magnétique nucléaire (RMN), car il est rapide, propose une interface graphique simple d'utilisation (via le serveur FELLS (Piovesan *et al.*, 2017) et permet l'analyse en parallèle d'une multitude de séquences. L'analyse de la séquence de URT1 via ESpritz-NMR prédit les 413 premiers acides-aminés de la protéine comme étant intrinsèquement désorganisés, avec une valeur de désordre supérieur à 0.5 pour plus de 90% des résidus (Fig. 15). Si l'on considère la séquence dans son entièreté, 50% (±6,95%) de la protéine URT1 est prédite comme intrinsèquement désorganisée. Ce résultat est confirmé par l'utilisation de 4 autres prédicteurs (Annexe 2).

1.2.2 La région intrinsèquement désorganisée de URT1 est conservée chez l'ensemble des Archaeplastida

Nous avons utilisé ESpritz-NMR afin d'analyser le profil de désorganisation des protéines homologues à URT1 dans plusieurs espèces d'*Archaeplastida*. La figure présentée en Annexe 3 montre le pourcentage de résidus considérés comme étant intrinsèquement désorganisés (D, rouge) ou organisés (O, bleu) pour la région N-terminale et C-terminale de chacune des protéines considérées. Pour faciliter l'analyse bio-informatique, nous avons défini la région N-terminale (NTER) correspondant exactement à la première moitié des résidus de chacune des protéines et la région C-terminale (CTER) à la seconde moitié. Cette analyse montre que la région N-terminale

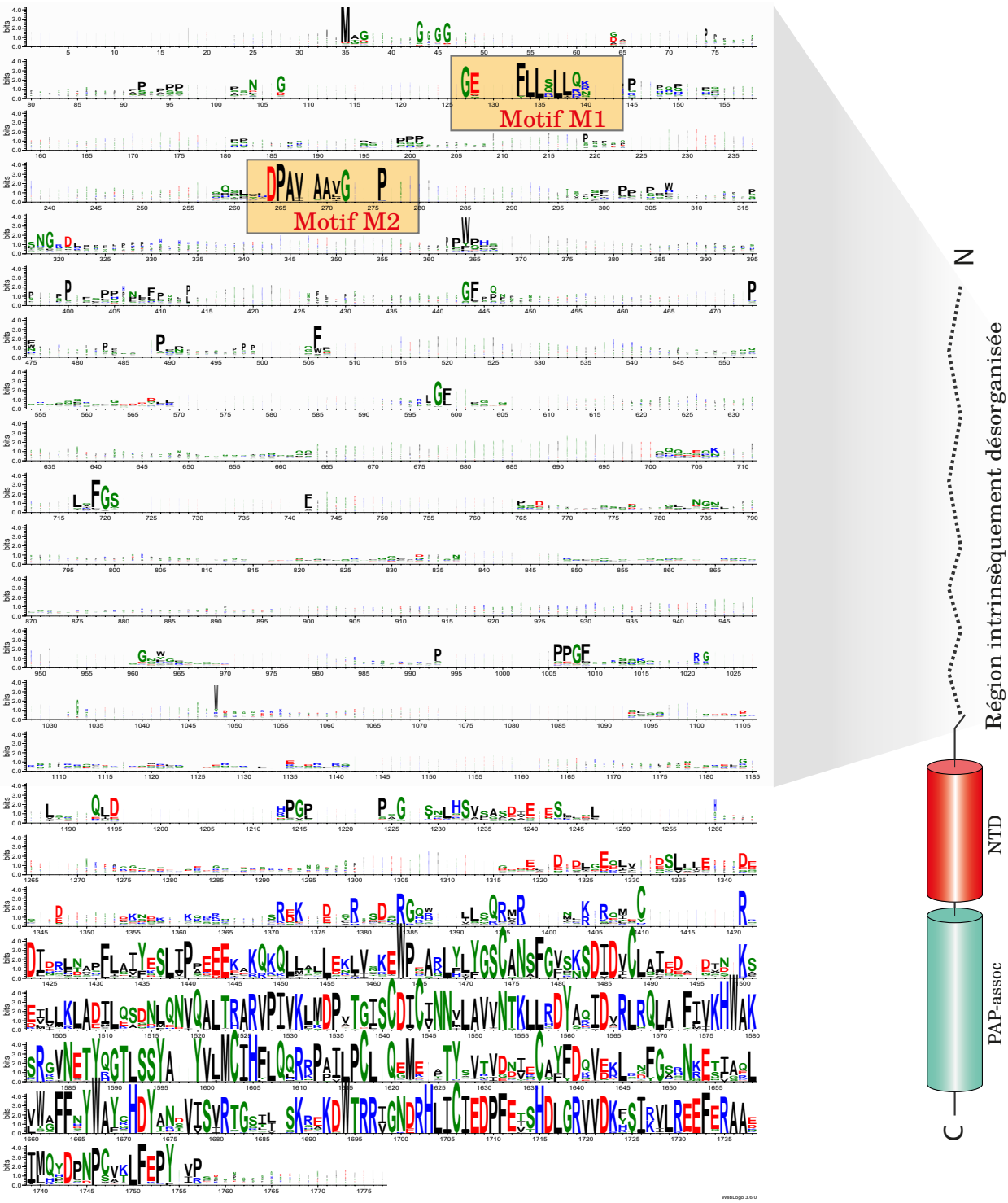


Figure 16 | Weblogo de URT1 et ses homologues chez les Embryophytes
Le logo a été réalisé avec WebLogo3 (https://weblogo.berkeley.edu/logo.cgi) à partir d'un alignement de 221 séquences homologues à URT1 chez les embryophytes (Les séquences utilisées pour la création du logo sont indiqués en Annexe 2.). L'alignement a été réalisé avec Jalview en utilisant l'algorithme Muscle. La région C-terminale contenant le domaine catalytique est extrêmement bien conservée. Le logo illustre bien l'évolution rapide de la région intrinsèquement désorganisée de URT1. L'alignement révèle la présence, au sein de l'IDR de URT1, de deux motifs hautement conservés, le motif 1 (M1) et le motif 2 (M2) (rectangles jaunes). La hauteur de chaque acide aminé représente sa conservation à la position considérée. Le maximum de conservation par site est égal à $\log_2 20$ soit ~4.32 bits.

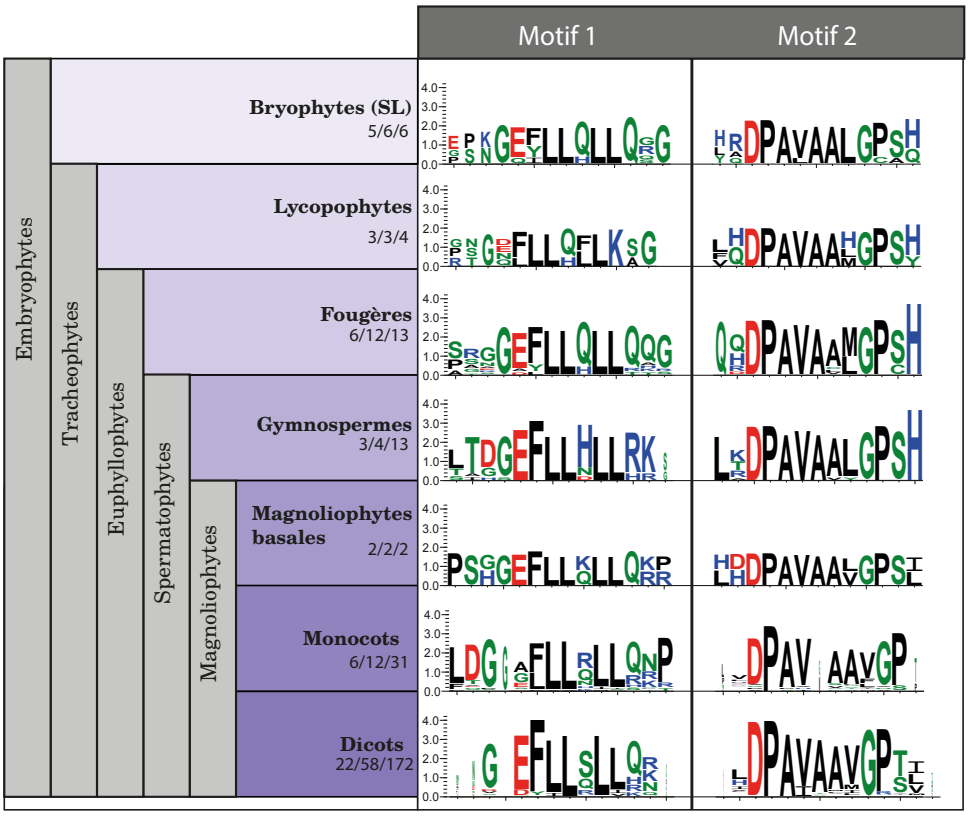


Figure 17 | Conservation des motifs des motifs 1 et 2 au cours de l'évolution des embryophytes
La séquence des logos sont exprimés en bits et calculé à partir d'un alignement de protéines homologues à URT1 pour l'ensemble des divisions des embryophytes. Le nombre d'ordre/familles/séquences utlisés pour chaque alignement est également indiqué. La couleur correspond à la classe des acides aminés: acides en rouge (D,E), basiques en bleu (K,R,H), hydrophobes (A,V,L,I,P,W,F,M) en noir et polaires en vert (G, S, T, Y, C, Q, N).

est de nature désordonnée pour l'ensemble des protéines analysées. Ceci suggère que la propension à la désorganisation de la région N-terminale de URT1 est une caractéristique conservée des algues aux plantes supérieurs.

1.3 Caractérisation de motifs conservés dans la région intrinsèquement désorganisée de URT1

URT1 possède une longue région N-terminale intrinsèquement désorganisée et son maintien au cours de l'évolution des plantes étaye l'idée de l'importance fonctionnelle de cette dernière. La conservation de la désorganisation n'implique en revanche pas une conservation au niveau de la séquence primaire. En effet, puisque les IDRs ne subissent aucune contrainte conformationelle, il existe, dans ces régions, une plus forte tolérance aux mutations et les résidus qui forment ces segments intrinsèquement désorganisés évoluent très rapidement. De fait, des régions intrinsèquement désorganisées semblent évoluer de manière neutre, suggérant que leur caractéristique de désorganisation peut être maintenue tout en tolérant des changements de séquence (Daughdrill *et al.*, 2007). En revanche, il est possible de retrouver au sein de ces régions dynamiques des résidus qui subissent une pression de sélection positive et qui forment des motifs fonctionnels et donc conservés. De tels motifs avaient été identifiés chez URT1 par Emilie Ferrier au cours de sa thèse dans notre laboratoire. Cette analyse avait alors été réalisée avec un nombre restreint de séquences de dicotylédones et de monocotylédones. Pour étendre cette étude, j'ai réalisé avec Weblogo3 un logo de séquences à partir d'un alignement de 221 séquences homologues à URT1 chez les embryophytes (les séquences utilisées pour cet alignement sont indiquées en Annexe 4). Le logo obtenu, représenté en Figure 16, révèle une très pauvre conservation de séquence dans la région intrinsèquement désorganisée de URT1 comparée à la région qui comprend le cœur catalytique. Plus frappant encore est la présence, au sein de cette IDR, de **deux motifs courts M1 et M2,** qui semblent, à l'inverse, avoir subi une **forte pression positive** au cours de l'évolution. Ces derniers sont en effet extrêmement bien conservés, puisqu'ils sont présents chez les bryophytes (sens large), les lycophytes, les fougères, les gymnospermes, les monocotylédones et les dicotylédones (Figure 17). Ils pourraient également être conservés chez certains organismes streptophytes puisque les motifs M1 et M2 sont également retrouvés chez l'algue *Klebsormidum flaccidum*. Aucun des deux motifs n'a, en revanche, pu être identifié chez les glaucophytes, rhodophytes et chlorophytes.

URT1 possède une région N-terminale enrichie en certains acides-aminés, prédite comme étant intrinsèquement désorganisée et qui contient deux motifs linéaires courts. La forte conservation au cours de l'évolution de cette désorganisation intrinsèque et des deux motifs conservés qui y sont contenus implique que cette région N-terminale joue un rôle fondamental dans la fonction moléculaire de URT1. Notre hypothèse est que les motifs M1 et M2 pourraient contribuer à promouvoir l'interaction de URT1 avec des partenaires protéiques. L'identification de partenaires

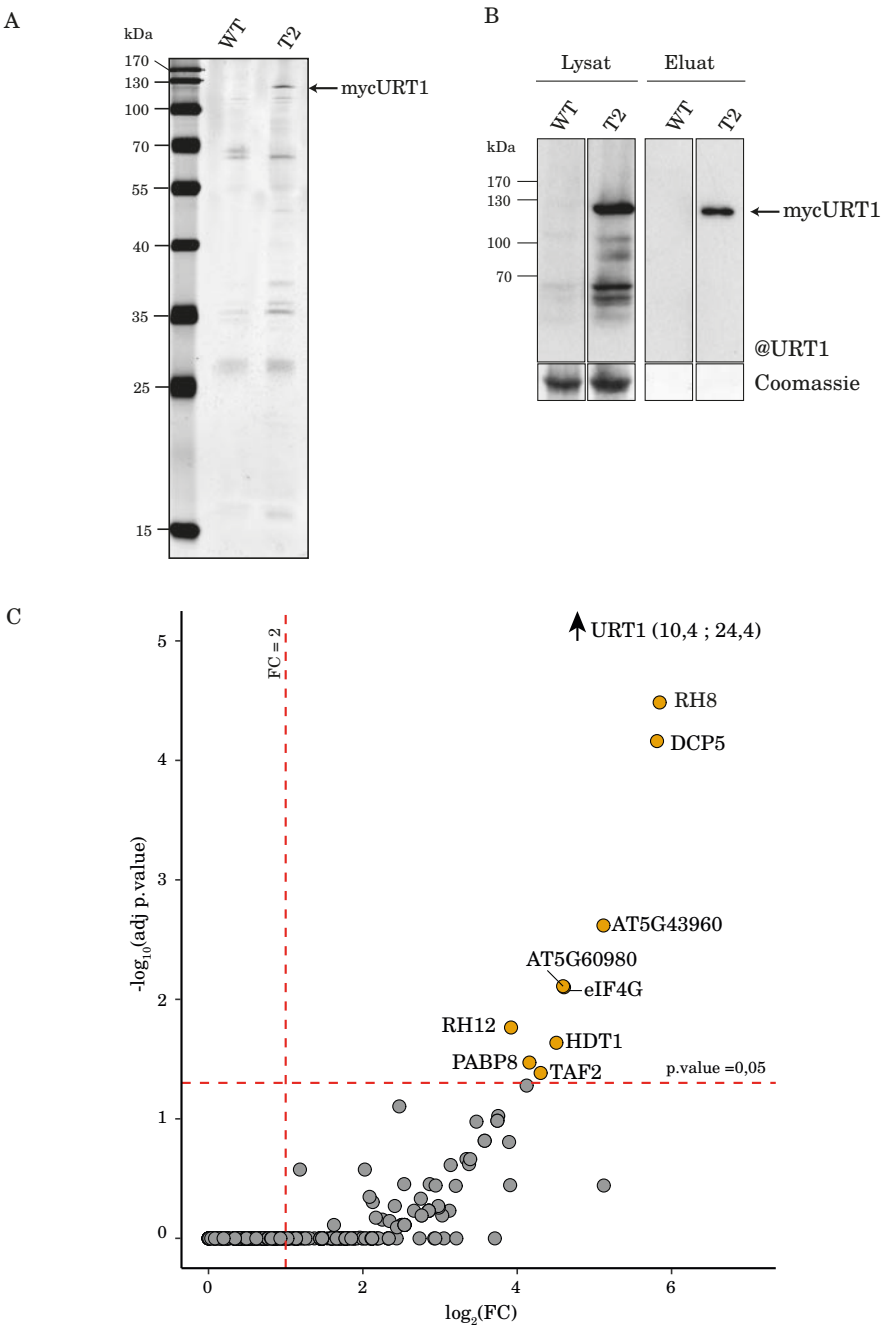


Figure 18 | Analyses des partenaires de URT1 par co-immunoprécipitation et analyse par spectrométrie de masse

A & B| Analyse par gel de coloration à l'argent (A) et par western blot révelé avec un anticorps dirigé contre URT1 (B) de myc-URT1 après expérience de co-immunoprécipitation en utilisant la lignée transgénique T2. La lignée T2 est mutante pour URT1 et surexprime URT1 avec 10 épitopes myc en N-terminal, (Zuber et al, 2016).
C| 4 fractions d'élution, dont 2 issues d'expériences de co-IP réalisées à partir de fleurs et 2 à partir de plantules ont été analysées par spectrométrie de masse avec 4 contrôles (IP) réalisées dans un fond sauvage n'exprimant pas myc-URT1. Le volcano plot indique les protéines enrichies dans l'IP myc-URT1. L'abscisse indique la valeur d'enrichissement (FC (Fold Change)) par rapport aux IP contrôles et l'ordonnée indique la p-value ajustée. La droite rouge horizontale représente une p-value = 0.05 et la droite verticale un FC = 2.

éventuels d'URT1 pourrait apporter de nouvelles informations quant au rôle joué par URT1 dans le métabolisme des ARNm. Nous avons dès lors voulu identifier quels étaient les partenaires protéiques de URT1 *in vivo*.

2. Recherche des partenaires protéiques de URT1

2.1 Choix des lignées transgéniques

Afin d'analyser l'interactome de URT1 nous avons adopté une approche de **co-immunoprécipitation** (co-IP) de URT1 et de ses partenaires éventuels couplée à une analyse par spectrométrie de masse. Pour ce faire nous avons utilisé les lignées T1 et T2 présentées en partie 1. Pour rappel ces lignées sont issues de la transformation de plantes mutantes *urt1* avec une construction permettant l'expression de URT1 marqué avec 10 épitopes myc (10myc-URT1) sous la dépendance du promoteur CaMV 35S. La lignée T1 exprime mycURT1 à un niveau inférieur au niveau endogène et la lignée T2 surexprime la protéine de fusion. Le niveau d'uridylation est partiellement et complètement restauré dans la lignée T1 et la lignée T2, respectivement. Ce qui indique que la protéine de fusion mycURT1 est active *in vivo*.

Des expériences de co-IP réalisées avec des billes magnétiques couplées à des anticorps dirigés contre l'épitope myc ont été réalisées à partir de fleurs issues de la lignée T1 et T2. Les co-IP réalisées à partir de la lignée T1 n'ont cependant abouti qu'à la purification d'une très faible quantité de myc-URT1, difficilement détectable en spectrométrie de masse et n'ont pas permis l'identification de partenaires protéiques. La lignée T2 permet quant à elle la purification d'une quantité suffisante de mycURT1. L'analyse des fractions d'élutions par gel de coloration à l'argent (Fig. 18A) et par western blot (Fig. 18B) montre un enrichissement de myc-URT1 dans la fraction d'élution de deux expériences indépendantes réalisées à partir de la lignée T2. Nous avons réalisé 4 expériences de co-IP de myc-URT1 à partir de la lignée T2 et analysé les fractions par spectrométrie de masse LC/MS-MS.

2.2 Les partenaires protéiques de URT1

Les résultats des expériences de co-IP sont représentés sous la forme d'un volcano plot en Figure 18C. **9 protéines** sont retrouvées enrichies de façon significative (p-value > 0.5, FC >2) dans les co-IP de myc-URT1 par rapport aux IP contrôles (IP à partir de plantes sauvages). Le nombre exact de spectres retrouvés pour chaque protéine et chaque expérience est donné en Figure 19.

Parmi ces 9 protéines on identifie deux protéines nucléaires, TAF2 et HDT1, dont la fonction ne semble pas reliée à celle de URT1 et qui sont susceptibles d'être des artefacts expérimentaux. TAF2 (TBP-ASSOCIATED FACTOR 2) est un composant du facteur de transcription IID (TFIID) et HDT1 est impliqué dans la régulation de la transcription via la désacétylation de l'histone H3 (Lawrence *et al.*, 2004). A noter tout de même que, chez l'homme, les histones désacétylases HDAC1 et HDAC2 ont été montrées comme étant impliquées dans la désacétylation et la régulation de la déadénylase CAF1 (Sharma *et al.*, 2016), une activité probablement permise par le fait que CCR4-

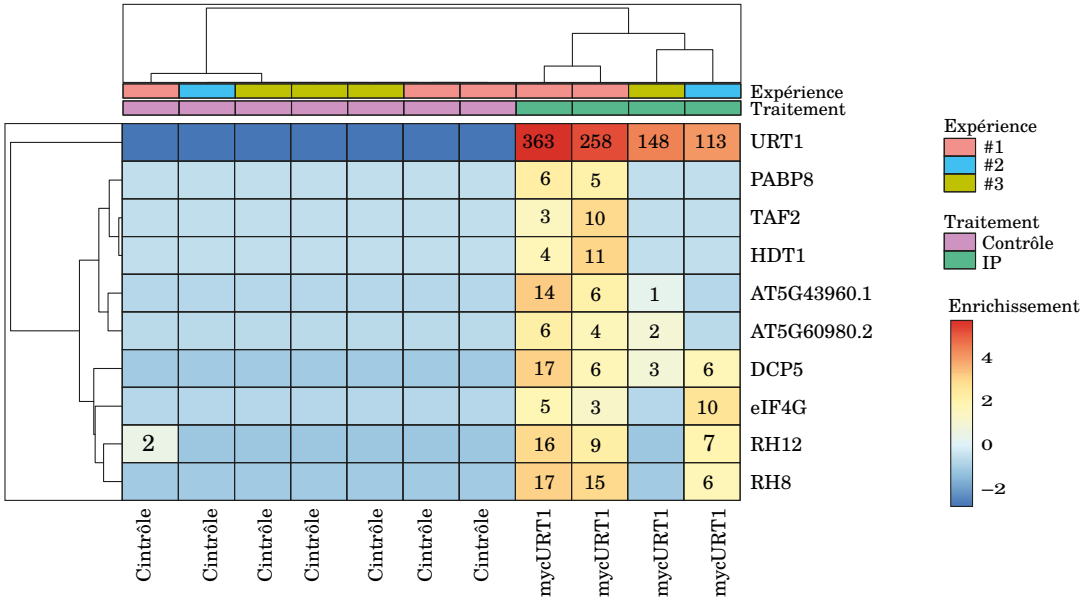


Figure 19 | Nombre de spectres identifiés par spectrométrie de masse pour les protéines co-immunoprécipitant avec mycURT1 avec une p-value < 0,05 et un FC >2

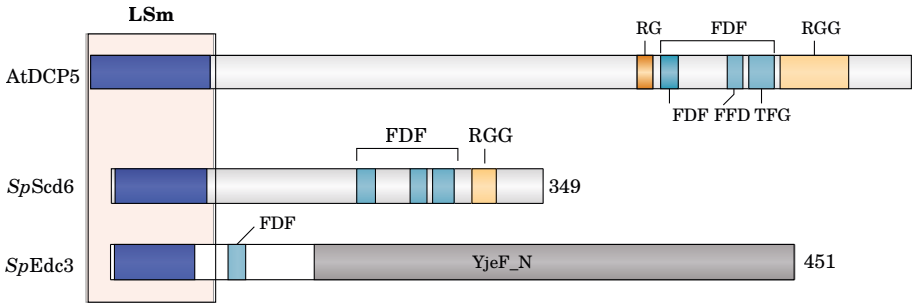
Heatmap indiquant le nombre de spectres identifiés pour les protéines significativement enrichies dans les co-IP de mycURT1. Chaque ligne correspond à une protéine différente, chaque colonne correspond à un échantillon différent analysé par spectrométrie de masse. L'expérience #1 a été réalisée sur des fleurs de la lignée T2. Les expériences #2 et #3 sur des plantules de 2 semaines cultivées *in vitro*. Les expériences 1, 2 et 3 ont été réalisées indépendamment. Les IP contrôles ont été réalisées sur des plantes sauvages n'exprimant pas mycURT1. Les numéros indiqués dans chacune des cases correspondent aux nombres de spectres identifiés pour chaque protéine. Les cases vides indiquent qu'aucun spectre de la protéine correspondante n'a été retrouvé. L'intensité de la couleur de chaque case est relative à la valeur d'enrichissement de la protéine dans les IP mycURT1 par rapport aux IP contrôles.

CAF-NOT peut également être localisé dans le noyau (Collart and Panasenko, 2012). Chez Arabidopsis aucune activité similaire n'a pour l'instant été démontrée et la localisation exclusivement cytoplasmique décrite de URT1 (Sement *et al.*, 2013) ne permet pour l'instant pas de spéculer sur une possible relation entre URT1 et HDAC1/2.
Les expériences d'immunoprécipitation résultent également en la purification de deux protéines faisant partie de la famille des facteurs de transport nucléaire NTF2, G3BP7 et G3BP8 (AT5G43960.1 et AT5G60980.2). Ces deux protéines font partie d'une famille de 8 protéines homologues à la protéine humaine G3BP1 (Ras-**G**TPase–activating protein SH**3**-omain–**B**inding **P**roteins) qui localise dans les granules de stress et contribue à leur assemblage (Abulfaraj *et al.*, 2018). Leur présence dans les IP de URT1 peut-être la conséquence de la présence sur ces deux protéines d'un domaine de liaison à l'ARN (RRM) ou de la localisation de URT1 dans les stress granules (Sement *et al.*, 2013).

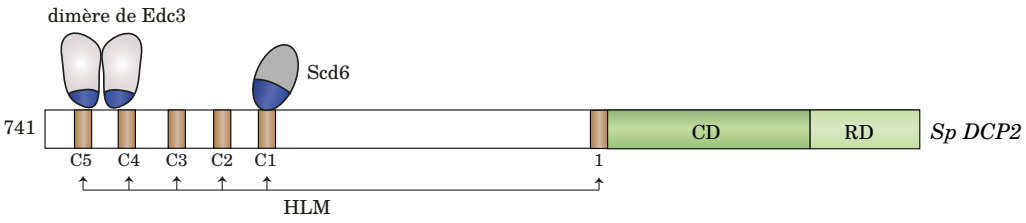
Nous retrouvons dans ces expériences, 5 protéines impliquées dans le métabolisme des ARN, parmi ces dernières, la protéine homologue à LSM14/Scd6, **DCP5** (ATG1G26110), deux homologues à DDX6/Dhh1, **RH8** (AT4G00660) et **RH12** (AT3G61240), le facteur d'initiation de la traduction **eIF4G** (AT3G60240) et une protéine de liaison aux queues poly(A), **PABP8** (AT1G49760).
eIF4G est un des composants du complexe eIF4F qui comportent également eIF4E et eIF4A. En interagissant avec eIF4E et les PAPB, eIF4G est impliqué dans la circularisation et l'initiation de la traduction des ARNm. PAB8 est l'une des 8 PABP cytoplasmiques codées par le génome d'Arabidopsis. Parmi celles-ci, PAB2, PAB4 et PAB8 sont exprimées dans les tissus végétatifs. Chez l'homme et la levure les homologues à DCP5 (LSM14/Scd6) et aux protéines RH8 et 12 (DDX6/Dhh1) sont des cofacteurs de décoiffage également impliqués dans la répression de la traduction et localisés dans les P-bodies. En outre, DDX6/Me31b interagît avec le domaine FDF de LSM14/Tral via son domaine RecA C-terminal chez l'homme et la drosophile (Tritschler *et al.*, 2008; Brandmann *et al.*, 2018). Les domaines impliqués dans cette interaction sont conservés chez Arabidopsis, ce qui suggère que cette interaction pourrait également être conservée chez la plante et justifierait le fait de trouver à la fois DCP5 et deux protéines RH dans nos expériences de co-immunoprécipitation.

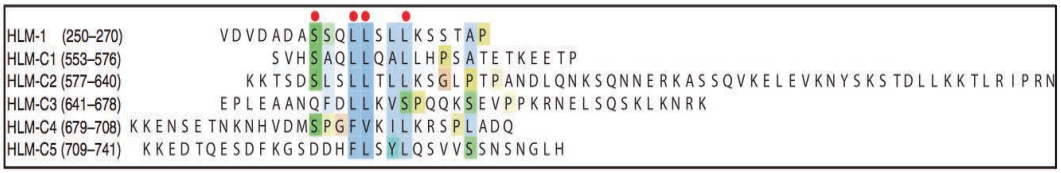
La co-immunoprécipitation de DCP5 avec myc-URT1 a tout particulièrement retenu notre attention du fait de la présence, dans cette protéine, d'un domaine N-terminal LSm (Fig. 20A). Chez *S. pombe*, les domaines LSm de Edc3 et de Scd6 interagissent avec des motifs en hélice α riche en leucine (HLM) présents sur la longue région C-terminale de Dcp2 (Fig. 20B et C) (Fromm *et al.*, 2012). Une interaction du même type a également été démontrée chez la drosophile entre Tral (Scd6) et un motif HLM présent sur DCP1 (Fromm *et al.*, 2012). Ces motifs HLM ne sont pas sans rappeler le motif M1 extrêmement bien conservé et retrouvé dans la région N-terminale intrinsèquement désorganisée de URT1. En effet le motif M1 de URT1 est riche en leucines (voir

A



B





C

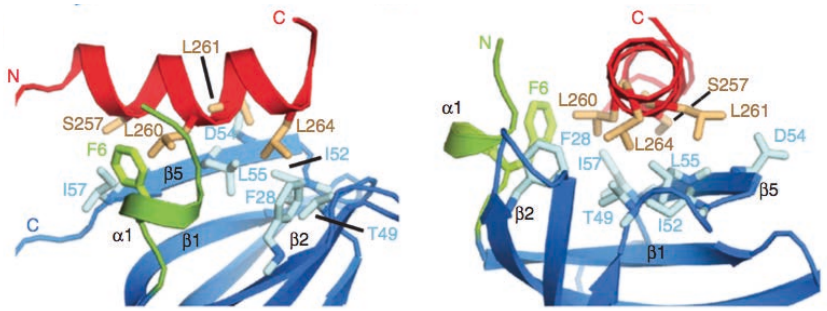


Figure 20 | Interaction entre les domaines LSm de Scd6 et Edc3 et le motif HLM de Dcp2 chez la levure

A | Organisation en domain de DCP5 comparée à celle Scd6 et à Edc3. Les 3 protéines possèdent un domaine LSm. At, *Arabidopsis thaliana; Sp Schizosaccharmyces pombe* **B |** Chez *S. pombe* les domaines LSm de Edc3 et Scd6 interargissent avec des motifs HLM (Helical Leucine-rich) de consensus Sxx**LL**xx**LL** présents sur la région C-terminale de Dcp2 (motifs HLM). CD, domaine catalytique; RD domaine régulateur de Dcp2 **C |** Structure résolue du complexe d'interaction entre le motif HLM de Dcp2 (rouge) et le domaine LSM de Edc3 (bleu et vert) chez *S. pombe*. Figure adaptée de Fromm et al. 2012.

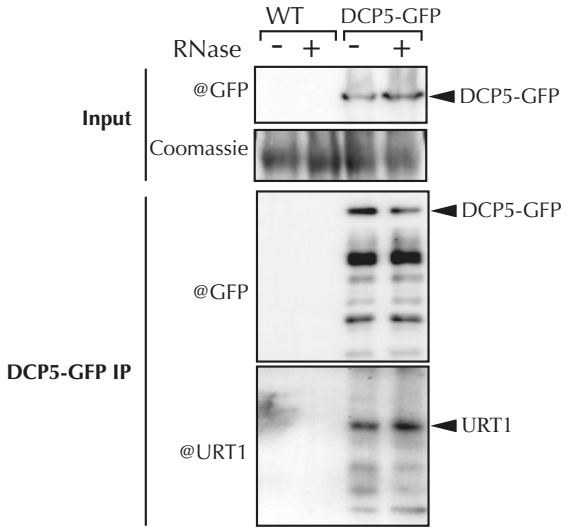


Figure 21 | URT1 co-immunoprécipite avec DCP5-GFP dans des expériences de co-IP réciproques
Expériences de co-IP de DCP5-GFP à partir de plantes de la lignée *dcp5-1* complémentée avec DCP5-GFP. Les expériences ont été réalisées en présence ou en absence de RNase A. URT1 est detectée dans les fractions d'élutions des co-IP de DCP5-GFP grâce à un anticorps dirigé contre la protéine endogène. DCP5-GFP est détectée grâce à un anticorps dirigé contre l'épitope GFP.

Fig. 16 et 17) et prédit comme adoptant une conformation en structure α par le prédicteur de structure secondaire PSIPRED.

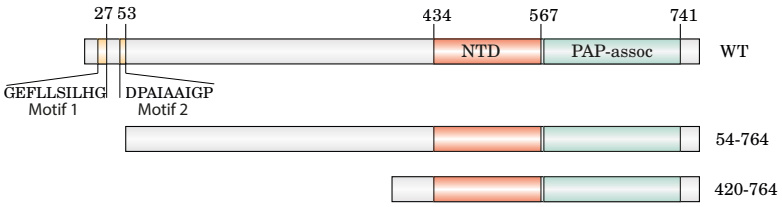
Notre hypothèse est que URT1 pourrait être engagée dans une interaction directe avec DCP5 et que cette interaction pourrait être permise par l'association du motif M1 conservé de URT1 avec le domaine LSm de DCP5 à l'instar de ce qui observée chez *S. pombe* entre Edc3 et Scd6 avec Dcp2 (Fig. 20B et C)

3. URT1 est engagée dans une interaction directe avec l'homologue de LSM14/Scd6, DCP5

Afin de confirmer que DCP5 est un bon candidat pour une interaction directe avec URT1, nous avons dans un premier temps validé la co-immunoprécipitation de DCP5 avec URT1 en réalisant une expérience de co-IP réciproque.
Pour ce faire, nous avons produit des lignées permettant l'expression de DCP5 sous la dépendance d'un promoteur ubiquitine et marqué avec un épitope GFP en C-terminal. Ne disposant pas d'anticorps dirigés contre DCP5, nous ne connaissons pas le niveau d'expression de DCP5-GFP par rapport au niveau d'expression endogène de DCP5. Des anticorps récemment obtenus par l'équipe de Cécile Bousquet-Antonelli à Perpignan pourront être utilisés afin de déterminer le niveau d'expression de DCP5-GFP dans ces lignées.
Les résultats présentés en Figure 21 montre que URT1 co-immunoprécipite avec DCP5 lors d'une expérience de co-IP réverse menée avec DCP5-GFP. De surcroit cette interaction est indépendante de la présence d'ARN puisque URT1 est retrouvé dans ces mêmes expériences de co-immunoprécipitation réalisées en présence de RNase A, une endoribonucléase qui clive les ARN au niveau des pyrimidines (uracile et cytosine). Du fait de la complexité du réseau d'interaction au sein des particules ribonucléoprotéiques, contenant DCP5, ces expériences n'excluent pas complètement la possibilité d'une interaction indirecte entre les deux protéines.

Dans l'optique de déterminer si URT1 et DCP5 sont engagées dans une interaction directe et si la région N-terminale et les motifs M1 et M2 sont impliqués dans cette interaction, nous avons adopté une stratégie de co-purification *in vitro* (*in vitro pull down*). Pour ce faire nous avons produit en système bactérien des protéines recombinantes correspondant à URT1 ou différentes versions tronquées de la protéine et DCP5. Les constructions URT1 ont été fusionnées à un épitope N-terminal 6xhis-GST et DCP5 à un épitope N-terminal 6xhisMBP.
Les expériences de co-purification *in vitro* montrent que hisGST-URT1 co-purifie avec hisMBP-DCP5 *in vitro* et que cette association est fortement diminuée lors de l'utilisation des versions de URT1 dépourvus de la région N-terminale (420-764) ou des motifs M1 et M2 (54-764) (Figure 22).

A



B

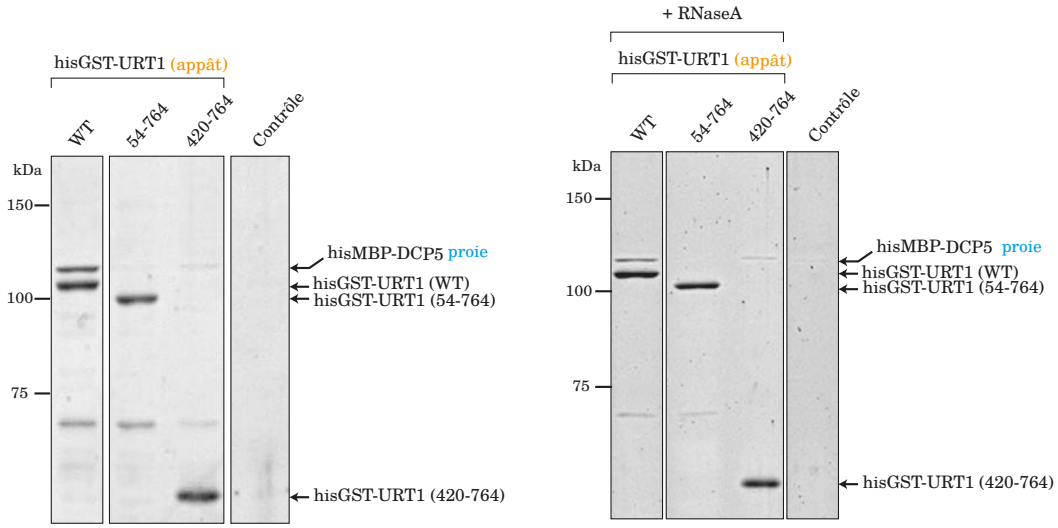


Figure 22 | URT1 interagit avec DCP5 et cette interaction est favorisée par la présence de la région N-terminale de URT1

A | Représentation schématique de la composition en motifs et domaines de URT1 et des constructions utilisées pour l'expérience de co-purification *in vitro* **B |** Analyse par SDS-PAGE coloré au SYPRO ruby des échantillons correspondant aux expériences de co-immunoprécipitation réalisées avec hisGST-URT1, hisGST-URT1 54-764 ou hisGST-URT1 420-764 immobilisées sur une résine glutathion et incubées avec hisMBP-DCP5.

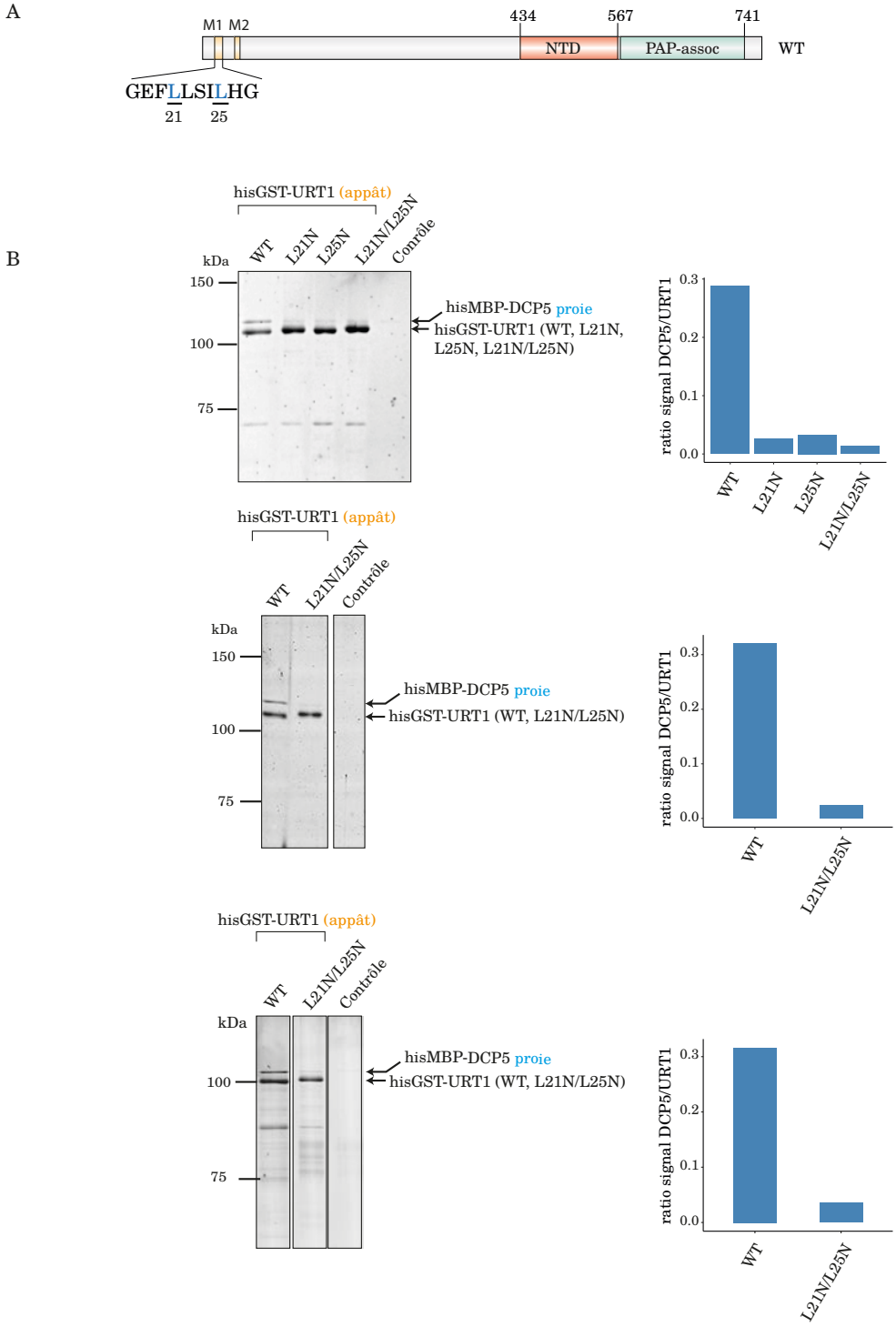


Figure 23 | Les leucines du motif M1 de URT1 sont impliquées dans l'interaction avec DCP5
A| Représentation schématique des leucines du motifs M1 de URT1 **B|** A gauche, analyse par SDS-PAGE coloré au SYPRO ruby des échantillons des expériences de co-immunoprécipitation, réalisées avec hisGST-URT1, hisGST-URT1 L21N, hisGST-URT1 L25N ou hisGST-URT1 L21N/L25N immobilisées sur une résine glutathion et incubées avec hisMBP-DCP5. A droite, valeur du ratio du signal correspondant à hisMBP-DCP5 par rapport au signal correspondant à hisGST-URT1 L21N, L25N ou L21N/L25N. L'ensemble des expériences ont été réalisées en présence de RNase A. Les contrôles correspondent à l'incubation de la résine glutathion avec hisMBP-DCP5.

Les résultats obtenus en présence de RNase A sont identiques aux résultats précédents et excluent la possibilité que l'interaction observée entre URT1 et DCP5 soit indirectement liée à la présence d'ARN dans nos échantillons (Figure 22).
Nous notons tout de même la présence d'un faible signal correspondant à DCP5 dans l'expérience de pull down réalisé avec hisGST-URT1 420-764. Ce signal, à la limite du seuil de détection, pourrait être due à une liaison aspécifique de hisMBP-DCP5 à la résine, à l'instar de ce qui est observé dans le contrôle de l'expérience en présence de RNase. Cela pourrait également suggérer que l'association de URT1 avec DCP5 soit également permise par des interactions faibles avec des acides aminés contenus dans le domaine catalytique de URT1.

Dans la mesure où le motif M1 de URT1 ressemble de très près aux motifs HLM de Dcp2 chez la levure et que ces motifs permettent l'interaction avec le domaine LSm de Scd6 nous avons, dans un second temps, testé si les leucines présentes dans le motif M1 étaient importantes pour l'interaction de URT1 avec DCP5. Dans cette optique nous avons réalisé différentes mutagenèses conduisant au remplacement des leucines (L) du motif M1 par des asparagines (N). Les résultats des expériences de co-purifications *in vitro* montrent que la mutation individuelle de la leucine 21 ou de la leucine 25 suffisent à réduire de manière considérable l'interaction observée entre URT1 et DCP5. En outre, les substitutions en tandem des deux leucines confortent les résultats observés puisqu'on observe une forte diminution du signal correspondant à hisMBP-DCP5 lorsque le pull-down est réalisé avec hisGST-URT1 L21N/L25N (Figure 23).

De manière générale, l'ensemble de ces données confirme l'association directe de URT1 avec DCP5 et valide la présence de DCP5 dans nos expériences de co-immunoprécipitation de myc-URT1 *in vivo*. Elles nous confortent, également, dans l'idée que la région N-terminale de URT1, et plus spécifiquement le motif M1, un motif riche en leucines conservé et prédit comme ayant une structure en hélice α est impliqué dans l'interaction de URT1 avec DCP5. La présence de signaux résiduels correspondant à hisMBP-DCP5 lors d'expérience de c-immunoprécipitation avec hisGST-URT1 (420-764) et hisGST-URT1 L21N/L25N semble tout de même indiquer que la région C-terminale de URT1 pourrait contribuer en partie à cette interaction.

4. L'interactome de DCP5

Les interactants directs et indirects de DCP5 sont peu connus chez Arabidopsis. Afin de mieux caractériser le réseau d'interactions impliquant DCP5, nous avons réalisé des expériences de co-IP de DCP5-GFP couplées à des analyses par spectrométrie de masse. Ces expériences comprennent 3 co-IP réalisées à partir de lignées exprimant DCP5-GFP dans un fond génétique dérégulé pour *dcp5* (knock-down) (la lignée *dcp5* est celle utilisée dans (Xu and Chua, 2009) et une expérience réalisée à partir de lignées exprimant DCP5-GFP dans un fond génétique sauvage (Figure 24A). Les résultats de ces expériences étant similaires, ils sont représentés ensemble sous la forme d'un

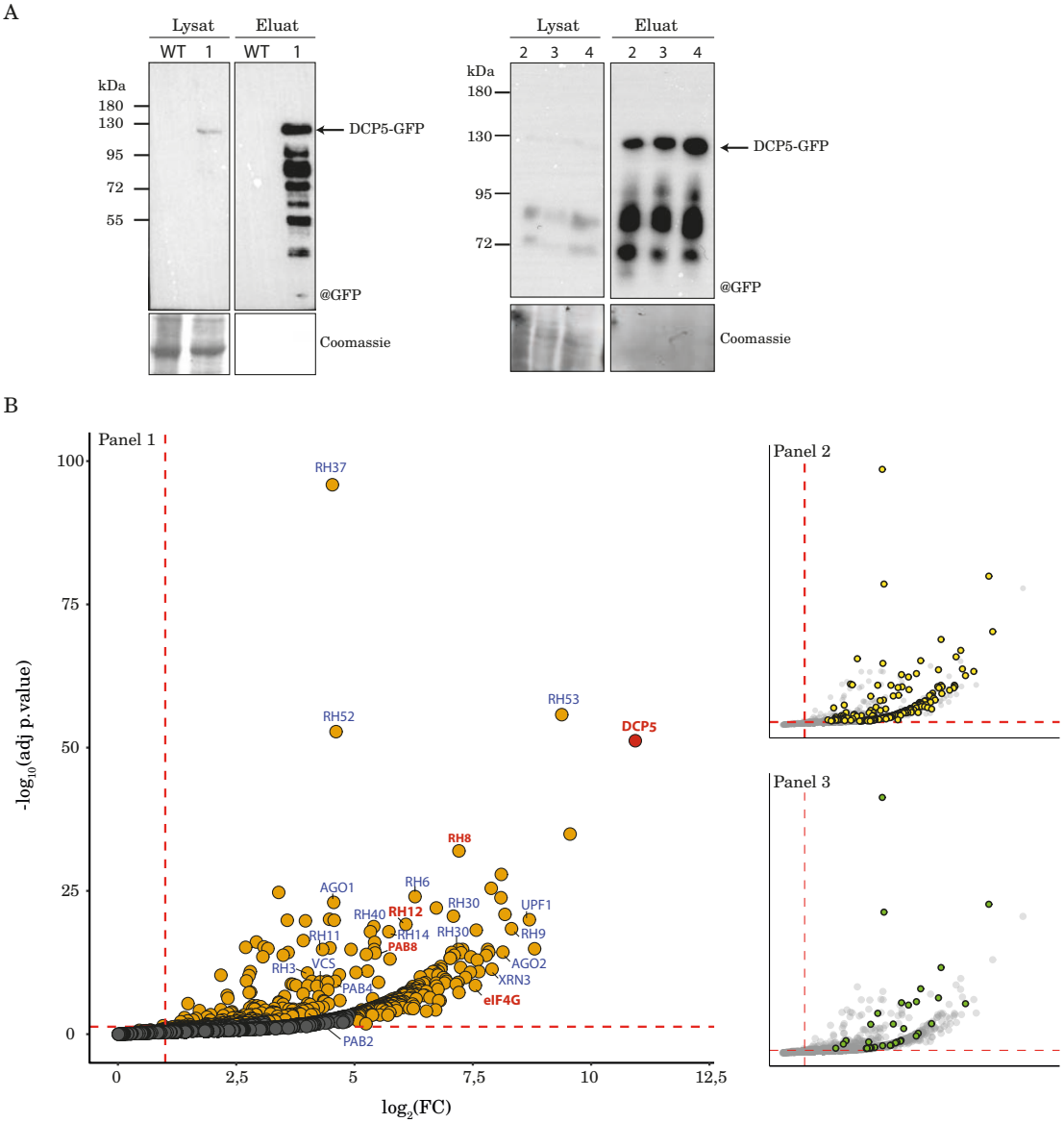


Figure 24 | Analyses des partenaires de DCP5 par co-immunoprécipitation de DCP5-GFP et spectrométrie de masse

A| Analyse par western blot avec un anticorps dirigé contre la GFP de DCP5-GFP après expériences de co-immunoprécipitation en utilisant des lignées exprimant DCP5-GFP dans un fond sauvage (panel de gauche) ou dans un fond *dcp5-1* (panel de droite) **B|** Les 4 fractions d'élutions et 4 contrôles (co-immunoprécipitation réalisées dans un fond WT n'exprimant pas le transgène DCP5-GFP) ont été analysées par spectrométrie de masse. Le volcano plot indique les protéines enrichies dans l'IP DCP5-GFP par rapport aux IP contrôles. L'abscisse donne la valeur d'enrichissement par rapport aux IP contrôles (FC (Fold Change)) et l'ordonnée donne la p-value ajustée. La droite rouge horizontale représente une p-value = 0.05 et la droite verticale un FC = 2. Les points représentés en orange sont ceux pour lesquels le nombre moyen de spectres est supérieur à 5. Les deux panels de gauche représentent l'enrichissement en protéines classées en tant que protéines de liaison à l'ARN (GO : GO:0003723) (en jaune, panel 2), ou l'enrichissement en hélicases à ARN (GO:0004004) (en vert, panel 3).

volcano plot en Figure 24B. DCP5 co-purifie avec un très large panel de protéines. En effet, 796 protéines sont retrouvées enrichies (FC >2 ; p-value < 0,05) dans les co-IP de DCP5-GFP par rapport aux IP contrôles (listées en Annexe 5), sur ces 793 protéines 438 ont été identifiées avec une moyenne de spectres (sur les 4 IP) supérieure à 5 (représentées en orange sur le volcano plot, panel 1). Une analyse de Gene Ontology révèle qu'un large panel de protéines fortement enrichies dans les IP DCP5 correspondent à des protéines de liaison à l'ARN (GO : GO:0003723) dont un nombre considérable sont des hélicases (Figure 24B, panel 2 et 3). Parmi ces hélicases nous notons la présence de 3 protéines homologues à DDX3. DDX3 est une protéine multifonctionnelle impliquée dans la transcription, l'épissage, l'export des ARNm vers le noyau, la biogenèse des ribosomes et la dégradation des ARN. La protéine est également impliquée dans l'initiation de la traduction de par son interaction avec le ribosome 80S et les facteurs d'initiations de la traduction (Lee *et al.*, 2008; Geissler *et al.*, 2012) et s'associe aux granules de stress (Lai *et al.*, 2008; Shih *et al.*, 2012). DCP5 co-purifie également avec des facteurs de silençage AGO1 et AGO2 et des facteurs impliqués dans la dégradation des ARN tels que DCP1, VCS, XRN3, XRN4 et le facteur du NMD, UPF1. Le réseau d'interaction commun de DCP5 et de UPF1 est détaillé d'une publication dans Plant Journal, publication que je signe en deuxième auteur (Annexe 6). Je ne présente à la suite que les résultats en relation directe avec l'étude de URT1.

De manière surprenante, URT1 n'est pas retrouvée significativement enrichie par l'analyse statistique des résultats de spectrométrie de masse des expériences de co-IP de DCP5, la protéine n'est en effet que très faiblement détectée dans une seule des 4 fractions d'élutions analysées. Il est difficile de conclure sur un résultat négatif dans ce type d'expériences mais l'absence de spectres de masse correspondant à URT1 peut être la résultante d'au moins trois causes, d'origine technique ou biologique :

(i) De manière formelle et puisque URT1 n'est pas détectée par spectrométrie de masse dans les co-IP de DCP5, nous ne pouvons occulter l'hypothèse que URT1 ne soit pas un interactant de DCP5. Cependant, nous excluons cette possibilité dans la mesure où une interaction directe entre ces deux protéines a été confirmée *in vitro*. De plus, URT1 est détectée dans les co-IP de DCP5-GFP en utilisant des anticorps dirigés contre URT1 (Fig. 21). La non-détection de URT1 par spectrométrie de masse dans ces fractions pourrait donc être due à une cause technique liée à la limite de détection de notre protocole actuel. Pour rappel, la lignée transgénique T1 exprimant faiblement le transgène myc-URT1 par rapport à la lignée T2 n'a pas pu être utilisée pour identifier les partenaires d'URT1 par spectrométrie de masse en raison de sa faible détection en tant qu'appât (voir § 1.2.1). A fortiori, URT1 semble difficilement détectable en tant que proie.

(ii) L'interaction entre URT1 et DCP5 pourrait être transitoire et, de ce fait, difficilement identifiable**.** Pour des raisons techniques notamment dues à la présence de larges régions intrinsèquement désorganisées et à la difficulté de purifier de manière intègre les protéines recombinantes URT1 et DCP5, nous n'avons pu déterminer les constantes d'affinité de la formation du complexe URT1/DCP5. Cependant, nous pensons que cette interaction, qui implique au moins en partie un motif linéaire court dans la région N-terminale de URT1, pourrait être une interaction

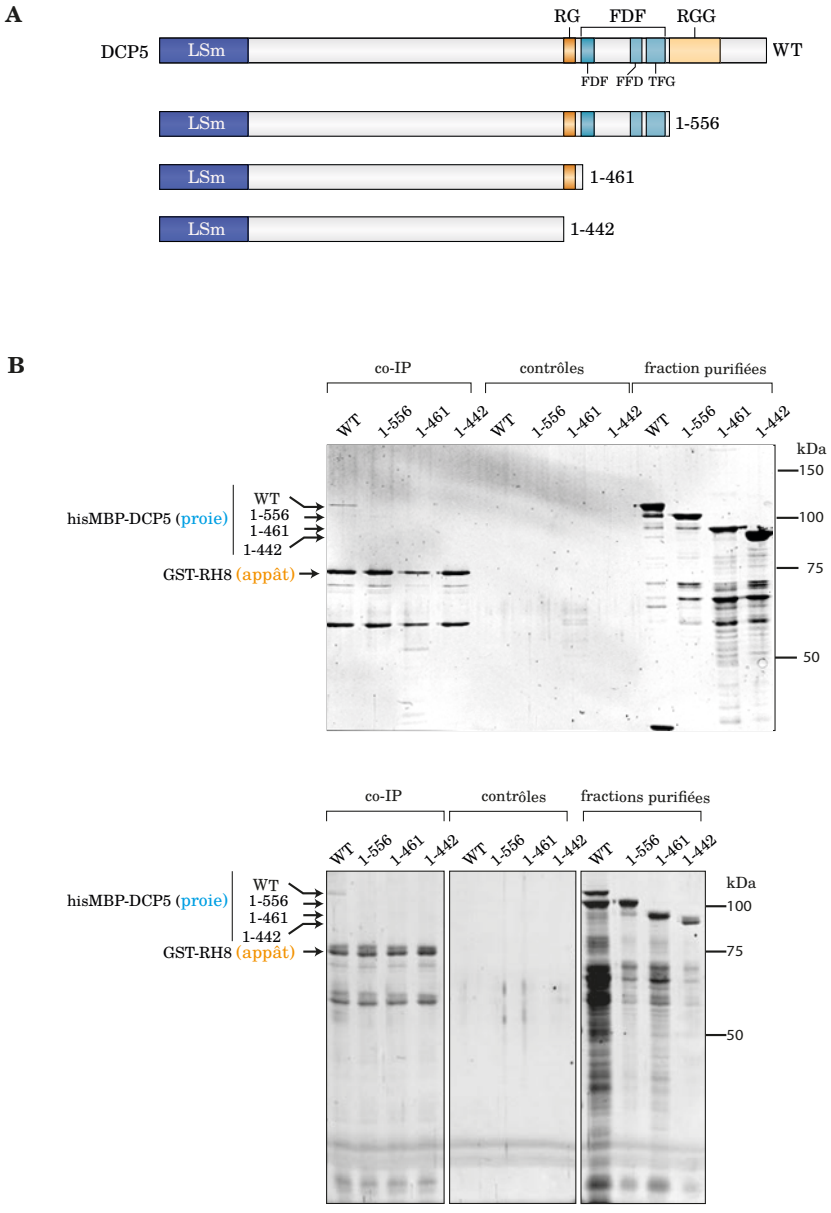


Figure 25 | RH8 interagit avec DCP5 et cette interaction implique le motif RGG de DCP5.
A| représentation schématique de la composition en motifs et domaines de DCP5 et des constructions utilisées pour les expériences de co-immunoprécipitation **B|** Analyse SDS-PAGE coloré au SYPRO ruby de deux expériences de co-immunoprécipitation indépendantes réalisées avec GST-DLH2 immobilisée sur une résine glutathion et incubée avec hisMBP-DCP5 WT, hisMBP-DCP5 1-556, hisMBP-DCP5 1-461 ou hisMBP-DCP5 1-442.

transitoire difficilement détectable *in vivo* en utilisant DCP5 comme proie. Cette interaction a tout de même pu être détectée *in vivo*, mais, à cette fin, nous avons utilisé des lignées de surexpression de mycURT1.

(iii) Enfin, le fait que DCP5 co-purifie avec un large nombre de protéines peut occulter la détection d'interactants potentiellement transitoires (ou mineurs d'un point de vue quantitatif) tel que URT1. Un protocole de co-immunoprécipitation incluant une étape de pontage chimique entre protéines (cross-link) est en cours d'implémentation en collaboration avec la plateforme de protéomique associée à notre institut. Ce protocole pourra très prochainement contribuer à augmenter la sensibilité de nos expériences de co-IP et pourrait permettre la détection de URT1 dans des co-IP de DCP5.

En attendant cette amélioration technique, les résultats actuels du réseau d'interaction de DCP5 révèlent des informations particulièrement intéressantes dans le contexte de l'étude de URT1. Premièrement, les résultats des IP DCP5 révèlent un fort enrichissement des 3 homologues à l'hélicase et répresseur de la traduction DDX6, RH6, RH8 et RH12. Les hélicases à ARN RH8 et RH12 font partie des quelques protéines enrichies de manière significative dans les co-IP URT1 (Fig. 18). Cet enrichissement pourrait être dû non pas à une interaction directe entre RH8/12 et URT1, mais plutôt résulter d'une interaction indirecte URT1-DCP5-RH8/12. En effet les homologues respectifs de DCP5 et RH6/8/12 chez l'homme, LSM14A et DDX6, interagissent. Pour tester une éventuelle conservation de cette interaction chez Arabidopsis, nous avons réalisé des expériences de co-immunoprécipitation *in vitro* menée avec une version de RH8 fusionnée à un épitope GST en N-terminal et des protéines de fusion MBP-DCP5 correspondant soit à la protéine DCP5 entière soit à des versions tronquées chez lesquelles les régions RGG, FDF et RG situées en C-terminal sont éliminées (Fig. 25A). Ces expériences révèlent une interaction directe entre RH8 et DCP5, étayant l'idée d'une conservation de cette interaction chez Arabidopsis. En revanche, la région C-terminale contenant les motifs RGG semble requise pour l'interaction de RH8 avec DCP5 (Fig. 25B). Cette observation doit être confirmée mais en l'état, elle diffère de ce qui est observé chez l'homme et la drosophile chez lesquels l'interaction de LSM14A/Tral avec DDX6/Me31b est permise par son domaine FDF (Tritschler *et al.*, 2009; Fromm *et al.*, 2012).

Une deuxième information intéressante découlant de l'analyse des expériences de co-IP de DCP5 (Fig. 24) est l'enrichissement significatif de eIF4G et PAB8 dans les IP DCP5. Ces protéines font partie des interactants également identifiés dans les co-IP de URT1. Les domaines RGG ont été impliqués dans plusieurs interactions, et notamment chez la levure avec eIF4G. Bien que cela reste à être expérimentalement testé chez Arabidopsis, une interaction entre la région C-terminale de DCP5 avec eIF4G est possible. L'interaction entre eIF4G et les PABP est quant à elle conservée chez les plantes (Cheng and Gallie, 2007). Il est également possible que la co-purification de PABP avec URT1 et DCP5 soit due à une liaison commune aux ARNm, plutôt qu'à un réseau d'interaction protéine-protéine.

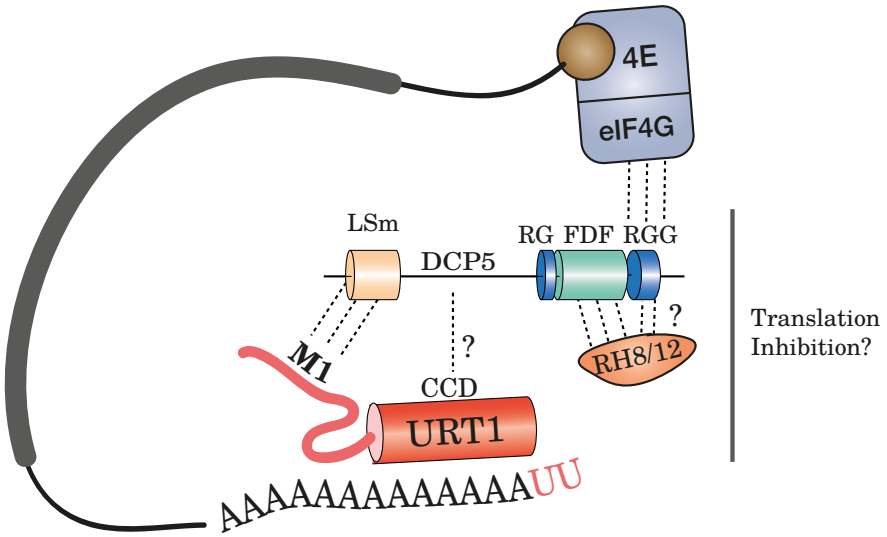


Figure 26 | Schéma récapitulant le réseau d'interaction de URT1 et le lien possible entre uridylation des ARNm et inhibition de la traduction.

5. Modèle du réseau d'interactions impliquant URT1 et DCP5

L'ensemble des résultats actuels des expériences de co-précipitations *in vitro* et *in vivo* suggèrent un réseau d'interactions impliquant notamment URT1, DCP5 et RH8/RH12. Ce réseau d'interactions est schématisé en Figure 26. Le motif M1 de URT1 interagit avec le motif LSm de DCP5. Le domaine catalytique (CCD) pourrait éventuellement contribuer à l'interaction avec DCP5. A l'instar de ses homologues chez d'autres eucaryotes, DCP5 recrute les hélicases ARN homologues à DDX6/Dhh1, et notamment RH8 et RH12. Des expériences préliminaires indiquent l'importance du domaine RGG de DCP5 dans cette interaction avec RH8. Ce résultat doit être confirmé et n'exclut pas l'implication du domaine FDF dans cette interaction. La co-précipitation de eIF4G dans les IPs URT1 et DCP5 est probablement expliquée par une interaction conservée entre DCP5 et eIF4G via le domaine RGG de DCP5. Un réseau d'interactions connectant la TUTase URT1 à des facteurs, tels que DCP5 et RH8/RH12 qui sont connus pour être impliqués dans l'inhibition de la traduction, ouvre des perspectives très intéressantes et soulève notamment des interrogations quant à un lien possible entre l'uridylation des ARNm et l'inhibition de la traduction.

Quatrième partie :

L'uridylation modifie les profils de polyadénylation d'un ARNm rapporteur et inhibe sa traduction

Le réseau d'interaction de URT1 est compatible avec un rôle de URT1 dans l'inhibition de la traduction et/ou dans la déstabilisation des ARNm via l'activation de l'élimination de la coiffe. Il convient donc de tester expérimentalement un impact éventuel de URT1 dans la traduction et/ou la stabilité d'ARN messagers. Deux stratégies sont développées au laboratoire : (i) une première approche pour détecter l'impact de l'uridylation sur la stabilité et la traduction des ARNm, et tester un éventuel rôle de l'interaction URT1-DCP5 dans ces processus dans un contexte biologique naturel, c'est à dire chez Arabidopsis. Cette approche a été initiée mais reste à un stade d'avancement préliminaire et ne sera pas détaillée dans ce manuscrit. J'ai participé à la production d'une partie du matériel biologique nécessaire à ces études en cours et qui incluent :

- la lignée double mutante *urt1-1 dcp5-1*
- des lignées *urt1-1* complémentées par expression de URT1 étiquetée en position C-terminale (URT1myc), et des versions tronquées ou mutées de la protéine de fusion. Les versions mutées comprennent des versions de URT1myc avec des mutations dans le domaine catalytique (CCD) qui abolissent ou réduisent l'activité de la protéine (URT1myc D491A/D493A, URT1myc H717L, URT1myc P618L) ou des mutations dans les motifs conservées M1 et M2 de la région N-terminale (URT1myc L21/25N, URT1myc A49/50K, URT1 L21/25N & A49/50K). J'ai également produit des lignées complémentées avec des versions de URT1 ne comprenant que le domaine catalytique sous sa forme active (URT1myc CCD) ou inactive (URT1myc CCD D491A/D493A)
- Des lignées *dcp5-1* complémentées par expression de DCP5 étiquetée avec myc ou la GFP en position C-ter ou N-ter. J'ai également produit des lignées complémentées qui expriment des versions de DCP5 tronquées pour le domaine RGG, RGG et FDF ou pour les trois domaines RGG, FDF, et RG ou pour le domaine LSm (pour une visualisation des domaines de DCP5, voir Fig. 14)

Une partie de ce matériel est utilisée dans le cadre d'une collaboration en cours avec l'équipe dirigée par Cécile Bousquet-Antonelli afin d'analyser les populations de transcrits associées aux monosomes et polysomes dans les mutants *urt1* et *dcp5*. De plus, j'ai obtenu des données transcriptomiques par analyses RNA-seq pour deux réplicas biologiques de plantes Col0, *urt1*-1, *dcp5-1* ainsi que le double mutant *urt1-1 dcp5-1*. Ces données devront être croisées avec des données futures qui seront obtenues par TAIL-seq, une fois l'optimisation du protocole validée afin d'effectuer une analyse globale et gène-à gène chez Arabidopsis. En effet, l'identification d'une interaction directe entre URT1 et DCP5 justifie de tester un éventuel impact sur le niveau d'uridylation des ARNm en absence de DCP5. Deux scénarios sont à considérer : (i) URT1 pourrait recruter DCP5 et indirectement RH8 et RH12 pour favoriser l'inhibition de la traduction ou la dégradation des ARNm uridylés ou (ii) à l'inverse, URT1 pourrait être recrutée sur des ARNm ciblés par DCP5. Dans l'hypothèse où la protéine DCP5 serait requise pour recruter URT1 sur les ARNm nous devrions observer une diminution de l'uridylation dans les mutant *dcp5*. Afin de réaliser un premier test de cette hypothèse, les extrémités 3' de trois transcrits *BAM3*, *AT1G24160* et *CAF1A* ont été analysées par 3'RACE-seq à partir d'ARN de fleurs récoltées sur des plantes WT,

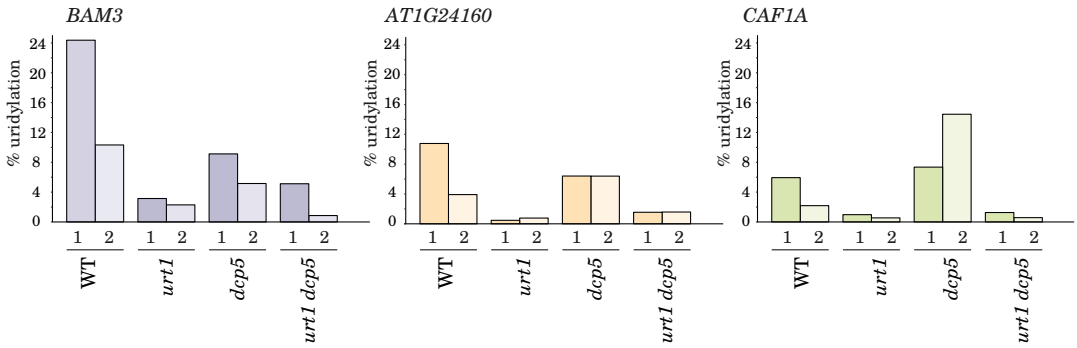


Figure 27 | L'uridylation par URT1 n'est pas abolie dans le mutant *dcp5-1* pour 3 ARN candidats
Diagrammes en barre représentant le pourcentage de transcrits *BAM* (pourpre), *AT1G24160* (orange) et *CAF1A* (vert) uridylés déterminés par 3'RACE-seq. Les expériences ont été réalisées à partir d'ARN de fleurs de plantes WT, *urt1*, *dcp5* et *urt1 dcp5* et pour deux réplicas biologiques.

urt1-1 et *dcp5-1*. J'ai également inclus dans cette analyse des échantillons provenant de plantes doubles mutantes *urt1 dcp5*, afin de tester un éventuel impact de l'absence de DCP5 sur l'uridylation des ARNm par une TUTase autre que URT1. Le taux d'uridylation dans Col0, *urt1*, *dcp5* et le double mutant *urt1 dcp5* pour les transcrits *BAM, AT1G24160 et CAF1A* est représenté sous la forme d'histogramme en Figure 27. Les résultats obtenus sont relativement variables. Cependant, cette analyse révèle clairement que l'uridylation des ARNm n'est pas abolie dans le mutant *dcp5-1*. Il est à noter que le mutant *dcp5-1* est un mutant hypomorphe, la mutation nulle étant létale (Xu and Chua, 2009). Nous ne pouvons donc pas conclure que les ARNm sont toujours uridylés en absence totale de DCP5, mais seulement que la mutation hypomorphe *dcp5-1* n'abolit pas l'uridylation de ces 3 ARNm testés. Ces ARNm n'ont pas été sélectionnés en raison d'un éventuel rôle de DCP5 dans leur métabolisme : Les transcrits *BAM3* et *AT1G24160* sont couramment analysés en raison de la détection aisée de l'uridylation de ces ARNm. Les ARNm *CAF1A* ont quant à eux été inclus dans l'analyse car ils faisaient partie d'une étude indépendante menée en parallèle au laboratoire. Le but était donc de tester si DCP5 avait un impact mesurable sur l'uridylation des ARNm quels qu'ils soient. Cependant, les résultats présentés en Figure 27 suggèrent un impact différentiel de la mutation *dcp5-1* sur l'uridylation de ces 3 ARNm. Idéalement, il conviendra donc de déterminer le taux d'uridylation des ARNm à une échelle globale et par une méthode permettant une analyse gène-à-gène, une fois l'optimisation de la méthode TAIL-seq validée pour son utilisation chez Arabidopsis.

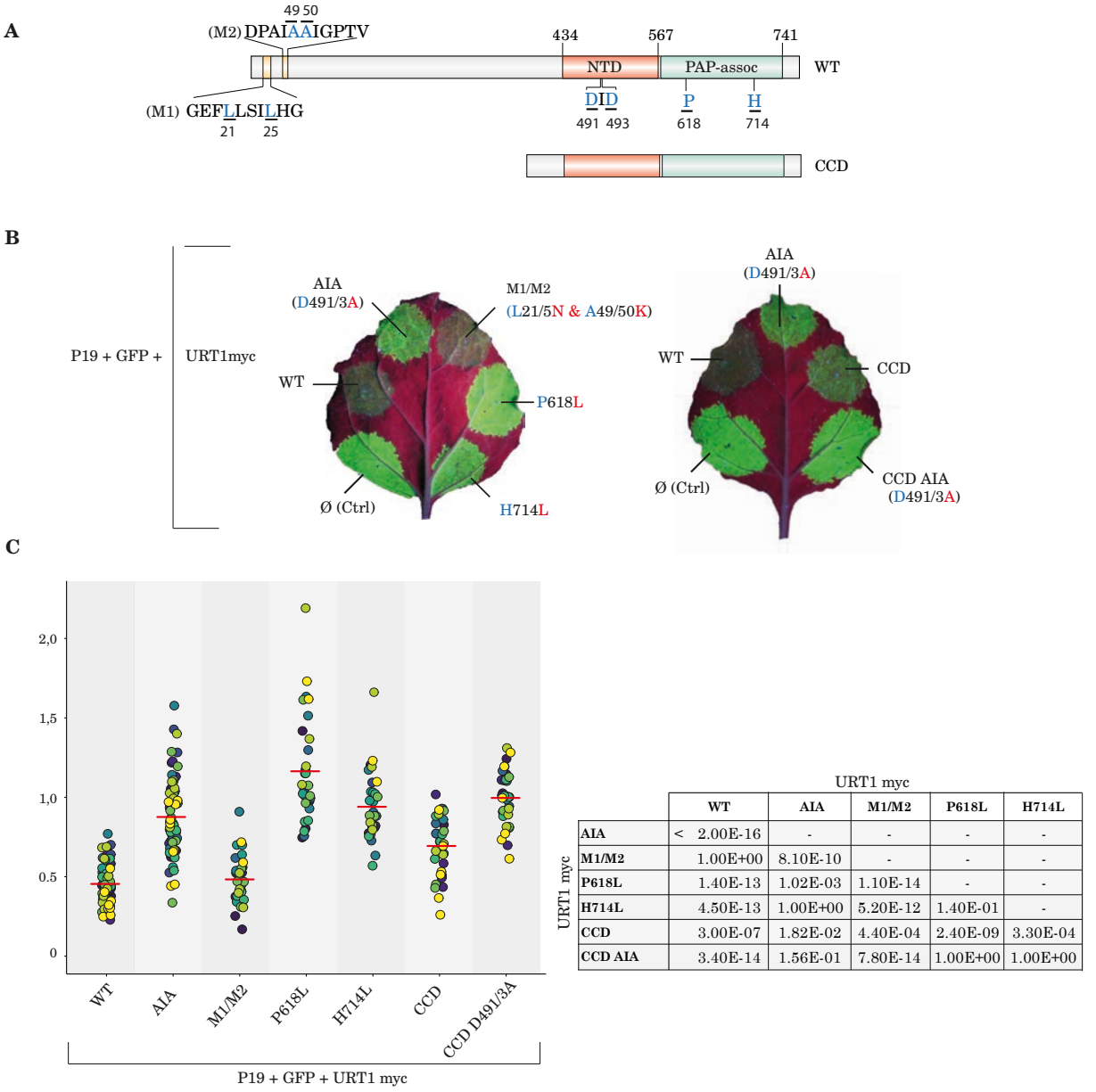
(ii) La deuxième approche que j'ai plus particulièrement utilisée pour analyser les rôles de l'uridylation par URT1 dans le métabolisme des ARNm a été l'utilisation d'un système d'expression transitoire en feuilles de *Nicotiana benthamiana* pour tester l'impact de l'expression de URT1 sur la production d'un rapporteur GFP. Ces résultats contribuent à une meilleure compréhension des rôles multiples jouées par l'uridylation dans le métabolisme des ARNm et sont détaillés à la suite.

1. La surexpression de URT1 inhibe la traduction d'un reporter GFP

1.1 La surexpression transitoire de URT1 conduit à une diminution de la production de GFP

Le système d'expression en feuilles de tabac *Nicotiana benthamiana* est basé sur l'agroinfiltration de cultures d'*Agrobactérium tumefaciens* préalablement transformées avec des plasmides permettant l'expression transitoire de gènes d'intérêt. Plusieurs constructions peuvent être infiltrées en parallèle, nous avons profité de ce système pour co-exprimer URT1myc ou différentes versions mutantes et tronquées de la protéine avec un rapporteur GFP et ce, en présence de l'inhibiteur de silençage P19.

Le rapporteur GFP utilisé est produit à partir d'un plasmide contenant la séquence codante de la **GFP5**, un variant de la GFP visible sous lumière bleue ou lumière UV, associée à une 5'UTR contenant un **intron du gène codant pour la chalcone synthase** et deux répétitions en tandem



URT1 myc

URT1 myc	WT	AIA	M1/M2	P618L	H714L
AIA	< 2.00E-16	-	-	-	-
M1/M2	1.00E+00	8.10E-10	-	-	-
P618L	1.40E-13	1.02E-03	1.10E-14	-	-
H714L	4.50E-13	1.00E+00	5.20E-12	1.40E-01	-
CCD	3.00E-07	1.82E-02	4.40E-04	2.40E-09	3.30E-04
CCD AIA	3.40E-14	1.56E-01	7.80E-14	1.00E+00	1.00E+00

Figure 28 | La co-expression de URT1myc entraine une diminution de la fluorescence de la protéine GFP

A| Représentation schématique de URT1 et des acides aminés ciblés pour réaliser les différentes constructions mutantes de URT1myc utilisées en partie B **B| Panel du haut,** expérience de co-infiltration en feuilles de tabac du rapporteur GFP et de différentes constructions de URT1 avec 4 épitopes myc en C-ter (URT1myc): URT1myc WT, AIA (D491A / D493A), M1/M2 (L21/5N & A49/50K), P618L, H714L, CCD, CCD AIA (D491A / D493A). L'inhibiteur de silençage P19 a été co-infiltré en parallèle du rapporteur GFP et des constructions URT1. Le contrôle consiste en une co-infiltration de P19 avec le rapporteur GFP seul. **C| Panel gauche**, scatterplot indiquant la valeur de fluorescence normalisée de chaque patch quantifiée par Image J et normalisée par rapport à la fluorescence du patch contrôle. Chaque couleur correspond à un réplica biologique indépendant. Le trait horizontal rouge indique la moyenne de l'ensemble des mesures. **Panel de droite**, p-values obtenus en réalisant un test de wilcoxon sur des données appariées.

de la séquence correspondant au **leader oméga** du virus de la mosaïque du tabac (TMV, « Tobacco Mosaic Virus »). La séquence oméga permet d'augmenter l'expression traductionnelle des ARNm coiffés (Gallie *et al.*, 1987; Gallie *et al.*, 1989; Gallie, 2002; Agalarov *et al.*, 2011). Pour information, des résultats similaires à ceux présentés à la suite ont été obtenus avec des constructions GFP que j'ai réalisées et qui ne contiennent ni la séquence oméga, ni l'intron du gène de la chalcone synthase.

Nous avons comparé l'expression du rapporteur GFP lors de le co-expression de ce dernier avec différentes versions de URT1. Les différentes constructions de URT1 utilisées correspondent à la version sauvage de URT1 (WT) (Fig. 28A), URT1 mutée pour les motifs M1 et M2 (URT1 L21/25N & A49/50K, abrégé URT1 M1/M2) et une version inactive de URT1 mutée pour deux des trois aspartates de la triade catalytique (URT1 D491A/D493A, la séquence originelle DID étant mutée en AIA, j'utiliserai cette dénomination AIA). Nous avons également produit deux versions de URT1 contenant des mutations dans le domaine poly(A) polymérase associé, URT1 P618L, et URT1 H714L. La mutation P618L entrave fortement l'activité de la protéine et diminue la localisation de URT1 dans les foci cytoplasmiques (Wang *et al.*, 2015). La mutation de l'histidine analogue à l'histidine H714 chez Cid1 réduit mais n'abolit pas complètement l'activité de Cid1 (Munoz-Tello *et al.*, 2012), elle a donc été choisie afin d'obtenir un niveau d'uridylation intermédiaire, qui reste cependant à être déterminé par des expériences *in vitro*. Nous avons également utilisé dans cette expérience des constructions de URT1 ne contenant que le cœur catalytique (URT1 420-764, CCD) (Fig. 28A) sans ou avec les deux mutations des aspartates de la triade catalytique, CCD et CCD AIA. L'ensemble des constructions correspondant à URT1 et ses versions mutées et tronquées sont exprimées sous le contrôle du promoteur CaMV 35S.

Les résultats des expériences de co-infiltration de P19, le rapporteur GFP et des différentes constructions de URT1 sont montrés en Fig. 28B. De multiples réplicas biologiques ont été réalisés et la fluorescence des différents patchs quantifiée via Image J (Fig. 28C). La co-infiltration du rapporteur GFP avec la version sauvage de la protéine URT1 résulte en une diminution significative de la production de GFP en comparaison de ce qui est observé dans le contrôle (GFP + P19 seules). Cette diminution n'est pas observée en présence de URT1 inactif, URT1 P618L ou URT1 H714L ce qui implique qu'elle est la conséquence de l'activité d'uridylation de URT1. Elle est en revanche maintenue lors de la co-infiltration du rapporteur avec URT1 M1/M2. Les motifs hautement conservés de la région N-terminale de URT1 et impliqués dans l'interaction avec DCP5 ne sont en conséquent pas requis pour l'effet observé de diminution de la fluorescence de la GFP. La co-infiltration du rapporteur avec la version de URT1 tronquée pour le domaine N-terminal (CCD) résulte également en une diminution de fluorescence. Cette baisse est plus faible que celle observée avec URT1 WT mais significative par rapport à ce qui observé en présence du CCD inactif (CCD AIA) (p-value = $1{,}7 \times 10^{-7}$). Cette observation semble indiquer que le domaine N-terminal de URT1 n'est pas impliqué dans la diminution de la fluorescence due à la GFP. La différence d'effet

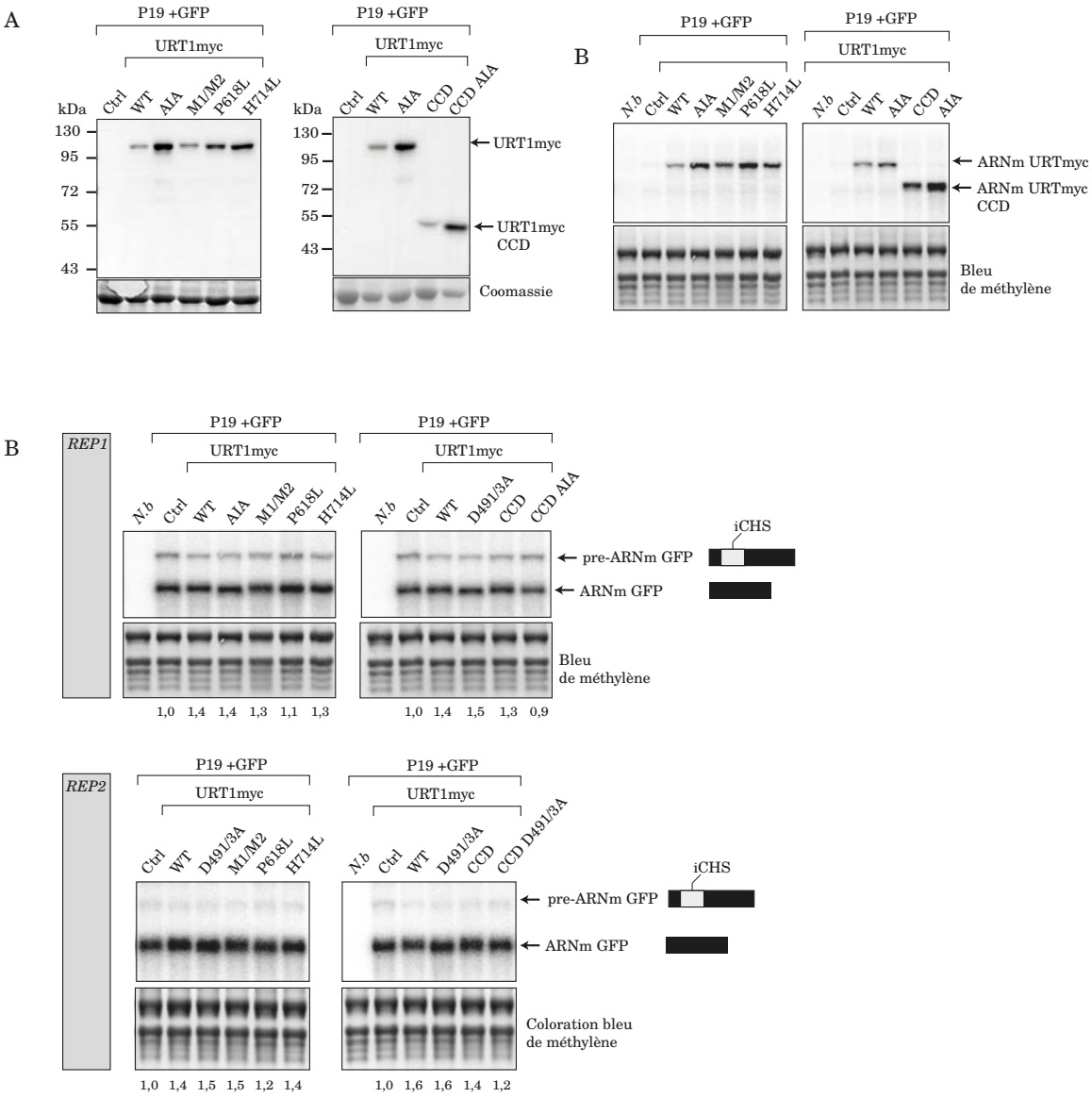


Figure 29 | Niveau d'expression de URT1myc et de l'ARNm GFP dans les différents patchs co-infiltrés.

A| Analyse par western blot avec un anticorps dirigé contre l'épitope myc, du niveau d'accumulation de URT1myc WT, M1/M2, AIA, P618L, H714L, CCD et CCD AIA lors de leur co-infiltration avec P19 et le rapporteur GFP **B|** Analyse par northern blot du niveau d'accumulation de l'ARNm URT1myc ou de ses différentes versions mutantes ou tronquées dans les différents patchs infiltrés en présence de P19, GFP et dans le patch contrôle. **C|** Analyse par northern blot dans deux réplicas indépendant, du niveau d'accumulation de l'ARNm GFP et du pré-ARNm GFP dans les patch infiltrés avec P19, GFP et les différents versions tronquées et mutées de URT1myc et dans le patch contrôle. Pour rappel, la construction permettant l'expression du rapporteur GFP, contient l'intron de la chalcone synthase (iCHS) dans la 5'UTR de la GFP. La valeur indiquée en bas de chaque northern blot correspond au ratio du signal ARNm / pré-ARNm normalisé par rapport au même ratio calculé dans le le contrôle. Le contrôle (Ctrl) correspond à des échantillons de protéines ou d'ARN extraits à partir de patchs infilitrés avec P19 et GFP seuls. Les pistes N. b, correspondent à des ARN extraits à partir de feuilles de *Nicotiana benthamiana* non infiltrés.

entre URT1 WT et CCD est la conséquence probable d'une plus faible activité de la version tronquée de URT1 par rapport à la version sauvage (thèse Emilie Ferrier).

Le niveau d'expression de chacune des constructions URT1 a été vérifié par un western blot avec un anticorps dirigé contre l'épitope myc (Figure. 29A). La diminution de la quantité de GFP lors de la co-infiltration de URT1myc WT, M1/M2 ou CCD ne s'explique pas par une suraccumulation de ces trois protéines par rapport aux versions de URT1myc AIA, P618L, H714L, ou CCD AIA. Nous observons à l'inverse une plus forte accumulation des versions de URT1 mutantes par rapport aux versions dont l'activité n'a pas été altérée. Cette augmentation corrèle avec une augmentation au niveau du messager URT1 (Figure. 29B). La diminution de fluorescence de la GFP est donc due à l'activité catalytique d'uridylation de URT1.

1.2 La diminution de GFP observée en présence de URT1 myc WT, M1/M2 ou URT1 CCD n'est pas la conséquence d'une diminution de l'ARNm

Nous avons voulu dans un second temps vérifier si les variations observées au niveau de la protéine GFP n'étaient pas la conséquence d'une dégradation de l'ARN messager. Nous avons dès lors comparé le niveau d'ARNm GFP par northern blot pour chacun des patchs co-infiltrés avec les différentes versions de URT1myc. Les résultats obtenus pour deux réplicas biologiques sont représentés en Figure. 29B. Au bas de chaque piste est indiqué la valeur du ratio du signal ARNm GFP / pre-ARNm GFP normalisé par rapport au même ratio calculé pour le contrôle (P19 + GFP seuls). Nous sommes partis du postulat que le niveau du pre-ARNm ne devrait pas être affecté par la co-infiltration de URT1, puisque cette dernière a une localisation cytoplasmique, et que la normalisation du signal de l'ARNm par rapport au pré-ARNm constitue un bon moyen de tenir compte de variations dues à l'infiltration. Un ratio inférieur à 1 serait le signe d'une diminution du signal correspond à l'ARNm GFP par rapport au signal observé dans le contrôle P19 + GFP. Nous observons une légère augmentation du ratio lors de l'infiltration des différentes versions de URT1 en comparaison du contrôle.

Aucun des ratios calculés n'est inférieur à 1, et ce même pour les patchs co-infiltrés avec des constructions conduisant à l'inhibition de la production de la protéine GFP. De surcroît les ratios calculés pour les patchs infiltrés avec URT1 myc WT (1,45 ± 0,15) et M1/M2 (1,4 ± 0,1), pour lesquels nous observions une diminution de la quantité de GFP, sont similaires aux ratios calculés pour les patchs infiltrés avec URT1 myc AIA (1,5 ± 0,1) P618L (1,15 ± 0,05) et H714L (1,35 ± 0,05). De même le ratio calculé pour les patchs infiltrés avec CCD (1,35 ± 0,05) est similaire voire supérieur à celui calculé pour les patch co-infiltrés avec CCD AIA (1,05 ± 0,15).

La diminution de la quantité de protéine GFP précédemment observée n'est donc pas la conséquence d'une diminution au niveau de l'ARNm GFP et pourrait être due à une inhibition au niveau traductionnel.

Au vu des résultats présentés en partie B, cette inhibition traductionnelle pourrait être la conséquence du recrutement, par URT1, de DCP5 et indirectement des inhibiteurs de traduction

A

URT1myc	Reads	exp #
Ctrl	63044	1
Ctrl	42381	2
Ctrl	50801	3
WT	73883	1
WT	48088	2
WT	56897	3
AIA	85669	1
AIA	43905	2
AIA	56913	3
M1/M2	85380	1
M1/M2	47613	2
H714L	54539	1
H714L	31868	2
CCD	59615	2
CCD	36117	3
CCD AIA	60169	3
CCD AIA	37329	3

B

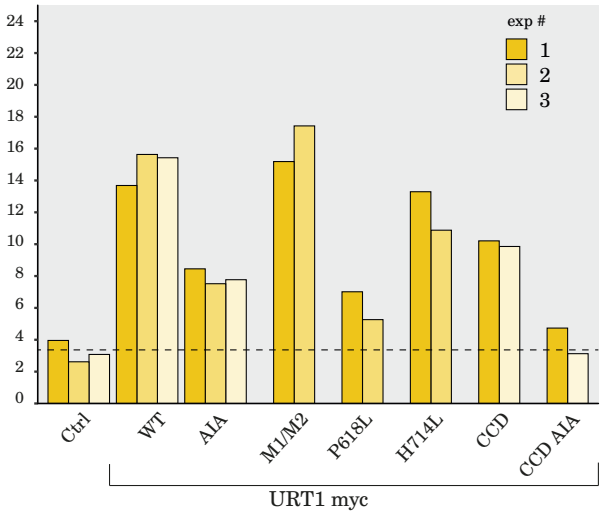


Figure 30 | Analyse du taux d'uridylation dans les patchs co-infiltrés avec les différentes versions de URT1myc

Afin d'analyser le taux d'uridylation du rapporteur GFP, les ARN des différents patchs infiltrés avec P19, le rapporteur GFP et les différentes constructions de URT1myc ont été extraits et utilisés pour préparer des librairies de 3'RACE-seq et permettre l'analyse des extrémités 3' de l'ARNm GFP. **A |** Tableau récapitulant le nombre de reads obtenus pour chaque banque et chaque replica biologique. La première colonne du tableau indique quelle contruction de URT1myc a été co-infiltrée avec P19 et le rapporteur GFP, la seconde le nombre de reads obtenus pour chaque banque et le numéro de l'expérience. 3 réplicas biologiques ont ainsi été réalisés pour le contrôle, URT1myc et URT1myc AIA et 2 réplicas biologiques pour URT1myc M1/M2, H714L, CCD et CCD AIA. **B |** Histogramme représentant le nombre de reads uridylés pour chaque banque de 3'RACE-seq analysée.

RH8 et RH12. Cependant, l'inhibition traductionnelle du rapporteur GFP est également observée lors de la co-infiltration d'une version de URT1 dépourvue des leucines impliquées dans l'interaction de URT1 avec DCP5 (URT1myc M1/M2) ou avec URT1 dépourvu de sa région N-terminale (URT1myc CCD). L'effet sur la traduction semble en réalité dû à l'activité d'uridylation *per se* puisque n'importe laquelle des mutations réalisées dans le cœur catalytique de URT1 semble affecter la capacité de URT1 à inhiber la traduction du rapporteur GFP.

2. Analyse du niveau d'uridylation et de la taille de la queue poly(A) du messager GFP

Dans l'optique de mieux comprendre pourquoi la traduction du messager GFP est inhibée lors de la surexpression de URT1myc en feuilles de tabac, nous avons voulu déterminer le profil d'uridylation et de taille des queues poly(A) du rapporteur GFP lors de l'infiltration des différentes constructions de URT1myc. Les extrémités 3' de trois réplicas biologiques dans le cas des patchs co-infiltrés avec URT1myc WT et AIA et de deux réplicas pour les patchs infiltrés avec URT1myc P618L, H714L, CCD et CCD AIA ont ainsi été analysées par 3'RACE-seq. Le nombre de reads obtenus pour chaque librairie est indiqué en Figure 30A.

2.1 Absence de corrélation entre le pourcentage d'uridylation et le niveau inhibition de la production de GFP

L'analyse du pourcentage d'uridylation des ARNm GFP co-exprimés avec les différentes versions de URT1 a produit certains résultats escomptés mais a également conduit à des conclusions surprenantes (Figure 30B). Parmi les résultats attendus, les 3,2% d'uridylation des ARNm GFP présents dans le patch contrôle, c-à-d en l'absence de URT1myc, sont la conséquence probable de l'activité de l'homologue de URT1 chez le tabac. La co-expression de URT1myc WT entraine une augmentation du pourcentage d'uridylation du transcrit GFP (~14,9% ± 1,1), ce qui indique qu'une version de URT1 étiquetée en C-terminal est capable d'uridyler des ARNm *in planta*. Une augmentation de l'uridylation est également observée avec M1/M2 (~16,3% ± 1,2) comparés au patch contrôle (~3,2% ± 0,7). Ceci suggère que l'uridylation par URT1 ne requiert pas une interaction avec DCP5, en accord avec les résultats exposés précédemment pour les 3 ARNm analysés dans le mutant *dcp5-1*. Dans ces tests d'agroinfiltration, URT1myc WT et URT1myc M1/M2 ont un niveau d'activité similaire, ce qui suggère que la mutation des deux motifs conservés de la région N-terminale n'a aucun impact sur l'activité d'uridylation de la protéine, telle qu'indiquée dans ce type d'essai. L'intégralité du domaine N-terminal est en revanche indispensable *in vivo* pour une activité maximale de URT1, puisque le pourcentage de transcrits uridylés en présence de URT1myc CCD n'égale pas celui obtenu en présence de la protéine sauvage. La proportion d'ARNm GFP uridylés est tout de même plus élevée en présence de URT1myc CCD (10,0 ± 0,2 %) par rapport au contrôle (~3,2% ± 0,7) et à la version inactive de URT1myc CCD AIA (3,9 ± 0,2 %) ce qui indique que URT1myc CCD est en capacité d'uridyler les ARNm mais que son

activité est diminuée par rapport à la version sauvage de URT1, à l'instar de ce qui avait été montré *in vitro* par Emilie Ferrier.

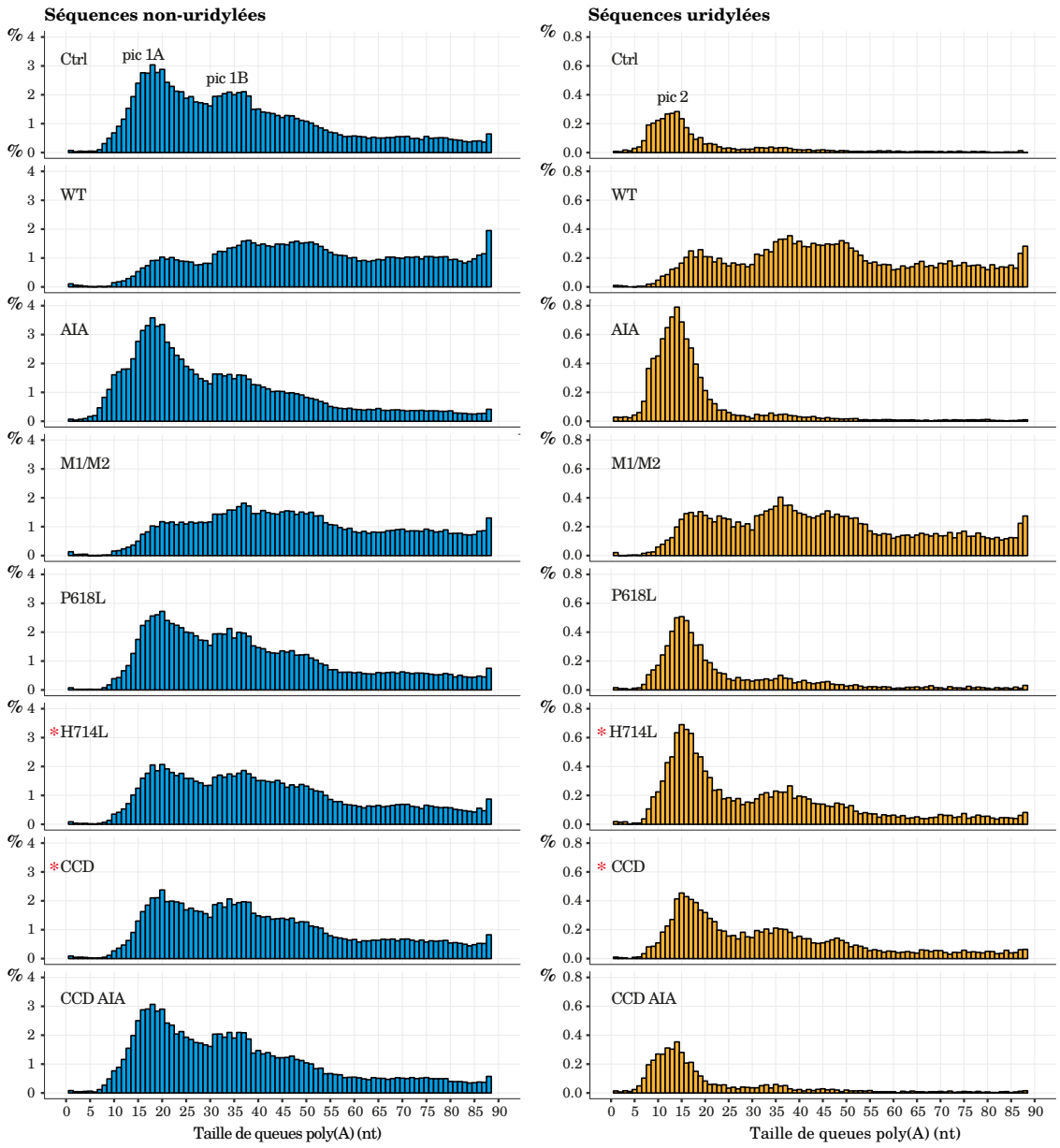
En marge de ces résultats attendus, des constatations plus surprenantes ont été observées. De manière contre-intuitive, l'expression de URT1myc inactif AIA résulte en une augmentation du taux d'uridylation (7,9% ± 1,2) comparés au patch contrôle (~3,2% ± 0,7). Cette augmentation est plus faible que celle observée avec URT1myc actif, mais elle est systématique. De même les deux mutations dans le domaine PAP-assoc, P618L (6,1% ± 1,2 %) et H714L (12,1 ± 1,7 %), entrainent une augmentation de l'uridylation des ARNm GFP (Figure. 30B). En revanche, cette augmentation n'est pas observée avec la version tronquée inactive CCD AIA (3,9 ± 0,2 %). La région N-terminale de URT1 semble donc être requise pour cette augmentation de l'uridylation due à la co-expression de versions inactives (AIA) ou présentant une activité réduite (P618L) et probablement affectée (H714L). Une explication possible à ce phénomène est que la protéine inactive puisse être bloquée ou du moins interagir plus longuement sur ses substrats et, de ce fait, entraverait leur dégradation, ce qui conduirait à l'accumulation de transcrits uridylés non dégradés. Une hypothèse alternative est que la surexpression de versions inactives de URT1 entraine une déplétion d'un facteur(s) impliqué(s) dans la dégradation des ARNm uridylés (voir § 3.3 de la discussion). Cette déplétion possible requerrait la partie N-terminale de URT1.

Un second résultat surprenant est apparu en comparant les niveaux d'uridylation des ARNm GFP avec le degré d'inhibition de production de la GFP dans les patchs respectifs. Une bonne corrélation est observée entre l'augmentation du taux d'uridylation en présence de URT1myc WT, URT1myc M1/M2 et URTmyc CCD, et la diminution de la fluorescence de la GFP dans les feuilles infiltrées. Puisque les niveaux d'ARNm de la GFP ne diminuent pas dans ces patchs, cette diminution de la fluorescence pourrait être due à une inhibition de la traduction de la GFP. Cette inhibition de traduction est abolie lors de l'introduction de mutations qui altèrent l'activité d'uridylation de ces versions d'URT1. Cependant, un taux d'uridylation important est observé avec URT1 H714L sans que la production de GFP ne soit affectée. L'augmentation du taux d'uridylation seule ne peut donc expliquer l'inhibition de la traduction de la GFP dans le cas précis de URT1 H714L et URT1myc CCD. . Comme détaillé dans le paragraphe suivant, des changements dans les populations d'ARNm uridylés et non-uridylés en cas d'expression de versions de URT1 inactives par rapport aux versions inactives semblent associés à la répression de l'expression de la GFP.

2.2 L'inhibition de la production de GFP corrèle avec une forte altération du profil des queues poly(A) des transcrits uridylés et non uridylés.

Nous avons profité de la profondeur conséquente de nos expériences de 3'RACE-seq pour analyser la distribution des queues poly(A) des transcrits GFP uridylés et non uridylés lors de la co-infiltration de URT1myc et de ses différents mutants (Fig. 31A et B). Les analyses par réplicas montrent des profils extrêmement reproductibles entre réplicas pour les échantillons

A



B

Taille moyenne des queues poly(A) (nt)

URT1myc	Ctrl	WT	AIA	M1/M2	P618L	H714L	CCD	CCD AIA
tanscrits **non uridylés**	37	52	33	49	39	42	41	36
transcrits **uridylés**	20	49	18	46	25	32	34	21

Figure 31 | Analyse de la taille des queues poly(A) des transcrits GFP uridylés et non uridylés dans les patchs co-infiltrés avec les différentes versions de URT1myc

A| Tableau indiquant la taille moyenne des queues poly(A) pour les transcrits GFP uridylés et non uridylés dans les différents patchs infiltrés avec P19, le rapporteur GFP et les différentes constructions de URT1myc **B|** Diagrammes en barres représentant la distribution de la taille des queues poly(A) pour les transcrits GFP uridylés et non uridylés . Le pourcentage de reads à chaque position a été normalisé par rapport au nombre total de reads correspondant à des tailles de queues poly(A) comprises entre 1 et 89A. Les réplicas des échantillons marqués avec un astérisque sont variable (voir Annexe 7), ils sont présentés à titre d'information mais doivent être validés par l'analyse de réplicas supplémentaires.

correspondant au patchs infiltrés par URT1myc WT, AIA, M1/M2, P618L et URT1myc CCD AIA (Annexe 7). En revanche, les profils sont très dissimilaires entre les deux réplicas analysés pour URT1myc H714L et URT1myc CCD (Annexe 7), et ce même si le taux d'uridylation des transcrits GFP est similaire entre les réplicas. Cette variabilité pourrait avoir une origine biologique et pourrait refléter une activité intermédiaires de ces deux versions de URT1myc. A ce stade, nous ne pouvons aussi pas exclure une cause technique à la variabilité observée uniquement pour ces deux échantillons analysés par 3' RACE-seq. L'analyse de réplicas supplémentaires est donc nécessaire pour comprendre ou apprécier l'étendue de la variabilité associée aux échantillons URT1myc H714L et URT1myc CCD. Pour cette raison, dans la suite de ce paragraphe nous nous focaliserons notre raisonnement sur l'analyse de la taille des queues poly(A) dans les autres échantillons, validés par des profils identiques entre réplicas.

Les profils de chaque réplica sont consultables en Annexe 7 mais pour faciliter la présentation des résultats de 3'RACE-seq, les résultats pour les différents réplicas sont présentés sous forme combinée en Figure 31. Un astérisque marque les profils URT1myc H714L et URT1myc CCD pour rappeler de manière visuelle que ces profils doivent être validés par des réplicas supplémentaires.

Les analyses par 3'RACE-seq révèlent un effet remarquable de la surexpression de URT1myc WT et de URT1myc M1/M2 sur le profil de distribution des tailles des queues poly(A) du transcrit GFP (Fig. 31A) et la moyenne (Fig. 31B) .

Dans le contrôle, le profil de distribution des transcrits non uridylés présente un pic majoritaire à ~17-18nt (pic 1A, Fig. 31A) suivi d'un second pic, moins proéminent, entre 30 et 40 nucléotides (pic 1B, Fig. 31A). La co-infiltration de l'une ou l'autre des deux versions catalytiquement actives URTmyc WT et M1/M2 entraine une forte diminution du pic à 17-18nt et un décalage du profil vers des queues poly(A) plus longues. Ceci résulte en une augmentation de la moyenne globale de la taille des queues poly(A) de 37nt dans le contrôle à 52 et 49nt dans URTmyc WT et URT1myc M1/M2 respectivement (Fig. 31B). Ces changements ne sont pas observés en présence des autres versions de URT1myc.

L'expression de URT1myc WT ou M1/M2 ont un effet encore plus drastique si l'on considère uniquement les transcrits uridylés. Dans le contexte contrôle, ainsi que pour les constructions pour lesquelles l'activité d'uridylation abolie ou altérée, la distribution de la taille de queue poly(A) est centrée autour d'un unique pic à ~15nt (pic 2, Fig. 31A). Ce pic disparait complètement en présence de URT1myc WT et M1/M2 et laisse place à une distribution diffuse jusqu'à 89 nucléotides, à l'instar de ce qui observé pour les transcrits non-uridylés. Une accumulation de queues poly(A) > 30, intermédiaire entre ce qui est observé pour le contrôle et URT1myc WT et M1/M2, est également visible en présence de URT1myc H714L et URT1myc CCD, ce qui corrèle avec la plus faible activité proposée pour URT1myc CCD par Emilie Ferrier, mais ces profils, variables entre réplicas, doivent encore être validés par l'analyse de réplicas supplémentaires.

Enfin nous remarquons que l'augmentation a priori contre-intuitive du pourcentage d'uridylation dans URT1myc AIA est essentiellement due à une accumulation importante de transcrits uridylés ayant une queue poly(A) d'environ 15nt. L'augmentation de cette population est également

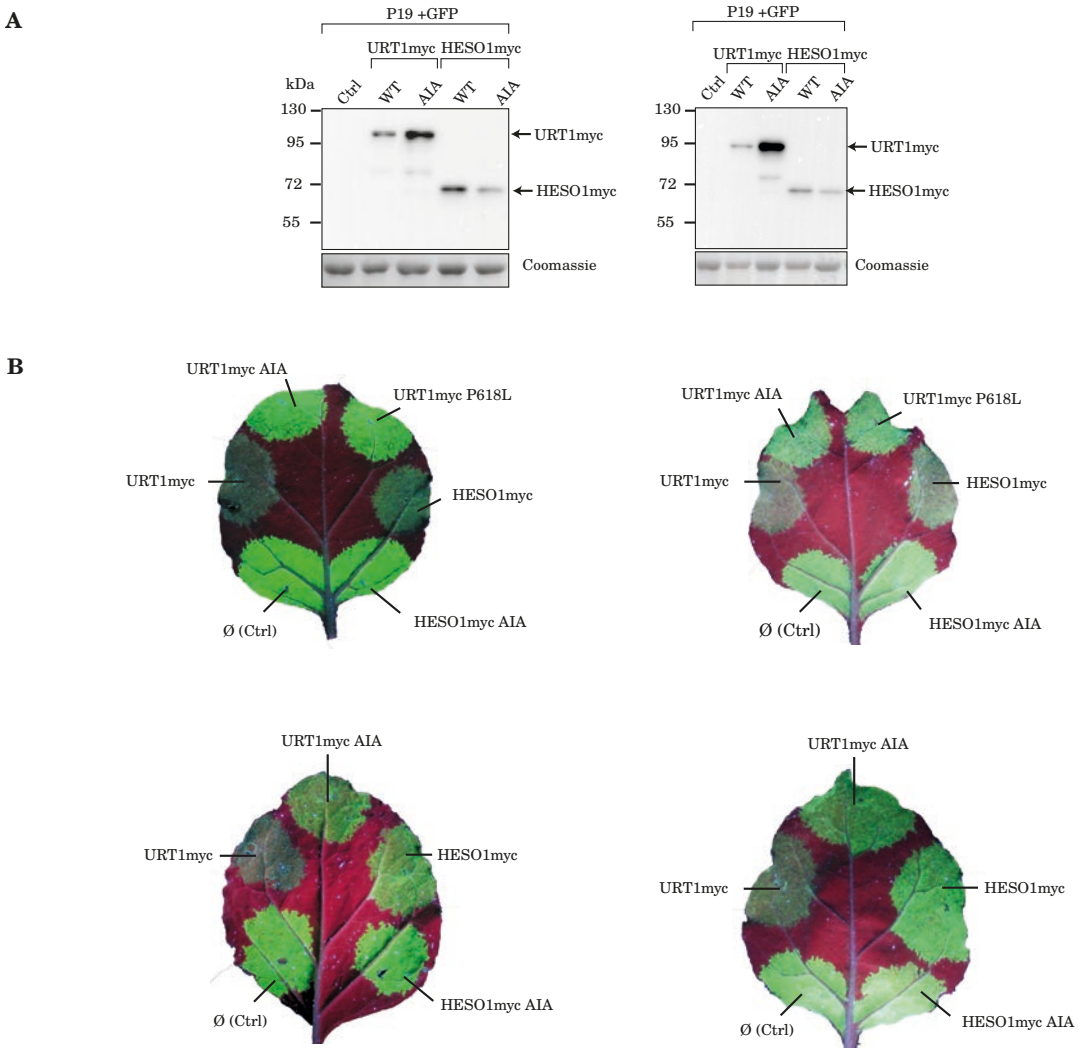


Figure 32 | La co-expression de HESO1myc a un effet variable sur le niveau de GFP
A| Analyse par western blot avec un anticorps dirigé contre l'épitope myc, du niveau d'accumulation de URT1 myc WT, URT1myc AIA, HESO1myc et HESO1myc AIA dans deux réplicas indépendants **B|** Expériences de co-infiltration en feuilles de tabac du rapporteur GFP, de P19 et de différentes versions de URT1myc (URT1myc WT, AIA ou P618 (pour les deux premières feuilles uniquement) ou de HESO1my (HESO1myc ou HESO1myc AIA). Les 4 photographies proviennent de 4 expériences indépendantes.

observée, dans une moindre mesure, pour les constructions URT1myc P618L et H714L. Ces observations concordent avec la possibilité que des versions URT1myc affectées dans l'activité catalytique préviennent la dégradation de cette population de transcrits uridylés, soit en diminuant leur accessibilité à des facteurs de dégradation soit en déplétant un facteur important pour leur dégradation.

Ces résultats montrent que l'inhibition de la traduction du rapporteur GFP corrèle avec une augmentation générale de la taille moyenne des tailles de queues poly(A) des transcrits uridylés et non uridylés.

3. Effet de surexpression de HESO1, la seconde TUTase présente chez *Arabidopsis thaliana*, sur la répression du rapporteur GFP

Afin de voir si l'effet d'inhibition de la traduction est spécifique de l'activité de URT1 nous avons réalisé des expériences de co-infiltration du rapporteur GFP en présence de la seconde TUTase présente chez Arabidopsis, HESO1. HESO1 est l'enzyme majoritairement responsable de l'uridylation des petits ARN chez Arabidopsis (Zhao *et al.*, 2012; Ren *et al.*, 2012). Des études menées dans le laboratoire par Hélène Zuber suggèrent que la protéine serait également impliquée dans l'uridylation des ARNm. Nous avons analysé l'effet de la co-infiltration de HESO1myc et HESO1myc AIA sur l'accumulation du rapporteur GFP. A l'instar des expériences précédentes, l'ensemble des constructions ont été infiltrées en présence de l'inhibiteur de silençage P19.
L'analyse par western blot montre que le niveau d'expression de HESO1 est sensiblement égal à celui de URT1 dans deux réplicas biologiques indépendants (Fig. 32A). Les résultats de l'agroinfiltration de HESO1 en présence de P19 et GFP sont illustrés en Figure 32B pour 4 expériences indépendantes. La co-infiltration de HESO1 semble, au même titre que URT1myc, affecter la production de la GFP même si cet effet semble variable d'une expérience à l'autre puisqu'il n'est observé que dans 2 réplicas sur 4. Cet effet n'est pas observé en présence de HESO1 AIA et pourrait donc être attribuable à l'activité d'uridylation de HESO1, bien que cette conclusion soit à tempérer puisque HESO1myc AIA s'accumule à niveau inférieur à HESO1myc (Fig. 32A). La variabilité observée entre les réplicas dans le cas de HESO1 pourrait être expliquée par une plus faible activité de HESO1myc par rapport à URT1myc. Une plus faible activité de HESO1myc pourrait par exemple résulter sur une diminution retardée du niveau de la protéine GFP (non visible 4 jours après infiltration, c.-à-d. au moment auquel ont été prises les photo présentés en Figure. 32B.

Ces expériences sont très préliminaires, mais le fait que la co-infiltration de HESO1 peut au moins dans certaines expériences, résulter en une diminution de la GFP, suggère que l'uridylation *per se* et non la nature de la protéine impliquée soit responsable de cette diminution. Ces observations

préliminaires n'excluent bien sûr pas que la diminution de la GFP observée lors de la co-infiltration de URT1myc et HESO1myc soit due à deux mécanismes différents.

4. Conclusion générale sur les expériences de co-infiltration en feuilles de tabac

Les résultats des expériences de co-infiltration en feuilles de tabac indiquent que la surexpression de URT1myc prévient la production de la GFP. Cette inhibition de l'expression de ce gène rapporteur n'est pas due à la déstabilisation de ses ARNm et pourrait donc résulter d'une inhibition traductionnelle. Dans les conditions de ces expériences, l'inhibition traductionnelle observée semble être indépendante de l'interaction de URT1 avec DCP5 puisqu'elle est toujours observée en présence d'une version de URT1 mutée pour les leucines du motif M1 et dans une moindre mesure, avec une version de URT1 dépourvue du domaine N-terminal. Grâce à l'utilisation de la méthode 3'RACE-seq, les profils des queues poly(A) uridylées et non-uridylées ont été déterminés à haute résolution. Ces analyses révèlent deux observations inédites, liées à la co-expression de formes actives de URT1 (1) l'accumulation de transcrits uridylés ayant de longues queues poly(A) et (2) la disparition d'une population d'ARNm avec des oligo(A) d'environ 16 nt accompagnée d'une augmentation générale de la taille des queues poly(A). Nous ne savons pour l'instant pas avec certitude si un de ces deux effets est impliqué dans l'inhibition traductionnelle observée. Nous pensons en revanche que l'uridylation en soi, plutôt que la nature de la TUTase et son réseau d'interactants, est responsable de la modulation des profils des queues poly(A) uridylés ou non-uridylés. En effet, bien que moins robuste, l'inhibition de la production de la GFP peut être observé en présence de la seconde TUTase présente chez Arabidopsis, HESO1, et ne serait donc pas spécifique de l'activité de URT1myc. L'ensemble de ces résultats suggère des pistes intéressantes pour élucider les différents rôles moléculaires de l'uridylation, discutés ci-après.

DISCUSSION

Les résultats principaux de mes travaux de thèse ont été (i) l'identification d'une chaine d'interactions connectant URT1 avec des facteurs impliqués dans l'inhibition de la traduction et potentiellement dans l'élimination de la coiffe et (ii) la caractérisation de conséquences moléculaires de l'uridylation sur les profils de polyadénylation des ARNm d'un gène rapporteur, et de son expression. Après discussion en regard d'informations bibliographiques chez Arabidopsis ou d'autres eucaryotes, ces résultats sont mis en perspective dans un modèle des rôles de l'uridylation dans le métabolisme des ARNm chez les plantes.

1. Implications de l'interaction entre URT1 et DCP5

1.1 Les partenaires connus des TUTases chez les eucaryotes

Les premiers réseaux d'interactions impliquant des TUTases ont été décrits pour RET1 et RET2 dans les mitochondries du parasite *Trypanosoma brucei*. RET1 forme avec DSS1 et 3 protéines de fonctions inconnues le MPsome (Mitochondrial 3' Processome) un complexe requis pour la maturation des ARN guides (Aphasizhev *et al.*, 2016). La protéine interagit également de manière transitoire avec des protéines PPR (pentatricopeptide repeat), et avec une Poly(A) polymérase mitochondriale. La TUTase RET2 est quant à elle un membre du complexe RECC (RNA Editing Core Complex) qui catalyse l'édition des ARNm codés par le génome mitochondrial (Aphasizheva *et al.*, 2014).

Dans le cytoplasme des cellules de mammifères, seule l'interaction directe entre TUT4/TUT7 et LIN28A a été caractérisée. TUT4 et TUT7 interagissent avec le domaine de liaison au zinc (Zinc kunckle) de LIN28A via sa région N-terminale (Faehnle *et al.*, 2017; Wang *et al.*, 2017) ce qui promeut l'oligouridylation des précurseurs des miARN let-7 (pré-miARN let7) et leur dégradation par Dis3L2 (Heo *et al.*, 2008; Hagan *et al.*, 2009; Heo *et al.*, 2009; Piskounova *et al.*, 2011; Thornton *et al.*, 2012; Ustianenko *et al.*, 2013; Faehnle *et al.*, 2014). Aucune interaction de ces deux TUTases avec des facteurs impliqués dans le métabolisme des ARNm n'a encore été décelée. Par contre, chez la drosophile, une interaction directe a été rapportée récemment entre la TUTase Tailor et l'exoribonucléase de 3' en 5' Dis3L2. Cette association entre Tailor et Dis3L2 forme le complexe TRUMP pour Terminal RNA Uridylation-Mediated Processing complex (Reimão-Pinto *et al.*, 2016). Le complexe TRUMP promeut l'uridylation et la dégradation d'un grand nombre d'ARN non-codants incluant l'ARNr 5S, des ARN de transfert, des petits ARN nucléaires et nucléolaires (snRNA, snoRNA) ou encore des tiges boucles de type pre-miARN générés par épissage alternatif (Mirtrons).

L'interaction de URT1 avec DCP5 décrite en partie 3 est la seule interaction connue à ce jour d'une TUTase impliquée dans l'uridylation des ARNm avec un inhibiteur de traduction, un facteur de décapping et un composant des P-bodies. Cette interaction pourrait avoir plusieurs implications sur le rôle moléculaire de l'uridylation par URT1 chez Arabidopsis.

1.2 Possibles rôles moléculaires de l'interaction entre URT1 et DCP5

1.2.1 DCP5 pourrait connecter URT1 et ses substrats au complexe de decapping DCP1/DCP2

Chez la plante un rôle de DCP5 dans l'enlèvement de la coiffe est compatible avec le fait que DCP5 co-immunoprécipite avec DCP1/DCP2 lors de test d'interactions *in vitro*, que la dérégulation de DCP5 dans le mutant *dcp5-1* (mutant knock-down) conduit à l'augmentation du temps de demi-vie des ARNm SEN1, DCP2 et EXPL1 et enfin, que les transcrits EXPL1 coiffés s'accumulent dans le mutant *dcp5-1* (Xu and Chua, 2009). DCP1 a également été détecté dans nos immunoprécipitations de DCP5, sans que nous sachions si cette interaction est directe ou indirecte (Chicois et al). Une interaction de DCP5 avec DCP1/DCP2 est évolutivement conservée puisque les protéines homologues à DCP5 chez *S. cerevisiae* (Scd6), *S. pombe* (Sum2) et la drosophile (Tral) se lient avec Dcp2 et Dcp1, respectivement (Nissan *et al.*, 2010; Fromm *et al.*, 2012). A noter que Scd6 est également connectée avec Edc3 (Nissan *et al.*, 2010), un facteur connu pour directement stimuler l'activité du complexe (Fromm *et al.*, 2012; Mugridge *et al.*, 2018, p.3), ce qui supporte un rôle pour cette protéine en tant que modulateur du complexe de décapping. Il reste à démontrer formellement si, *in vivo*, DCP5 et ses différents homologues stimulent l'enlèvement de la coiffe des ARNm directement via leur interaction avec DCP1/DCP2, ou par l'intermédiaire de facteurs associés (comme par exemple Pat1, Edc3 et Dhh1 dans le cas de Scd6) ou si cet effet est la conséquence indirecte de leur rôle dans l'inhibition de la traduction.

Chez *S. pombe* et la drosophile l'interaction entre Sum2/Tral et le complexe DCP1/DCP2 est permise par une interaction entre le domaine LSm des deux protéines avec les motifs de HLM de Dcp2 (dans le cas de Sum2) ou de Dcp1 (dans le cas de Tral) (Fromm *et al.*, 2012). A contrario, les travaux de Xu et Chua suggèrent que l'interaction de DCP5 avec DCP1/DCP2 est permise par la région C-terminale de DCP5 qui comprend les motifs RG et RGG séparés par le domaine FDF (Xu and Chua, 2009). La région C-terminale de DCP5 est également impliquée dans la dimérisation de DCP5 (Xu and Chua, 2009) et dans l'interaction entre DCP5 et RH8 (Fig. 25). Tous ces résultats ont été obtenus *in vitro* et doivent donc être confirmés. Mais sur la base de ces données actuelles, et dans la mesure où URT1 interagit avec le domaine LSm de DCP5, nous ne pouvons donc pas exclure que DCP5 puisse simultanément interagir avec URT1 et DCP1/DCP2. A noter cependant qu'à ce jour nous n'avons obtenu aucun résultat démontrant une interaction commune entre DCP5, URT1 et DCP1/DCP2 puisque ni DCP1, ni DCP2 ne sont détectées par spectrométrie de masse dans nos expériences. Bien entendu, ce résultat négatif ne constitue pas une preuve absolue.

Au cas où une chaine d'interactions entre URT1-DCP5-DCP1-DCP2 existe *in vivo*, cette interaction pourrait avoir deux conséquences : **(1)** URT1 pourrait indirectement recruter DCP1 et DCP2 via DCP5 et cette interaction pourrait favoriser le décapping et la dégradation de 5' en 3' des ARN oligoadénylés et uridylés par URT1. Une observation en faveur de ce scénario est que 98% des ARNm LOM1 dépourvus de coiffe sont uridylés contre 1% des transcrits coiffés (Sement *et al.*, 2013). Bien que cette analyse doit être étendue à un nombre plus conséquent de transcrits, ceci pourrait indiquer qu'à l'instar de ce qui est observé chez *S. pombe* (Rissland and Norbury, 2009),

l'uridylation précède l'enlèvement de la coiffe chez Arabidopsis. **(2)** l'hypothèse alternative serait que URT1 soit indirectement recrutée par DCP1/DCP2 via DCP5 pour ralentir la déadénylation des extrémités 3' des ARNm décoiffés. Nous n'observons cependant pas d'effet majeur de la dérégulation de DCP5 sur l'uridylation par URT1 des transcrits *BAM3*, *AT1G24160* et *CAF1A*. Il serait nécessaire d'étendre cette analyse à l'étude d'un panel plus large de transcrits pour conclure mais les résultats obtenus sur ces 3 transcrits montrent tout de même que DCP5 n'est pas requis pour l'uridylation par URT1.

Enfin, nous ne pouvons exclure de manière formelle que l'interaction *in vitro* observée de DCP1/DCP2 avec le domaine RGG de DCP5 soit dû un artefact expérimental. Si l'interaction entre DCP5 et DCP1/DCP2 implique le domaine LSm de DCP5 à l'instar de la situation chez d'autres eucaryotes, deux possibilités sont à considérer. (i) les interactions DCP5-URT1 et DCP5-DCP1-DCP2 sont mutuellement exclusives (ii) la chaine d'interactions URT1-DCP5-DCP1/DCP2 est tout de même possible du fait de la dimérisation de DCP5 via sa région C-terminale (Xu et Chua, 2009). Les motifs ou domaines permettant les interactions de DCP5 et DCP1/DCP2 se doivent d'être mieux caractérisées chez Arabidopsis afin de définitivement conclure quant à un possible réseau d'interaction comprenant URT1, DCP5 et DCP1/DCP2.

1.2.2 DCP5 pourrait connecter URT1 au processus d'inhibition de la traduction

DCP5 réprime la traduction de l'ARNm OLEO lors de tests de traduction *in vitro* menés dans un système acellulaire de germes de blé (Xu and Chua, 2009). Chez *S. cerevisiae* et le Xénope, Scd6 et xRAP55 sont également capable d'inhiber la traduction d'un messager rapporteur *in vitro* (Tanaka *et al.*, 2006; Rajyaguru *et al.*, 2012). Scd6 empêche la formation du complexe de pré-initiation 43S en se liant au facteur d'initiation de la traduction eIF4G via son domaine RGG (Nissan *et al.*, 2010; Rajyaguru *et al.*, 2012). Les répétitions RGG présentes chez Scd6 sont conservées chez DCP5 et ses homologues chez les eucaryotes (à l'exception de CITH (CAR-I/Trailer Hitch) chez *P. berghei*) ce qui suggère que l'interaction avec eIF4G est une interaction conservée chez les eucaryotes (Roy and Rajyaguru, 2018). Nous observons par ailleurs que eIF4G co-immunoprécipite avec DCP5 dans nos expériences de Co-IP de DCP5-GFP. Il est donc envisageable que DCP5 inhibe la traduction des ARNm par un mécanisme similaire à ce qui est décrit pour Scd6. A noter tout de même que, contrairement à Scd6, DCP5 contient deux domaines riches en arginines et glycine, le domaine RG et le domaine RGG (Fig. 20), il n'est pas à exclure que eIF4G puisse également interagir avec le domaine RG chez Arabidopsis.

En plus de eIF4G, DCP5 co-purifie avec les répresseurs de la traduction et homologues de Dhh1/DDX6 chez la levure et l'homme, RH6, RH8 et RH12. L'interaction directe de RH6 et DCP5 a été validée par des expériences de double-hybride (Chicois *et al.*, 2018) et celle de RH8 et DCP5 par des expériences de pull-down *in vitro* (Fig. 25). Nous avons par ailleurs montré que la délétion du domaine RGG de DCP5 abolit l'interaction avec RH8 in *vitro*. Ce résultat est cependant surprenant. D'une part parce qu'il pourrait exclure la possibilité d'une interaction mutuelle entre

DCP5/RH8 et DCP5/eIF4G et d'autre part parce qu'il est en contradiction avec des travaux récents menés chez l'homme montrant que ce sont les motifs FFD et TFG du domaine FDF qui sont impliqués dans l'interaction de LSM14A (l'homologue à DCP5) avec DDX6 (l'homologue à RH6, 8, 12) (Brandmann *et al.*, 2018, p.14). Il serait judicieux de tester l'interaction directe de DCP5 avec les protéines RH6, RH8 et RH12 et des versions des trois protéines contenant des mutations ponctuelles dans les motifs RGG et le domaine FDF pour conclure sur cette interaction chez Arabidopsis.

L'interaction de URT1 avec DCP5 et celle de DCP5 avec les protéines RH et eIF4G pourraient conduire à l'inhibition traductionnelle des ARNm oligoadénylés et uridylés par URT1. Ce scénario est compatible avec le fait que nous identifions RH8 et eiF4G dans les expériences de Co-IP de myc-URT1.

1.2.3 DCP5 pourrait favoriser la localisation de URT1 dans les P-bodies

L'interaction de URT1 avec DCP5 et indirectement avec les protéines RH6, 8 et 12 pourrait également contribuer à la localisation de URT1 avec ou sans ses ARNm cibles au niveau des P-bodies. Des expériences de localisation de URT1 dans des feuilles de tabac et des cellules BY2 montrent que URT1 est localisée de manière diffuse dans le cytosol et également dans des granules, qui co-localisent avec des marqueurs de P-bodies ou de granules de stress (Sement *et al.*, 2013). En marge de ces résultats obtenus dans le contexte d'expression transitoire des transgènes, j'ai pu confirmer dans des plantes d'Arabidopsis transformées et exprimant YFP-URT1 que cette protéine de fusion est localisée de manière diffuse dans le cytosol et également au sein de granules cytosoliques. La nature exacte de ces granules reste à confirmer par co-expression de marqueurs adéquats mais ces observations confirment la propension de URT1 à être présente dans des granules cytosoliques. En utilisant le système d'expression transitoire en feuilles de *N. benthamiana*, Emilie Ferrier a montré que les motifs M1 et M2 ne sont pas requis pour la localisation de URT1 dans les P-bodies. Cependant, nous ne pouvons pas exclure que les motifs M1 et M2 participent à un réseau d'interactions multiples aboutissant au recrutement de URT1 dans les P-bodies et de stocker/séquestrer URT1 au sein des P-Bodies. Cette contribution des motifs M1/M2 pourrait participer à la robustesse de la localisation de URT1, sans être déterminante. La séquestration de URT1 au sein des P-bodies, structures *a priori* inertes d'un point de vue catalytique, pourrait contribuer à réguler l'activité cellulaire de URT1. De manière intéressante, TUT4 est également présente dans les P-bodies dans des cellules humaines (Hubstenberger *et al.*, 2017; Warkocki *et al.*, 2018, p.1). Les raisons sous-jacentes à la conservation de localisation de TUTases au sein des P-bodies sont inconnues. Ces raisons pourraient être identiques ou divergentes selon les eucaryotes.

En conclusion, la description d'un réseau d'interactions directes et indirectes connectant URT1 à des facteurs d'inhibition de la traduction et/ou activateurs de l'élimination de la coiffe est inédite

et ouvre des perspectives intéressantes quant à la détermination des rôles moléculaires joués par URT1 dans l'expression des gènes.

2. ARNm cibles de l'uridylation par URT1

URT1 a la capacité d'uridyler les miRNAs (du moins quand la méthyltransférase HEN1 et la TUTase HESO1 sont absentes (Wang *et al.*, 2015) mais ses cibles principales connues à ce jour sont les ARNm. Une étape essentielle dans la compréhension des rôles joués par URT1 dans la régulation de l'expression des gènes est l'identification de ses cibles ARNm.

2.1 Variation intergénique du niveau d'uridylation

L'uridylation par URT1 a tout d'abord été détectée systématiquement sur des ARNm candidats analysés au laboratoire, soit environ une quinzaine de transcrits (Sement et al, Zuber et al., ces travaux). L'uridylation des ARNm *CCR2* a également été rapportée par une autre équipe (Morozov *et al.*, 2012) . En outre, l'analyse globale des reads obtenus par TAIL-seq et provenant de 2716 gènes révèle que 32% des reads correspondants à des ARNm ayant une queue poly(A) inférieure à 31 A sont uridylés chez Arabidopsis. Ce taux d'uridylation chute à 7% en l'absence de URT1. L'uridylation par URT1 est une modification qui semble donc affecter un nombre important, voire l'ensemble des ARNm. La profondeur des analyses gène-à-gène par 3'RACE-seq révèle cependant un certain degré de variation intergénique du niveau d'uridylation et de la taille des queues poly(A). Un enjeu majeur pour l'étude des rôles de l'uridylation des ARNm est de déterminer si le taux d'uridylation et la taille des queues poly(A) corrèle avec le temps de demi-vie, la traductibilité, la fonction biologique ou la localisation cellulaire des transcrits analysés chez Arabidopsis. Ceci nécessite une analyse globale par TAIL-seq avec une profondeur permettant une analyse gène-à-gène d'un grand nombre de transcrits. La profondeur requise n'est pas atteinte avec le protocole actuel, qui est en cours d'optimisation au laboratoire.

2.2 URT1 uridyle préférentiellement des ARNm oligoadénylés

Au sein d'une population d'ARNm transcrits d'un même gène, URT1 uridyle préférentiellement les ARNm ayant une queue poly(A) courte. En effet, si nous observons en 3' RACE-seq une distribution étendue de la taille des queues poly(A) non uridylés, le profil des transcrits uridylés n'est en revanche composé que d'un seul pic majoritaire correspondant à des transcrits ayant des queues poly(A) ~14/16A.

Cette préférence de URT1 pour les queues oligoadénylées pourrait être la conséquence d'une propriété intrinsèque de la protéine à mesurer la taille de la queues poly(A) à l'image des TUTases humaines TUT4 et TUT7 qui semblent uridyler moins efficacement des substrats avec des queues poly(A) de 25 et 50A que des substrats avec 0A ou 10A *in vitro* (Lim *et al.*, 2014). Nous observons cependant que la surexpression de URT1 en feuille de tabac a pour conséquence l'uridylation de

queues poly(A) de tailles hétérogènes, ce qui implique que la protéine est en réalité intrinsèquement capable d'uridyler des queues poly(A) supérieures à ~12-14A. Il n'en demeure pas moins qu'à des niveaux physiologiques d'expression de URT1 nous n'observons qu'une proportion très négligeable de queues poly(A) longues uridylées.
La raison probable à l'uridylation préférentielle des queues poly(A) courtes est que les TUTases n'accèdent à l'extrémité 3' des transcrits que lorsque les queues poly(A) ne sont plus liées par des PABP, ou du moins quand cette liaison est moins stable en raison du raccourcissement de la queue poly(A) par les déadénylases. La fixation de PABP pourrait créer un encombrement stérique peu propice à l'activité des TUTases. Cette hypothèse corrèle avec le fait que la PABP PAB2 chez Arabidops bloque l'ajout de U par URT1 *in* vitro (Zuber *et al.*, 2016) et que chez l'homme la déplétion de PABC1 induit une augmentation du taux d'uridylation et l'uridylation de queues poly(A) > 25A (Yi *et al.*, 2018)

Il est possible que l'uridylation des queues poly(A) longues par les TUTases soit également limitée par l'activité processive des déadénylases CCR4. Chez l'homme et la levure les PABP interagissent avec CCR4 et favorise son activité de déadénylation. CCR4 n'est pas actif en l'absence de PABP sur la queue poly(A). Si un phénomène similaire est conservé chez Arabidopsis cela impliquerait que le raccourcissement des queues poly(A) jusqu'à un seuil critique (qui semble être chez Arabidopsis~ 14 /16nt) ait deux conséquences, la perte de fixation des PABP et la perte d'activité de CCR4A/B.
Dans ce contexte les extrémités 3' des ARNm seraient alors libres pour l'uridylation par URT1. L'uridylation par URT1 serait donc dépendante de la déadénylation par CCR4A/B et expliquerait pourquoi la surexpression de URT1 dans des lignées d'Arabidopsis ne conduit pas à l'augmentation massive du taux d'uridylation (Zuber *et al.*, 2016). Il serait intéressant d'analyser le profil d'uridylation des transcrits dans un fond double mutant *ccr4a/b* et dans un fond double mutant *ccr4a/b* dans lequel URT1 est surexprimé.
Si la déadénylation est limitante pour l'uridylation par URT1 comment justifier cependant l'apparition des longues queues poly(A) uridylés lors de la surexpression de URT1 en feuille de tabac ? Il est probable que cette uridylation des longues queues poly(A) soit une résultante directe du niveau de surexpression des TUTases actives. Cependant, nous ne pouvons exclure que la production en masse des ARNm correspondant à la GFP, à URT1 et à P19 saturent le niveau cellulaire de PABP et donc que la densité de PABP puisse diminuer sur les extensions polyadényliques de certains ARNm. Cette déplétion, combinée avec la surexpression de TUTases actives, pourrait expliquer l'uridylation des longues queues poly(A). Cette hypothèse de l'impact de déplétion des PAPBP pourrait être testée en agroinfiltrant PAB2 en parallèle des autres constructions. La surexpression de PAB2 pourrait contrecarrer l'uridylation par URT1 des queues poly(A) longues. Ces expériences pourraient être réalisées avec les différentes PABP exprimées chez Arabidopsis pour tester un éventuel impact différentiel de l'une de ces protéines.
A noter qu'il existe au moins une hypothèse alternative (et non-mutuellement exclusive) à l'uridylation préférentielle par URT1 des queues oligoadénylées. Il se pourrait en réalité que URT1

uridyle l'ensemble des populations de queues poly(A) *in vivo* mais que la population de queues poly(A) longues uridylés ne soit peu visible lors d'une analyse des ARNm (« steady state »). Il est en effet possible que cette uridylation soit transitoire (c'est à dire rapidement éliminée) ou qu'elle conduise à la dégradation rapide des transcrits uridylés. En tout état de cause, cette population d'ARNm avec de longues queues poly(A) uridylées ne s'accumule pas dans une plante sauvage.

En conclusion, si il ne peut être exclu que des facteurs en trans recrutent sélectivement URT1 sur certains ARNm, nos résultats actuels indiquent que l'uridylation des ARNm oligoadénylés est un processus par défaut, qui pourrait affecter tous les ARNm ayant subi une étape de déadénylation. L'uridylation par URT1 semble contrôlée par l'accessibilité à ses substrats, accessibilité dépendant de l'action des déadénylases et de la perte de liaison des PABP. Le taux d'uridylation d'un ARNm donné pourrait donc être la résultante des facteurs contrôlant l'efficacité de déadénylation de cet ARNm.

3. Caractérisation de rôles moléculaires de l'uridylation des ARNm grâce à la surexpression de URT1 en feuilles de tabac

Nous avons utilisé un système d'expression transitoire en feuilles de *N. benthamiana* afin de décrypter certaines fonctions moléculaires de l'uridylation des ARNm. En choisissant notre système expérimental d'expression transitoire, nous avons décidé de ne pas ancrer URT1 sur un ARNm cible. Nous ne voulions pas par exemple induire une processivité artificielle de URT1 en stabilisant son interaction avec sa cible. Par contre, nous voulions obtenir une surexpression suffisante pour déplacer l'équilibre résultant du niveau d'expression de cette TUTase pour en étudier les effets. En raison de la stratégie expérimentale retenue, les ARNm du gène rapporteur *GFP* ne sont pas ciblés de manière spécifique et URT1 est capable d'uridyler tous les ARNm (et possiblement d'autres ARN substrats). En plus des ARNm GFP, des analyses par 3' RACE-seq sont en cours pour un ARNm endogène. Ces résultats, bien que préliminaires, apparaissent similaires à ceux obtenus pour les ARNm GFP.

3.1 Changement massif de la distribution des queues poly(A) des transcrits uridylés et non uridylés

La surexpression de versions actives de URT1myc résulte en une augmentation du taux d'uridylation du messager GFP. De manière intéressante, cette augmentation de l'uridylation ne provoque pas une diminution de l'accumulation des ARNm mais un changement massif de la distribution des queues poly(A) pour les transcrits uridylés et non uridylés. Ce changement est caractérisé par (1) une quasi-absence du pic à 16nt-18nt pour les transcrits non uridylés. (2) et l'allongement global de la taille des queues poly(A) des transcrits uridylés et non uridylés.

Ce phénomène a deux explications possibles. La première est que l'uridylation massive des queues oligoadénylées (pic à 16nt) par URT1 promeut leur dégradation et celle-ci entrainerait un biais dans la distribution vers les queues poly(A) longues. Cette hypothèse reste possible mais nous ne la favorisons pas à ce stade dans la mesure où nous n'observons pas de diminution notable de l'ARNm GFP lors de nos analyses en northern blot.

La seconde possibilité, en lien avec le rôle de URT1 dans la protection des extrémités 3', est que l'uridylation par URT1 des queues poly(A) longues ralentirait leur déadénylation. Cette protection « physique » par les uridines prévient le raccourcissement des queues poly(A) et ainsi la production du pic à 16nt ou 18nt. Cette hypothèse est étayée par des résultats récents obtenus au laboratoire par Caroline de Almeida. L'ajout d'une uridine en 3' d'un substrat ARN oligoadénylé est suffisant pour ralentir l'activité d'une déadénylase (CAF1B) *in vitro*. Ces résultats indiquent que l'uridylation est intrinsèquement capable de réduire l'activité d'une déadénylase, ce qui est en accord avec la nature même de ce type d'exoribonucléases qui favorisent l'élimination d'adénosines en 3' des ARNm. L'ensemble de ces résultats, obtenus *in planta* et *in vitro*, conforte l'idée qu'un des rôles de l'uridylation des ARNm est de limiter une déadénylation excessive des ARNm. Si l'uridylation des longues queues poly(A) est possible *in vivo*, l'uridylation pourrait participer au contrôle de la vitesse de déadénylation et ainsi participer à réguler la vitesse de dégradation des ARNm.

3.2 Inhibition de la traduction du messager GFP

La surexpression de formes actives de URT1 résulte en l'inhibition de la traduction du messager GFP. De plus, les mêmes expériences effectuées en présence la seconde TUTase chez Arabidopsis, HESO1 laissent apparaître un effet similaire, bien que moins robuste, sur la production de la GFP. Ces résultats révèlent que l'inhibition de la traduction du messager GFP est probablement due à l'activité d'uridylation *per se*. Cette conclusion explique pourquoi la surexpression de versions de URT1 mutées pour les motifs M1 et M2 ou tronquées pour la région N-terminale intrinsèquement désorganisée conduit également à l'inhibition de la traduction de la GFP. En effet, dans ces expériences d'expression transitoire, l'inhibition de la production de GFP ne requiert pas le recrutement de DCP5 par URT1. Il est important de souligner que nous ne pouvons exclure à ce stade que le recrutement de DCP5 par URT1 dans des conditions physiologiques (c'est à dire dans une plante sauvage d'Arabidopsis) puisse aboutir à un effet inhibiteur de l'uridylation.

Au vu des résultats obtenus par 3'RACE-seq l'inhibition de la traduction du messager GFP pourrait être la conséquence directe de l'absence d'ARNm avec des queues poly(A) courtes (pic à 16-18 nt), de l'allongement des queues poly(A) et/ou de l'accumulation de longues queues poly(A) uridylées. Il est par exemple possible que l'accumulation de transcrits uridylés déplète un facteur requis pour la traduction efficace des ARNm. Concernant la taille des queues poly(A), s'il paraît à première vue contre-intuitif d'imaginer que l'allongement des queues poly(A) puisse inhiber la traduction du transcrit GFP, des études récentes menée chez *C. elegans* suggèrent cependant que ce sont en réalité des ARNm avec des queues poly(A) d'environ ~30nt qui sont les plus efficacement traduits.

Si la liaison d'une seule PABP est suffisante et optimale pour la traduction chez Arabidopsis également, l'inhibition de la traduction observée en présence de URT1myc et URTmyc M1/M2 et dans une moindre mesure avec URT1myc CCD pourrait être la conséquence de l'absence d'une population d'ARNm avec des queues poly(A) courtes et/ou l'apparition massive, de queues poly(A) bien supérieures à 30A. Cette hypothèse pourra être testée chez Arabidopsis en comparant la taille des queues poly(A) entre les ARNm les plus associés aux polysomes avec ceux les moins associés. L'obtention d'un jeu de données TAIL-seq avec une profondeur permettant une réelle analyse gène-à-gène est là aussi requise. Alternativement, il faudra résoudre les problèmes techniques associés à la dégradation des queues poly(A) dans les lysats cellulaires pour comparer de manière fiable les queues poly(A) d'ARNm d'une fraction ARN total par rapport à ceux associés aux polysomes.

Nous ne pouvons exclure que ce phénomène d'inhibition de la traduction pourrait également résulter de conséquences indirectes liées à l'utilisation du système d'expression transitoire chez *N. benthamiana*.
En plus du rapporteur GFP nous avons également analysés l'effet de la surexpression de URT1 sur l'ARNm endogène PR2. Nous observons également un allongement de la taille des queues poly(A) uridylés du transcrit PR2 ce qui suggère que l'effet observé lors la surexpression de URT1 n'est pas spécifique de la GFP et qu'il est susceptible d'affecter l'ensemble des transcrits chez *Nicotiana benthamiana*. L'uridylation par URT1 pourrait par exemple conduire à la dérégulation de transcrits codant pour les facteurs impliqués dans les processus de traduction.
L'inhibition de la traduction pourrait également être due à l'uridylation d'autres ARN que les ARNm, comme par exemple les ARN ribosomiques ou les ARNt, ce qui pourrait conduire à leur déstabilisation et donc à l'inhibition de la traduction. Un changement global de l'abondance des ARNr ou ARNt n'est pas observé dans nos échantillons, mais nous ne pouvons exclure que des effets ciblés aient lieu (uridylation d'ARNt spécifiques ou d'autres ARN comme l'ARN 7S par exemple).

Enfin, la co-expression de URT1 diminue l'expression de la GFP en présence du suppresseur de silençage P19, ainsi qu'en présence d'autres suppresseurs d'origine virale. Cependant, l'activité de URT1 affecte également la production de ces protéines. Afin d'éliminer un éventuel rôle de la voie de silençage post-transcriptionnel (PTGS), les expériences de co-expression de URT1 et GFP seront réalisées dans une lignée de *N. benthamiana* chez laquelle l'expression de RDR6, un acteur majeur du PTGS, est fortement diminuée (Dalmay *et al.*, 2000).

3.3 Le région N-terminale de URT1 est impliquée dans l'accumulation de transcrits GFP oligoadénylés

La surexpression de URT1 AIA et dans une certaine mesure également celle de URT1 P618L et H714L conduit à une augmentation du pic à 16nt-18nt pour les transcrits non uridylés. L'accumulation de cette population d'ARNm oligoadénylés est possiblement due au fait que

l'inactivation de URT1 ralentit sa dissociation du substrat ARN. L'activité de URT1 est distributive *in vitro* ce qui implique qu'elle catalyse l'ajout d'une uridine et se détache de sa matrice. Ceci corrèle avec le fait que URT1 n'ajoute en moyenne qu'un ou deux U à ses substrats. Les versions de URT1 mutées pour l'activité de la protéine pourraient être toujours en mesure de se lier à leur substrat mais elles sont incapables de catalyser l'ajout d'uridines. Ceci pourrait augmenter le temps d'interaction entre URT1 AIA, P618L et H714L et les substrats oligoadenylés et indirectement protéger ces derniers de l'activité d'exoribonucléases 3'-5' ou de l'interaction avec d'autres facteurs de dégradation. De manière intéressante la surexpression de URT1myc CCD AIA ne conduit pas à l'accumulation de transcrit oligoadénylés ce qui implique le région N-terminale de URT1 est impliquée dans l'effet observé. La région N-terminale de URT1 pourrait notamment être impliquée dans la liaison de URT1 à ses substrats.
Une hypothèse alternative serait que la suraccumulation des versions inactives de URT1 en comparaison de URT1myc WT, M1/M2 et CCD, comme l'atteste le western blot présenté en Figure 29A, conduise à la déplétion de facteurs impliqués dans la dégradation des ARNm oligoadenylés. Au vu des résultats obtenus en partie 2 sur l'interaction entre URT1 et DCP5, il est légitime de penser que la surexpression des versions inactives de URT1 puisse entrainer la déplétion de DCP5 puisque DCP5 interagit avec le motif M1 de URT1. Cette hypothèse expliquerait également pourquoi la version URT1myc CCD AIA ne provoque par l'accumulation des transcrits oligoadénylés à l'instar des autres versions inactives de URT1myc. Il a été montré que DCP5 était impliqué dans les mécanismes d'enlèvement de la coiffe de par son interaction avec DCP1/DCP2, sa déplétion pourrait conduire à l'accumulation de transcrits oligoadénylés.
Il serait dès lors intéressant de tester si la co-expression d'une version de URT1myc mutée pour le motif M1 et mutée pour les aspartates de la triade catalytique (AIA), P618L ou H714L induit toujours une augmentation du niveau des ARNm oligoadenylés.
Enfin la forte expression de versions inactives de URT1 pourrait également induire la formation de particules ribonucléoprotéiques dans lesquels seraient inclus les ARNm oligoadénylés ciblés par URT1.

En conclusion, les expériences de surexpression de URT1 et de ses versions mutées ne correspondent pas à des conditions physiologiques d'uridylation des ARNm mais permettent néanmoins d'exacerber les conséquences de l'uridylation des ARNm *in planta* afin d'en déduire des rôles moléculaires liés à cette modification. Ces expériences révèlent notamment que l'uridylation *per se* pourrait inhiber ou ralentir la déadénylation des longues queues poly(A), probablement en inhibant les activités responsables de la déadénylation.

L'ensemble des rôles moléculaires de l'uridylation dans le métabolisme des ARNm sont intégrés dans le paragraphe suivant qui conclut cette discussion.

4. Conséquences moléculaires de l'uridylation des ARNm par URT1

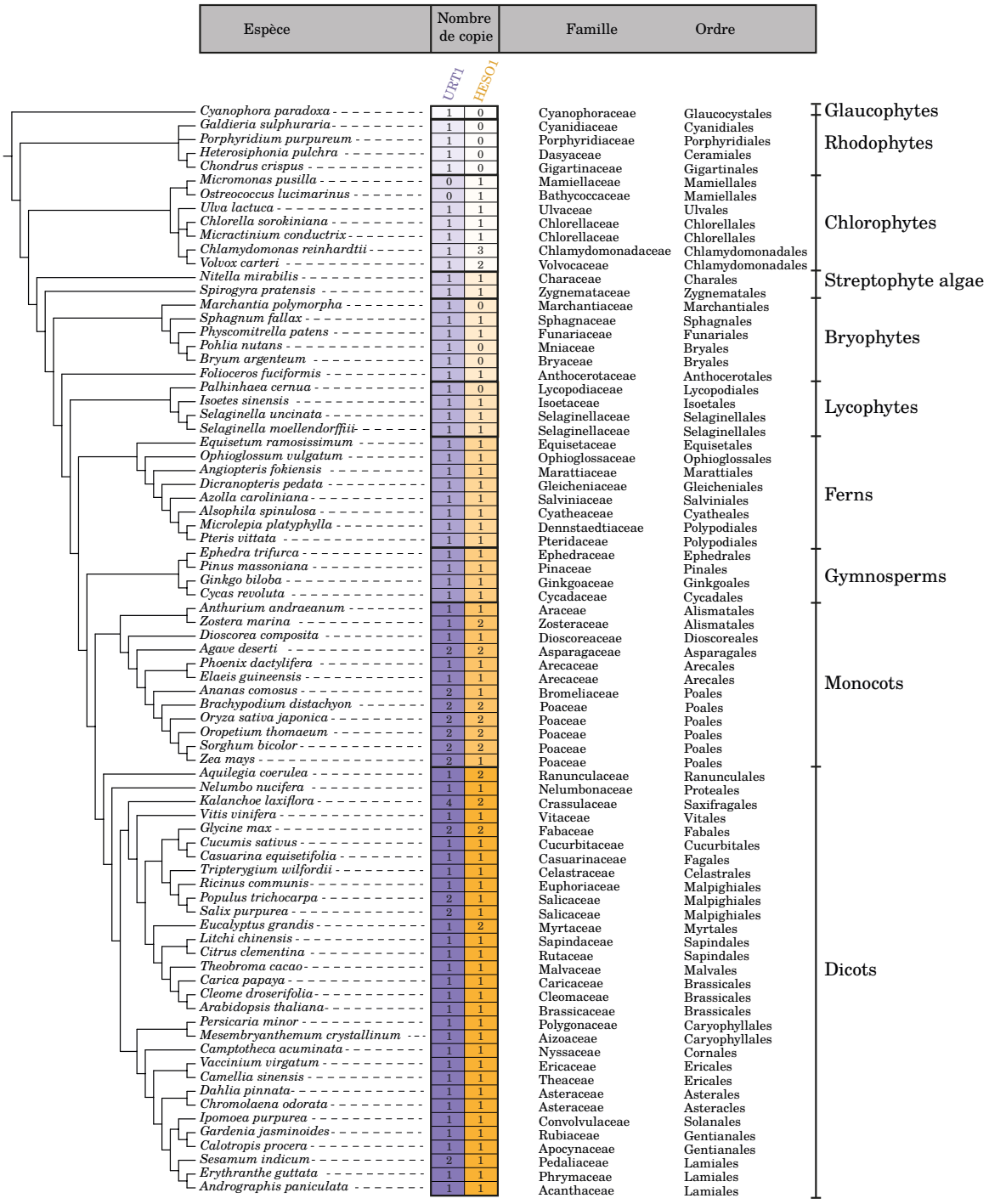
Il existe au moins deux conséquences moléculaires de l'uridylation des ARNm chez Arabidopsis.
(1) Celle qui à l'instar de l'uridylation par TUT4/TUT7 chez l'homme et par Cid1 chez *S. pombe*, promeut la dégradation des ARNm. Chez Arabidopsis cette uridylation, qui implique URT1 et une autre TUTase, est visible dans un fond mutant *xrn4* et affecte des queues poly(A) excessivement déadénylés (<10) (partie 1). Nous pensons que la seconde TUTase impliquée dans l'uridylation des ARNm chez Arabidopsis est HESO1. Des expériences de 3'RACE-seq sont en cours laboratoire pour déterminer si l'uridylation résiduelle observée dans *urt1* est abolie dans un fond double mutant *urt1 heso1* et si l'accumulation de transcrits excessivement déadénylés et uridylés est diminuée dans un fond mutant *heso1 xrn4* et abolie dans un fond triple mutant *urt1 heso1 xrn4*. A l'instar des autres eucaryotes comme l'homme ou *S. pombe*, l'ajout d'une ou deux uridines en 3' d'un queue oligo(A) courte (<10) chez Arabidopsis pourrait favoriser la liaison du complexe LSm1-7, qui lui-même permet le recrutement du complexe de décapping (Norbury, 2013; Rissland and Norbury, 2009). Alternativement, l'addition de queues oligo(U) plus longues pourraient initier une dégradation de 3' en 5' par SOV, l'orthologue de Dis3L2 chez Arabidopsis, voire par l'exosome. A noter que cette population d'ARNm avec des queues oligo(A) courtes (<10) s'accumule aussi spécifiquement quand URT1 inactive (URT1 AIA) est surexprimée dans les feuilles de *N. benthamiana* (Fig. 31), confortant l'idée que le processus de dégradation des ARNm est affecté par la surexpression de cette version inactive de URT1. Dans ces échantillons, la population d'ARNm uridylés comportant des oligo(A) d'environ 16 nt est aussi accumulée, révélant qu'elle est probablement aussi sujette à dégradation en conditions physiologiques (c'est à dire sans surexpression de URT1 AIA). Si cette population d'ARNm uridylés est sujette à la voie de dégradation de 5' en 3', l'élimination de la coiffe pourrait être favorisé par le recrutement de DCP1/DCP2 via l'interaction URT1-DCP5.

(2) La seconde conséquence de l'uridylation est celle résultant principalement de URT1 et qui prévient d'une déadénylation excessive. Deux observations présentées dans Sement et al. et Zuber et al. sont en faveur du rôle de URT1 dans la protection des extrémités 3': le fait que l'inactivation de *urt1* résulte en l'accumulation de transcrits excessivement déadénylés et que l'uridylation par URT1 restaure la taille des queues poly(A) à 16nt, permettant la liaison d'une PABP. Les expériences présentées dans cette thèse montrent de plus que l'accumulation de queues poly(A) longues uridylées due à la surexpression de URT1, entraine l'absence d'une population de queues poly(A) courtes centrées sur 16 nt. De plus, des expériences menées dans le laboratoire par Caroline De Almeida montre que l'activité de la déadénylase CAF1B est fortement affectée par la présence d'uridines à l'extrémité 3' d'un substrat oligoadénylé *in vitro*. Bien qu'il reste encore à être démontré que cet effet est également observé en présence de CCR4, l'ensemble de ces expériences rendent totalement plausible un modèle dans lequel l'uridylation *per se* ralenti l'activité des déadénylases,

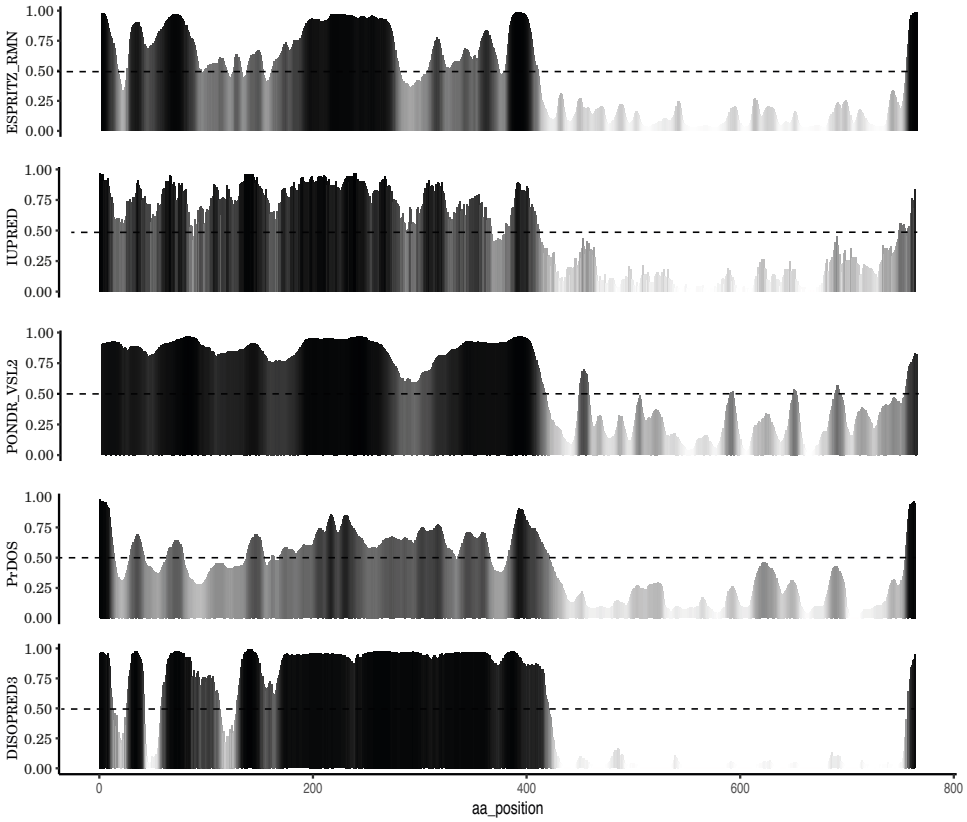
que ce soit pour éviter une déadénylation excessive des extensions oligo(A) courtes ou pour participer au contrôle de la déadénylation des queues longues.

Les deux rôles principaux de l'uridylation, que sont l'activation de l'élimination de la coiffe et la prévention d'une déadénylation excessive, ne sont pas contradictoires. En effet, leur combinaison participe certainement à promouvoir une polarité de dégradation des ARNm de 5' en 3'. Chez *S. cerevisiae*, le complexe LSm1-7 remplit exactement ces deux fonctions : promotion de l'élimination de la coiffe grâce au recrutement de Dcp1/-Dcp2 via Pat1 et prévention de la formation d'ARNm tronqués en 3', probablement par encombrement stérique contrant une activité 3'-5' inconnue. Établir ou renforcer la polarité de dégradation des ARNm de 5' en 3' est probablement crucial dans le cas de la dégradation co-traductionnelle des ARNm, décrite chez la levure et Arabidopsis (Pelechano *et al.*, 2015; Yu *et al.*, 2016). De plus, chez les plantes, prévenir la formation d'ARNm complètement déadénylés ou même tronqués en 3' est essentiel. En effet, ces ARNm complètement déadénylés ou tronqués représentent des ARN « aberrants », détectés chez les plantes par voie du silençage post-transcriptionnel ou PTGS. De manière intéressante, une des caractéristiques intrinsèques de la RNA polymérase RNA dépendante RDR6, un facteur principal du PTGS, est de pouvoir reconnaître les ARN dépourvus d'extensions adényliques. Si ces ARNm « aberrants » sont détectés pas la machinerie du PTGS, les siRNAs en résultant vont cibler les ARNm endogènes, déclencher leur dégradation et *in fine* aboutir à la mort de la plante. Si un effet létal d'une mutation simple *urt1* n'est pas observée chez Arabidopsis, du moins en conditions classiques de culture, cet effet est révélé dans un fond mutant *xrn4*. En effet, une double mutation *urt1 xrn4* induit une mort prématurée des plantes accompagnée de l'accumulation de siRNAs ciblant les ARNm. Cette observation participe à révéler l'uridylation des ARNm comme une étape importante du processus de dégradation des ARNm.

Annexes



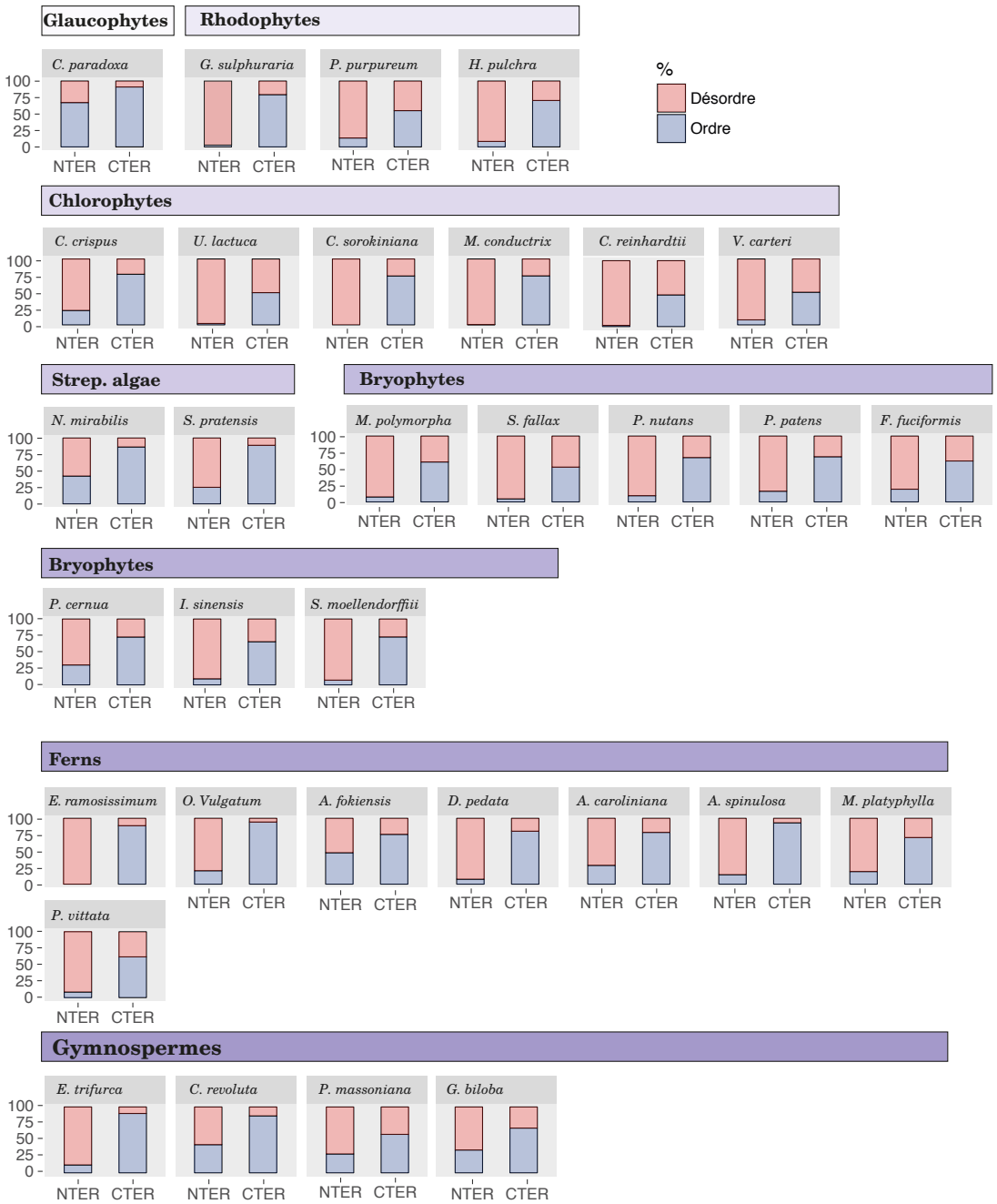
Annexe 1 | **Nombre de copie de URT1 et HESO1 dans 79 espèces représentatives d'Archaeplastida**
Relation phylogénétique, obtenu en utilisant Phylostatic, entre les 79 espèces utilisées pour réaliser l'arbre phylogénétique présenté en Figure 5. La taxonomie (famille et ordre) et le nombre de copie de chaque protéine homologue à URT1 et HESO1 sont indiqués à droite de chaque espèce.



Annexe 2 | **La région N-terminale de URT1 est prédite comme étant intrinsèquement désorganisée** La région N-terminale de URT1 est prédite comme étant intrinsèquement désorganisée par 5 prédicteurs indépendants: ESPRITZ RMN (http://protein.bio.unipd.it/fells/news), IUPRED (http://iupred.elte.hu), PONDR VSL2 (http://www.pondr.com), PrDOS (http://prdos.hgc.jp/cgi-bin/top.cgi) et DISOPRED3 (http://bioinf.cs.ucl.ac.uk). La ligne horizontale pointillée représente une valeur de désordre égale à 0,5.

Annexe 3 | **Pourcentage de désorganisation de la région N-terminale de URT1 et C-terminale des homologues à URT1 chez les Archaeplastida**

Pourcentage de résidus considérés comme étant intrinsèquement désorganisés (D, rouge) ou organisés (O, bleu) pour la région N-terminale et C-terminale des homologues à URT1 chez les Archeaplastida. La région N-terminale (NTER) correspondant exactement à la première moitié de résidus de chacun des protéines et la région C-terminale (CTER) à la seconde moitié. Le pourcentage de désorganisation a été calculé en utilisant le prédicteur ESPritz-NMR du serveur FELLS (http://protein.bio.unipd.it/fells).




Monocots
A. andraeanum
D. composita
Z. marina
A. deserti
A. deserti
P. dactylifera
E. guineensis
A. comosus A
A. comosus B
B. distachyon A
B. distachyon B
O. sativa A
O. sativa B
O. thomaeum A
O. thomaeum B
S. bicolor A
S. bicolor B
Z. mays A
Z. mays B
100
75
50
25
0
NTER
CTER
%
Désordre
Ordre
Dicots
A. coerulea
N. nucifera
K. laxiflora
K. laxiflora 2
K. laxiflora 3
K. laxiflora
V. vinifera
G. max
G. max
C. sativus
C. equisetifolia
T. wilfordii
R. communis
P. trichocarpa
P. trichocarpa
S. purpurea
S. purpurea
E. grandis
L. chinensis
C. clementina
T. cacao
C. papaya
C. droserifolia
A. thaliana
P. minor
M. crystallinum
C. acuminata
V. virgatum
C. sinensis
D. pinnata
C. odorata
I. purpurea
G. jasminoides
C. procera
S. indicum
S. indicum
E. guttata
A. paniculata

Annexe 4 | Identifiants NCBI ou Phytozome des séquences utilisées pour réaliser les logos présentés en figure 16 et 17. Le séparateur | est utilisé pour indiquer plusieurs séquences lors qu'il existe plusieurs copies de URT1.

	Order	family		Source	base de données
Bryophytes	Marchantiales	Marchantiaceae	*Marchantia polymorpha*	Mapoly0079s0057.1	Phytozome
	Bryales	Bryaceae	*Bryum argenteum*	GCZP01036087/GCZP01001333/GCZP01001334/CZP01001335	NCBI
	Bryales	Mniaceae	*Pohlia nutans*	GACA01018571 / GACA01001037 /GACA01015304	NCBI
	Funariales	Funariaceae	*Physcomitrella patens*	Pp3c2_23720	Phytozome
	Sphagnales	Sphagnaceae	*Sphagnum fallax*	Sphfalx0013s0255.1	Phytozome
	Anthocerotales	Anthocerotaceae	*Folioceros fuciformis*	(You et al., 2017)	
Lycophytes	Isoetales	Isoetaceae	*Isoetes sinensis*	(You et al., 2017)	
	Selaginellales	Selaginellaceae	*Selaginella uncinata*	(You et al., 2017)	
	Selaginellales	Selaginellaceae	*Selaginella moellendorffii*	XP_002973804.2	NCBI
	Lycopodiales	Lycopodiaceae	*Lycopodium cernuum*	(You et al., 2017)	
Fougères	Polypodiales	Polypodiaceae	*Phymatosorus cuspidatus*	(You et al., 2017)	
	Polypodiales	Dryopteridaceae	*Cyrtomium fortunei*	(You et al., 2017)	
	Polypodiales	Athyriaceae	*Diplazium esculentum*	(You et al., 2017)	
	Polypodiales	Dennstaedtiaceae	*Microlepia platyphylla*	(You et al., 2017)	
	Polypodiales	Pteridaceae	*Pteris vittata*	(You et al., 2017)	
	Polypodiales	Lindsaeaceae	*Sphenomeris chinensis*	(You et al., 2017)	
	Polypodiales	Pteridaceae	*Ceratopteris richardii*	GBGN01047417	NCBI
	Salviniales	Salviniaceae	*Azolla caroliniana*	(You et al., 2017)	
	Salviniales	Salviniaceae	*Salvinia molesta*	(You et al., 2017)	
	Osmundales	Osmundaceae	*Osmunda vachellii*	(You et al., 2017)	
	Mariattales	Marattiaceae	*Angiospteris fokiensis*	(You et al., 2017)	
	Ophioglossales	Ophioglossaceae	*Ophioglossum Vulgatum*	(You et al., 2017)	
	Psilotales	Psilotaceae	*Psilotum nudum*	(You et al., 2017)	
Gymnospermes	Pinales (conifers)	Pinaceae	*Cephalotaxus hainanensis*	GBHQ01023156.1	NCBI
	Pinales (conifers)	Pinaceae	*Pinus patula*	GECO01045160.1	NCBI
	Pinales (conifers)	Pinaceae	*Pinus monticola*	GDQD01003526.1	NCBI
	Pinales (conifers)	Pinaceae	*Pseudotsuga menziesii*	GFFY01002157.1	NCBI
	Pinales (conifers)	Pinaceae	*Pinus albicaulis*	GFLL01092022.1	NCBI
	Pinales (conifers)	Pinaceae	*Picea glauca*	GCHX01458501.1	NCBI
	Pinales (conifers)	Pinaceae	*Pinus lambertiana*	GEUZ01000427.1	NCBI
	Pinales (conifers)	Pinaceae	*Pinus massoniana*	GAQR01043425.1	NCBI
	Pinales (conifers)	Pinaceae	*Pinus canariensis*	GBLJ01047138.1	NCBI
	Pinales (conifers)	Araucariaceae	*Wollemia nobilis*	GCHU01006864.1	NCBI
	Pinales (conifers)	Araucariaceae	*Araucaria cunninghamii*	GCKF01036911.1	NCBI
	Ginkgoales	Ginkgoaceae	*Ginkgo biloba*	GBYR01026913.1 + GBKT01032776.1	NCBI
	Cycadales	Cycadaceae	*Cycas revoluta*	GBJU01018207.1 + GBJU01013174 + SRX661923	NCBI
Magnoliophytes basales	Amborellales	Amborellaceae	*Amborella trichopoda*	evm_27.model.AmTr_v1.0_scaffold00001.433	Phytozome
	Pinales (conifers)	Aristolochiaceae	*Aristolochia fimbriata*	FD757980/FD754571/FD757424/FD753531/FD758745/FD757923/ FD751341	NCBI
Monocots	Alismatales	Arecaceae	*Anthurium andraeanum*	TSA:GBKP01037740	NCBI
	Alismatales	Hydrocharitaceae	*Elodea nuttallii*	GBEN01086983/GBEN01040210	NCBI
	Alismatales	Posidoniaceae	*Posidonia oceanica*	TSA:GEMD01011778	NCBI
	Alismatales	Zosteraceae	*Zostera marina*	KMZ76115	NCBI
	Dioscoreales	Dioscoreaceae	*Dioscorea composita*	GBJW01022598/GBJW01050363	NCBI
	Dioscoreales	Dioscoreaceae	*Dioscorea zingiberensis*	GBCR01081279	NCBI
	Asparagales	Amaryllidaceae	*Allium sativum*	V260904	NCBI
	Asparagales	Amaryllidaceae	*Allium cepa*	GAAN01000053	NCBI
	Asparagales	Amaryllidaceae	*Allium fistulosum*	FX554720	NCBI
	Asparagales	Asparagaceae	*Agave deserti*	GAHT01034983 | GAHT01037312	NCBI
	Asparagales	Orchidaceae	*Phalaenopsis equestris*	WGS: APLD01038720	NCBI
	Arecales	Arecaceae	*Phoenix dactylifera*	XM_008785369	NCBI
	Arecales	Arecaceae	*Elaeis guineensis*	GBSV01006673	NCBI
	Arecales	Arecaceae	*Landoltia punctata*	TSA:GAQZ01006949	NCBI
	Arecales	Arecaceae	*Zantedeschia aethiopica*	TSA:GAOT01015084	NCBI
	Poales	Bromeliaceae	*Ananas comosus*	Aco017738	Phytozome
	Poales	Poaceae	*Oropetium thomaeum*	URT1A: LFJQ01000016	NCBI
	Poales	Poaceae	*Zea mays*	URT1A: XP_00864488	NCBI
	Poales	Poaceae	*Sorghum bicolor*	URT1A: Sobic.001G261900.1	Phytozome
	Poales	Poaceae	*Panicum hallii*	URT1A: Pahal.B02309 + TSA:JR065624	NCBI/Phytozome
	Poales	Poaceae	*Panicum virgatum*	URT1A1: Chr9a.Pavir.Ia02772 + EST FL814274.1 | URT1A2:	Phytozome
	Poales	Poaceae	*Setaria italica*	URT1A : K4A5Y1 + AGNK01005691	NCBI
	Poales	Poaceae	*Setaria viridis*	URT1A : Sevir.9G280700	Phytozome
	Poales	Poaceae	*Oryza sativa Japonica*	URT1A: AACV01021025	NCBI
	Poales	Poaceae	*Brachypodium stacei*	URT1A : Brast03G108700.1	Phytozome
	Poales	Poaceae	*Brachypodium distachyon*	URT1A: Bradi3g22908.1	Phytozome
	Zingiberales	Zingiberaceae	*Zingiber officinale*	DY375774/DY375773/DY354865	NCBI
	Ranunculales	*Ranunculaceae*	*Aquilegia coerulea*	Aquca_013_00489	Phytozome
	Ranunculales	*Berberidaceae*	*Sinopodophyllum hexandrum*	GBJY01004579 + GBJY01004578	NCBI
	proteales	*Proteaceae*	*Banksia hookeriana*	GBXB01005448	NCBI
	saxifragales	*Crassulaceae*	*Kalanchoe laxiflora*	Kalax.0099s0075 | Kalax.0017s0041 | Kalax.0032s0067 | Kalax.0327s0017	Phytozome
	vitales	*Vitaceae*	*Vitis vinifera*	XP_010661598.1	NCBI
	celastrales	*Celastraceae*	*Tripterygium wilfordii*	GAVZ01050527	NCBI
	malpighiales	*Euphorbiaceae*	*Ricinus communis*	XP_015584508.1	NCBI
	malpighiales	*Euphorbiaceae*	*Hevea brasiliensis*	AJJZ010484251	NCBI
	malpighiales	*Euphorbiaceae*	*Jatropha curcas*	XP_012083529.1	NCBI
	malpighiales	*Euphorbiaceae*	*Manihot esculenta*	Manes.05G042700.1	Phytozome
	malpighiales	*Linaceae*	*Linum usitatissimum*	Lus10015854 | Lus10009309	Phytozome

	malpighiales	*Salicaceae*	*Populus trichocarpa*	XP_002301312.2 | XP_006375316.1	NCBI
	malpighiales	*Salicaceae*	*Populus euphratica*	XP_011002088.1 / AOFL01009891 | XP_011034817.1 / AOFL010010	NCBI
	malpighiales	*Salicaceae*	*Salix purpurea*	SapurV1A.0655s0100.1 | SapurV1A.0024s0620.1	Phytozome
	malpighiales	*Salicaceae*	*Viola uliginosa*	GCAB01038156 / GCAB01038158 / GCAB01038162 / GCAB01038160	NCBI
	fabales	*Fabaceae*	*Acacia koa*	GBYE01007739	NCBI
	fabales	*Fabaceae*	*Arachis duranensis*	GBJK01051113	NCBI
	fabales	*Fabaceae*	*Arachis hypogaea*	GAER01037812	NCBI
	fabales	*Fabaceae*	*Arachis ipaensis*	GBIW01004467	NCBI
	fabales	*Fabaceae*	*Cajanus cajan*	AGCT01012539 + EZ672457	NCBI
	fabales	*Fabaceae*	*Cercis gigantea*	GAOK01035584	NCBI
	fabales	*Fabaceae*	*Cicer arietinum*	XP_004510903.1	NCBI
	fabales	*Fabaceae*	*Glycine max*	XP_003529982.1 | XP_003521938.1	NCBI
	fabales	*Fabaceae*	*Lathyrus sativus*	GBSS01010210	NCBI
	fabales	*Fabaceae*	*Lotus corniculatus*	GACB01028791/GACB01028488	NCBI
	fabales	*Fabaceae*	*Lupinus angustifolius*	AOCW01126410+GBRP01000555/GBRP01000556/GBRP01057188	NCBI
	fabales	*Fabaceae*	*Medicago truncatula*	JR373276 + APNO01003854	NCBI
	fabales	*Fabaceae*	*Phaseolus vulgaris*	XP_007134916.1	NCBI
	fabales	*Fabaceae*	*Pisum sativum*	GAMJ01001219	NCBI
	fabales	*Fabaceae*	*Senna alexandrina*	GCVR01017170 + GCVR01010478	NCBI
	fabales	*Fabaceae*	*Trifolium pratense*	FAXK01002099	NCBI
	fabales	*Fabaceae*	*Vicia faba*	GASA01007352	NCBI
	fabales	*Fabaceae*	*Vigna angularis*	KOM47856.1+ BAT97824.1	NCBI
	fabales	*Fabaceae*	*Vigna radiata*	XP_014522958.1	NCBI
	fabales	*Fabaceae*	*Zenia insignis*	GDGH01044098	NCBI
	rosales	*Cannabaceae*	*Cannabis sativa*	JP479310	NCBI
	rosales	*Cannabaceae*	*Humulus lupulus*	LA464455	NCBI
	rosales	*Moraceae*	*Morus notabilis*	XP_010106745.1	NCBI
	rosales	*Rosaceae*	*Crataegus pinnatifida*	GALU01006064	NCBI
	rosales	*Rosaceae*	*Fragaria vesca*	AEMH01003044	NCBI
	rosales	*Rosaceae*	*Malus domestica*	XP_008342981.1	NCBI
	rosales	*Rosaceae*	*Prunus armeniaca*	GAWA01072980	NCBI
	rosales	*Rosaceae*	*Prunus mume*	XP_008232824.1	NCBI
	rosales	*Rosaceae*	*Prunus persica*	XP_007220905.1	NCBI
	rosales	*Rosaceae*	*Pyrus communis*	GBXL01006357	NCBI
	rosales	*Rhamnaceae*	*Ziziphus jujuba*	JREP01016398	NCBI
	cucurbitales	*Cucurbitaceae*	*Citrullus lanatus*	AGCB01033509 + AGCB01033508	NCBI
	cucurbitales	*Cucurbitaceae*	*Cucumis sativus*	XP_011656690.1	NCBI
	cucurbitales	*Cucurbitaceae*	*Cucumis melo*	XP_008460141.1	NCBI
	cucurbitales	*Cucurbitaceae*	*Cucurbita pepo*	GBZI01010791	NCBI
	cucurbitales	*Cucurbitaceae*	*Lagenaria siceraria*	ATBX01353038 + ATBX01354226 + ATBX01356791+ ATBX0136487	NCBI
	cucurbitales	*Cucurbitaceae*	*Siraitia grosvenorii*	JL554419	NCBI
	fagales	*Betulaceae*	*Betula platyphylla*	KA241906	NCBI
	fagales	*Betulaceae*	*Corylus avellana*	KA430417+ KA423844	NCBI
	fagales	*Casuarinaceae*	*Casuarina equisetifolia*	GEBJ01034182	NCBI
	myrtales	*Myrtaceae*	*Eucalyptus camaldulensis*	BADO01119091 + BADO01119089	NCBI
	myrtales	*Myrtaceae*	*Eucalyptus grandis*	Eucgr.D02430.1, corrigé manuellement avec la séquence génomique	Phytozome
	myrtales	*Lythraceae*	*Punica granatum*	GBGR01002208	NCBI
	sapindales	*Rutaceae*	*Citrus clementina*	XP_006445207.1	NCBI
	sapindales	*Rutaceae*	*Citrus sinensis*	XP_006490961.1	NCBI
	sapindales	*Rutaceae*	*Citrus trifoliata*	GCVN01032700	NCBI
	sapindales	*Anacardiaceae*	*Mangifera indica*	GBCV01008339	NCBI
	sapindales	*Sapindaceae*	*Litchi chinensis*	GCAD01013316	NCBI
	sapindales	*Meliaceae*	*Azadirachta indica*	AMWY02051988 + AMWY02051989 + AMWY02086455 + AMWY020	NCBI
	malvales	*Malvaceeae*	*Theobroma cacao*	XP_007051991.1	NCBI
	malvales	*Malvaceeae*	*Gossypium raimondii*	XP_012437624.1 | XP_012489804.1	NCBI
	malvales	*Malvaceeae*	*Hibiscus cannabinus*	GEED01029686	NCBI
	malvales	*Malvaceeae*	*Kosteletzkya virginica*	GCJL01028853	NCBI
	malvales	*Malvaceeae*	*Corchorus capsularis*	GBSD01004910	NCBI
	malvales	*Thymelaeaceae*	*Aquilaria agallochum*	JMHV01004207	NCBI
	brassicales	*Brassicaceae*	*Arabidopsis lyrata*	AL4G44790.t1	Phytozome
	brassicales	*Brassicaceae*	*Arabidopsis thaliana*	AT2G45620.1	Phytozome
	brassicales	*Brassicaceae*	*Boechera stricta*	Bostr.25993s0212.1	Phytozome
	brassicales	*Brassicaceae*	*Brassica oleracea capitata*	Bol021741	Phytozome
	brassicales	*Brassicaceae*	*Brassica rapa*	Brara.D02756.1	Phytozome
	brassicales	*Brassicaceae*	*Capsella grandiflora*	Cagra.0050s0138.1	Phytozome
	brassicales	*Brassicaceae*	*Capsella rubella*	Carubv10024989m	Phytozome
Dicots	brassicales	*Brassicaceae*	*Eutrema salsugineum*	Thhalv10001324m	Phytozome
	brassicales	*Brassicaceae*	*Camelina sativa*	XP_010518183.1 | XP_010506546.1 | XP_010508136.1	NCBI
	brassicales	*Brassicaceae*	*Brassica napus*	BnaA04g26310D | BnaC04g50360D	Phytozome
	brassicales	*Brassicaceae*	*Raphanus raphanistrum*	JRQH01015498.1	NCBI
	brassicales	*Brassicaceae*	*Raphanus sativus*	FX668040.1	NCBI
	brassicales	*Brassicaceae*	*Arabis alpina*	KFK37353.1	NCBI
	brassicales	*Brassicaceae*	*Arabis nordmanniana*	LNCG01107004 | LNCG01107001	NCBI
	brassicales	*Brassicaceae*	*Noccaea caerulescens*	GASZ01006115	NCBI
	brassicales	*Brassicaceae*	*Thlaspi arvense*	GAKE01006259.1	NCBI
	brassicales	*Brassicaceae*	*Aethionema carneum*	GDQO01036706.1	NCBI
	brassicales	*Brassicaceae*	*Moriera spinosa*	GDRX01046722	NCBI
	brassicales	*Brassicaceae*	*Boechera gunnisoniana*	GBAD01083801.1	NCBI
	brassicales	*Brassicaceae*	*Isatis tinctoria*	GDAU01006658.1	NCBI
	brassicales	*Brassicaceae*	*Leavenworthia alabamica*	ASXC01002471	NCBI
	brassicales	*Brassicaceae*	*Sisymbrium irio*	ASZH01020081	NCBI
	brassicales	*Caricaceae*	*Carica papaya*	evm.model.supercontig_18.161	Phytozome

	brassicales	*Pentadiplandraceae*	*Pentadiplandra brazzeana*	GDGJ01000333.1	NCBI
	brassicales	*Cleomaceae*	*Cleome droserifolia*	GDRJ01039446.1	NCBI
	brassicales	*Cleomaceae*	*Tarenaya hassleriana*	GFML01025624.1	NCBI
	brassicales	*Koeberliniaceae*	*Koeberlinia spinosa*	GDGB01000307.1	NCBI
	Caryophyllales	*Aizoaceae*	*Mesembryanthemum crystallin*	GBLK01003189	NCBI
	Caryophyllales	*Cactaceae*	*Opuntia ficus-indica*	GEUK01018652.1	NCBI
	Caryophyllales	*Amarathaceae*	*Amaranthus hypochondriacus*	AHYPO_017165-RA	Phytozome
	Caryophyllales	*Amarathaceae*	*Amaranthus tricolor*	GAPZ01004765	NCBI
	Caryophyllales	*Amarathaceae / Caryophyl*	*Beta vulgaris*	JP487242.1	NCBI
	Caryophyllales	*Amarathaceae*	*Bienertia sinuspersici*	A9W2KF0201R	NCBI
	Caryophyllales	*Amarathaceae*	*Spinacia oleracea*	XM_022000808.1	NCBI
	Caryophyllales	*Amarathaceae*	*Suaeda fruticosa*	GCVE01025130.1	NCBI
	Caryophyllales	*Caryophyllaceae*	*Colobanthus quitensis*	GCIB01008204 + GCIB01038727	NCBI
	Caryophyllales	*Caryophyllaceae*	*Dianthus caryophyllus*	BAUD01000037.1	NCBI
	Caryophyllales	*Caryophyllaceae*	*Silene latifolia*	LHUT01005313.1	NCBI
	Caryophyllales	*Tamaricaceae*	*Reaumuria trigyna*	JR252802.1 + JR247381.1	NCBI
	Caryophyllales	*Chenopodiaceae*	*Oxybasis rubra*	GEEQ01023290.1	NCBI
	Caryophyllales	*Nepenthaceae*	*Nepenthes ampullaria*	GFAE01077501.1	NCBI
	Caryophyllales	*Polygonaceae*	*Persicaria minor*	GEFX01093057	NCBI
	Caryophyllales	*Polygonaceae*	*Rumex palustris*	HAAL01031596	NCBI
	Ericales	*Actinidiaceae*	*Actinidia chinensis*	AONS01016125	NCBI
	Ericales	*Actinidiaceae*	*Actinidia deliciosa*	GEYI01045322	NCBI
	Ericales	*Ebenaceae*	*Diospyros lotus*	GBSJ01303514 + GBSJ01387439 + GBSJ01264820 + GBSJ0138917	NCBI
	Ericales	*Ericaceae*	*Opuntia ficus-indica*	GBFL01023990.1	NCBI
	Ericales	*Primulaceae*	*Primula veris*	JTKG01000242	NCBI
	Ericales	*Theaceae*	*Camellia sinensis*	GBKQ01019767 + GARM01007214 + GARM01007215	NCBI
	Ericales	*Polemoniaceae*	*Phlox drummondii*	GFFR01032505	NCBI
	Cornales	*Cornaceae*	*Camptotheca acuminata*	GACF01049846	NCBI
	Solanales	*Convolvulaceae*	*Ipomoea purpurea*	GABG01006718	NCBI
	Solanales	*Convolvulaceae*	*Cuscuta pentagona*	GAON01258943	NCBI
	Solanales	*Solanaceae*	*Capsicum annuum*	XP_016574627.1	NCBI
	Solanales	*Solanaceae*	*Nicotiana tomentosiformis*	XP_009626263	NCBI
	Solanales	*Solanaceae*	*Nicotiana sylvestris*	XP_009799247.1	NCBI
	Solanales	*Solanaceae*	*Solanum chacoense*	GEDG01002903	NCBI
	Solanales	*Solanaceae*	*Solanum pennellii*	XP_015062211.1	NCBI
	Solanales	*Solanaceae*	*Solanum lycopersicum*	XP_004229872.1	NCBI
	Solanales	*Solanaceae*	*Solanum torvum*	GBEG01030054.1	NCBI
	Solanales	*Solanaceae*	*Solanum tuberosum*	XP_015167051.1	NCBI
	Solanales	*Solanaceae*	*Solanum melongena*	GBGZ01084204.1	NCBI
	Solanales	*Solanaceae*	*Withania somnifera*	GBHJ01033800	NCBI
	Gentianales	*Apocynaceae*	*Calotropis procera*	GBZK01029009	NCBI
	Gentianales	*Apocynaceae*	*Catharanthus roseus*	GACD01036855.1	NCBI
	Gentianales	*Apocynaceae*	*Rauvolfia serpentina*	GACE01041214	NCBI
	Gentianales	*Apocynaceae*	*Rhazya stricta*	GAMW01010160.1	NCBI
	Gentianales	*Rubiaceae*	*Coffea canephora*	CDP20563.1	NCBI
	Gentianales	*Rubiaceae*	*Gardenia jasminoides*	GAQP01055138	NCBI
	Gentianales	*Rubiaceae*	*Neolamarckia cadamba*	GASC01025167 + GASC01053571	NCBI
	Lamiales	*Acanthaceae*	*Andrographis paniculata*	GBJB01014055	NCBI
	Lamiales	*Gesneriaceae*	*Sinningia speciosa*	GCZU01001875	NCBI
	Lamiales	*Lamiaceae*	*Ocimum tenuiflorum*	AYJT01081657 | AYJT01044637	NCBI
	Lamiales	*Lentibulariaceae*	*Genlisea aurea*	EPS59851.1	NCBI
	Lamiales	*Oleaceae*	*Fraxinus excelsior*	CBXU010006534 | CBXU010025724	NCBI
	Lamiales	*Oleaceae*	*Olea europaea*	GCJV01018264 + GABQ01080245	NCBI
	Lamiales	*Pedaliaceae*	*Sesamum indicum*	XP_011091205.1 | XP_011093188.1	NCBI
	Lamiales	*Phrymaceae*	*Erythranthe guttata*	XP_012843633.1	NCBI
	Apiales	*Araliaceae*	*Panax ginseng*	GDQW01116961 + GDQW01116959	NCBI
	Asterales	*Asteraceae*	*Chromolaena odorata*	GACH01014035	NCBI
	Asterales	*Asteraceae*	*Cynara cardunculus*	KVH91189.1	NCBI
	Asterales	*Asteraceae*	*Dahlia pinnata*	KVH91189.1	NCBI
	Asterales	*Asteraceae*	*Gerbera hybrid*	GACN01010715	NCBI
	Asterales	*Asteraceae*	*Karelinia caspia*	GANI01031186	NCBI
	Asterales	*Asteraceae*	*Lactuca sativa*	JI577008	NCBI
	Asterales	*Asteraceae*	*Lactuca serriola*	JO023605	NCBI

Annexe 5 | Liste des protéines enrichies dans les IP de DCP5-GFP.

Le Log2(FoldChange), la p-value ajusté, ainsi que le nombre de spectres moyens sur les 4 IP contrôles et les 4 IP DCP5-GFP sont indiqués pour chaque protéine co-immunoprécipitants avec DCP5-GFP. Seuls les protéines ayant un FC > 2 et un une p-value ajustée (adjp) <0.05 sont indiquées ici. Les protéines sont classées par ordres de p-value ajustée.

AGI	LogFC	adjp	Spect. Ctrl	Spect. IP	Nom	description
AT2G42520.1	4.531	1.33E-96	4	184	RH37	DEAD-box ATP-dependent RNA helicase 37 [Source:UniProtKB/Swiss-Prot;Acc:Q84W89]
AT3G22330.1	9.379	1.75E-56	0	104.75	RH53	DEAD-box ATP-dependent RNA helicase 53, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9LUW5]
AT3G58570.1	4.610	1.59E-53	3.6	179.25	RH52	DEAD-box ATP-dependent RNA helicase 52 [Source:UniProtKB/Swiss-Prot;Acc:Q9M2F9]
AT1G26110.1	10.932	6.15E-52	0	322.5	DCP5	Protein decapping 5 [Source:UniProtKB/Swiss-Prot;Acc:Q9C658]
AT5G40490.1	9.556	1.21E-35	0	113.75		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:Q9FM47]
AT4G00660.2	7.207	1.06E-32	0.6	209.25	RH8	DEAD-box ATP-dependent RNA helicase 8 [Source:UniProtKB/Swiss-Prot;Acc:Q8RXK6]
AT5G61020.1	8.102	1.31E-28	0	46	ECT3	ECT3 [Source:UniProtKB/TrEMBL;Acc:A0A178UM32]
AT1G43190.1	7.889	3.52E-26	0	39	PTB3	Polypyrimidine tract-binding protein homolog 3 [Source:UniProtKB/Swiss-Prot;Acc:Q6ICX4]
AT3G57150.1	3.397	1.87E-25	3.2	68.25	CBF5	Uncharacterized protein At3g57150 (Fragment) [Source:UniProtKB/TrEMBL;Acc:C0SVF3]
AT2G45810.1	6.278	9.86E-25	1.2	210	RH6	DEAD-box ATP-dependent RNA helicase 6 [Source:UniProtKB/Swiss-Prot;Acc:Q94BV4]
AT2G34160.1	8.094	1.57E-24	0	44.5		Uncharacterized protein At2g34160 [Source:UniProtKB/Swiss-Prot;Acc:O22969]
AT1G48410.2	4.560	1.03E-23	2.2	111.25	AGO1	ICU9 [Source:UniProtKB/TrEMBL;Acc:A0A178WL72]
AT4G28440.1	6.728	9.50E-23	0.2	60.5		Uncharacterized protein At4g28440 [Source:UniProtKB/Swiss-Prot;Acc:O49453]
AT5G04280.1	8.178	1.25E-21	0	47	RZ1C	Glycine-rich RNA-binding protein RZ1C [Source:UniProtKB/Swiss-Prot;Acc:Q8RWN5]
AT5G63120.2	7.090	2.51E-21	0.2	77.75	RH30	DEAD-box ATP-dependent RNA helicase 30 [Source:UniProtKB/Swiss-Prot;Acc:Q8W4R3]
AT4G34980.1	4.471	8.83E-21	2.4	113	SBT1.6	Subtilisin-like protease SBT1.6 [Source:UniProtKB/Swiss-Prot;Acc:O49607]
AT5G47010.1	8.695	1.04E-20	0	70	UPF1	Regulator of nonsense transcripts 1 homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9FJR0]
AT4G02290.1	3.579	1.31E-20	2	50.25	AtGH9B13	Endoglucanase 17 [Source:UniProtKB/Swiss-Prot;Acc:O81416]
AT5G51750.1	4.570	1.31E-20	1.8	88.25	SBT1.3	Subtilisin-like protease SBT1.3 [Source:UniProtKB/Swiss-Prot;Acc:Q9FLI4]
AT3G60190.1	3.963	1.68E-20	1.8	57.75	DRP1E	Dynamin-related protein 1E [Source:UniProtKB/Swiss-Prot;Acc:Q9FNX5]
AT3G61240.1	6.086	6.92E-20	1.2	185.5	RH12	DEAD-box ATP-dependent RNA helicase 12 [Source:UniProtKB/Swiss-Prot;Acc:Q9M2E0]
AT3G06480.1	5.407	1.65E-19	0.4	42	RH40	DEAD-box ATP-dependent RNA helicase 40 [Source:UniProtKB/Swiss-Prot;Acc:Q9SQV1]
AT3G22310.1	8.318	3.97E-19	0	50.5	RH9	DEAD-box ATP-dependent RNA helicase 9 [Source:UniProtKB/Swiss-Prot;Acc:Q9LUW6]
AT1G33430.2	7.573	6.92E-19	0	31		Hexosyltransferase [Source:UniProtKB/TrEMBL;Acc:F4HR76]
AT3G01540.2	5.725	1.32E-18	0.6	77.5	RH14	DEAD-box ATP-dependent RNA helicase 14 [Source:UniProtKB/Swiss-Prot;Acc:Q8H136]
AT1G29250.1	5.336	1.32E-18	1.4	97		Alba DNA/RNA-binding protein [Source:UniProtKB/TrEMBL;Acc:Q9LP53]
AT1G01300.1	3.917	4.58E-17	2.2	69.25	APF2	Aspartyl protease family protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9LNJ3]
AT2G06850.1	2.926	8.25E-17	4.2	67	XTH4	Xyloglucan endotransglucosylase/hydrolase protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q39099]
AT1G76010.1	5.424	9.24E-17	0.6	59.75		Alba DNA/RNA-binding protein [Source:UniProtKB/TrEMBL;Acc:Q93VA8]
AT4G25630.1	3.087	5.93E-16	3.6	57	MED36A	Mediator of RNA polymerase II transcription subunit 36a [Source:UniProtKB/Swiss-Prot;Acc:Q94AH9]
AT3G20820.1	2.699	7.07E-16	7.6	102.25		Leucine-rich repeat (LRR) family protein [Source:UniProtKB/TrEMBL;Acc:Q9LT39]
AT1G20220.1	4.480	8.91E-16	0.8	41		Alba DNA/RNA-binding protein [Source:UniProtKB/TrEMBL;Acc:Q944A2]
AT5G52470.1	3.171	1.15E-15	3.6	61.5	MED36B	Probable mediator of RNA polymerase II transcription subunit 36b [Source:UniProtKB/Swiss-Prot;Acc:Q9FEF8]
AT2G27100.1	8.804	1.28E-15	0	74.5	SE	SE [Source:UniProtKB/TrEMBL;Acc:A0A178VZD4]
AT3G07030.1	7.284	1.65E-15	0	25.75		Alba DNA/RNA-binding protein [Source:UniProtKB/TrEMBL;Acc:A0A1I9LSC0]
AT5G53060.1	7.814	1.67E-15	0	35.5	RCF3	RNA-binding KH domain-containing protein RCF3 [Source:UniProtKB/Swiss-Prot;Acc:Q8W4B1]
AT2G27880.1	7.197	1.67E-15	0	25.5	AGO5	Protein argonaute 5 [Source:UniProtKB/Swiss-Prot;Acc:Q9SJK3]
AT3G13460.1	4.926	1.67E-15	0.8	56	ECT2	ECT2 [Source:UniProtKB/TrEMBL;Acc:A0A178VBC2]
AT3G58510.1	4.325	1.67E-15	4.2	186.75	RH11	DEAD-box ATP-dependent RNA helicase 11 [Source:UniProtKB/Swiss-Prot;Acc:Q8LA13]
AT1G14830.1	5.415	1.71E-15	1	72.25	DRP1C	DRP1C [Source:UniProtKB/TrEMBL;Acc:A0A178W4V6]
AT1G31280.1	8.135	4.79E-15	0	47.5	AGO2	Protein argonaute 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9SHF3]
AT4G21670.1	7.837	4.87E-15	0	38.5	CPL1	RNA polymerase II C-terminal domain phosphatase-like 1 [Source:UniProtKB/Swiss-Prot;Acc:Q5YDB6]
AT5G63420.1	7.055	5.51E-15	0	23.5	emb2746	RNA-metabolising metallo-beta-lactamase family protein [Source:UniProtKB/TrEMBL;Acc:Q84W56]
AT5G62390.1	3.596	5.51E-15	4.8	103.25	BAG7	BAG family molecular chaperone regulator 7 [Source:UniProtKB/Swiss-Prot;Acc:Q9LVA0]
AT3G13224.2	7.254	6.33E-15	0	23.75		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:Q9LHL2]
AT1G49760.1	5.435	6.66E-15	0.4	43.25	PAB8	Polyadenylate-binding protein 8 [Source:UniProtKB/Swiss-Prot;Acc:Q9FXA2]
AT3G25150.2	5.249	1.21E-14	0.4	38.25		Nuclear transport factor 2 (NTF2) family protein with RNA binding (RRM-RBD-RNP motifs) domain-containing prote
AT5G65260.1	7.105	1.65E-14	0	22.75	PABN2	Polyadenylate-binding protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9FJN9]
AT1G59610.1	3.492	1.71E-14	4.2	88.25	DRP2B	Dynamin-2B [Source:UniProtKB/Swiss-Prot;Acc:Q9LQ55]
AT5G58470.1	7.784	1.87E-14	0	33.75	TAF15B	Transcription initiation factor TFIID subunit 15b [Source:UniProtKB/Swiss-Prot;Acc:Q94KD0]
AT5G42080.1	3.062	3.13E-14	7.2	118.25	DRP1A	RSW9 [Source:UniProtKB/TrEMBL;Acc:A0A178UEJ4]
AT1G70710.1	7.456	3.23E-14	0	29	CEL1	Endoglucanase 8 [Source:UniProtKB/Swiss-Prot;Acc:Q9CAC1]
AT2G42570.1	7.021	5.50E-14	0	21.5	TBL39	Protein trichome birefringence-like 39 [Source:UniProtKB/Swiss-Prot;Acc:Q9SIN2]
AT2G26280.1	7.022	5.56E-14	0	22	CID7	Polyadenylate-binding protein-interacting protein 7 [Source:UniProtKB/Swiss-Prot;Acc:O64843]
AT2G17870.1	5.749	7.70E-14	0.2	29	CSP3	Cold shock domain-containing protein 3 [Source:UniProtKB/Swiss-Prot;Acc:Q94C69]
AT5G48650.1	7.592	1.21E-13	0	31.5		Nuclear transport factor 2 (NTF2) family protein with RNA binding (RRM-RBD-RNP motifs) domain-containing prote
AT3G03950.3	6.824	1.57E-12	0	19.25	ECT1	AT3G03950 protein [Source:UniProtKB/TrEMBL;Acc:Q3MK94]
AT5G37720.1	7.241	2.21E-12	0	24.25	ALY4	THO complex subunit 4D [Source:UniProtKB/Swiss-Prot;Acc:Q6NQ72]
AT1G75660.1	7.904	4.55E-12	0	40.25	XRN3	5'-3' exoribonuclease 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9FQ03]
AT2G33845.1	6.824	4.57E-12	0	18.75		Expressed protein [Source:UniProtKB/TrEMBL;Acc:Q8RYC3]
AT2G34810.1	6.726	7.67E-12	0	18.75		Berberine bridge enzyme-like 16 [Source:UniProtKB/Swiss-Prot;Acc:O64745]
AT5G04430.2	5.274	9.86E-12	0.2	22.5	BTR1	binding to TOMV RNA 1L (long form) [Source:TAIR;Acc:AT5G04430]
AT3G13860.1	7.628	1.21E-11	0	33.25	HSP60-3A	Chaperonin CPN60-like 2, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q93ZM7]
AT1G55150.1	6.775	1.26E-11	0	17.75		DEA(D/H)-box RNA helicase family protein [Source:TAIR;Acc:AT1G55150]
AT3G13060.2	6.685	1.26E-11	0	19	ECT5	Evolutionarily conserved C-terminal region 5 [Source:UniProtKB/TrEMBL;Acc:Q0WR25]
AT5G03740.1	6.657	1.26E-11	0	17.25	HDT3	HDT3 [Source:UniProtKB/TrEMBL;Acc:A0A178UP12]
AT4G22010.1	2.797	1.26E-11	5.2	77.5	sks4	At4g22010 [Source:UniProtKB/TrEMBL;Acc:O65449]
AT1G64390.1	7.446	1.65E-11	0	29.25	AtGH9C2	Endoglucanase 6 [Source:UniProtKB/Swiss-Prot;Acc:Q42059]
AT5G14610.1	5.032	1.73E-11	0.6	49		DEAD box RNA helicase family protein [Source:UniProtKB/TrEMBL;Acc:F4K6V1]
AT5G26742.2	4.005	2.19E-11	5.6	224.75	RH3	RH3 [Source:UniProtKB/TrEMBL;Acc:A0A178UT03]
AT2G36870.1	6.745	3.40E-11	0	17.5	XTH32	Probable xyloglucan endotransglucosylase/hydrolase protein 32 [Source:UniProtKB/Swiss-Prot;Acc:Q9SJL9]
AT4G38680.1	4.679	4.46E-11	1.6	128	CSP2	GRP2 [Source:UniProtKB/TrEMBL;Acc:A0A178UYV3]
AT1G17580.1	6.559	4.86E-11	0	17	XI-1	Myosin-5 [Source:UniProtKB/Swiss-Prot;Acc:Q39160]
AT1G04680.1	6.577	4.90E-11	0	16		Pectate lyase [Source:UniProtKB/TrEMBL;Acc:A0A178W464]
AT5G44780.1	6.763	5.04E-11	0	19.75	MORF4	Multiple organellar RNA editing factor 4, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:O48582]
AT5G25980.2	2.176	5.04E-11	75	656	TGG2	Myrosinase 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9C5C2]
AT4G30190.2	2.769	6.11E-11	2.4	34.25	HA2	Plasma membrane ATPase [Source:UniProtKB/TrEMBL;Acc:F4JPJ7]
AT3G15000.1	6.389	1.04E-10	0	13.75	MORF8	Multiple organellar RNA editing factor 8, chloroplastic/mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9LKA5]
AT5G20490.1	7.345	1.30E-10	0	29.25	XI-K	Myosin-17 [Source:UniProtKB/Swiss-Prot;Acc:F4K5J1]
AT3G26420.1	6.876	1.39E-10	0	19	RZ1A	Glycine-rich RNA-binding protein RZ1A [Source:UniProtKB/Swiss-Prot;Acc:Q9LIN3]
AT4G27060.1	6.927	2.93E-10	0	21.5	TOR1	Microtubule-associated protein TORTIFOLIA1 [Source:UniProtKB/Swiss-Prot;Acc:Q9T041]
AT2G18960.1	2.715	5.13E-10	3.2	44.25	AHA1	Plasma membrane ATPase [Source:UniProtKB/TrEMBL;Acc:A0A178VSN5]
AT4G29010.1	4.409	5.68E-10	0.8	40.25	AIM1	Peroxisomal fatty acid beta-oxidation multifunctional protein AIM1 [Source:UniProtKB/Swiss-Prot;Acc:Q9ZPI6]
AT2G23350.1	4.589	5.96E-10	0.6	36	PAB4	Polyadenylate-binding protein 4 [Source:UniProtKB/Swiss-Prot;Acc:O22173]
AT2G44710.1	6.938	5.96E-10	0	21.5		At2g44720/F16B22.21 [Source:UniProtKB/TrEMBL;Acc:Q8RWQ1]
AT3G06810.1	4.098	6.05E-10	0.6	25.75	IBR3	Probable acyl-CoA dehydrogenase IBR3 [Source:UniProtKB/Swiss-Prot;Acc:Q8RWZ3]
AT3G12800.1	4.275	7.59E-10	0.4	18.75	SDRB	Uncharacterized protein At3g12790 [Source:UniProtKB/TrEMBL;Acc:Q0WMN4]
AT1G27090.1	5.506	9.04E-10	1.4	177		Glycine-rich protein [Source:UniProtKB/TrEMBL;Acc:Q9LFX8]
AT4G26450.1	7.187	9.25E-10	0	25		Uncharacterized protein At4g26450 [Source:UniProtKB/Swiss-Prot;Acc:P0CB21]
AT2G21060.1	4.433	9.50E-10	0.6	31.25	CSP4	GRP2B [Source:UniProtKB/TrEMBL;Acc:A0A178VU21]
AT5G55670.1	6.495	1.00E-09	0	15		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:Q9FM71]
AT4G31770.1	6.905	1.22E-09	0	19.25	DBR1	Lariat debranching enzyme [Source:UniProtKB/Swiss-Prot;Acc:Q94K01]
AT3G07010.1	6.726	1.22E-09	0	16.75		Probable pectate lyase 8 [Source:UniProtKB/Swiss-Prot;Acc:Q9M8Z8]

AT4G36020.1	6.455	1.48E-09	0	15	CSP1	Cold shock protein 1 [Source:UniProtKB/Swiss-Prot;Acc:O65639]
AT1G10290.1	3.663	1.86E-09	5	94.5	DRP2A	DRP2A [Source:UniProtKB/TrEMBL;Acc:A0A178WBB3]
AT1G14710.1	7.147	1.93E-09	0	24.25		Hydroxyproline-rich glycoprotein family protein [Source:UniProtKB/TrEMBL;Acc:F4HWB0]
AT3G61820.1	3.702	2.07E-09	0.8	23.25		Eukaryotic aspartyl protease family protein [Source:UniProtKB/TrEMBL;Acc:Q9M356]
AT1G48480.1	6.362	2.34E-09	0	14.5	RKL1	Probable inactive receptor kinase At1g48480 [Source:UniProtKB/Swiss-Prot;Acc:Q9LP77]
AT3G60240.4	7.555	3.12E-09	0	33.5	EIF4G	Eukaryotic translation initiation factor 4G [Source:UniProtKB/Swiss-Prot;Acc:Q76E23]
AT1G72970.1	3.765	3.22E-09	0.8	26	HTH	HTH [Source:UniProtKB/TrEMBL;Acc:A0A178WG21]
AT5G27120.1	4.012	3.80E-09	0.6	20.75	NOP5-1	Probable nucleolar protein 5-1 [Source:UniProtKB/Swiss-Prot;Acc:O04658]
AT2G43970.1	6.731	3.97E-09	0	17.75	LARP6B	La-related protein 6B [Source:UniProtKB/Swiss-Prot;Acc:O80567]
AT3G13300.1	4.203	4.20E-09	0.6	26.5	VCS	Enhancer of mRNA-decapping protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT8]
AT2G32080.1	6.351	4.90E-09	0	13.5	PURA1	PUR ALPHA-1 [Source:UniProtKB/TrEMBL;Acc:A0A178VXA3]
AT2G29190.1	6.230	4.90E-09	0	13.5	APUM2	Pumilio homolog 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9ZW06]
AT5G15270.2	6.301	9.92E-09	0	13		RNA-binding KH domain-containing protein [Source:UniProtKB/TrEMBL;Acc:Q0WNX3]
AT2G29140.1	6.193	1.35E-08	0	13	APUM3	Pumilio homolog 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9ZW02]
AT1G10170.1	6.764	1.54E-08	0	19.75	ATNFXL1	NF-X-like 1 [Source:TAIR;Acc:AT1G10170]
AT4G25550.1	4.434	1.90E-08	0.6	33	CFIS2	Pre-mRNA cleavage factor Im 25 kDa subunit 2 [Source:UniProtKB/Swiss-Prot;Acc:Q8GXS3]
AT4G24780.1	6.418	1.98E-08	0	15		Probable pectate lyase 18 [Source:UniProtKB/Swiss-Prot;Acc:Q9C5M8]
AT1G32790.2	6.244	2.12E-08	0	12.5	CID11	CTC-interacting domain 11 [Source:UniProtKB/TrEMBL;Acc:F4IEC7]
AT5G64960.1	6.405	2.29E-08	0	14.25	CDKC-2	Cyclin-dependent kinase C-2 [Source:UniProtKB/Swiss-Prot;Acc:Q8W4P1]
AT5G48900.1	6.375	2.29E-08	0	13.75		Probable pectate lyase 20 [Source:UniProtKB/Swiss-Prot;Acc:Q93WF1]
AT3G15010.1	6.163	2.96E-08	0	12	UBA2C	UBP1-associated protein 2C [Source:UniProtKB/Swiss-Prot;Acc:Q9LKA4]
AT3G54230.2	6.331	3.01E-08	0	13.5	SUA	suppressor of abi3-5 [Source:TAIR;Acc:AT3G54230]
AT3G04610.1	6.518	3.34E-08	0	15.75	FLK	Flowering locus K homology domain [Source:UniProtKB/Swiss-Prot;Acc:Q9SR13]
AT4G39260.1	2.802	4.63E-08	6.2	81	RBG8	Glycine-rich RNA-binding protein 8 [Source:UniProtKB/Swiss-Prot;Acc:Q03251]
AT1G16720.1	6.239	4.69E-08	0	14	HCF173	high chlorophyll fluorescence phenotype 173 [Source:TAIR;Acc:AT1G16720]
AT3G06980.1	6.167	5.21E-08	0	12	RH50	DEAD-box ATP-dependent RNA helicase 50 [Source:UniProtKB/Swiss-Prot;Acc:Q8GUG7]
AT4G08350.1	7.205	5.31E-08	0	27.75	GTA02	Putative transcription elongation factor SPT5 homolog 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9STN3]
AT1G75940.1	2.774	5.59E-08	7.2	102.75	BGLU20	Beta-glucosidase 20 [Source:UniProtKB/Swiss-Prot;Acc:Q84WV2]
AT2G44590.3	6.120	5.77E-08	0	13	DRP1D	At2g44590 [Source:UniProtKB/TrEMBL;Acc:B5X4Z5]
AT2G29200.1	6.056	6.33E-08	0	11.75	APUM1	Pumilio homolog 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9ZW07]
ATCG00800.1	6.060	8.23E-08	0	11.5	RPS3	30S ribosomal protein S3, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P56798]
AT3G51950.1	6.023	8.88E-08	0	11.25		Zinc finger CCCH domain-containing protein 46 [Source:UniProtKB/Swiss-Prot;Acc:Q9SV09]
AT2G28950.1	6.177	8.92E-08	0	12	EXPA6	EXPA6 [Source:UniProtKB/TrEMBL;Acc:A0A178VZL6]
AT3G13290.1	3.912	9.47E-08	0.4	15.75	VCR	Varicose-related protein [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT9]
AT3G04620.1	6.324	9.58E-08	0	13		Alba DNA/RNA-binding protein [Source:UniProtKB/TrEMBL;Acc:Q9SR12]
AT5G24710.1	6.129	9.98E-08	0	11.25		Transducin/WD40 repeat-like superfamily protein [Source:UniProtKB/TrEMBL;Acc:F4KIH8]
AT4G24270.2	6.018	1.03E-07	0	11.5	EMB140	EMBRYO DEFECTIVE 140 [Source:UniProtKB/TrEMBL;Acc:F4JQ75]
AT2G15560.1	6.035	1.16E-07	0	11		At2g15560 [Source:UniProtKB/TrEMBL;Acc:Q6NQ97]
AT5G13590.1	6.033	1.28E-07	0	11.75		unknown protein; Has 150 Blast hits to 121 proteins in 42 species: Archae - 0; Bacteria - 8; Metazoa - 80; Fungi - 5
AT1G71770.1	6.306	1.63E-07	0	13.25	PAB5	Polyadenylate-binding protein 5 [Source:UniProtKB/Swiss-Prot;Acc:Q05196]
AT1G31850.1	6.115	1.63E-07	0	11		Probable methyltransferase PMT20 [Source:UniProtKB/Swiss-Prot;Acc:Q9C6S7]
AT1G30680.1	6.048	1.64E-07	0	11		Twinkle homolog protein, chloroplastic/mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:B5X582]
AT1G30350.1	6.484	1.64E-07	0	14.5		Probable pectate lyase 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9C8G4]
AT1G01100.1	6.275	1.64E-07	0	12	RPP1A	60S acidic ribosomal protein P1-1 [Source:UniProtKB/Swiss-Prot;Acc:Q8LCW9]
AT1G13190.1	6.456	1.68E-07	0	14.75		F3F19.21 protein [Source:UniProtKB/TrEMBL;Acc:Q9SAF2]
AT3G05060.1	2.308	1.74E-07	4.6	40	NOP5-2	Probable nucleolar protein 5-2 [Source:UniProtKB/Swiss-Prot;Acc:Q9MAB3]
AT4G39680.1	6.195	1.82E-07	0	12.75		SAP domain-containing protein [Source:UniProtKB/TrEMBL;Acc:O65655]
AT1G75930.1	3.523	2.63E-07	1.2	28.25	EXL6	GDSL esterase/lipase EXL6 [Source:UniProtKB/Swiss-Prot;Acc:Q93X94]
AT5G59950.5	4.247	2.97E-07	0.4	19.25		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:F4JXE3]
AT4G26630.1	6.221	4.19E-07	0	13		DEK domain-containing chromatin associated protein [Source:UniProtKB/TrEMBL;Acc:Q9SUA1]
AT1G33680.1	6.591	4.20E-07	0	15.75		KH domain-containing protein [Source:UniProtKB/TrEMBL;Acc:Q0WLY0]
AT5G43810.1	5.918	4.26E-07	0	11.25	AGO10	Protein argonaute 10 [Source:UniProtKB/Swiss-Prot;Acc:Q9XGW1]
AT4G08950.1	5.910	5.17E-07	0	10.25	EXO	Protein EXORDIUM [Source:UniProtKB/Swiss-Prot;Acc:Q9ZPE7]
AT5G57110.1	5.866	5.24E-07	0	10.25	ACA8	Calcium-transporting ATPase [Source:UniProtKB/TrEMBL;Acc:Q0WV19]
AT5G08610.1	6.078	5.38E-07	0	11.25	RH26	PDE340 [Source:UniProtKB/TrEMBL;Acc:A0A178UFQ6]
AT2G21660.1	2.324	5.39E-07	15.8	183.5	RBG7	Glycine-rich RNA-binding protein 7 [Source:UniProtKB/Swiss-Prot;Acc:Q03250]
AT3G61690.1	6.191	5.66E-07	0	12.75		Nucleotidyltransferase [Source:UniProtKB/TrEMBL;Acc:F4JFF2]
AT4G21710.1	6.135	5.82E-07	0	13	NRPB2	DNA-directed RNA polymerase II subunit 2 [Source:UniProtKB/Swiss-Prot;Acc:P38420]
AT1G13020.1	5.383	6.56E-07	0.2	23.5	EIF4B2	eIF4B2 [Source:UniProtKB/TrEMBL;Acc:A0A178W345]
AT1G07310.1	6.303	7.03E-07	0	12.75		Calcium-dependent lipid-binding (CaLB domain) family protein [Source:UniProtKB/TrEMBL;Acc:Q9LNV0]
AT3G07810.2	5.810	7.55E-07	0	10.25		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:F4JFN7]
AT1G69250.1	5.835	7.93E-07	0	10.5		At1g69250 [Source:UniProtKB/TrEMBL;Acc:Q9LDI9]
AT3G01780.1	5.846	8.00E-07	0	9.75	TPLATE	Protein TPLATE [Source:UniProtKB/Swiss-Prot;Acc:F4J8D3]
AT1G50920.1	5.787	8.76E-07	0	9.5		Nucleolar GTP-binding protein 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9C6I8]
AT5G26670.1	5.941	9.50E-07	0	9.75	PAE10	Pectin acetylesterase 10 [Source:UniProtKB/Swiss-Prot;Acc:Q66GM8]
AT2G33410.1	4.688	1.30E-06	0.2	14.75		Putative RNA-binding protein [Source:UniProtKB/TrEMBL;Acc:O22791]
AT1G13730.1	5.821	1.36E-06	0	9.5		At1g13730 [Source:UniProtKB/TrEMBL;Acc:Q9LMX6]
AT1G14840.1	6.808	1.38E-06	0	19	MAP70.4	Microtubule-associated protein 70-4 [Source:UniProtKB/Swiss-Prot;Acc:Q9LQU7]
AT1G22760.1	6.781	1.44E-06	0	18.5	PAB3	Polyadenylate-binding protein 3 [Source:UniProtKB/Swiss-Prot;Acc:O64380]
AT3G52500.1	5.847	1.44E-06	0	9.5		At3g52500 [Source:UniProtKB/TrEMBL;Acc:Q9SVD1]
AT2G38110.1	4.349	1.44E-06	0.2	11.75	GPAT6	Glycerol-3-phosphate 2-O-acyltransferase 6 [Source:UniProtKB/Swiss-Prot;Acc:O80437]
AT5G52530.1	5.718	1.75E-06	0	9.5		Dentin sialophosphoprotein-like protein [Source:UniProtKB/TrEMBL;Acc:Q9FGR1]
AT4G09040.1	5.865	1.81E-06	0	9.75		At4g09040 [Source:UniProtKB/TrEMBL;Acc:Q6NPL0]
AT2G39990.1	6.048	2.05E-06	0	10.75	TIF3F1	Eukaryotic translation initiation factor 3 subunit F [Source:UniProtKB/Swiss-Prot;Acc:O04202]
480.1-AT2G33370.1-AT3G	4.360	2.05E-06	0.2	11.75	NA	NA
AT2G31810.1	6.044	2.07E-06	0	11.5		Acetolactate synthase small subunit 2, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q93YZ7]
\T4G37510.1-AT2G48121.1	5.760	2.29E-06	0	9.5	NA	NA
AT2G13540.1	5.996	2.36E-06	0	10.75	ABH1	Nuclear cap-binding protein subunit 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SIU2]
AT2G28790.1	3.398	2.47E-06	0.8	19		Pathogenesis-related thaumatin superfamily protein [Source:TAIR;Acc:AT2G28790]
AT4G03110.1	5.967	3.01E-06	0	10.25	BRN1	RNA-binding protein BRN1 [Source:UniProtKB/Swiss-Prot;Acc:Q8LFS6]
AT3G44110.1	3.402	3.47E-06	0.6	15.25	ATJ3	DnaJ protein homolog atj3 [Source:UniProtKB/TrEMBL;Acc:Q0WW92]
AT2G43680.1	5.883	3.52E-06	0	10	IQD14	IQD14 [Source:UniProtKB/TrEMBL;Acc:A0A178VRE6]
AT4G10610.1	5.829	3.56E-06	0	9.25	CID12	RBP37 [Source:UniProtKB/TrEMBL;Acc:A0A178UWM0]
AT5G67240.1	5.826	3.56E-06	0	10.25	SDN3	SDN3 [Source:UniProtKB/TrEMBL;Acc:A0A178UGK8]
AT5G49360.1	5.630	3.56E-06	0	9.25	BXL1	Beta-D-xylosidase 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9FGY1]
AT5G12940.1	5.701	3.76E-06	0	9.25		Leucine-rich repeat (LRR) family protein [Source:UniProtKB/TrEMBL;Acc:Q9LXU5]
AT3G21100.2	5.662	3.82E-06	0	9		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:A0A1I9LQK5]
AT1G75560.1	5.663	3.84E-06	0	8.5		DNA-binding protein [Source:UniProtKB/TrEMBL;Acc:Q8GXC5]
AT5G26710.1	5.818	3.90E-06	0	8.75		Glutamate--tRNA ligase, cytoplasmic [Source:UniProtKB/Swiss-Prot;Acc:O82462]
AT1G54490.1	5.662	4.22E-06	0	9.5	XRN4	XRN4 [Source:UniProtKB/TrEMBL;Acc:A0A178WFK9]
AT3G02830.1	5.640	4.50E-06	0	8.75	ZFN1	Zinc finger CCCH domain-containing protein 33 [Source:UniProtKB/Swiss-Prot;Acc:Q8GXX7]
AT3G25920.1	5.818	4.77E-06	0	10	RPL15	50S ribosomal protein L15, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P25873]
AT5G02530.1	4.007	4.77E-06	0.8	36	ALY2	THO complex subunit 4B [Source:UniProtKB/Swiss-Prot;Acc:Q8L719]
AT5G54900.1	5.591	5.08E-06	0	8.5	RBP45A	Polyadenylate-binding protein RBP45A [Source:UniProtKB/Swiss-Prot;Acc:Q9FPJ8]
AT3G14100.1	4.451	5.22E-06	0.2	12.75	UBP1C	Oligouridylate-binding protein 1C [Source:UniProtKB/Swiss-Prot;Acc:Q9LJH8]
AT2G38610.1	5.598	5.24E-06	0	8.25		KH domain-containing protein At2g38610 [Source:UniProtKB/Swiss-Prot;Acc:Q9ZVI3]
AT5G47690.3	5.615	5.58E-06	0	8.25		Binding protein [Source:UniProtKB/TrEMBL;Acc:B3H5K3]
AT3G50370.1	6.228	5.68E-06	0	14.25		unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOC
AT3G23450.1	3.488	5.99E-06	0.8	22.75		Transmembrane protein [Source:UniProtKB/TrEMBL;Acc:Q9LW52]
AT5G61970.1	5.710	6.21E-06	0	9.25		Signal recognition particle subunit SRP68 [Source:UniProtKB/TrEMBL;Acc:Q9FH46]

AT5G22060.1	5.727	6.50E-06	0	9.25	ATJ2	At5g22060 [Source:UniProtKB/TrEMBL;Acc:Q0V7U1]
AT1G02813.1	4.061	6.77E-06	0.4	15		Protein of unknown function, DUF538 [Source:TAIR;Acc:AT1G02813]
AT3G17840.1	5.687	7.11E-06	0	9.5	RLK902	Probable inactive receptor kinase RLK902 [Source:UniProtKB/Swiss-Prot;Acc:Q9LVI6]
AT1G54560.1	5.976	7.39E-06	0	11.25	XI-E	Myosin-11 [Source:UniProtKB/Swiss-Prot;Acc:F4HWY6]
AT5G54600.1	5.592	7.67E-06	0	8.25	RPL24	SVR8 [Source:UniProtKB/TrEMBL;Acc:A0A178U954]
AT5G55230.2	5.552	8.08E-06	0	8	ATMAP65-1	Microtubule-associated proteins 65-1 [Source:UniProtKB/TrEMBL;Acc:F4K3E4]
AT1G24260.2	5.649	1.05E-05	0	8.25		At1g24260 [Source:UniProtKB/TrEMBL;Acc:B4F7R9]
AT1G75910.1	2.852	1.05E-05	5.6	102.75	EXL4	GDSL esterase/lipase EXL4 [Source:UniProtKB/Swiss-Prot;Acc:Q0WUV7]
AT1G12310.1	2.941	1.10E-05	1	16.5	CML13	Probable calcium-binding protein CML13 [Source:UniProtKB/Swiss-Prot;Acc:Q94AZ4]
AT3G13750.1	5.809	1.12E-05	0	9.5	BGAL1	Beta-galactosidase 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SCW1]
AT3G29310.1	5.996	1.14E-05	0	10.5	BAG1	BAG family molecular chaperone regulator 8, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9LIB3]
AT3G49390.1	5.618	1.14E-05	0	8	CID10	Polyadenylate-binding protein-interacting protein 10 [Source:UniProtKB/Swiss-Prot;Acc:Q9SG10]
AT1G47500.1	5.542	1.22E-05	0	8	RBP47C'	Polyadenylate-binding protein RBP47C' [Source:UniProtKB/Swiss-Prot;Acc:Q9SX80]
AT1G08370.1	5.475	1.25E-05	0	8	DCP1	mRNA-decapping enzyme-like protein [Source:UniProtKB/Swiss-Prot;Acc:Q9SJF3]
AT3G28500.1	5.482	1.31E-05	0	8	RPP2C	60S acidic ribosomal protein P2-3 [Source:UniProtKB/Swiss-Prot;Acc:Q9LH85]
AT5G22860.1	6.256	1.34E-05	0	13.5		Prolylcarboxypeptidase-like protein [Source:UniProtKB/TrEMBL;Acc:Q9FFC2]
AT1G55500.2	5.509	1.34E-05	0	8.5	ECT4	evolutionarily conserved C-terminal region 4 [Source:TAIR;Acc:AT1G55500]
AT3G12410.1	5.776	1.39E-05	0	9.25		Polynucleotidyl transferase, ribonuclease H-like superfamily protein [Source:UniProtKB/TrEMBL;Acc:Q9LHG7]
AT1G78060.1	3.020	1.60E-05	1	18.25	BXL7	Probable beta-D-xylosidase 7 [Source:UniProtKB/Swiss-Prot;Acc:Q9SGZ5]
AT5G55660.1	5.602	1.62E-05	0	9		DEK domain-containing chromatin associated protein [Source:UniProtKB/TrEMBL;Acc:F4K4Y5]
AT3G53420.1	5.566	1.63E-05	0	7.75	PIP2-1	PIP2A [Source:UniProtKB/TrEMBL;Acc:A0A178V8P4]
AT4G26650.1	5.549	1.63E-05	0	8.5		AT4g26650/T15N24_100 [Source:UniProtKB/TrEMBL;Acc:Q8W569]
AT1G01080.2	5.462	1.83E-05	0	8		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:F4HQH8]
AT5G43960.1	5.470	1.84E-05	0	7.25		Gb [Source:UniProtKB/TrEMBL;Acc:Q9FND0]
AT3G44750.1	5.537	1.95E-05	0	7.5	HDT1	Histone deacetylase HDT1 [Source:UniProtKB/Swiss-Prot;Acc:Q9FVE6]
AT1G28290.1	4.333	1.97E-05	0.2	11.5	AGP31	Non-classical arabinogalactan protein 31 [Source:UniProtKB/Swiss-Prot;Acc:Q9FZA2]
AT1G70180.2	5.671	2.04E-05	0	8.25		At1g70180 [Source:UniProtKB/TrEMBL;Acc:Q8VZ94]
AT1G47490.1	5.433	2.04E-05	0	8	RBP47C	Polyadenylate-binding protein RBP47C [Source:UniProtKB/Swiss-Prot;Acc:Q9SX79]
AT5G51120.2	5.454	2.10E-05	0	7.75	PABN1	Polyadenylate-binding protein 1 [Source:UniProtKB/TrEMBL;Acc:F4KBV3]
AT3G49120.1	4.237	2.20E-05	0.2	10.75	PER34	Peroxidase 34 [Source:UniProtKB/Swiss-Prot;Acc:Q9SMU8]
AT5G07510.1	6.013	2.30E-05	0	10	GRP14	Glycine-rich protein [Source:UniProtKB/TrEMBL;Acc:Q1PDZ0]
AT1G48650.2	5.385	2.30E-05	0	7.5		DEA(D/H)-box RNA helicase family protein [Source:UniProtKB/TrEMBL;Acc:F4HYJ6]
AT4G36960.1	5.858	2.50E-05	0	9.5		AT4g36960/C7A10_400 [Source:UniProtKB/TrEMBL;Acc:Q94AC3]
AT5G64270.1	5.458	2.51E-05	0	7.75		Nuclear protein-like [Source:UniProtKB/TrEMBL;Acc:Q9FMF9]
AT2G22400.1	5.560	2.56E-05	0	9		S-adenosyl-L-methionine-dependent methyltransferases superfamily protein [Source:UniProtKB/TrEMBL;Acc:Q8
AT3G60830.1	5.471	2.56E-05	0	8.25	ARP7	Actin-related protein 7 [Source:UniProtKB/Swiss-Prot;Acc:Q8L4Y5]
AT3G52150.1	5.410	2.70E-05	0	7.75	PSRP2	PSRP2 [Source:UniProtKB/TrEMBL;Acc:A0A178VFQ5]
AT5G22040.1	5.397	2.78E-05	0	7		At5g22040 [Source:UniProtKB/TrEMBL;Acc:Q9C584]
AT1G24360.1	2.783	2.82E-05	1.2	17.75		3-oxoacyl-[acyl-carrier-protein] reductase, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P33207]
AT4G35890.1	5.689	2.85E-05	0	9	LARP1C	La-related protein 1C [Source:UniProtKB/Swiss-Prot;Acc:Q94K80]
AT5G10270.1	5.521	2.94E-05	0	8.75	CDKC-1	CDKC [Source:UniProtKB/TrEMBL;Acc:A0A178UJP9]
AT5G09440.1	3.480	3.01E-05	0.4	11.25	EXL4	EXL4 [Source:UniProtKB/TrEMBL;Acc:A0A178UJ96]
AT5G08180.1	6.162	3.01E-05	0	10.75		H/ACA ribonucleoprotein complex subunit 2-like protein [Source:UniProtKB/Swiss-Prot;Acc:Q9LEY9]
AT5G55920.1	5.379	3.10E-05	0	7	OLI2	Nucleolar protein-like [Source:UniProtKB/TrEMBL;Acc:Q9FG73]
AT1G54080.2	4.124	3.11E-05	0.2	10	UBP1A	Oligouridylate-binding protein 1A [Source:UniProtKB/TrEMBL;Acc:F4HV67]
AT5G62190.1	2.883	3.11E-05	1	16	RH7	DEAD-box ATP-dependent RNA helicase 7 [Source:UniProtKB/Swiss-Prot;Acc:Q39189]
AT1G08730.1	5.674	3.19E-05	0	9	XI-C	Myosin-9 [Source:UniProtKB/Swiss-Prot;Acc:F4HXP9]
AT5G67500.2	5.326	3.40E-05	0	7	VDAC2	voltage dependent anion channel 2 [Source:TAIR;Acc:AT5G67500]
AT5G15090.1	3.302	3.40E-05	0.6	14	VDAC3	Mitochondrial outer membrane protein porin 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9SMX3]
AT1G47128.1	5.546	3.47E-05	0	8.25	RD21A	RD21A [Source:UniProtKB/TrEMBL;Acc:A0A178WIH7]
AT2G01750.2	6.239	3.47E-05	0	12.75	MAP70-3	Microtubule-associated protein 70-3 [Source:UniProtKB/TrEMBL;Acc:F4IPB2]
AT1G11650.2	3.934	3.63E-05	0.2	9	RBP45B	RBP45B [Source:UniProtKB/TrEMBL;Acc:A0A178WLD6]
AT1G30360.1	5.342	3.92E-05	0	6.75	ERD4	Hyperosmolality-gated Ca2+ permeable channel 3.1 [Source:UniProtKB/TrEMBL;Acc:A0A097NUQ9]
AT4G39200.1	2.853	4.00E-05	0.8	12.75	RPS25E	40S ribosomal protein S25-4 [Source:UniProtKB/Swiss-Prot;Acc:Q9T029]
AT1G30460.1	5.594	4.03E-05	0	8	CPSF30	30-kDa cleavage and polyadenylation specificity factor 30 [Source:UniProtKB/Swiss-Prot;Acc:A9LNK9]
AT1G62820.1	3.706	4.24E-05	0.4	13.25	CML14	Probable calcium-binding protein CML14 [Source:UniProtKB/Swiss-Prot;Acc:Q8VZ50]
AT5G66680.1	3.627	4.39E-05	0.6	19	OST48	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit [Source:UniProtKB/Swiss-Prot;Acc
AT3G61130.1	5.374	4.48E-05	0	6.75	GAUT1	Hexosyltransferase (Fragment) [Source:UniProtKB/TrEMBL;Acc:W8PUP8]
AT3G61760.1	3.859	4.51E-05	0.4	15	DRP1B	DL1B [Source:UniProtKB/TrEMBL;Acc:A0A178VE87]
AT5G48375.1	2.887	4.55E-05	3.6	60	TGG3	Putative myrosinase 3 [Source:UniProtKB/Swiss-Prot;Acc:Q3E8E5]
AT2G35920.1	5.362	4.64E-05	0	7.75		RNA helicase family protein [Source:TAIR;Acc:AT2G35920]
AT3G13990.1	5.599	4.84E-05	0	8.5		AT3g13990/MDC16_11 [Source:UniProtKB/TrEMBL;Acc:Q8L7T4]
AT5G64260.1	3.613	4.84E-05	0.4	12.25	EXL2	Protein EXORDIUM-like 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9FE06]
AT4G32460.1	5.298	4.88E-05	0	7		At4g32460 [Source:UniProtKB/TrEMBL;Acc:Q8LAR0]
AT1G60650.1	5.492	4.93E-05	0	7	RZ1B	Glycine-rich RNA-binding protein RZ1B [Source:UniProtKB/Swiss-Prot;Acc:O22703]
AT5G11350.1	5.283	5.14E-05	0	7.25		DNAse I-like superfamily protein [Source:TAIR;Acc:AT5G11350]
AT4G34660.1	5.518	5.24E-05	0	7.5	SH3P2	SH3 domain-containing protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q8VWF1]
AT3G14750.1	5.274	5.31E-05	0	6.5	FLXL1	Protein FLX-like 1 [Source:UniProtKB/Swiss-Prot;Acc:Q93V84]
AT5G51550.1	5.532	5.35E-05	0	7.5	EXL3	Protein EXORDIUM-like 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9FHM9]
AT1G56110.1	2.176	5.42E-05	8	50	NOP56	At1g56110/T6H22_9 [Source:UniProtKB/TrEMBL;Acc:Q9SGT7]
AT3G12860.1	3.899	5.77E-05	0.2	9.5		NOP56-like pre RNA processing ribonucleoprotein [Source:UniProtKB/TrEMBL;Acc:Q9LTV0]
AT4G11050.1	5.587	5.78E-05	0	8.75	AtGH9C3	Endoglucanase 19 [Source:UniProtKB/Swiss-Prot;Acc:Q8L7I0]
AT5G18550.1	5.301	5.81E-05	0	6.75		Zinc finger CCCH domain-containing protein 58 [Source:UniProtKB/Swiss-Prot;Acc:Q6NPN3]
AT5G20950.1	5.149	5.81E-05	0	6.25		Glycosyl hydrolase family protein [Source:TAIR;Acc:AT5G20950]
AT5G07560.1	6.713	5.94E-05	0	15.5	GRP20	GRP20 [Source:UniProtKB/TrEMBL;Acc:A0A178UJU8]
AT4G27500.1	3.112	6.44E-05	1.2	24	PPI1	Proton pump-interactor 1 [Source:UniProtKB/Swiss-Prot;Acc:O23144]
AT2G40660.1	3.690	6.75E-05	0.4	10.75		Nucleic acid-binding, OB-fold-like protein [Source:UniProtKB/TrEMBL;Acc:Q93VB0]
AT2G15430.1	4.016	6.79E-05	0.2	9.25	NRPB3	RPB35.5A [Source:UniProtKB/TrEMBL;Acc:A0A178W202]
AT1G67750.1	5.252	6.90E-05	0	6.25		Probable pectate lyase 5 [Source:UniProtKB/Swiss-Prot;Acc:Q9FXD8]
AT5G38600.1	5.886	7.76E-05	0	10.5		Proline-rich spliceosome-associated (PSP) family protein / zinc knuckle (CCHC-type) family protein [Source:UniPro
AT1G68060.1	6.519	9.56E-05	0	16.25	MAP70.1	Microtubule-associated protein 70-1 [Source:UniProtKB/Swiss-Prot;Acc:Q9C9X0]
AT5G12440.1	5.193	9.88E-05	0	6		CCCH-type zinc fingerfamily protein with RNA-binding domain [Source:TAIR;Acc:AT5G12440]
AT3G62310.1	3.914	9.89E-05	0.2	8.75		Probable pre-mRNA-splicing factor ATP-dependent RNA helicase DEAH2 [Source:UniProtKB/Swiss-Prot;Acc:Q9LZ(
AT3G50590.1	5.243	0.000105344	0	6.5		Transducin/WD40 repeat-like superfamily protein [Source:UniProtKB/TrEMBL;Acc:F4J0P2]
AT2G47250.1	5.201	0.000105344	0	6.75		Probable pre-mRNA-splicing factor ATP-dependent RNA helicase DEAH3 [Source:UniProtKB/Swiss-Prot;Acc:O228
AT2G32000.1	5.255	0.000114231	0	6.75		DNA topoisomerase 3-beta [Source:UniProtKB/Swiss-Prot;Acc:F4ISQ7]
AT4G33200.1	5.658	0.000116861	0	10	XI-I	myosin, putative [Source:TAIR;Acc:AT4G33200]
AT1G66410.2	2.409	0.000123489	1.6	18.25	CAM4	Calmodulin 4 [Source:UniProtKB/TrEMBL;Acc:F4IEU4]
AT2G40290.1	3.258	0.000130161	0.6	13.75		Eukaryotic translation initiation factor 2 subunit alpha homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9SIZ2]
AT1G49670.2	5.236	0.000142303	0	6.75	NQR	ARP protein (REF) [Source:TAIR;Acc:AT1G49670]
AT3G62250.1	3.291	0.000153417	3	62.75	RPS27AC	Ubiquitin-40S ribosomal protein S27a-3 [Source:UniProtKB/Swiss-Prot;Acc:P59233]
AT3G49240.1	5.252	0.000161081	0	7.25	EMB1796	Pentatricopeptide repeat-containing protein At3g49240, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9M3A
AT5G62670.1	3.329	0.000162977	1.2	26.75	AHA11	ATPase 11, plasma membrane-type [Source:UniProtKB/Swiss-Prot;Acc:Q9LV11]
AT2G46020.2	5.746	0.000172932	0	10.25	BRM	ATP-dependent helicase BRM [Source:UniProtKB/Swiss-Prot;Acc:Q6EVK6]
AT5G58040.1	5.144	0.000172932	0	6	FIPS5	FIP1[V]-like protein [Source:UniProtKB/Swiss-Prot;Acc:F4KDH9]
AT3G12130.1	5.299	0.000175797	0	6.25		Zinc finger CCCH domain-containing protein 36 [Source:UniProtKB/Swiss-Prot;Acc:Q9C7C3]
AT5G26000.1	2.389	0.000177094	17	163.5	TGG1	TGG1 [Source:UniProtKB/TrEMBL;Acc:A0A178U7L7]
AT3G54670.3	5.092	0.000178049	0	6	NA	NA
AT4G29510.1	5.091	0.000180311	0	6.25	PRMT11	Protein arginine N-methyltransferase 1.1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SU94]
AT2G07360.2	3.807	0.000199292	0.2	8		SH3 domain-containing protein [Source:UniProtKB/TrEMBL;Acc:F4IL68]

AT2G24590.1	3.035	0.000201054	0.6	10.75	RSZ22A	Serine/arginine-rich splicing factor RSZ22A [Source:UniProtKB/Swiss-Prot;Acc:Q9SJA6]
AT5G21150.1	5.180	0.000223687	0	6.75	AGO9	Protein argonaute 9 [Source:UniProtKB/Swiss-Prot;Acc:Q84VQ0]
AT2G30800.1	5.128	0.000224814	0	6	HVT1	DExH-box ATP-dependent RNA helicase DExH6 [Source:UniProtKB/Swiss-Prot;Acc:F4INY4]
AT4G31580.1	3.346	0.000229682	0.6	11.5	RSZ22	Serine/arginine-rich splicing factor RSZ22 [Source:UniProtKB/Swiss-Prot;Acc:O81126]
AT3G60770.1	2.334	0.000251777	1.6	17.75	RPS13A	40S ribosomal protein S13-1 [Source:UniProtKB/Swiss-Prot;Acc:P59223]
AT5G14320.1	5.103	0.000263743	0	6.25	RPS13	30S ribosomal protein S13, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P42732]
AT5G17170.1	5.103	0.00027934	0	5.5	ENH1	Rubredoxin family protein [Source:UniProtKB/TrEMBL;Acc:Q9FFJ2]
AT2G19170.1	4.985	0.000281031	0	5.5	SBT2.5	Subtilisin-like protease SBT2.5 [Source:UniProtKB/Swiss-Prot;Acc:O64481]
AT1G62390.1	3.829	0.000281031	0.2	8.75	Phox2	Protein CLMP1 [Source:UniProtKB/Swiss-Prot;Acc:O48802]
AT4G27000.1	3.224	0.000281031	0.4	9.25	RBP45C	Polyadenylate-binding protein RBP45C [Source:UniProtKB/Swiss-Prot;Acc:Q93W34]
AT5G10350.1	5.167	0.00028681	0	6	PABN3	Polyadenylate-binding protein 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9LX90]
AT4G28880.1	5.100	0.000319751	0	6	CKL3	Ckl3 [Source:UniProtKB/TrEMBL;Acc:A0A178UXY3]
AT3G50670.1	5.036	0.000322865	0	5.75	RNU1	U1SNRNP [Source:UniProtKB/TrEMBL;Acc:A0A178VDW4]
AT5G35180.4	3.863	0.000322865	0.2	8.75		Protein of unknown function (DUF1336) [Source:TAIR;Acc:AT5G35180]
AT1G80270.1	5.066	0.000326613	0	6	PPR596	PPR596 [Source:UniProtKB/TrEMBL;Acc:A0A178W3Z6]
AT5G19690.1	3.687	0.000328663	0.2	7.5	STT3A	STT3A [Source:UniProtKB/TrEMBL;Acc:A0A178URP5]
AT3G49490.1	5.109	0.000330807	0	6.25		Uncharacterized protein T9C5.90 [Source:UniProtKB/TrEMBL;Acc:Q9SCK9]
AT3G11964.1	5.101	0.00033525	0	6.25	RRP5	rRNA biogenesis protein RRP5 [Source:UniProtKB/Swiss-Prot;Acc:F4J8K6]
AT1G10590.3	4.975	0.000343536	0	5.5		Nucleic acid-binding, OB-fold-like protein [Source:UniProtKB/TrEMBL;Acc:Q9XIJ0]
AT1G17370.1	4.998	0.00034787	0	5.25	UBP1B	Oligouridylate-binding protein 1B [Source:UniProtKB/Swiss-Prot;Acc:Q9LQI9]
AT5G47700.1	5.054	0.000354847	0	5.25	RPP1C	60S acidic ribosomal protein P1-3 [Source:UniProtKB/Swiss-Prot;Acc:Q8LEQ0]
AT1G66260.1	5.358	0.000358014	0	6.75	ALY3	THO complex subunit 4C [Source:UniProtKB/Swiss-Prot;Acc:Q94EH8]
AT4G32720.1	4.145	0.000360757	0.2	10.75	LA1	La1 [Source:UniProtKB/TrEMBL;Acc:A0A178V700]
AT5G35970.1	5.045	0.000361199	0	6		P-loop containing nucleoside triphosphate hydrolases superfamily protein [Source:UniProtKB/TrEMBL;Acc:F4K1G7
AT4G18100.1	3.667	0.000366896	0.2	7.25	RPL32A	60S ribosomal protein L32-1 [Source:UniProtKB/Swiss-Prot;Acc:P49211]
AT3G06610.1	5.043	0.000371245	0	5.25		DNA-binding enhancer protein-like protein [Source:UniProtKB/TrEMBL;Acc:Q9C901]
AT1G02800.1	4.975	0.000373639	0	5.75	CEL2	Endoglucanase 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SRX3]
AT5G57350.1	2.803	0.000387083	0.6	10.75	AHA3	Plasma membrane ATPase [Source:UniProtKB/TrEMBL;Acc:A0A178UHY0]
AT2G45820.1	3.663	0.000398296	0.2	7.25	DBP	Remorin [Source:UniProtKB/Swiss-Prot;Acc:O80837]
AT3G01280.1	2.920	0.000398296	1	17	VDAC1	Mitochondrial outer membrane protein porin 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SRH5]
AT4G25340.1	4.969	0.000404842	0	5.25	FKBP53	Peptidyl-prolyl cis-trans isomerase FKBP53 [Source:UniProtKB/Swiss-Prot;Acc:Q93ZG9]
AT4G28250.1	2.917	0.000406212	0.8	13.5	EXPB3	EXPB3 [Source:UniProtKB/TrEMBL;Acc:A0A178USU9]
AT2G47850.1	5.073	0.000410073	0	5.25		Zinc finger CCCH domain-containing protein 32 [Source:UniProtKB/Swiss-Prot;Acc:Q84W91]
AT3G07050.1	5.069	0.000413737	0	6	NSN1	Guanine nucleotide-binding protein-like NSN1 [Source:UniProtKB/Swiss-Prot;Acc:Q9M8Z5]
AT5G22770.1	5.138	0.000420458	0	6.25	ALPHA-ADR	AP-2 complex subunit alpha-1 [Source:UniProtKB/Swiss-Prot;Acc:Q8LPL6]
AT4G24680.1	3.599	0.000429224	0.2	7.5	MOS1	modifier of snc1 [Source:TAIR;Acc:AT4G24680]
AT1G78630.1	5.024	0.000436515	0	5.25	RPL13	50S ribosomal protein L13, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9SYL9]
AT3G27360.1-AT5G10390.1	4.971	0.000439908	0	5.75	NA	NA
AT5G65350.1	4.971	0.000439908	0	5.75	HTR11	Histone H3-like 5 [Source:UniProtKB/Swiss-Prot;Acc:Q9FKQ3]
AT5G40200.1	4.945	0.00044126	0	5.25	DEGP9	Protease Do-like 9 [Source:UniProtKB/Swiss-Prot;Acc:Q9FL12]
AT2G37990.1	4.970	0.000448197	0	5.75		Ribosome biogenesis regulatory protein homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9SH88]
AT4G34110.1	4.083	0.000448197	1	43.5	PAB2	Polyadenylate-binding protein [Source:UniProtKB/TrEMBL;Acc:A0A178UWB3]
AT5G20160.2	2.671	0.000448197	0.8	10.75		Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein [Source:UniProtKB/TrEMBL;Acc:F4K455]
AT3G05910.1	5.539	0.000453706	0	7.75	PAE12	Pectin acetylesterase 12 [Source:UniProtKB/Swiss-Prot;Acc:Q9SFF6]
AT2G28290.1	5.706	0.000467192	0	10.25	SYD	P-loop containing nucleoside triphosphate hydrolases superfamily protein [Source:UniProtKB/TrEMBL;Acc:A0A1P
AT5G43900.3	5.349	0.000467192	0	7.5	MYA2	Myosin 2 [Source:UniProtKB/TrEMBL;Acc:F4K7C5]
AT3G50000.1	5.118	0.000468259	0	5.5	CKA2	Casein kinase II subunit alpha-2 [Source:UniProtKB/Swiss-Prot;Acc:Q08466]
AT4G24770.1	3.585	0.000480515	0.2	7	CP31A	31 kDa ribonucleoprotein, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q04836]
AT3G19960.2	5.378	0.000492113	0	8	ATM1	Myosin 1 [Source:UniProtKB/TrEMBL;Acc:F4JCF9]
AT2G27170.1	4.991	0.000507829	0	5.75	SMC3	Structural maintenance of chromosomes protein 3 [Source:UniProtKB/Swiss-Prot;Acc:Q56YN8]
AT5G63260.2	4.914	0.000520854	0	5.5		Zinc finger C-x8-C-x5-C-x3-H type family protein [Source:UniProtKB/TrEMBL;Acc:F4K9A6]
AT4G23490.1	4.955	0.000522606	0	5		At4g23490/F16G20_190 [Source:UniProtKB/TrEMBL;Acc:Q8H0S2]
AT3G19760.1	2.762	0.000566108	1.8	26.75	EIF4A3	Eukaryotic initiation factor 4A-III homolog [Source:UniProtKB/Swiss-Prot;Acc:Q94A52]
AT2G07690.1	4.935	0.00057037	0	5.5	MCM5	DNA replication licensing factor MCM5 [Source:UniProtKB/Swiss-Prot;Acc:O80786]
AT1G07910.1	4.903	0.00057037	0	5.5	ATRNL	RNAligase [Source:TAIR;Acc:AT1G07910]
AT5G22650.1	2.227	0.00057561	1.6	16	HDT2	Histone deacetylase HDT2 [Source:UniProtKB/Swiss-Prot;Acc:Q56WH4]
AT1G07320.1	4.959	0.000579715	0	5	RPL4	RPL4 [Source:UniProtKB/TrEMBL;Acc:A0A178WQG5]
AT3G20890.1	4.879	0.000585143	0	4.75		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:F4IWB2]
AT5G53620.1	4.983	0.000612499	0	5.75		At5g53620 [Source:UniProtKB/TrEMBL;Acc:Q93XY1]
AT2G46780.1	4.905	0.000647231	0	5.25		At2g46780 [Source:UniProtKB/TrEMBL;Acc:Q501B0]
AT4G19120.1	5.099	0.000661743	0	5.25	ERD3	Probable methyltransferase PMT21 [Source:UniProtKB/Swiss-Prot;Acc:Q94II3]
AT2G37230.1	4.999	0.000664725	0	5.75		Pentatricopeptide repeat-containing protein At2g37230 [Source:UniProtKB/Swiss-Prot;Acc:Q9ZUU3]
AT4G13710.1	4.914	0.000669167	0	5.5		Probable pectate lyase 15 [Source:UniProtKB/Swiss-Prot;Acc:Q944R1]
AT5G51980.2	4.883	0.000670358	0	5		Transducin/WD40 repeat-like superfamily protein [Source:TAIR;Acc:AT5G51980]
AT4G16210.1	5.022	0.000731485	0	5.5	ECHIA	Probable enoyl-CoA hydratase 1, peroxisomal [Source:UniProtKB/Swiss-Prot;Acc:Q6NL24]
AT4G28990.2	4.805	0.000739398	0	4.75		RNA-binding protein-like protein [Source:UniProtKB/TrEMBL;Acc:F4JM55]
AT2G38040.1	3.517	0.000739398	1.6	38.25	CAC3	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha, chloroplastic [Source:UniProtKB/Swiss-Prot;A
AT2G34040.1	3.691	0.000753634	0.6	19.25		Apoptosis inhibitory protein 5 (API5) [Source:UniProtKB/TrEMBL;Acc:O22957]
AT5G60980.2	3.369	0.000771722	1	25.5		AT5g60980/MSL3_100 [Source:UniProtKB/TrEMBL;Acc:Q9FME2]
AT4G09150.2	4.972	0.000803983	0	6		T-complex protein 11 [Source:UniProtKB/TrEMBL;Acc:F4JJB2]
AT2G35410.1	4.765	0.000815225	0	4.75		Putative chloroplast RNA binding protein [Source:UniProtKB/TrEMBL;Acc:O82299]
AT5G46630.2	4.943	0.000825978	0	5.25		Clathrin adaptor complexes medium subunit family protein [Source:UniProtKB/TrEMBL;Acc:F4KHJ7]
AT3G58660.1	4.902	0.000860938	0	5.25		Ribosomal protein L1p/L10e family [Source:UniProtKB/TrEMBL;Acc:Q9LXT5]
AT2G44530.1	4.853	0.000885886	0	5.25	PRS5	Ribose-phosphate pyrophosphokinase 5, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:O64888]
AT3G62330.1	4.783	0.000926989	0	4.5		At3g62330 [Source:UniProtKB/TrEMBL;Acc:Q9LZQ7]
AT1G63640.1	4.915	0.00094685	0	5.5		P-loop nucleoside triphosphate hydrolases superfamily protein with CH (Calponin Homology) domain [Source:TAI
AT1G04170.1	4.778	0.000956919	0	4.5	EIF2 GAMMA	Eukaryotic translation initiation factor 2 gamma subunit [Source:UniProtKB/TrEMBL;Acc:O64490]
AT2G21580.1	2.581	0.001007653	0.8	10.25	RPS25B	40S ribosomal protein S25-2 [Source:UniProtKB/Swiss-Prot;Acc:Q9SIK2]
AT1G61730.1	4.805	0.001022477	0	5		Probable transcription factor At1g61730 [Source:UniProtKB/Swiss-Prot;Acc:Q9SYA9]
AT3G49720.1	4.836	0.001062432	0	4.5	CGR2	Uncharacterized protein At3g49720 [Source:UniProtKB/Swiss-Prot;Acc:Q9M2Y6]
AT3G62600.1	4.745	0.001062432	0	4.75	ERDJ3B	DnaJ protein ERDJ3B [Source:UniProtKB/Swiss-Prot;Acc:Q9LZK5]
AT4G24500.1	4.828	0.001089499	0	4.5	SIC	Protein SICKLE [Source:UniProtKB/Swiss-Prot;Acc:Q9SB47]
AT5G21160.3	4.939	0.001104659	0	5.75		LA RNA-binding protein [Source:UniProtKB/TrEMBL;Acc:F4K704]
AT3G59990.1	4.719	0.001111809	0	4.5	MAP2B	Methionine aminopeptidase 2B [Source:UniProtKB/Swiss-Prot;Acc:Q56Y85]
AT4G23400.1	4.826	0.001161351	0	4.75	PIP1-5	PIP1D [Source:UniProtKB/TrEMBL;Acc:A0A178V129]
AT5G03280.1	4.907	0.001190921	0	5.75	EIN2	Ethylene-insensitive protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9S814]
AT1G04990.2	4.706	0.001202302	0	4.5		Zinc finger CCCH domain-containing protein 3 [Source:UniProtKB/Swiss-Prot;Acc:Q94AD9]
AT3G15590.1	4.693	0.001202769	0	4.75		Pentatricopeptide repeat-containing protein At3g15590, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9LRP6
AT2G40610.1	4.763	0.001205836	0	5	EXPA8	EXPA8 [Source:UniProtKB/TrEMBL;Acc:A0A178VTJ9]
AT1G07820.1-AT2G28740.1	2.468	0.001205836	1.8	21.75	NA	NA
AT3G23820.1	4.743	0.001260047	0	4.25	GAE6	UDP-glucuronate 4-epimerase 6 [Source:UniProtKB/Swiss-Prot;Acc:Q9LIS3]
AT1G14900.1	4.685	0.001283266	0	4.25	HMGA	HMG-Y-related protein A [Source:UniProtKB/Swiss-Prot;Acc:Q43386]
AT5G47190.1	4.685	0.001283266	0	4.25		50S ribosomal protein L19-2, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q8RXX5]
AT5G50150.1	4.746	0.001320426	0	4.5		AT5g50150/MPF21_17 [Source:UniProtKB/TrEMBL;Acc:Q9FG96]
AT3G28220.1	4.682	0.001368908	0	4.25		AT3g28220/T19D11_3 [Source:UniProtKB/TrEMBL;Acc:Q9LHA6]
AT5G63140.1	4.718	0.001420412	0	4.75	PAP29	Probable inactive purple acid phosphatase 29 [Source:UniProtKB/Swiss-Prot;Acc:Q9FMK9]
AT3G12610.1	4.714	0.001470324	0	4.75	DRT100	DRT100 [Source:UniProtKB/TrEMBL;Acc:A0A178VMS3]
AT3G06400.3	4.912	0.00147926	0	5.75	CHR11	Chromatin-remodeling protein 11 [Source:UniProtKB/TrEMBL;Acc:F4JAV9]
AT5G55550.3	4.681	0.00147926	0	4.75		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:TAIR;Acc:AT5G55550]
AT3G03920.1	2.420	0.001496386	2.2	16.5		Putative H/ACA ribonucleoprotein complex subunit 1-like protein 1 [Source:UniProtKB/Swiss-Prot;Acc:Q8VZT0]

AT3G19130.1	4.698	0.001515074	0	4.75	RBP47B	Polyadenylate-binding protein RBP47B [Source:UniProtKB/Swiss-Prot;Acc:Q0WW84]
AT2G33800.1	2.554	0.001533767	0.8	10.5	RPS5	30S ribosomal protein S5, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P93014]
AT5G18590.1	4.651	0.001559537	0	4.5		Galactose oxidase/kelch repeat superfamily protein [Source:UniProtKB/TrEMBL;Acc:Q0WWX4]
AT3G59770.3	4.738	0.001600916	0	5	SAC9	SacI homology domain-containing protein / WW domain-containing protein [Source:UniProtKB/TrEMBL;Acc:F4J9
AT5G63670.1	4.724	0.001612686	0	4.25	SPT42	Transcription elongation factor SPT4 homolog 2 [Source:UniProtKB/Swiss-Prot;Acc:Q94C60]
AT1G47980.1	4.669	0.001612686	0	4.25		unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOC
AT3G52140.4	2.824	0.001612686	0.8	12.5		Clustered mitochondria protein homolog [Source:UniProtKB/TrEMBL;Acc:F4J5R9]
AT3G23700.1	4.626	0.001637755	0	4.5		At3g23700/MYM9_3 [Source:UniProtKB/TrEMBL;Acc:Q9LK47]
AT1G51510.1	4.640	0.001667856	0	4.25	Y14	Y14 [Source:UniProtKB/TrEMBL;Acc:A0A178WGD6]
AT3G18600.1	4.931	0.001702173	0	5.5	RH51	DEAD-box ATP-dependent RNA helicase 51 [Source:UniProtKB/Swiss-Prot;Acc:Q9LIH9]
AT5G13010.1	4.893	0.001718883	0	5	CUV	Pre-mRNA-splicing factor ATP-dependent RNA helicase DEAH7 [Source:UniProtKB/Swiss-Prot;Acc:F4K2E9]
AT5G23400.1	4.671	0.00176721	0	4.75		Disease resistance protein-like [Source:UniProtKB/TrEMBL;Acc:Q9FHL8]
AT3G27830.1	2.831	0.00186851	0.6	10.75	RPL12A	RPL12-A [Source:UniProtKB/TrEMBL;Acc:A0A178VJ19]
AT4G24290.2	4.589	0.001995116	0	4.25		MACPF domain-containing protein At4g24290 [Source:UniProtKB/Swiss-Prot;Acc:Q9STW5]
AT1G79350.1	4.688	0.002091305	0	4.75	FGT1	Protein FORGETTER 1 [Source:UniProtKB/Swiss-Prot;Acc:F4IF36]
AT1G79280.2	4.721	0.002145307	0	5	NUA	nuclear pore anchor [Source:TAIR;Acc:AT1G79280]
AT4G14430.1	5.077	0.002153787	0	5.5	ECI2	PEC12 [Source:UniProtKB/TrEMBL;Acc:A0A178UV55]
AT1G67680.1	4.570	0.002186945	0	4		Signal recognition particle subunit SRP72 [Source:UniProtKB/TrEMBL;Acc:Q9FXD4]
AT4G30440.1	4.537	0.002286129	0	4	GAE1	GAE1 [Source:UniProtKB/TrEMBL;Acc:A0A178UXF7]
AT1G12770.1	4.548	0.002316004	0	4.25	RH47	DEAD-box ATP-dependent RNA helicase 47, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q8W4E1]
AT5G13300.1	4.564	0.002413651	0	4.25	AGD3	ADP-ribosylation factor GTPase-activating protein AGD3 [Source:UniProtKB/Swiss-Prot;Acc:Q5W7F2]
AT3G54340.1	4.549	0.002440536	0	4	AP3	Floral homeotic protein APETALA 3 [Source:UniProtKB/Swiss-Prot;Acc:P35632]
AT5G60790.1	2.582	0.002440536	1.2	15.5	ABCF1	ABC transporter F family member 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9FJH6]
AT1G79090.1	4.887	0.002468357	0	5.25	PAT1	Protein PAT1 homolog [Source:UniProtKB/Swiss-Prot;Acc:Q0WPK4]
AT3G55460.1	2.858	0.002483791	0.4	7.25	SCL30	Serine/arginine-rich SC35-like splicing factor SCL30 [Source:UniProtKB/Swiss-Prot;Acc:Q8L3X8]
AT1G19880.1	4.538	0.002489113	0	4.25		Regulator of chromosome condensation (RCC1) family protein [Source:UniProtKB/TrEMBL;Acc:F4HQZ1]
AT5G24850.1	4.628	0.002526522	0	4.5	CRYD	Cryptochrome DASH, chloroplastic/mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q84KJ5]
AT3G63095.1	3.406	0.002545861	0.2	6		Tetratricopeptide repeat (TPR)-like superfamily protein [Source:UniProtKB/TrEMBL;Acc:A8MRN4]
AT5G43320.1	4.704	0.002594738	0	4.25	CKL8	Casein kinase 1-like protein 8 [Source:UniProtKB/Swiss-Prot;Acc:Q9LSX4]
AT5G41460.1	4.562	0.002594738	0	4.25		At5g41460 [Source:UniProtKB/TrEMBL;Acc:Q9FN55]
AT5G14520.1	4.500	0.002594738	0	4	PES	Pescadillo homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9LYK7]
AT1G18150.2	4.607	0.002637485	0	4.5	MPK8	Mitogen-activated protein kinase 8 [Source:UniProtKB/Swiss-Prot;Acc:Q9LM33]
AT5G14170.1	4.823	0.00272859	0	5.25	CHC1	SWI/SNF complex component SNF12 homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9FMT4]
AT2G18330.1	5.088	0.002839513	0	5.75		AAA-type ATPase family protein [Source:UniProtKB/TrEMBL;Acc:Q9ZPW5]
AT1G18450.1	3.136	0.002887287	0.4	9	ARP4	Actin-related protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q84M92]
AT1G43700.1	4.585	0.002912497	0	4.5	VIP1	Transcription factor VIP1 [Source:UniProtKB/Swiss-Prot;Acc:Q9MA75]
AT5G19010.1	4.517	0.003121514	0	4.25	MPK16	Mitogen-activated protein kinase 16 [Source:UniProtKB/Swiss-Prot;Acc:Q8W4J2]
AT5G63810.1	4.517	0.003121514	0	4.25	BGAL10	Beta-galactosidase 10 [Source:UniProtKB/Swiss-Prot;Acc:Q9FN08]
AT3G28860.1	4.463	0.003194907	0	3.75	ABCB19	ABC transporter B family member 19 [Source:UniProtKB/Swiss-Prot;Acc:Q9LJX0]
AT3G54210.1	4.463	0.003194907	0	3.75	RPL17	50S ribosomal protein L17, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9M385]
AT1G69840.1	4.502	0.003210998	0	4	HIR2	Hypersensitive-induced response protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9CAR7]
AT3G13390.1	3.537	0.003222523	0.2	6.75	sks11	L-ascorbate oxidase pectinesterase-like protein pollen-specific protein-like [Source:UniProtKB/TrEMBL;Acc:Q9LJF
AT1G05190.1	2.943	0.003263593	0.4	7.5	RPL6	50S ribosomal protein L6, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:O23049]
AT2G18510.1	4.447	0.003369454	0	3.75	emb2444	Emb2444 [Source:UniProtKB/TrEMBL;Acc:A0A178VZK5]
AT4G16143.1	2.221	0.003387945	1	10	IMPA2	Importin subunit alpha-2 [Source:UniProtKB/Swiss-Prot;Acc:F4JL11]
AT3G09100.2	4.547	0.003397703	0	4.25		mRNA capping enzyme family protein [Source:UniProtKB/TrEMBL;Acc:F4IYM6]
AT3G09980.1	4.503	0.003399997	0	3.75		Ankyrin repeat 30A-like protein (DUF662) [Source:UniProtKB/TrEMBL;Acc:Q8RXZ8]
AT5G14430.1	4.451	0.003399997	0	3.75		Probable methyltransferase PMT9 [Source:UniProtKB/Swiss-Prot;Acc:Q8VZV7]
AT5G49220.1	4.603	0.003454302	0	4		Uncharacterized protein At5g49220 [Source:UniProtKB/TrEMBL;Acc:Q8VZK6]
AT1G74970.1	4.526	0.003466349	0	4	RPS9	30S ribosomal protein S9, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9XJ27]
AT1G10270.1	4.659	0.003550703	0	4.75	GRP23	Pentatricopeptide repeat-containing protein At1g10270 [Source:UniProtKB/Swiss-Prot;Acc:Q9SY69]
AT4G39000.1	4.580	0.003618798	0	4.25	AtGH9B17	Endoglucanase 23 [Source:UniProtKB/Swiss-Prot;Acc:Q8GY58]
AT2G31370.3	4.841	0.003665392	0	5	POSF21	AT2G31370 protein [Source:UniProtKB/TrEMBL;Acc:B9DGB1]
AT1G14410.1	4.392	0.003704075	0	3.75	WHY1	Single-stranded DNA-binding protein WHY1, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9M9S3]
AT1G65440.1	4.800	0.003713207	0	5.5	SPT6	Transcription elongation factor SPT6 homolog [Source:UniProtKB/Swiss-Prot;Acc:A8MS85]
AT2G39800.1	4.372	0.003713207	0	3.75	P5CSA	Delta-1-pyrroline-5-carboxylate synthase A [Source:UniProtKB/Swiss-Prot;Acc:P54887]
AT3G60245.1	4.372	0.003713207	0	3.75	RPL37AC	60S ribosomal protein L37a-2 [Source:UniProtKB/Swiss-Prot;Acc:Q8RXU5]
AT4G01037.1	4.620	0.003735621	0	4.5	WTF1	WTF1 [Source:UniProtKB/TrEMBL;Acc:A0A178UW54]
AT2G39670.2	4.368	0.003797162	0	3.75		Radical SAM superfamily protein [Source:UniProtKB/TrEMBL;Acc:F4IVY6]
AT1G21630.2	4.439	0.00383693	0	3.75		Calcium-binding EF hand family protein [Source:UniProtKB/TrEMBL;Acc:F4HY32]
AT5G16750.1	4.363	0.003857639	0	3.75	TOZ	Transducin family protein / WD-40 repeat family protein [Source:UniProtKB/TrEMBL;Acc:Q9LFE2]
\T1G09590.1-AT1G09690.1	4.849	0.004182712	0	4	NA	NA
AT2G18740.1	4.645	0.004235832	0	3.5		Putative small nuclear ribonucleoprotein E [Source:UniProtKB/TrEMBL;Acc:Q9ZV45]
AT5G59310.1	4.701	0.004242405	0	3.75	LTP4	Non-specific lipid-transfer protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9LLR6]
AT1G72610.1	3.305	0.004259986	0.2	5.25	GLP1	Germin-like protein subfamily 3 member 1 [Source:UniProtKB/Swiss-Prot;Acc:P94040]
AT1G22015.1	4.451	0.004364761	0	4	B3GALT5	Hexosyltransferase [Source:UniProtKB/TrEMBL;Acc:A0A178WND4]
AT1G76810.1	4.746	0.004590351	0	5.25		Eukaryotic translation initiation factor 2 (eIF-2) family protein [Source:UniProtKB/TrEMBL;Acc:F4I420]
AT5G18620.2	4.575	0.004593511	0	4.5	CHR17	Chromatin remodeling factor17 [Source:UniProtKB/TrEMBL;Acc:F4JY25]
AT1G02140.1	3.265	0.004657988	0.2	5.5	MEE63	MEE63 [Source:UniProtKB/TrEMBL;Acc:A0A178WA48]
AT2G27710.1	2.164	0.004692308	1.2	12	RPP2B	AT2G27710 protein [Source:UniProtKB/TrEMBL;Acc:B9DGN3]
AT2G20020.1	4.562	0.004819293	0	4.25	CAF1	CRS2-associated factor 1, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9SL79]
AT5G02960.1	2.497	0.004837965	0.6	8	RPS23B	40S ribosomal protein S23-2 [Source:UniProtKB/Swiss-Prot;Acc:P49201]
AT5G47820.1	5.075	0.004925433	0	6.25	KIN4A	Kinesin-like protein KIN-4A [Source:UniProtKB/Swiss-Prot;Acc:Q8GS71]
AT3G23620.1	4.325	0.004992238	0	3.5		Ribosome production factor 2 homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9LUG5]
AT4G10480.1	2.768	0.005049513	0.4	7		Putative alpha NAC [Source:UniProtKB/TrEMBL;Acc:Q0WWN5]
ATCG00810.1	4.332	0.005103364	0	3.5	RPL22	50S ribosomal protein L22, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P56795]
AT5G46070.1	4.336	0.005154091	0	3.5		Guanylate-binding family protein [Source:UniProtKB/TrEMBL;Acc:F4KG14]
AT4G12600.2	2.596	0.005154091	0.8	8.25		Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein [Source:UniProtKB/TrEMBL;Acc:F4JRD3]
AT1G27430.1	4.419	0.005193111	0	4		GYF domain-containing protein [Source:UniProtKB/TrEMBL;Acc:F4HSW8]
ATCG00500.1	2.896	0.00522589	1	16.5	ACCD	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta, chloroplastic [Source:UniProtKB/TrEMBL;Acc:A
AT4G30890.1	4.778	0.005240812	0	5	UBP24	Ubiquitin carboxyl-terminal hydrolase 24 [Source:UniProtKB/Swiss-Prot;Acc:Q9FPS3]
AT4G36350.1	4.383	0.005247216	0	3.25	PAP25	Purple acid phosphatase 25 [Source:UniProtKB/Swiss-Prot;Acc:O23244]
AT3G22430.1	4.318	0.005424946	0	3.5		RNA recognition motif XS domain protein [Source:UniProtKB/TrEMBL;Acc:F4J099]
AT3G02220.1	4.346	0.005469297	0	3.25		Small acidic-like protein [Source:UniProtKB/TrEMBL;Acc:Q9SRT8]
AT1G04980.1	4.374	0.005494035	0	3.25	PDIL2-2	Protein disulfide-isomerase like 2-2 [Source:UniProtKB/Swiss-Prot;Acc:Q9MAU6]
AT3G04290.1	4.392	0.005669816	0	3.5	LTL1	LTL1 [Source:UniProtKB/TrEMBL;Acc:A0A178VLC7]
AT2G01720.1	3.124	0.005688929	0.2	5.25	OST1B	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1B [Source:UniProtKB/Swiss-Prot;Acc:Q9
AT2G23070.1	4.439	0.005906568	0	3.5	CKA4	Casein kinase II subunit alpha-4, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:O64816]
AT4G20980.1	2.993	0.005906568	0.6	12.75	TIF3D1	Eukaryotic translation initiation factor 3 subunit D [Source:UniProtKB/Swiss-Prot;Acc:P56820]
AT3G54540.1	3.322	0.005927363	0.2	6.25	ABCF4	ABC transporter F family member 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9M1H3]
AT2G19520.1	2.509	0.005927363	0.6	8	MSI4	NFC4 [Source:UniProtKB/TrEMBL;Acc:A0A178VX04]
AT4G16340.1	4.280	0.006024779	0	3.5	SPK1	Guanine nucleotide exchange factor SPIKE 1 [Source:UniProtKB/Swiss-Prot;Acc:Q8SAB7]
AT5G42390.1	4.768	0.006111385	0	4.75	SPP	Stromal processing peptidase, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9FIH8]
AT2G19870.1	4.305	0.006119593	0	3.5		Uncharacterized protein At2g19870; F6F22.10 [Source:UniProtKB/TrEMBL;Acc:Q8W497]
AT3G57660.1	4.273	0.006197042	0	3.5	NRPA1	DNA-directed RNA polymerase I subunit 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SVY0]
AT4G39960.1	4.266	0.00628915	0	3.25		AT4g39960/T5J17_130 [Source:UniProtKB/TrEMBL;Acc:Q940V1]
AT5G16140.1	4.266	0.00628915	0	3.25	CRS2B	Chloroplastic group IIB intron splicing facilitator CRS2-B, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9LF14]
AT2G34970.1	4.280	0.006503988	0	3.25		At2g34970 [Source:UniProtKB/TrEMBL;Acc:O64760]
AT1G73620.1	3.159	0.006503988	0.2	5		Pathogenesis-related thaumatin superfamily protein [Source:UniProtKB/TrEMBL;Acc:Q9C9U9]
AT1G24764.1	4.818	0.006612807	0	5.75	ATMAP70-2	microtubule-associated proteins 70-2 [Source:TAIR;Acc:AT1G24764]

AT3G53460.4	3.196	0.00669436	1.4	29.75	CP29	Chloroplast RNA-binding protein 29 [Source:UniProtKB/TrEMBL;Acc:F4JAF3]
AT3G23780.1	4.384	0.006881749	0	4	NRPD2	DNA-directed RNA polymerase subunit beta [Source:UniProtKB/TrEMBL;Acc:A0A178VA17]
AT5G67630.1	3.097	0.007029366	0.2	5.25		RuvB-like helicase [Source:UniProtKB/TrEMBL;Acc:Q9FJW0]
AT1G68720.1	4.259	0.007072571	0	3.5	TADA	tRNA(adenine(34)) deaminase, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9S7I0]
AT3G18390.1	4.263	0.007158181	0	3.5	EMB1865	CRS1 / YhbY (CRM) domain-containing protein [Source:UniProtKB/TrEMBL;Acc:Q9LS51]
AT5G42540.2	4.263	0.007158181	0	3.5	XRN2	Exoribonuclease 2 [Source:UniProtKB/TrEMBL;Acc:F4K1L3]
AT4G14360.1	4.226	0.007173103	0	3.25		Probable methyltransferase PMT3 [Source:UniProtKB/Swiss-Prot;Acc:Q93YV7]
AT4G13850.1	2.367	0.00734697	0.6	7.25	RBG2	Glycine-rich RNA-binding protein 2, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9SVM8]
AT5G04290.1	4.704	0.007427368	0	5.25	RDM3	Protein RNA-directed DNA methylation 3 [Source:UniProtKB/Swiss-Prot;Acc:F4JW79]
AT1G78920.1	4.295	0.007778063	0	3.5	AVPL1	Pyrophosphate-energized membrane proton pump 2 [Source:UniProtKB/Swiss-Prot;Acc:Q56ZN6]
AT4G15770.1	4.216	0.007778063	0	3		60S ribosome subunit biogenesis protein NIP7 homolog [Source:UniProtKB/TrEMBL;Acc:Q6NM52]
AT3G56910.1	4.344	0.007802335	0	3.25	PSRP5	50S ribosomal protein 5, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9LER7]
AT3G08620.1	4.242	0.007896642	0	3		KH domain-containing protein At3g08620 [Source:UniProtKB/Swiss-Prot;Acc:Q8GYR4]
AT3G26400.1	4.671	0.008137306	0	4.5	EIF4B1	Eukaryotic translation initiation factor 4B1 [Source:UniProtKB/Swiss-Prot;Acc:Q9LIN5]
AT2G31970.1	4.178	0.008441093	0	3.25	RAD50	DNA repair protein RAD50 [Source:UniProtKB/Swiss-Prot;Acc:Q9SL02]
AT4G10710.1	4.245	0.008530317	0	3.5	SPT16	FACT complex subunit SPT16 [Source:UniProtKB/Swiss-Prot;Acc:O82491]
AT3G55620.1	3.079	0.008530317	0.2	4.75	EIF6-2	Eukaryotic translation initiation factor 6-2 [Source:UniProtKB/Swiss-Prot;Acc:Q9M060]
AT3G24480.1	4.168	0.008602527	0	3	LRX4	Leucine-rich repeat extensin-like protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9LHF1]
AT3G23990.1	4.134	0.008729173	4	118.75	CPN60	Chaperonin CPN60, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:P29197]
AT1G60780.1	4.173	0.008843207	0	3.25	AP1M2	AP-1 complex subunit mu-2 [Source:UniProtKB/Swiss-Prot;Acc:O22715]
AT1G12920.1	3.019	0.009032091	0.2	4.75	ERF1-2	Eukaryotic peptide chain release factor subunit 1-2 [Source:UniProtKB/Swiss-Prot;Acc:Q9LPV8]
AT1G76940.1	4.165	0.009087264	0	3.25	NSRA	Nuclear speckle RNA-binding protein A [Source:UniProtKB/Swiss-Prot;Acc:A1A6K6]
AT3G07660.1	4.165	0.009087264	0	3.25		Flocculation protein (DUF1296) [Source:UniProtKB/TrEMBL;Acc:Q0WUY5]
AT5G13530.1	4.169	0.009189951	0	3.25	KEG	E3 ubiquitin-protein ligase KEG [Source:UniProtKB/Swiss-Prot;Acc:Q9FY48]
AT3G03060.1	4.756	0.009201424	0	4.5		P-loop containing nucleoside triphosphate hydrolases superfamily protein [Source:UniProtKB/TrEMBL;Acc:Q0WV
AT4G31340.1	2.606	0.009261818	0.4	6.25		AT4G31340 protein [Source:UniProtKB/TrEMBL;Acc:Q8GUN1]
065.1-ATCG01230.1-ATCG	4.180	0.009433069	0	3.25	NA	NA
AT3G63490.1	2.409	0.009538262	1.8	19.25	RPL1	50S ribosomal protein L1, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9LY66]
AT4G09730.1	4.194	0.009669922	0	3.25	RH39	RH39 [Source:UniProtKB/TrEMBL;Acc:A0A178V449]
AT5G20240.1	4.194	0.009669922	0	3.25	PI	Floral homeotic protein PISTILLATA [Source:UniProtKB/Swiss-Prot;Acc:P48007]
AT5G23570.1	4.194	0.009669922	0	3.25	SGS3	Protein SUPPRESSOR OF GENE SILENCING 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9LDX1]
AT3G11910.1	2.585	0.009718406	0.4	6.5	UBP13	UBP13 [Source:UniProtKB/TrEMBL;Acc:A0A178VEE0]
AT5G63550.2	4.445	0.009736902	0	3.5		DEK domain-containing chromatin associated protein [Source:UniProtKB/TrEMBL;Acc:Q9FFQ8]
AT3G13570.1	2.667	0.009892691	0.4	6.25	SCL30A	Serine/arginine-rich SC35-like splicing factor SCL30A [Source:UniProtKB/Swiss-Prot;Acc:Q9LHP2]
AT1G18250.2	3.095	0.010277809	0.2	4.75	ATLP-1	Pathogenesis-related thaumatin superfamily protein [Source:UniProtKB/TrEMBL;Acc:F4IAP1]
AT5G30510.1	2.782	0.010289598	0.8	12.25	RPS1	30S ribosomal protein S1, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q93VC7]
AT1G17220.1	4.153	0.010335303	0	3.25	FUG1	Translation initiation factor IF-2, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9SHI1]
ATCG00830.1-ATCG01310.	2.299	0.010605822	0.8	9.25	NA	NA
AT5G06160.1	4.147	0.010713196	0	3	ATO	Splicing factor SF3a60 homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9FG01]
AT5G27720.1	4.147	0.010713196	0	3	LSM4	LSM4 [Source:UniProtKB/TrEMBL;Acc:A0A178UIB9]
AT1G68875.1	4.141	0.010713196	0	2.75		Uncharacterized protein At1g68875 [Source:UniProtKB/TrEMBL;Acc:Q84WC1]
AT4G25730.1	4.110	0.010713196	0	3		Putative rRNA methyltransferase [Source:UniProtKB/TrEMBL;Acc:F4JTD2]
AT3G56150.1	2.846	0.010713196	1.2	18	TIF3C1	Eukaryotic translation initiation factor 3 subunit C [Source:UniProtKB/Swiss-Prot;Acc:O49160]
AT2G33430.1	4.147	0.010813306	0	2.75	NA	NA
AT5G62200.1	4.147	0.010813306	0	2.75	ATS3B	Embryo-specific protein ATS3B [Source:UniProtKB/Swiss-Prot;Acc:Q6NPM5]
AT1G01510.1	2.694	0.010889851	0.4	5.5	AN	C-terminal binding protein AN [Source:UniProtKB/Swiss-Prot;Acc:O23702]
AT2G27030.3-AT3G56800.	2.202	0.010918033	1.6	17	NA	NA
AT3G02880.1	3.015	0.010999068	0.2	4.75		Probable inactive receptor kinase At3g02880 [Source:UniProtKB/Swiss-Prot;Acc:Q9M8T0]
ATCG00190.1	4.103	0.011077214	0	2.75	RPOB	DNA-directed RNA polymerase subunit beta [Source:UniProtKB/TrEMBL;Acc:A0A1B1W4T9]
AT1G55325.2	4.155	0.011173023	0	3.25	GCT	RNA polymerase II transcription mediators [Source:TAIR;Acc:AT1G55325]
AT1G80750.1	4.136	0.011524477	0	2.75	RPL7A	60S ribosomal protein L7-1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SAI5]
AT2G04280.1	4.075	0.011596968	0	3		At2g04280 [Source:UniProtKB/TrEMBL;Acc:Q9SI07]
AT4G12250.1	4.075	0.011596968	0	3	GAE5	GAE5 [Source:UniProtKB/TrEMBL;Acc:A0A178UUD1]
AT5G10810.1	4.140	0.011606476	0	3	ATER	Enhancer of rudimentary homolog [Source:UniProtKB/Swiss-Prot;Acc:Q96319]
AT2G43610.1	4.063	0.011928484	0	3		Endochitinase At2g43610 [Source:UniProtKB/Swiss-Prot;Acc:O22842]
AT5G65110.1	4.063	0.011928484	0	3	ACX2	Acyl-coenzyme A oxidase [Source:UniProtKB/TrEMBL;Acc:A0A178UA38]
AT4G16280.2	4.385	0.012013304	0	3.75	FCA	Flowering time control protein FCA [Source:UniProtKB/Swiss-Prot;Acc:O04425]
AT3G23300.1	4.067	0.012013304	0	3		S-adenosyl-L-methionine-dependent methyltransferases superfamily protein [Source:UniProtKB/TrEMBL;Acc:A0
AT4G19210.1	3.012	0.01205789	0.2	5	ABCE2	ABC transporter E family member 2 [Source:UniProtKB/Swiss-Prot;Acc:Q8LPJ4]
AT4G25740.1	2.571	0.012089509	0.4	6	RPS10A	40S ribosomal protein S10-1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SW09]
AT1G48900.1	4.079	0.012226869	0	3	SRP-54C	Signal recognition particle 54 kDa protein 3 [Source:UniProtKB/Swiss-Prot;Acc:P49967]
AT1G66270.1	4.083	0.012286243	0	3	BGLU21	BGLU21 [Source:UniProtKB/TrEMBL;Acc:A0A178WA93]
AT1G67170.1	4.052	0.012334974	0	2.75	FLXL2	Protein FLX-like 2 [Source:UniProtKB/Swiss-Prot;Acc:Q84TD8]
AT1G69530.3	4.088	0.012340445	0	3	ATEXPA1	AT1G69530 protein [Source:UniProtKB/TrEMBL;Acc:C0Z241]
AT1G20880.1	4.060	0.012393854	0	2.75		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:F4HUQ5]
AT4G36680.1	4.077	0.012522307	0	2.75		Pentatricopeptide repeat-containing protein At4g36680, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9M06
AT1G17260.1	2.550	0.012696556	0.4	6.25	AHA10	Plasma membrane ATPase [Source:UniProtKB/TrEMBL;Acc:A0A178W776]
AT1G10850.1	4.053	0.012849463	0	3		Leucine-rich repeat protein kinase family protein [Source:UniProtKB/TrEMBL;Acc:Q940B9]
AT1G48090.1	4.055	0.013282243	0	3		calcium-dependent lipid-binding family protein [Source:TAIR;Acc:AT1G48090]
AT3G19390.1	4.078	0.01339175	0	3	RD21C	Probable cysteine protease RD21C [Source:UniProtKB/Swiss-Prot;Acc:Q9LT78]
AT2G46280.1	2.114	0.013492376	1.2	11.5	TIF3I1	Eukaryotic translation initiation factor 3 subunit I [Source:UniProtKB/Swiss-Prot;Acc:Q38884]
AT1G32380.1	4.094	0.013608536	0	3	PRS2	Ribose-phosphate pyrophosphokinase 2, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q42583]
AT1G18660.4	4.046	0.013648948	0	3		Zinc finger (C3HC4-type RING finger) family protein [Source:UniProtKB/TrEMBL;Acc:F4ICB5]
AT2G38280.1	4.046	0.013648948	0	3	AMPD	FAC1 [Source:UniProtKB/TrEMBL;Acc:A0A178VY66]
AT3G26618.1	2.923	0.013773658	0.2	4.75	ERF1-3	ERF1-3 [Source:UniProtKB/TrEMBL;Acc:A0A178VIS2]
AT5G28060.1	2.912	0.013793239	0.2	4.5	RPS24B	40S ribosomal protein S24-2 [Source:UniProtKB/Swiss-Prot;Acc:Q8LC83]
ATCG00170.1	4.086	0.013805926	0	2.75	RPOC2	DNA-directed RNA polymerase subunit beta'' [Source:UniProtKB/Swiss-Prot;Acc:P56764]
AT4G04920.1	4.055	0.013805926	0	3	MED16	Mediator of RNA polymerase II transcription subunit 16 [Source:UniProtKB/Swiss-Prot;Acc:F4JGZ1]
AT1G55170.1	4.177	0.013837201	0	3	FLXL3	Protein FLX-like 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9C717]
AT1G48610.1	4.038	0.014333765	0	2.75		Putative DNA-binding protein At1g48610 [Source:UniProtKB/Swiss-Prot;Acc:Q94AD1]
AT3G56860.3	4.199	0.014353085	0	3.5	UBA2A	UBP1-associated protein 2A [Source:UniProtKB/Swiss-Prot;Acc:Q9LES2]
AT5G44200.1	4.199	0.014353085	0	3.5	CBP20	Nuclear cap-binding protein subunit 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9XFD1]
AT1G21690.3	3.999	0.014409858	0	2.75	EMB1968	ATPase family associated with various cellular activities (AAA) [Source:UniProtKB/TrEMBL;Acc:F4HY43]
AT1G52930.1	4.151	0.014875519	0	2.75	BRIX1-2	Ribosome biogenesis protein BRX1 homolog 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9C928]
AT5G51280.1	4.030	0.015048358	0	2.5	RH35	DEAD-box ATP-dependent RNA helicase 35 [Source:UniProtKB/Swiss-Prot;Acc:Q9LU46]
AT2G02160.1	5.243	0.015532712	0	7		Zinc finger CCCH domain-containing protein 17 [Source:UniProtKB/Swiss-Prot;Acc:Q9ZUM0]
AT5G46580.1	3.951	0.016306021	0	2.75		Pentatricopeptide repeat-containing protein At5g46580, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9LS25]
AT5G64420.1	3.951	0.016306021	0	2.75		DNA polymerase V family [Source:UniProtKB/TrEMBL;Acc:Q9FGF4]
AT3G15190.1	2.538	0.016306021	0.4	6	RPS20	PRPS20 [Source:UniProtKB/TrEMBL;Acc:A0A178VMN1]
AT5G14910.1	3.956	0.016356186	0	2.75		AT5g14910/F2G14_30 [Source:UniProtKB/TrEMBL;Acc:Q93VK7]
AT5G62270.2	3.956	0.016356186	0	2.75		Gb [Source:UniProtKB/TrEMBL;Acc:Q9LVA9]
AT3G04920.1	2.886	0.016385842	0.2	4.25	RPS24A	40S ribosomal protein S24-1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SS17]
AT2G47680.1	3.964	0.016423575	0	2.75		DExH-box ATP-dependent RNA helicase DExH8 [Source:UniProtKB/Swiss-Prot;Acc:O22243]
AT5G50250.1	3.964	0.016423575	0	2.75	CP31B	RNA-binding protein CP31B, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9FGS0]
AT5G65900.1	3.964	0.016423575	0	2.75	RH27	DEAD-box ATP-dependent RNA helicase 27 [Source:UniProtKB/Swiss-Prot;Acc:Q9SB89]
AT4G18670.1	3.979	0.016537193	0	2.5	LRX5	Leucine-rich repeat extensin-like protein 5 [Source:UniProtKB/Swiss-Prot;Acc:Q9SN46]
AT1G74690.1	3.943	0.016537193	0	2.5	IQD31	Protein IQ-DOMAIN 31 [Source:UniProtKB/Swiss-Prot;Acc:Q8L4D8]
AT5G47210.1	2.060	0.016537193	4	35.5	RGGC	RGG repeats nuclear RNA binding protein C [Source:UniProtKB/Swiss-Prot;Acc:Q9LVT8]
AT1G23750.1	3.942	0.016647757	0	2.75		F5O8.30 protein [Source:UniProtKB/TrEMBL;Acc:Q9ZUC0]
AT2G32930.2	3.942	0.016647757	0	2.75	ZFN2	zinc finger nuclease 2 [Source:TAIR;Acc:AT2G32930]

AT2G33210.1	3.788	0.016647757	3.8	92.25	HSP60-2	Chaperonin CPN60-like 1, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q8L7B5]
ATCG00330.1	3.982	0.016708739	0	2.75	RPS14	30S ribosomal protein S14, chloroplastic [Source:UniProtKB/TrEMBL;Acc:A0A1B1W4T5]
AT5G02450.1	2.283	0.016708739	1	10.75	RPL36C	60S ribosomal protein L36-3 [Source:UniProtKB/Swiss-Prot;Acc:Q9LZ57]
AT3G15680.1	4.077	0.016785124	0	2.75		At3g15680 [Source:UniProtKB/TrEMBL;Acc:Q9LW11]
AT4G21450.3	4.238	0.016923368	0	3.25		PapD-like superfamily protein [Source:UniProtKB/TrEMBL;Acc:F4JJI7]
AT3G06410.1	3.970	0.01704979	0	2.75		Zinc finger CCCH domain-containing protein 34 [Source:UniProtKB/Swiss-Prot;Acc:Q9SQU4]
AT1G26770.2	3.931	0.01704979	0	2.5	ATEXPA10	Expansin A10 [Source:UniProtKB/TrEMBL;Acc:F4HPC1]
AT2G31900.1	4.055	0.017082573	0	3	XIF	myosin-like protein XIF [Source:TAIR;Acc:AT2G31900]
AT1G08450.1	3.971	0.017149821	0	2.5	CRT3	Calreticulin-3 [Source:UniProtKB/Swiss-Prot;Acc:O04153]
AT1G67325.2	3.947	0.017355315	0	2.5		RanBP2-type zinc finger protein At1g67325 [Source:UniProtKB/Swiss-Prot;Acc:Q8GZ43]
AT1G02080.1	3.936	0.017379536	0	2.75		Transcription regulator [Source:UniProtKB/TrEMBL;Acc:F4HVV6]
AT4G12640.1	3.945	0.017555019	0	2.75		RNA recognition motif (RRM)-containing protein [Source:UniProtKB/TrEMBL;Acc:F4JRD9]
AT1G33811.1	3.960	0.017819731	0	2.75		GDSL esterase/lipase At1g33811 [Source:UniProtKB/Swiss-Prot;Acc:Q8L5Z1]
AT5G22330.1	4.012	0.018763165	0	3	RIN1	RuvB-like protein 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9FMR9]
AT1G78260.1	3.924	0.018835507	0	2.75		At1g78260 [Source:UniProtKB/TrEMBL;Acc:Q8LF90]
AT3G16810.1	3.924	0.018835507	0	2.75	APUM24	Pumilio homolog 24 [Source:UniProtKB/Swiss-Prot;Acc:Q9LRZ3]
AT3G55610.1	3.924	0.018835507	0	2.75	P5CSB	Delta-1-pyrroline-5-carboxylate synthase [Source:UniProtKB/TrEMBL;Acc:A0A178VF78]
AT4G26100.1	3.924	0.018835507	0	2.75	CKL1	CKL1 [Source:UniProtKB/TrEMBL;Acc:A0A178UYM7]
AT4G39520.1	2.863	0.018952437	0.2	4.25	DRG3	Developmentally-regulated G-protein 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9SVA6]
AT3G61260.1	3.983	0.019213233	0	2.75		Uncharacterized protein At3g61260 [Source:UniProtKB/Swiss-Prot;Acc:Q9M2D8]
AT2G36850.1	3.916	0.019264966	0	2.5	CALS10	Callose synthase 10 [Source:UniProtKB/Swiss-Prot;Acc:Q9SJM0]
AT3G09740.1	3.916	0.019264966	0	2.5	SYP71	At3g09740 [Source:UniProtKB/TrEMBL;Acc:Q2HIU8]
AT3G14240.1	3.969	0.019330378	0	2.75	SBT1.5	Subtilisin-like protease SBT1.5 [Source:UniProtKB/Swiss-Prot;Acc:Q9LUM3]
AT5G62740.1	3.965	0.019870328	0	2.5	HIR1	HIR4 [Source:UniProtKB/TrEMBL;Acc:A0A178U8N7]
AT1G79850.1	3.882	0.019919269	0	2.5	RPS17	30S ribosomal protein S17, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P16180]
AT1G10200.1	3.877	0.019941941	0	2.5	WLIM1	WLIM1 [Source:UniProtKB/TrEMBL;Acc:A0A178W4R1]
AT4G26480.1	3.877	0.019941941	0	2.5		KH domain-containing protein At4g26480 [Source:UniProtKB/Swiss-Prot;Acc:Q0WLR1]
AT2G17970.1	4.163	0.020294993	0	3.5		2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein [Source:UniProtKB/TrEMBL;Acc:Q9SL
AT4G01310.1	2.484	0.020294993	0.4	5.75	RPL5	50S ribosomal protein L5, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:O04603]
AT3G05970.1	3.816	0.020841412	0	2.25	LACS6	Long chain acyl-CoA synthetase 6, peroxisomal [Source:UniProtKB/Swiss-Prot;Acc:Q8LPS1]
AT5G61780.1	2.501	0.020841412	2	21.75	TSN2	Ribonuclease TUDOR 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9FLT0]
AT5G11170.1	2.416	0.021224166	1.6	17.25	RH15	DEAD-box ATP-dependent RNA helicase 56 [Source:UniProtKB/Swiss-Prot;Acc:Q9LFN6]
AT1G72440.1	3.880	0.021280918	0	2.5	EDA25	CCAAT-binding factor [Source:UniProtKB/TrEMBL;Acc:F4IDC2]
AT4G15000.1	2.288	0.021333747	0.6	6.75	RPL27C	60S ribosomal protein L27 [Source:UniProtKB/TrEMBL;Acc:Q0WRB8]
AT1G63780.1	3.872	0.021445641	0	2.5	IMP4	IMP4 [Source:UniProtKB/TrEMBL;Acc:A0A178WDJ3]
AT3G58640.1	4.159	0.021508539	0	3.25		AT3g58640/F14P22_230 [Source:UniProtKB/TrEMBL;Acc:Q94AB2]
AT1G32580.1	3.801	0.021588029	0	2.25	MORF5	Multiple organellar RNA editing factor 5, chloroplastic/mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9C7Y2]
AT4G35785.5	3.897	0.02172897	0	2.25		At4g35785 [Source:UniProtKB/TrEMBL;Acc:Q494N5]
AT3G22270.1	4.131	0.021852995	0	3.25	PAT1H1	Protein PAT1 homolog 1 [Source:UniProtKB/Swiss-Prot;Acc:F4J077]
AT2G40360.1	3.962	0.022127228	0	2.75	BOP1	Ribosome biogenesis protein BOP1 homolog [Source:UniProtKB/Swiss-Prot;Acc:F4IH25]
AT1G06070.1	3.955	0.022127228	0	2.75		At1g06070 [Source:UniProtKB/TrEMBL;Acc:Q9LNE0]
AT3G02110.1	3.821	0.022131818	0	2.5	SCPL25	Serine carboxypeptidase-like 25 [Source:UniProtKB/Swiss-Prot;Acc:Q8L9Y0]
AT1G22910.3	3.830	0.022204445	0	2.5		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:F4I321]
AT5G10370.1	3.830	0.022204445	0	2.5		ATP-dependent RNA helicase DEAH12, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:F4KGU4]
AT1G24450.1	3.835	0.022238252	0	2.5	NFD2	Protein NUCLEAR FUSION DEFECTIVE 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9FYL8]
AT2G27040.1	2.374	0.022262265	1.2	13.5	AGO4	Protein argonaute 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9ZVD5]
AT2G35390.2	3.998	0.022397692	0	2.75	PRS1	Ribose-phosphate pyrophosphokinase 1, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q42581]
AT5G57460.1	3.793	0.022626267	0	2.25		At5g57460 [Source:UniProtKB/TrEMBL;Acc:Q9FKM4]
ATCG01240.1-ATCG00900.1	2.127	0.022638337	0.6	6.25	NA	NA
AT1G49740.1	3.815	0.022990267	0	2.5		F14J22.5 protein [Source:UniProtKB/TrEMBL;Acc:Q9FXA0]
AT1G77810.1	3.815	0.022990267	0	2.5		Hexosyltransferase [Source:UniProtKB/TrEMBL;Acc:A0A178WEM8]
AT3G18520.2	3.815	0.022990267	0	2.5	HDA15	Histone deacetylase [Source:UniProtKB/TrEMBL;Acc:F4J8S1]
AT2G04390.1	2.465	0.022990267	0.4	6	RPS17A	40S ribosomal protein S17-1 [Source:UniProtKB/Swiss-Prot;Acc:P49205]
AT1G62330.1	3.825	0.023116518	0	2.5	OFUT15	O-fucosyltransferase 15 [Source:UniProtKB/Swiss-Prot;Acc:F4HYR4]
AT4G02990.1	3.845	0.02325789	0	2.5	MTERF4	Transcription termination factor MTERF4, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9ZT96]
AT5G46420.1	3.845	0.02325789	0	2.5		16S rRNA processing protein RimM family [Source:UniProtKB/TrEMBL;Acc:Q9FHG3]
AT1G29350.1	2.869	0.023278047	0.2	4.25		Kinase-related protein of unknown function (DUF1296) [Source:TAIR;Acc:AT1G29350]
AT3G56460.1	2.426	0.023278047	0.4	5.5		GroES-like zinc-binding alcohol dehydrogenase family protein [Source:UniProtKB/TrEMBL;Acc:Q9LXZ4]
AT5G62350.1	2.793	0.023391656	0.2	4.25		Plant invertase/pectin methylesterase inhibitor superfamily protein [Source:UniProtKB/TrEMBL;Acc:Q9LVA4]
AT3G48710.1	4.099	0.023465532	0	3.25		At3g48710 [Source:UniProtKB/TrEMBL;Acc:Q9SMM8]
AT3G12400.1	3.873	0.023493133	0	2.5	ELC	Ubiquitin-conjugating enzyme/RWD-like protein [Source:UniProtKB/TrEMBL;Acc:A0A1I9LRG3]
AT1G56360.1	3.810	0.023508433	0	2.5	PAP6	Purple acid phosphatase 6 [Source:UniProtKB/Swiss-Prot;Acc:Q9C510]
AT4G17560.1	3.810	0.023508433	0	2.5		50S ribosomal protein L19-1, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q8W463]
AT5G40480.1	3.810	0.023508433	0	2.5	GB210	Nuclear pore complex protein GP210 [Source:UniProtKB/Swiss-Prot;Acc:F4KHD8]
AT2G03510.1	3.805	0.023508433	0	2.5		Expressed protein [Source:UniProtKB/TrEMBL;Acc:Q9ZQ87]
AT3G26560.1	3.805	0.023508433	0	2.5		Probable pre-mRNA-splicing factor ATP-dependent RNA helicase DEAH5 [Source:UniProtKB/Swiss-Prot;Acc:Q389
AT4G00450.1	3.805	0.023508433	0	2.5	MED12	Mediator of RNA polymerase II transcription subunit 12 [Source:UniProtKB/Swiss-Prot;Acc:H3K2Y6]
AT1G73670.1	3.907	0.023563343	0	2.75	MPK15	Mitogen-activated protein kinase 15 [Source:UniProtKB/Swiss-Prot;Acc:Q9C9U4]
AT1G64880.1	3.841	0.023752168	0	2.5		At1g64880 [Source:UniProtKB/TrEMBL;Acc:Q6GKU7]
AT2G33730.1	3.841	0.023752168	0	2.5	RH21	DEAD-box ATP-dependent RNA helicase 21 [Source:UniProtKB/Swiss-Prot;Acc:P93008]
AT1G14650.1	3.880	0.024179954	0	2.5		Probable splicing factor 3A subunit 1 [Source:UniProtKB/Swiss-Prot;Acc:Q8RXF1]
AT4G22380.1	3.106	0.024179954	0.2	4		At4g22380 [Source:UniProtKB/TrEMBL;Acc:Q8LCC7]
AT1G21700.1	3.977	0.024694358	0	3	SWI3C	SWI/SNF complex subunit SWI3C [Source:UniProtKB/Swiss-Prot;Acc:Q9XI07]
AT4G10110.1	4.027	0.0247893	0	3		At4g10100 [Source:UniProtKB/TrEMBL;Acc:Q8GW25]
AT1G29670.1	2.515	0.0247893	2	24.5		GDSL-like Lipase/Acylhydrolase superfamily protein [Source:TAIR;Acc:AT1G29670]
AT5G44320.1	2.741	0.024822499	0.2	4.25		Eukaryotic translation initiation factor 3 subunit D [Source:UniProtKB/TrEMBL;Acc:Q9FKV6]
AT1G15440.1	3.763	0.025719909	0	2.25	PWP2	Periodic tryptophan protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q8VYZ5]
AT1G17880.1	3.763	0.025719909	0	2.25	BTF3	Basic transcription factor 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9SMW7]
AT2G42800.1	2.059	0.026579969	0.8	7.5	AtRLP29	RLP29 [Source:UniProtKB/TrEMBL;Acc:A0A178VT48]
AT3G06530.2	3.804	0.026702367	0	2.5		ARM repeat superfamily protein [Source:UniProtKB/TrEMBL;Acc:A0A1I9LP28]
AT2G37170.1	3.811	0.026725962	0	2.5	PIP2B	plasma membrane intrinsic protein 2 [Source:TAIR;Acc:AT2G37170]
AT3G04740.1	3.811	0.026725962	0	2.5	MED14	Mediator of RNA polymerase II transcription subunit 14 [Source:UniProtKB/Swiss-Prot;Acc:Q9SR02]
AT2G47580.1	3.867	0.026966472	0	2.5	U1A	U1A [Source:UniProtKB/TrEMBL;Acc:A0A178VZ67]
AT5G08120.1	4.097	0.027223995	0	3.25	MPB2C	MPB2C [Source:UniProtKB/TrEMBL;Acc:A0A178UL06]
AT4G04940.1	3.783	0.027601629	0	2.25		Putative WD-repeat membrane protein [Source:UniProtKB/TrEMBL;Acc:Q93YS7]
AT5G16930.1	3.891	0.027909232	0	2.25		AAA-type ATPase family protein [Source:UniProtKB/TrEMBL;Acc:Q8RXI0]
AT3G50360.1	3.734	0.028191486	0	2.25	CML20	Probable calcium-binding protein CML20 [Source:UniProtKB/Swiss-Prot;Acc:O82659]
AT3G14010.1	3.879	0.028449456	0	2.75	CID4	CTC-interacting domain 4 [Source:UniProtKB/TrEMBL;Acc:A0A1I9LLC0]
AT1G04160.1	3.780	0.028538839	0	2.5	XI-B	Myosin-8 [Source:UniProtKB/Swiss-Prot;Acc:F4I460]
AT4G31430.2	3.780	0.028538839	0	2.5		unknown protein; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 s
AT5G27700.1	2.735	0.029010428	0.2	4	RPS21C	40S ribosomal protein S21 [Source:UniProtKB/TrEMBL;Acc:A0A178UBS9]
AT1G52040.1	2.118	0.02965351	1.4	12.25	MBP1	Myrosinase-binding protein 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9SAV0]
AT1G31290.1	4.042	0.029937213	0	3.25	AGO3	Protein argonaute 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9SHF2]
AT5G45330.1	3.970	0.030126756	0	3	DCP5-L	Decapping 5-like protein [Source:UniProtKB/Swiss-Prot;Acc:Q9FH77]
AT5G61960.1	3.917	0.030126756	0	2.5	ML1	Protein MEI2-like 1 [Source:UniProtKB/Swiss-Prot;Acc:Q8W4I9]
AT1G16870.1	3.688	0.030126756	0	2.25		Mitochondrial 28S ribosomal protein S29-like protein [Source:UniProtKB/TrEMBL;Acc:Q8W4K2]
AT3G47800.1	3.688	0.030126756	0	2.25		Aldose 1-epimerase [Source:UniProtKB/TrEMBL;Acc:Q9STT3]
AT4G13340.1	3.688	0.030126756	0	2.25	LRX3	Leucine-rich repeat extensin-like protein 3 [Source:UniProtKB/Swiss-Prot;Acc:Q9T0K5]
ATCG00740.1	3.688	0.030126756	0	2.25	RPOA	DNA-directed RNA polymerase subunit alpha [Source:UniProtKB/Swiss-Prot;Acc:P56762]
AT5G08450.1	4.044	0.030887088	0	3.25		Uncharacterized protein At5g08450 [Source:UniProtKB/TrEMBL;Acc:Q8RWV9]

AT3G27180.1	3.704	0.030887088	0	2.25		At3g27180 [Source:UniProtKB/TrEMBL;Acc:B3DNP3]
AT3G51050.1	3.704	0.030887088	0	2.25		FG-GAP repeat-containing protein [Source:UniProtKB/TrEMBL;Acc:F4J381]
AT5G65810.1	3.704	0.030887088	0	2.25	CGR3	Probable pectin methylesterase CGR3 [Source:UniProtKB/Swiss-Prot;Acc:Q0WPN7]
AT1G25260.1	3.698	0.030887088	0	2.25		Ribosome assembly factor mrt4 [Source:UniProtKB/TrEMBL;Acc:Q94AK8]
AT1G61790.1	3.698	0.030887088	0	2.25	OST3B	Probable dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 3B [Source:UniProtKB/Swiss-Pro
AT4G27870.1	3.698	0.030887088	0	2.25		Vacuolar iron transporter (VIT) family protein [Source:TAIR;Acc:AT4G27870]
AT5G19580.1	3.693	0.030887088	0	2.25		Glyoxal oxidase-related protein [Source:UniProtKB/TrEMBL;Acc:F4K172]
AT2G38810.1	3.682	0.030887088	0	2.25	HTA8	Probable histone H2A variant 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9SII0]
AT2G39700.1	3.682	0.030887088	0	2.25	EXPA4	Expansin-A4 [Source:UniProtKB/Swiss-Prot;Acc:O48818]
AT2G41040.1	3.682	0.030887088	0	2.25		Uncharacterized methyltransferase At2g41040, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q0WPT7]
AT1G19360.1	3.678	0.030887088	0	2.25	RRA3	Arabinosyltransferase RRA3 [Source:UniProtKB/Swiss-Prot;Acc:Q9LN62]
AT1G65590.1	3.678	0.030887088	0	2.25	HEXO3	Beta-hexosaminidase 3 [Source:UniProtKB/Swiss-Prot;Acc:Q8L7S6]
AT2G39740.1	3.678	0.030887088	0	2.25	HESO1	Protein HESO1 [Source:UniProtKB/Swiss-Prot;Acc:Q5XET5]
AT1G20970.1	3.673	0.030887088	0	2.25		FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma
ATCG00650.1	3.673	0.030887088	0	2.25	RPS18	30S ribosomal protein S18, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P56807]
AT3G10690.1	2.685	0.030887088	0.2	4	GYRA	DNA gyrase subunit A, chloroplastic/mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9CAF6]
AT3G61860.1	2.822	0.030924664	0.2	4	RS31	RSP31 [Source:UniProtKB/TrEMBL;Acc:A0A178VG60]
AT4G21150.1	2.701	0.030988213	1.4	18.25	RPN2	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2 [Source:UniProtKB/Swiss-Prot;Acc:Q93
AT5G06360.1	3.668	0.031657602	0	2		Ribosomal protein S8e family protein [Source:UniProtKB/TrEMBL;Acc:Q9FNH2]
AT5G13510.1	2.352	0.03228511	0.4	5	RPL10	50S ribosomal protein L10, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9FY50]
AT3G62870.1	2.262	0.032645995	2.6	24.25	RPL7AB	60S ribosomal protein L7a-2 [Source:UniProtKB/Swiss-Prot;Acc:Q9LZH9]
AT1G29940.1	3.758	0.032769618	0	2.5	NRPA2	DNA-directed RNA polymerase subunit beta [Source:UniProtKB/TrEMBL;Acc:A0A178W5B4]
AT5G44400.1	3.758	0.032769618	0	2.5		Berberine bridge enzyme-like 26 [Source:UniProtKB/Swiss-Prot;Acc:Q9FKU8]
AT3G53740.2	2.101	0.032820658	1	10.25	RPL36B	60S ribosomal protein L36 [Source:UniProtKB/TrEMBL;Acc:A0A178VA59]
AT4G25970.1	3.673	0.032894298	0	2	PSD3	Phosphatidylserine decarboxylase proenzyme 3 [Source:UniProtKB/Swiss-Prot;Acc:A4GNA8]
AT2G20490.1	3.663	0.032894298	0	2	NOP10	H/ACA ribonucleoprotein complex subunit 3-like protein [Source:UniProtKB/Swiss-Prot;Acc:Q93XX8]
AT1G75350.1	3.659	0.032894298	0	2	RPL31	50S ribosomal protein L31, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9FWS4]
AT1G69120.1	3.655	0.032894298	0	2	AP1	AP1 [Source:UniProtKB/TrEMBL;Acc:A0A178W846]
AT4G21660.2	3.655	0.032894298	0	2		proline-rich spliceosome-associated (PSP) family protein [Source:TAIR;Acc:AT4G21660]
AT1G51690.3	3.711	0.032942555	0	2.25	ATB ALPHA	protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform [Source:TAIR;Acc:AT1G51690]
AT1G04110.1	3.660	0.032942555	0	2.25	SBT1.2	SDD1 [Source:UniProtKB/TrEMBL;Acc:A0A178WPN6]
AT4G05410.1	3.660	0.032942555	0	2.25	YAO	U3 snoRNP-associated protein-like YAO [Source:UniProtKB/Swiss-Prot;Acc:Q9M0V4]
AT5G57015.1	3.660	0.032942555	0	2.25	CKL12	Ckl12 [Source:UniProtKB/TrEMBL;Acc:A0A178UEU7]
ATCG00380.1	2.237	0.034243809	0.6	6.5	RPS4	30S ribosomal protein S4, chloroplastic [Source:UniProtKB/TrEMBL;Acc:A0A1B1W4U6]
AT1G24050.1	3.662	0.035379364	0	2.25		At1g24050/T23E23_11 [Source:UniProtKB/TrEMBL;Acc:Q8L466]
AT5G15550.1	3.598	0.035797241	0	2	WDR12	Ribosome biogenesis protein WDR12 homolog [Source:UniProtKB/TrEMBL;Acc:A0A178UL95]
AT1G06190.1	3.642	0.036154395	0	2		Rho termination factor [Source:TAIR;Acc:AT1G06190]
AT3G10650.1	3.766	0.037709315	0	2.5	NUP1	NUP1 [Source:UniProtKB/TrEMBL;Acc:A0A178V9H6]
AT1G23860.1	2.341	0.037971482	0.4	4.25	RSZ21	SRZ21 [Source:UniProtKB/TrEMBL;Acc:A0A178WC08]
AT1G55570.1	2.686	0.038057675	0.2	4	sks12	At1g55570/T5A14_1 [Source:UniProtKB/TrEMBL;Acc:Q9ZVV4]
AT1G15200.3	3.645	0.038151826	0	2		Protein-protein interaction regulator family protein [Source:UniProtKB/TrEMBL;Acc:F4HZI8]
AT5G07550.1	3.754	0.038278846	0	2	GRP19	Oleosin [Source:UniProtKB/TrEMBL;Acc:Q42574]
AT2G25110.1	2.414	0.039376546	0.4	4.25	SDF2	Stromal cell-derived factor 2-like protein [Source:UniProtKB/Swiss-Prot;Acc:Q93ZE8]
AT4G11420.1	2.847	0.039655996	1.4	21.25	TIF3A1	Eukaryotic translation initiation factor 3 subunit A [Source:UniProtKB/Swiss-Prot;Acc:Q9LD55]
AT1G72710.1	3.742	0.039737313	0	2.25	CKL2	Casein kinase 1-like protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9CAI5]
AT1G73460.1	3.638	0.039747884	0	2.25		Protein kinase superfamily protein [Source:UniProtKB/TrEMBL;Acc:F4HQ88]
AT3G63130.1	3.638	0.039747884	0	2.25	RANGAP1	RAN GTPase-activating protein 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9LE82]
AT4G28540.1	3.631	0.039747884	0	2.25	CKL6	Casein kinase 1-like protein 6 [Source:UniProtKB/Swiss-Prot;Acc:Q8LPJ1]
AT1G11475.1	3.632	0.039845836	0	2	NRPB10	NRPE10 [Source:UniProtKB/TrEMBL;Acc:A0A178WP73]
AT3G52660.1	3.632	0.039845836	0	2		At3g52660 [Source:UniProtKB/TrEMBL;Acc:Q9LXJ8]
AT2G02100.1	3.760	0.039887948	0	2	PDF2.2	PDF2.2 [Source:UniProtKB/TrEMBL;Acc:A0A178VVN9]
AT5G19950.1	3.674	0.040468635	0	2		AT5G19950 protein [Source:UniProtKB/TrEMBL;Acc:Q8L502]
AT3G53890.1	2.587	0.040468635	0.2	3.5	RPS21B	40S ribosomal protein S21-1 [Source:UniProtKB/Swiss-Prot;Acc:Q9M337]
AT5G66010.1	3.863	0.041698673	0	2.25		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:F4JZ11]
AT1G50360.1	3.726	0.04189977	0	2.5	VIII-A	Myosin-3 [Source:UniProtKB/Swiss-Prot;Acc:F4I507]
AT3G13160.1	3.540	0.04189977	0	2		Pentatricopeptide repeat-containing protein At3g13160, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9LK57
AT5G10200.1	3.540	0.04189977	0	2		ARM-repeat/Tetratricopeptide repeat (TPR)-like protein [Source:UniProtKB/TrEMBL;Acc:Q0WMT6]
AT5G55000.2	3.540	0.04189977	0	2	FIP2	FH protein interacting protein FIP2 [Source:UniProtKB/Swiss-Prot;Acc:Q9SE95]
AT2G03640.3	3.530	0.04189977	0	2		Nuclear transport factor 2 (NTF2) family protein with RNA binding (RRM-RBD-RNP motifs) domain-containing prote
AT2G42830.2	3.530	0.04189977	0	2	SHP2	K-box region and MADS-box transcription factor family protein [Source:TAIR;Acc:AT2G42830]
AT3G18580.1	3.530	0.04189977	0	2		Nucleic acid-binding, OB-fold-like protein [Source:UniProtKB/TrEMBL;Acc:Q9LII1]
AT3G28730.1	3.530	0.04189977	0	2	SSRP1	FACT complex subunit SSRP1 [Source:UniProtKB/TrEMBL;Acc:A0A178VCP0]
AT3G30300.1	3.530	0.04189977	0	2	OFUT27	O-fucosyltransferase 27 [Source:UniProtKB/Swiss-Prot;Acc:Q8GZ81]
AT5G66470.1	3.530	0.04189977	0	2		GTPase ERA-like, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q8VZ74]
AT1G08600.3	3.525	0.04189977	0	2	ATRX	CHR20 [Source:UniProtKB/TrEMBL;Acc:A0A178VZQ9]
AT2G44730.1	3.525	0.04189977	0	2		Alcohol dehydrogenase transcription factor Myb/SANT-like family protein [Source:UniProtKB/TrEMBL;Acc:O8051
AT4G38130.1	3.525	0.04189977	0	2	HDA19	Histone deacetylase 19 [Source:UniProtKB/Swiss-Prot;Acc:O22446]
AT4G40000.1	3.525	0.04189977	0	2		S-adenosyl-L-methionine-dependent methyltransferases superfamily protein [Source:UniProtKB/TrEMBL;Acc:F4J
AT4G31880.1	2.643	0.04189977	0.2	4		Transcriptional regulator [Source:UniProtKB/TrEMBL;Acc:Q8GUP3]
AT5G42820.2	2.207	0.04189977	0.4	4.5	U2AF35B	U2AF35B [Source:UniProtKB/TrEMBL;Acc:A0A178UM74]
AT5G50000.1	3.534	0.043401715	0	2		Protein kinase [Source:UniProtKB/TrEMBL;Acc:Q9FGB1]
AT3G03940.1	3.514	0.043441984	0	2		Protein kinase family protein [Source:UniProtKB/TrEMBL;Acc:Q9SQR8]
AT4G36940.1	3.514	0.043441984	0	2	NAPRT1	NAPRT1 [Source:UniProtKB/TrEMBL;Acc:A0A178V5W0]
AT5G37020.1	3.514	0.043441984	0	2	ARF8	Auxin response factor [Source:UniProtKB/TrEMBL;Acc:A0A178UFA9]
AT5G52650.1	2.241	0.043447812	0.4	5	RPS10C	40S ribosomal protein S10-3 [Source:UniProtKB/Swiss-Prot;Acc:Q9LTF2]
AT3G07320.1	3.553	0.043454201	0	1.75		O-Glycosyl hydrolases family 17 protein [Source:UniProtKB/TrEMBL;Acc:Q9SRT4]
AT4G18960.1	3.591	0.044095612	0	2	AG	K-box region and MADS-box transcription factor family protein [Source:TAIR;Acc:AT4G18960]
AT1G23280.1	3.569	0.044095612	0	2		Protein MAK16 homolog [Source:UniProtKB/TrEMBL;Acc:F4I4Q1]
AT4G25210.1	3.569	0.044095612	0	2	GEBPL	GLABROUS1 enhancer-binding protein-like [Source:UniProtKB/Swiss-Prot;Acc:Q9SB42]
AT3G53500.2	2.642	0.04468435	0.2	3.5	RS2Z32	RSZ32 [Source:UniProtKB/TrEMBL;Acc:A0A178V7M8]
AT5G60960.1	3.561	0.04490194	0	2	PNM1	PNM1 [Source:UniProtKB/TrEMBL;Acc:A0A178U8V9]
AT1G33470.1	3.500	0.045359399	0	1.75		RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL;Acc:Q67YS0]
AT1G52730.2	3.624	0.045417744	0	2.25		AT1G52730 protein [Source:UniProtKB/TrEMBL;Acc:Q9SSS7]
AT1G77470.1	3.517	0.045417744	0	2	RFC5	Replication factor C subunit 5 [Source:UniProtKB/Swiss-Prot;Acc:Q9CAQ8]
AT5G60460.1	3.505	0.045417744	0	2		Protein transport protein Sec61 subunit beta [Source:UniProtKB/TrEMBL;Acc:Q9FKK1]
AT2G29540.3	3.610	0.046233623	0	2	ATRPC14	RNApolymerase 14 kDa subunit [Source:UniProtKB/TrEMBL;Acc:A8MRK9]
AT5G01590.1	3.617	0.046246805	0	2	TIC56	Protein TIC 56, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q7Y1W1]
AT3G27700.1	3.984	0.047225701	0	3.25		Uncharacterized protein At3g27700 (Fragment) [Source:UniProtKB/TrEMBL;Acc:C0SVD1]
AT5G20710.1	2.034	0.047691091	1	7.25	BGAL7	Beta-galactosidase 7 [Source:UniProtKB/Swiss-Prot;Acc:Q9SCV5]
AT1G23240.1	3.714	0.047779991	0	2		Caleosin-related family protein [Source:TAIR;Acc:AT1G23240]
AT3G55170.1	3.696	0.047779991	0	1.75	RPL35C	60S ribosomal protein L35-3 [Source:UniProtKB/Swiss-Prot;Acc:Q9M3D2]
AT1G53050.1	3.552	0.047779991	0	2		F8L10.9 protein [Source:UniProtKB/TrEMBL;Acc:Q9LNN0]
ATCG00750.1	2.321	0.047826892	0.4	4.75	RPS11	30S ribosomal protein S11, chloroplastic [Source:UniProtKB/TrEMBL;Acc:A0A1B1W4X8]
AT5G11520.1	2.168	0.04801899	0.4	4.5	ASP3	Aspartate aminotransferase 3, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:P46644]
AT4G07410.1	3.525	0.048123478	0	2	PCN	WD repeat-containing protein PCN [Source:UniProtKB/Swiss-Prot;Acc:Q8RXU6]
AT2G40620.1	3.495	0.048348138	0	2	BZIP18	bZIP transcription factor 18 [Source:UniProtKB/Swiss-Prot;Acc:O22873]
AT5G45190.2	3.489	0.048348138	0	2		Cyclin family protein [Source:UniProtKB/TrEMBL;Acc:F4KD43]
AT3G13670.1	3.701	0.049393213	0	2.5		Casein kinase-like protein [Source:UniProtKB/TrEMBL;Acc:Q9LID3]
AT4G28860.1	3.701	0.049393213	0	2.5	CKL4	Casein kinase 1-like protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q8LPI7]
AT5G21274.1	2.925	0.049816735	0.4	6.5	CAM6	CAM6 [Source:UniProtKB/TrEMBL;Acc:A0A178UHI2]

Annexe 6 :

The UPF1 interactome reveals interaction networks between RNA degradation and translation repression factors in Arabidopsis

Chicois C, Scheer H, Garcia S, Zuber H, Mutterer J, Hammann P, Gagliardi D et Garcia D.

Plant J. 2018 Jul 7. doi: 10.1111/tpj.14022.

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2018) doi: 10.1111/tpj.14022

The UPF1 interactome reveals interaction networks between RNA degradation and translation repression factors in Arabidopsis

Clara Chicois[1], Hélène Scheer[1], Shahïnez Garcia[1], Hélène Zuber[1], Jérôme Mutterer[1], Johana Chicher[2], Philippe Hammann[2], Dominique Gagliardi[1] and Damien Garcia[1,*]

[1]*Institut de biologie moléculaire des plantes (IBMP), CNRS, Université de Strasbourg, 67000 Strasbourg, France, and*
[2]*Plateforme Protéomique Strasbourg-Esplanade, CNRS, Université de Strasbourg, 67000 Strasbourg, France*

Received 16 April 2018; revised 20 June 2018; accepted 26 June 2018.
*For correspondence (e-mail damien.garcia@ibmp-cnrs.unistra.fr).

SUMMARY

The RNA helicase UP-FRAMESHIFT (UPF1) is a key factor of nonsense-mediated decay (NMD), a mRNA decay pathway involved in RNA quality control and in the fine-tuning of gene expression. UPF1 recruits UPF2 and UPF3 to constitute the NMD core complex, which is conserved across eukaryotes. No other components of UPF1-containing ribonucleoproteins (RNPs) are known in plants, despite its key role in regulating gene expression. Here, we report the identification of a large set of proteins that co-purify with the Arabidopsis UPF1, either in an RNA-dependent or RNA-independent manner. We found that like UPF1, several of its co-purifying proteins have a dual localization in the cytosol and in P-bodies, which are dynamic structures formed by the condensation of translationally repressed mRNPs. Interestingly, more than half of the proteins of the UPF1 interactome also co-purify with DCP5, a conserved translation repressor also involved in P-body formation. We identified a terminal nucleotidyltransferase, ribonucleases and several RNA helicases among the most significantly enriched proteins co-purifying with both UPF1 and DCP5. Among these, RNA helicases are the homologs of DDX6/Dhh1, known as translation repressors in humans and yeast, respectively. Overall, this study reports a large set of proteins associated with the Arabidopsis UPF1 and DCP5, two components of P-bodies, and reveals an extensive interaction network between RNA degradation and translation repression factors. Using this resource, we identified five hitherto unknown components of P-bodies in plants, pointing out the value of this dataset for the identification of proteins potentially involved in translation repression and/or RNA degradation.

Keywords: nonsense-mediated decay, UPF1, RNA degradation, DCP5, translation repression, P-bodies, DDX6, RNA helicase, proteome, mass spectrometry.

INTRODUCTION

Eukaryotes possess three translation-dependent quality control pathways to reduce the accumulation of mRNAs showing aberrant ribosomal progression: no-go decay; non-stop decay; and nonsense-mediated decay (NMD; Shoemaker and Green, 2012). Among these pathways, NMD is specialized in the elimination of transcripts harboring premature translation termination codons (PTCs). PTCs can result from mutations, transcription errors or defective splicing, but also exist in alternatively spliced transcript isoforms, in mRNAs harboring long 3′-UTRs, or upstream open reading frames (ORF). In addition, an extensively studied NMD activating feature is the exon junction complex (EJC)-mediated NMD, involving the presence of an intron and the deposition of the EJC at least 50 nt downstream of the PTC (Chamieh *et al.*, 2008; Lindeboom *et al.*, 2016; Colombo *et al.*, 2017; Tian *et al.*, 2017). Recognition of PTCs requires the eukaryotic release factors (eRF1 and eRF3) and the activity of the UP-FRAMESHIFT (UPF) proteins. The direct interaction of UPF1 with the UPF2–UPF3 module is one of the key steps, which triggers the formation of the core NMD complex. This core complex is very well conserved throughout eukaryotes, and is present in yeast, insects, mammals as well as in plants (He *et al.*, 1997; Serin *et al.*, 2001; Kerényi *et al.*, 2008). NMD

activation is dependent on the phosphorylation of UPF1 by SMG1, required for the recognition by the 14-3-3-like domain of SMG7 that binds UPF1 phosphoserine-containing residues and triggers RNA decay (Lloyd and Davies, 2013). Despite the discovery of NMD more than 20 years ago, the precise molecular mechanisms by which a nonsense codon is recognized as premature and the individual roles of the UPF proteins are still poorly understood. Initially proposed to avoid the accumulation of potentially deleterious truncated proteins, NMD is now recognized as a genuine post-transcriptional regulation pathway (Nasif *et al.*, 2017). In mammals, NMD controls endoplasmic reticulum stress response by limiting the expression of IRE1a and repressing the unfolded protein response (Karam *et al.*, 2015). In fission yeast, UPF1 is essential for survival upon oxidative stress, and controls the expression of more than 100 genes transcriptionally induced in response to oxidative stress (Rodriguez *et al.*, 2006). In plants, NMD plays an important role in dampening the expression of immunity genes, which avoids the fitness cost of defense activation and modulates the output of bacterial infection (Rayson *et al.*, 2012; Riehs-Kearnan *et al.*, 2012; Gloggnitzer *et al.*, 2014). Many interactions between the NMD pathway and viral infections have also been observed across kingdoms. These interactions include the restriction of RNA virus infection in plants and mammals (Balistreri *et al.*, 2014; Garcia *et al.*, 2014), and suppression of the NMD pathway by the T-lymphotropic virus type 1 and the Rous sarcoma virus (Weil and Beemon, 2006; Mocquet *et al.*, 2012).

The ATP-dependent RNA helicase UPF1 is the key regulator of NMD, and its deregulation influences all the previously mentioned processes. Its activity induces a cascade of events leading to accelerated mRNA decay, by the action of deadenylases, decapping factors and exoribonucleases in mammals and yeast (Muhlrad and Parker, 1994; Lejeune *et al.*, 2003; Mitchell and Tollervey, 2003). In addition, UPF1 is localized in P-bodies in various organisms, including yeast, mammals and plants (Unterholzner and Izaurralde, 2004; Mérai *et al.*, 2013), although this localization is not required for NMD in animals (Stalder and Muhlemann, 2009). P-bodies are a class of membrane-less foci formed by a local phase transition process occurring in the presence of RNA-binding proteins and mRNAs. In human epithelial cells, a recent study combining fluorescence-activated particle sorting (FAPS) with genomic and proteomic studies identified proteins and mRNAs present in P-bodies (Hubstenberger *et al.*, 2017). Proteins linked to translation repression, miRNA function, mRNA decay and NMD were found prominent in the isolated P-bodies. The authors observed that mRNAs present in P-bodies are as abundant as other mRNAs, do not have shorter half-lives, and that their abundance is not changed when P-body formation is repressed. These observations, considered together with single-molecule imaging strategies used to follow mRNA degradation (Horvathova *et al.*, 2017; Tutucci *et al.*, 2018), suggest that P-bodies accumulate translationally repressed mRNAs rather than decayed mRNAs and are unlikely a major site for mRNA degradation (Hubstenberger *et al.*, 2017). This current view of P-bodies is supported by the study of *in vitro* produced P-body-like structures, in which the phase transition protects RNA against endonucleolytic cleavage or decapping activity (Schütz *et al.*, 2017). This evidence will need to be integrated in alternative models and interpretation of the presence of UPF1 in P-bodies. In mammals, a great deal of information has been gathered over the years on the UPF1 interaction network by strategies coupling immunoprecipitation (IP) to mass spectrometry (Isken *et al.*, 2008; Flury *et al.*, 2014; Schweingruber *et al.*, 2016). In contrast, in plants, little is known about the UPF1 interaction network beyond its expected association with UPF2 and UPF3 (Kerényi *et al.*, 2008). The identification and study of proteins interacting with core NMD factors is a pre-requisite to determine how the NMD is orchestrated, and to determine the links between UPF1 and P-bodies in plants.

In this work, we used an immunoaffinity approach coupled to mass spectrometry to identify proteins associated with Arabidopsis UPF1. We co-localized UPF1 with archetypal P-body components, including the LSM14A homolog DECAPPING 5 (DCP5). Half of UPF1 associated proteins also co-purified with this translation repressor. These include members of the plant DEAD-box helicase 6 (DDX6) translation repressors family, required for P-body formation in human cells. In addition, we tested the subcellular localization of proteins that co-purify with both UPF1 and DCP5. This study identified that UPF1 co-localized with selected partners in the cytosol and/or P-bodies, identifying five hitherto unknown plant P-body components. Our extended UPF1 and DCP5 interactome studies represent a unique dataset that can be used, as exemplified in this work, to identify previously unknown P-body components. Their study will facilitate the discovery of factors involved in translation repression and/or RNA degradation in plants.

RESULTS

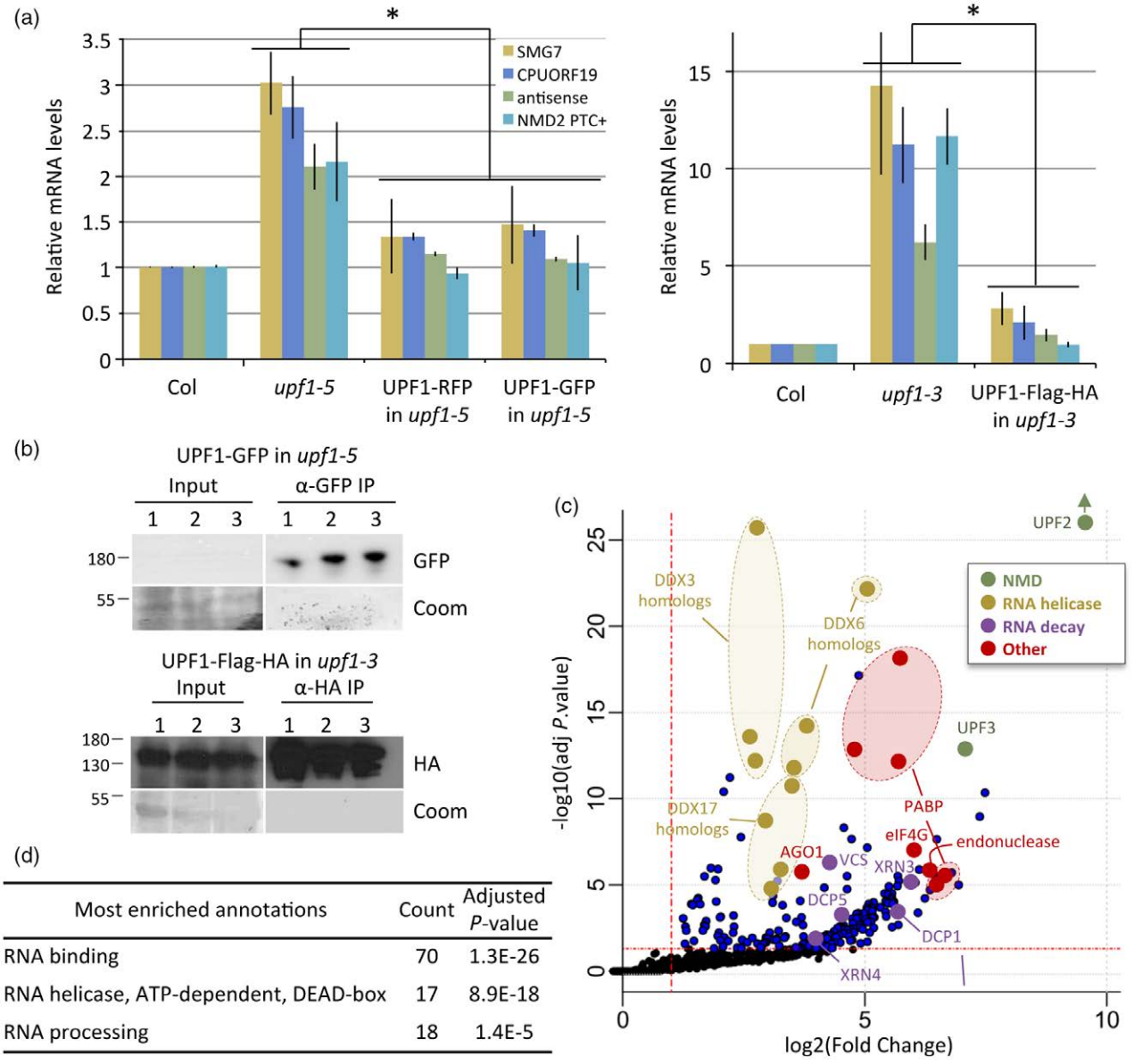
Isolation of UPF1-enriched fractions identifies a large set of RNA-binding factors

Early studies of NMD in plants were based on the analysis of mutants in homologs of UPF1 and UPF3. These mutants showed defects in flowering time, response to abiotic stresses and germination (Hori and Watanabe, 2005; Arciga-Reyes *et al.*, 2006; Yoine *et al.*, 2006). Major advances in understanding plant NMD were achieved in transient expression assays using NMD reporter genes (Kertész *et al.*, 2006; Kerényi *et al.*, 2008; Mérai *et al.*, 2013). Unfortunately, attempts to discover UPF1-

interacting proteins in plant, by yeast two-hybrid screens and transient expression strategies, could validate the conserved core UPF1–UPF2–UPF3 complex, but failed to identify additional partners of UPF1 (Kerényi *et al.*, 2008). Therefore, we performed co-IP experiments with green fluorescent protein (GFP) or double Flag-HA-tagged versions of UPF1. The fusion proteins were expressed under the control of the endogenous UPF1 promoter in either null (*upf1-3*) or hypomorphic (*upf1-5*) *upf1* mutant backgrounds. All transgenes restored the regulation of four distinct NMD targets with different NMD-inducing features, demonstrating that they encode functional UPF1 protein fusions (Figure 1a). These transgenic lines expressing UPF1–GFP and UPF1–Flag-HA were then used for IPs with antibodies directed against the GFP and HA tags, respectively. We observed a strong UPF1 enrichment in the IP fractions (Figure 1b). Mass spectrometric analysis of these fractions (Figure 1c; Table S1) identified 213 proteins significantly enriched in UPF1-IPs as compared with control IPs (adjusted P-value < 0.05). The presence of UPF2 and UPF3, the two known proteins of the UPF core complex in plants, among the most enriched UPF1 partners validated our approach (Figure 1c; Table S1). A global gene ontology terms analysis of the proteins detected in UPF1 IPs revealed a high proportion of RNA-binding proteins, including many RNA helicases, and proteins involved in RNA processing (Figure 1d). This result is coherent with the RNA-binding capacity of UPF1, and the fact that multiple RNA helicases and RNA-binding proteins interact with UPF1 in other organisms (Bond *et al.*, 2001; Flury *et al.*, 2014; Schweingruber *et al.*, 2016). Among the proteins significantly enriched in UPF1-IPs, we noticed the decapping associated factors DCP1, DCP5 and VCS. We confirmed the association of DCP5 with UPF1 by reverse IP experiments using plant expressing a tagged version of DCP5 (Figure S1a). Furthermore, UPF1 co-purified with the plant DDX6/Dhh1 homologs RH6, RH8 and RH12. DDX6/Dhh1 is an RNA helicase that is localized in P-bodies and required for their assembly in mammals and yeast (Ayache *et al.*, 2015). Human DDX6 was previously shown to associate with both the decapping complex and UPF1, but the molecular function of this association was not further investigated (Ayache *et al.*, 2015). However, both DCP5 (Scd6p in yeast, LSM14 in human) and DDX6 have been implicated in translation repression (Minshall *et al.*, 2009; Rajyaguru *et al.*, 2012; Presnyak and Coller, 2013). UPF1 co-purified also with homologs of two other RNA helicases, DDX3 and DDX17. DDX3 functions in stress granule assembly and translation inhibition, while DDX17 is involved in mRNA processing in mammals (Janknecht, 2010; Soto-Rifo and Ohlmann, 2013). As previously observed in mammals (Poulsen *et al.*, 2013), Argonaute 1 (AGO1), a major RNA silencing effector, also co-purified with plant UPF1. The interaction of UPF1 with AGO1 was confirmed by reverse IP experiments using Flag–AGO1 as bait (Figure S1b). Finally, UPF1 co-purified with eIF4G, already known to influence NMD in mammals (Lejeune *et al.*, 2003), and several members of the plant PolyA-binding protein (PABP) family. Taken together, our strategy identified a set of 213 proteins significantly enriched in the UPF1 purified fraction (Table S1). This study provides a unique dataset, and suggests that UPF1 is present in the cell together with a large set of key factors linked to cytosolic mRNA metabolism.

UPF1-associated proteins include RNA-dependent and RNA-independent partners

Overall, 48% of UPF1-associated proteins are also present in an experimentally identified Arabidopsis mRNA-binding proteome (At-RBP) established by 'mRNA interactome capture' (Figure 2a; Reichel *et al.*, 2016). This striking overlap suggests that their presence in the UPF1-IPs could at least partially be due to RNA-mediated interactions. To discriminate between RNA-dependent and RNA-independent interactions, we performed co-IP experiments with Flag-HA-tagged UPF1 in the presence or absence of RNase A. This experiment confirmed the co-purification of AGO1, DCP1 and PAB2 with UPF1–Flag-HA (Figure 2b). In addition, Western blot analysis revealed that the interaction of AGO1 with UPF1 is RNA-independent as similar amounts were detected in both treated and untreated IPs. PAB2 was also recovered upon RNase A treatment, suggesting that the interaction of PAB2 with UPF1 is not only mediated by RNA. By contrast, the association of UPF1 and DCP1 was completely disrupted upon RNase treatment. Mass spectrometric analysis of the immunoprecipitated fractions identified 66 proteins present in both RNase-treated and non-treated samples, indicating RNA-independent interactions with UPF1 (Figure 2c; Table S2). The specific recovery of the core UPF1–UPF2–UPF3 complex among the proteins classified as RNA-independent validated this approach. RNA-independent partners also included the RNA-induced silencing complex component AGO1, as well as the DDX6 homologs RH6, RH8, RH12, and the PABPs PAB2, PAB4, PAB5 and PAB8 (Figure 2c). Fifty-four proteins were identified solely in non-treated IPs but never detected upon RNase treatment, and were therefore classified as RNA-dependent partners of UPF1 (Figure 2c; Table S3). Such RNA-dependent partners included notably four proteins containing a RNA recognition motif domain, a putative endonuclease, and the decapping associated factors DCP5 and DCP1 (Figure 2c). Interestingly the action of UPF1 is coupled with the action of the decapping machinery in mammals (Lejeune *et al.*, 2003). The identification of DCP5 and DCP1 in UPF1 IPs suggests that a similar coupling between UPF1 and decapping could also occur in plants.



(d)

Most enriched annotations	Count	Adjusted *P*-value
RNA binding	70	1.3E-26
RNA helicase, ATP-dependent, DEAD-box	17	8.9E-18
RNA processing	18	1.4E-5

Figure 1. Analysis of the plant UPF1 interactome using tagged versions of UPF1.
(a) Reverse transcriptase-quantitative polymerase chain reaction (RT-qPCR) analysis of four nonsense-mediated decay (NMD) targets in *upf1* mutant lines complemented with pUPF1:UPF1-GFP, pUPF1:UPF1-RFP and pUPF1:UPF1-Flag-HA. Values are normalized to *ACTIN2* and expressed relative to Col-0 set as 1 ± SEM. Asterisks represent significant differences between mutants and the complemented lines, $P < 0.05$; $n = 3$.
(b) Western blot analysis of UPF1–GFP and UPF1–FlagHA immunoprecipitates (IPs). Two series of three replicates analyzed by mass spectrometry are shown. Molecular weight information is indicated in kDa.
(c) Volcano plot showing proteins enriched in UPF1 IPs (control IPs $n = 6$; UPF1 IPs $n = 8$). The adjusted *P*-value is presented on the *y*-axis, Fold-change (FC) in presented on the *x*-axis. The red dotted lines represent the significance thresholds of $P < 0.05$ and log2(FC) > 1.
(d) Most enriched gene ontology or UniProt terms among the UPF1 partners. This analysis was done with a total of 213 input proteins.

Analysis of insertion mutants in RNA decay factors and UPF1-associated proteins

To try to identify factors required for NMD in Arabidopsis, we tested mutants in some of the previously considered UPF1 protein partners, including decapping activators, for endogenous NMD targets accumulation. To this mutant list, we added the *xrn4* and *ski2* mutants in order to test factors involved in cytosolic 5′–3′ and 3′–5′ mRNA degradation, respectively. The four NMD targets tested had various NMD eliciting features, namely an upstream ORF, a long 3′-UTR and a PTC (in a specific splicing variant or a potential natural antisense gene). We expected to identify

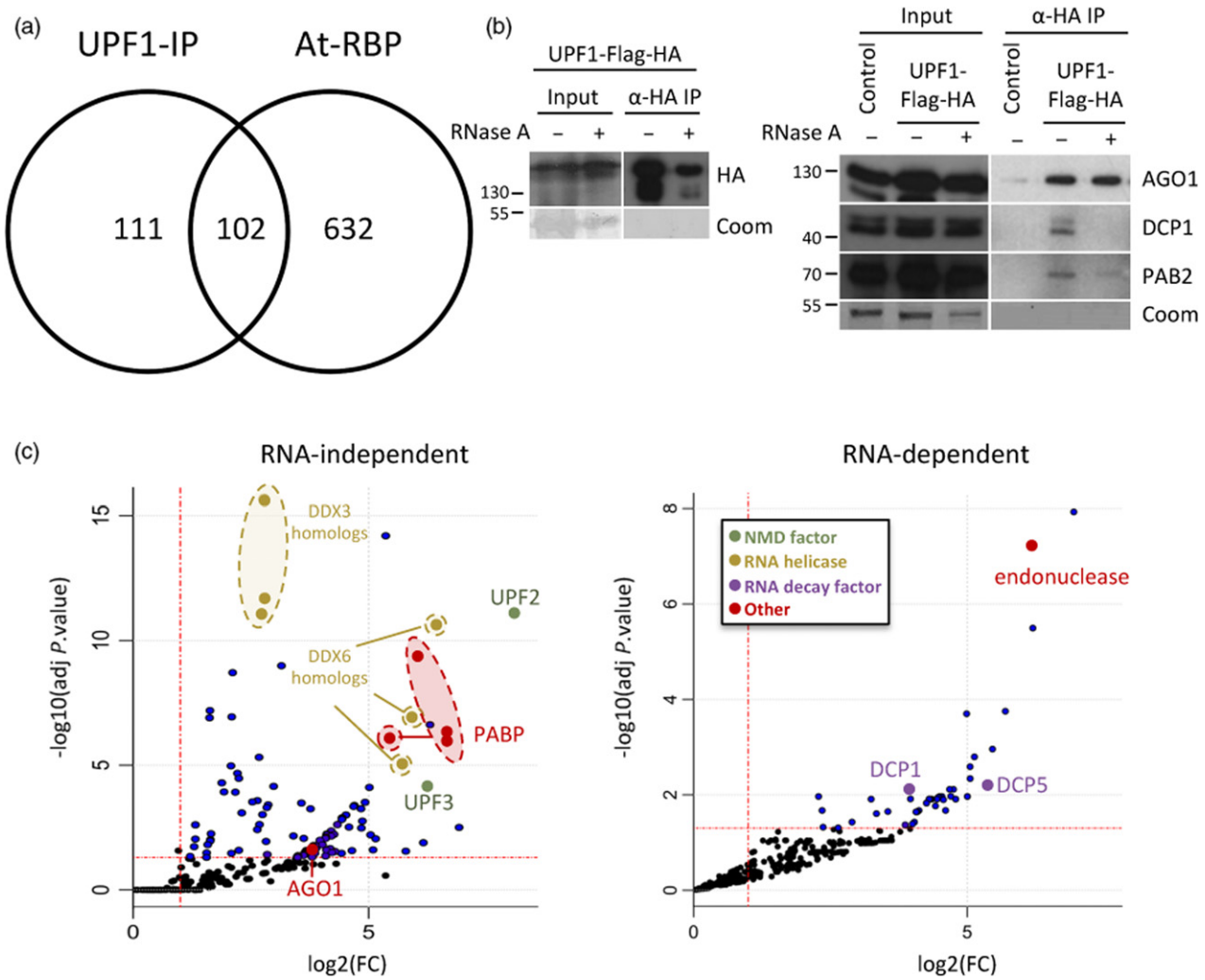


Figure 2. Analysis of RNA-dependent and RNA-independent UFP1 partners.
(a) Venn diagram showing the overlap between UPF1 immunoprecipitates (IPs) and the At-RBP proteome.
(b) Western blot analysis of the co-IPs of UPF1–Flag-HA with selected UPF1 partners, in both RNase A treated (+) and non-treated (–) samples. UPF1, AGO1, DCP1 and PAB2 protein levels were analyzed using antibodies raised against endogenous proteins. Protein molecular weight information is indicated in kDa. Coom, Coomassie staining of total proteins.
(c) Volcano plot showing proteins enriched in both RNase A-treated and non-treated UPF1–Flag-HA IP (RNA-independent) and proteins enriched in non-treated UPF1–Flag-HA IP only (RNA-dependent). The adjusted *P*-value is presented on the *y*-axis; fold-change (FC) on the *x*-axis. The red dotted lines represent the significance thresholds of $P < 0.05$ and log2(FC) > 1.

mutations that enhance the accumulation of NMD targets, as is the case for *upf1* and *upf3*, indicative of NMD activators. To our surprise, none of the mutants led to increased accumulation of the NMD targets tested (Figure 3), suggesting that none of these UPF1 partners or RNA degradation pathways is limiting for NMD. By contrast, we observed slight but significant decreased levels of at least one of the NMD targets in *dcp1*, *dcp5*, *rh12*, *elf4g and rh14* mutants (Figure 3). The observed effect could be due to reduced transcription or enhanced turnover of the selected targets. An interesting possibility is that some of these factors could in fact limit the degradation of a subset of NMD-targeted RNAs.

UPF1 and DCP5 are part of similar mRNPs and associate with DDX6 homologs

As UPF1 was shown to trigger translational repression during NMD in mammals (Isken *et al.*, 2008) and to lead to translation inhibition in plant hormone signaling (Li *et al.*, 2015), we decided to investigate the link between UPF1 and the translation repressors. For this purpose, we identified the proteins associated with the translation repressor DCP5 by IP coupled to mass spectrometry (Figure 4a,b). We identified 369 proteins significantly enriched in DCP5 IPs (Table S4). Global analysis of gene ontology terms revealed an enrichment of factors annotated as RNA

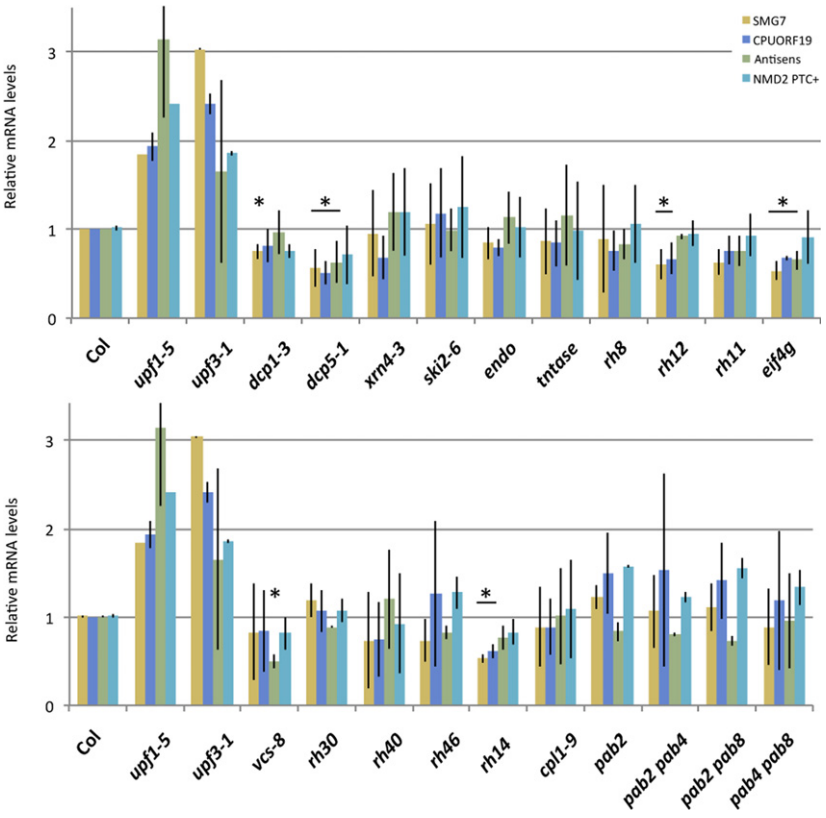
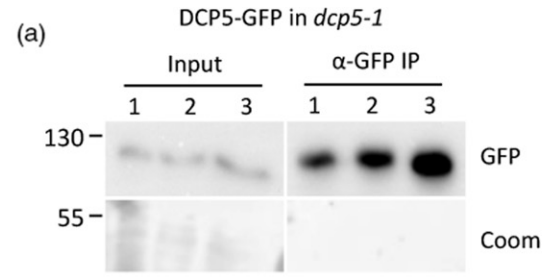


Figure 3. Accumulation of four archetypal nonsense-mediated decay (NMD) targets in insertion mutants in UPF1 partners.
Real-time reverse transcriptase-quantitative polymerase chain reaction (RT-qPCR) analysis of four known NMD target transcripts in mutants in UPF1-associated factors. Values are normalized to *ACTIN2* and expressed relative to Col-0 set as 1 ± SEM. Asterisk represents significant differences with the control Col-0, $P < 0.025$; $n = 3$.

binding, non-membrane bounded organelles, RNA helicases and RNA processing (Figure 4c). As observed for the proteins associated with UPF1, a large proportion of the proteins co-IP with DCP5 (37%) were present in the experimentally determined At-RBP proteome (Reichel *et al.*, 2016). Strikingly, UPF1 is among the most significantly enriched proteins co-purifying with DCP5 (Figure 4b), and 64% of the proteins co-purified with UPF1 were also detected in the DCP5 IPs (Figure 4d). This result indicates that both proteins are likely to participate to similar mRNP complexes. Among the common partners of DCP5 and UPF1, we noticed a number of RNA helicases including the three DDX6 homologs, all of which were among the most significantly enriched in both IPs (RH6, 8, 12; Figures 1c and 4b). DDX6 is a major component of P-bodies and is strictly required for their formation in human (Ayache *et al.*, 2015). The finding that DDX6 homologs are among the best-ranked UPF1 and DCP5 co-purifying proteins prompted us to test whether they directly bind to UPF1 or DCP5 using a yeast two-hybrid assay. As a positive control, we tested the known interaction between UPF1 and the C-terminal part of UPF2 (Figure 5a; He and Jacobson, 1995). In agreement with our previous finding that the association of UPF1 and DCP5 depends on the presence of RNA (Figure 2c), the yeast two-hybrid test suggested that DCP5 and UPF1 do not interact directly (Figure 5b). The assay recapitulated the interaction between UPF1 and UPF2, but no direct interaction between UPF1 and RH6 or DCP5 could be observed (Figure 5a,b). By contrast, we detected a direct interaction between DCP5 and RH6 (Figure 5b). In mammals, the DCP5 homolog LSM14 directly binds to DDX6 via both its FDF and TFG motifs (Brandmann *et al.*, 2018). Of note, the FDF and TFG motifs are present in the Arabidopsis DCP5 protein, suggesting that the mode of interaction between the LSM14 homolog DCP5 and the DDX6 homolog RH6 could be evolutionary conserved in plants.

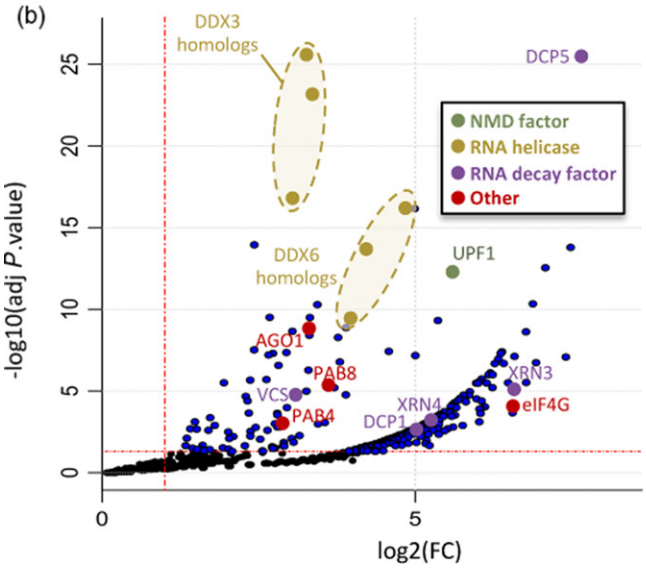
UPF1 co-localizes with its partners in the cytosol and in P-bodies

Next, we analyzed the subcellular localization of UPF1 and some of the proteins identified in UPF1-IPs by co-expressing fluorescent fusion proteins both transiently in *Nicotiana benthamiana* and stably in Arabidopsis. In agreement with previous studies, UPF1–red fluorescent protein (RFP) was detected both diffused in the cytosol and in cytoplasmic foci, likely P-bodies (Figure 6a). Because P-bodies are highly dynamic and their fast movement within the cytosol can have a blurring effect that may compromise the robustness of co-localization studies, we acquired our images with a multitrack line-by-line scanning mode instead of the commonly used frame-by-frame sequential acquisition of GFP and RFP signals (Figure S2a). With the



(c)

Most enriched annotations	Count	Adjusted *P*-value
mRNA binding	105	5.3E-35
Non-membrane-bounded organelle	71	1.2E-21
Helicase	27	1.6E-21
RNA processing	36	3.2E-14



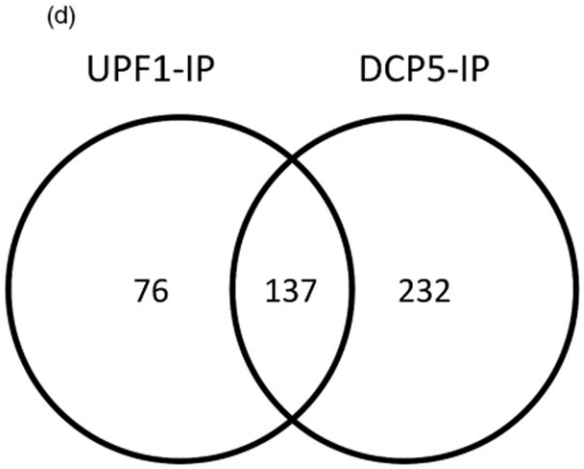


Figure 4. The DCP5 interactome shows extensive overlap with the UPF1 interactome.
(a) Western blot analysis of GFP immunoprecipitates (IPs) performed in triplicate, on extracts from DCP5-GFP *dcp5-1* complemented lines.
(b) Volcano plot showing proteins enriched in DCP5 IPs (control IPs $n = 7$; DCP5 IPs $n = 3$). The adjusted *P*-value is presented on the *y*-axis; fold-change (FC) on the *x*-axis. The red dotted lines represent the significance thresholds of $P < 0.05$ and log2(FC) > 1.
(c) Most enriched gene ontology or Uniprot terms among the DCP5 partners. This analysis was done with a total of 318 input proteins.
(d) Venn diagram showing the overlap between UPF1 and DCP5 co-purified partners.

multitrack line-by-line acquisition, we detected a near-perfect co-localization of UPF1 with the archetypal P-body markers DCP1 and DCP5 (Figure 6a). Similar results were obtained when UPF1 was transiently co-expressed with tagged versions of RH6, RH8 and RH12, the three Arabidopsis homologs of DDX6, required for P-bodies formation in mammals. This result confirms the recently observed P-body localization for RH12 (Bhullar *et al.*, 2017), and extends this localization to the two other DDX6 homologs, RH6 and RH8. The localization of UPF1 to P-bodies was further confirmed in time-lapse experiments monitoring the concomitant movement of UPF1 and the P-body marker DCP5 (Video S1). We quantified the co-localization of UPF1 with P-body markers in stable Arabidopsis transformants using an ImageJ macro that identifies the cytoplasmic foci in images and measures the extent of co-localization between foci in both GFP and RFP channels. Indeed, the approach reliably discriminated between a yellow fluorescent protein (YFP)-tagged version of the P-body marker DCP1 and the RFP-tagged stress granule marker PAB2, showing only 18% co-localization (Figure S2b). The analysis of stable Arabidopsis transformants co-expressing UPF1–RFP with GFP-tagged P-body markers revealed that 90, 95 and 100% of the foci that contained UPF1 also contained the P-body components DCP5, DCP1 or RH6, respectively (Figure 6). These data demonstrate that the majority of the UPF1-containing foci are indeed P-bodies.

We further analyzed the co-localization of UPF1 with five other factors that were enriched in UPF1 IPs. For this analysis, we chose either partners regulated by a mechanism coupling alternative splicing and NMD, the polypyrimidine tract-binding factor PTB3, and the glycine-rich protein GRP7 (Schöning *et al.*, 2008; Stauffer *et al.*, 2010). We also analyzed proteins with predicted functions linked to NMD in other species, including a putative endonuclease, a

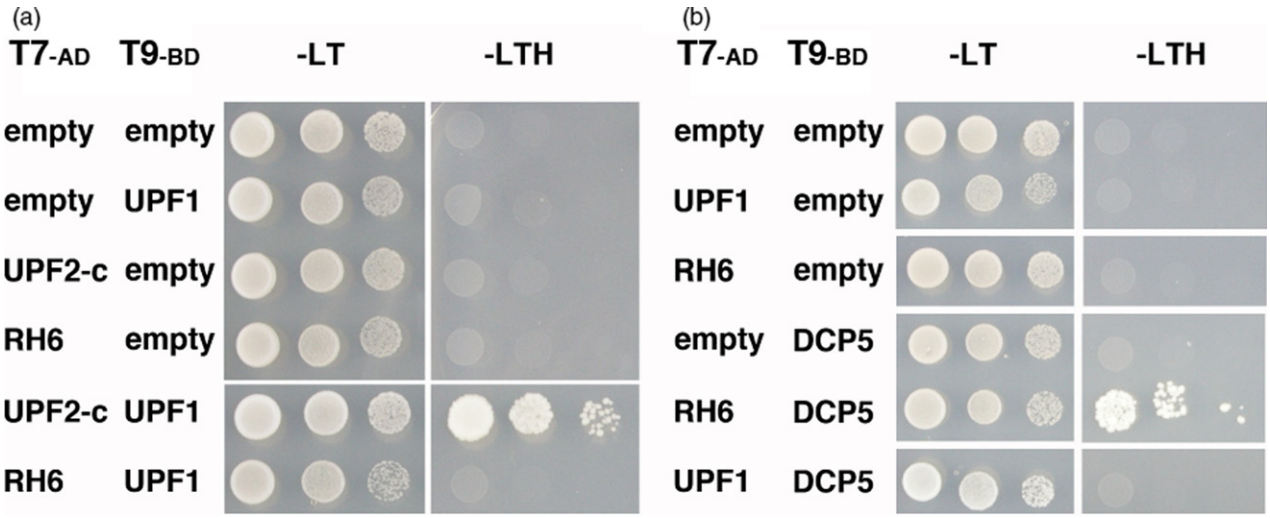


Figure 5. Yeast two-hybrid interaction assay between UPF1, DCP5 and the DDX6 homolog RH6.
(a) Interaction test between UPF1 and RH6 shows no growth on selective media. The UPF2 C-terminal domain (UPF2-C) was used as a positive control.
(b) Interaction test between DCP5, RH6 and UPF1 shows growth on selective media only for the DCP5–RH6 combination. 3-AT was used to avoid autoactivation, 70 mM in (a) and 5 mM in (b). –LT, minimal SD medium –Leucine, –Tryptophane; –LTH, minimal SD medium –Leucine, –Tryptophane, –Histidine. T7: pGADT7 AD (*LEU2*); T9: pGBT9 BD (*TRP1*). [Colour figure can be viewed at wileyonlinelibrary.com].

predicted terminal nucleotidyltransferase (TNTase) and RH11, a DEAD box RNA helicase (Eberle *et al.*, 2009; Morozov *et al.*, 2012; Hug and Cáceres, 2014). Upon transient expression in *N. benthamiana*, the RNA-binding proteins PTB3 and GRP7 co-localized with the fraction of UPF1 that is diffused in the cytosol, while only UPF1 but neither PTB3 nor GRP7 were enriched in P-bodies (Figure 6c). By contrast, the putative endonuclease, TNTase, and RH11 co-localized with UPF1 in both the cytosol and in P-bodies (Figure 6c). These data identify these three UPF1-associated proteins as previously unappreciated components of P-bodies.

Taken together, our co-localization studies demonstrate that Arabidopsis UPF1 co-localizes with proteins detected in UPF1 IPs in either the cytosol or in P-bodies, and identified a putative endonuclease, a TNTase, and the RNA helicases RH6, RH8 (DDX6 homologs) and RH11 (DDX3 homolog) as hitherto unknown components of P-bodies in plants.

DISCUSSION

Our work provides unbiased information about the molecular context present in the vicinity of the Arabidopsis UPF1, and illustrates its association with RNA-binding proteins, proteins linked to RNA decay, and translational repression. Furthermore, we demonstrate that several of these partners co-localize with UPF1 in P-bodies. These proteins include DCP1, DCP5 and homologs of the P-body organizing factor DDX6. The association of UPF1 with DCP5 depends on the presence of RNA, indicating that DCP5 and UPF1 could be present on the same RNA substrates. Indeed, the interactomes of UPF1 and DCP5 largely overlap and consist mostly of mRNA-binding proteins and RNA helicases previously characterized in the Arabidopsis mRNA-binding proteome (Reichel *et al.*, 2016). Among the proteins interacting with both UPF1 and DCP5, we validated five P-body components. This result demonstrates that our work provides an interesting resource for the discovery of hitherto unknown components of cytoplasmic RNA granules, with potential roles in RNA stability, translation inhibition or RNA storage.

Evolutionary conserved hallmarks of the plant UPF1 interactome

In this study we determined RNA-dependent and RNA-independent partner proteins of the key NMD factor UPF1 in Arabidopsis. This analysis confirmed the existence of the NMD core complex UPF1–UPF2–UPF3, and revealed that the association of UPF1 with decapping factors, PABPs and RNA helicases of the DDX6 family, known in mammals, is conserved in plants (Lejeune *et al.*, 2003; Ivanov *et al.*, 2008; Ayache *et al.*, 2015). We also detected an association of Arabidopsis UPF1 with AGO1, a key RNA-silencing component in plants (Borges and Martienssen, 2015). Interestingly, hUPF1 was also reported to interact with hAGO1 and hAGO2. The reason for this interaction remained unclear as hUPF1 knockdown did not affect miRNA targeted reporter constructs (Höck *et al.*, 2007). Likewise, the comparison of miRNA and NMD targets in Arabidopsis at the genome-wide scale did not reveal much overlap (Zhang *et al.*, 2013). Nevertheless, this interaction could have other impacts on the miRNA pathway, a recent

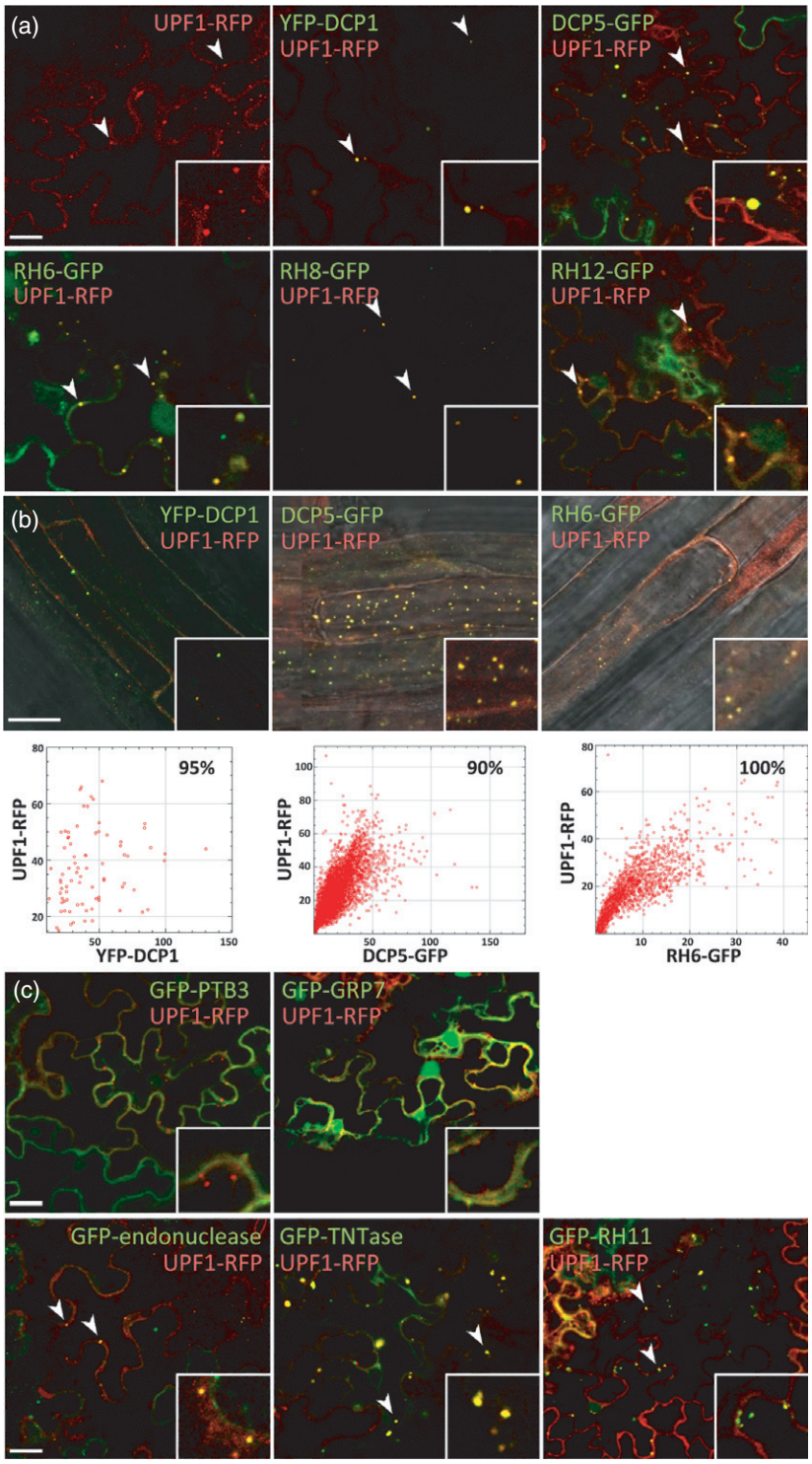


Figure 6. Subcellular co-localization of UPF1 with selected UPF1 partners identifies previously unknown P-body components.
(a) Confocal microscopy images showing the co-localization of UPF1–RFP with the P-body markers DCP1, DCP5 and RH12 fused to GFP as well as the other DDX6 homologs, RH6 and RH8, transiently co-expressed in *Nicotiana benthamiana* leaves.
(b) Confocal microscopy images of Arabidopsis epidermal root cells stably expressing UPF1–RFP with YFP–DCP1, DCP5–GFP or RH6–GFP. Insets on the right corner show details of the cytosolic foci at higher magnification. Dot plots showing the quantification of the signal in foci (GFP/YFP in *x*-axis, RFP in *y*-axis) are presented under each confocal image. Each plot was acquired following foci analysis of 10 randomly chosen confocal sections. The co-localization percentage between P-body markers and UPF1–RFP is indicated for each analysis.
(c) Images of the localization of UPF1–RFP and chosen UPF1 partners fused to GFP as in (a). Scale bars represent 10 μm. The arrows point towards P-bodies.

report in humans described a function of UPF1 in inducing Tudor-SN-mediated miRNA decay (Elbarbary *et al.*, 2017). Such a pathway has not yet been described in plants, but the identification of Arabidopsis AGO1 as potential direct UPF1 partner suggests that similar links between UPF1 and silencing may also exist in plants.

Our co-purification analysis also identified several previously unknown and possibly plant-specific partners of UPF1. Among them, a putative TNTase (AT3G61690) and a putative endonuclease (AT2G15560) co-localized with UPF1 in P-bodies. The association of both proteins with UPF1 suggests that they could act on either NMD targets or other UPF1-associated mRNAs. Proteins with similar enzymatic activities have been associated with NMD in other species. In *Aspergillus nidulans*, the 3′ nucleotide addition by the TNTase CutA and CutB is induced by NMD and promotes ribosome dissociation (Morozov *et al.*, 2012). Human histone transcripts, degraded in a UPF1-dependent manner in mammals, are also modified by oligouridylation (Slevin *et al.*, 2014). Regarding endonucleases, in mammals, two distinct endonucleases have been linked to UPF1 action, SMG6, in the degradation of NMD targets (Eberle *et al.*, 2009), and Regnase, in another UPF1-dependent RNA degradation pathway called Regnase-mediated decay (Mino *et al.*, 2015). Whereas we found that these factors do not seem to be limiting factors for NMD, the possible involvement of the P-body-associated TNTase and endonuclease proteins in NMD, and their molecular functions remains to be determined in plants.

The Arabidopsis UPF1 is associated with putative translational repressors

Among the proteins that associated with UPF1 in an RNA-independent manner were three homologs of the RNA helicase DDX6, RH6, RH8 and RH12. DDX6 is involved in translational repression and required for P-body assembly in mammals (Xu and Chua, 2009; Rajyaguru *et al.*, 2012; Presnyak and Coller, 2013; Ayache *et al.*, 2015). Each of the plant DDX6 homologs co-localized with UPF1 in P-bodies and co-purified in a RNA-independent manner, but yeast two-hybrid assays suggested that interaction between UPF1 and RH6 is not direct, whereas the known translation repressor DCP5 and RH6 directly interact (Figure 5). The association of UPF1 with DCP1 and DCP5 depends on the presence of RNA. In humans, co-purification of UPF1 with DCP1 and DCP2 was proposed to link NMD with downstream RNA degradation pathways. This was further supported as downregulating DCP2 decreased the degradation rate of NMD targets (Lykke-Andersen, 2002; Lejeune *et al.*, 2003). UPF1 triggers translational repression during NMD in mammals (Isken *et al.*, 2008), our evidence of co-purification of UPF1 with the translation repressor DCP5 and the DDX6 family homologs suggests the possibility of a similar activity in plants. In fact, there are two possible models, either UPF1 leads to translation inhibition after the first round of translation and leads to RNA decay as shown in mammals, or UPF1 directly represses translation without decay induction as shown for the EBF1/2 mRNAs in Arabidopsis (Li *et al.*, 2015; Merchante *et al.*, 2015).

Crossed proteomes of P-body components as a resource for P-body protein identification

The observation that UPF1 and DCP5 co-localize and co-purify with a similar set of mRNA-binding proteins suggested that both are recruited to similar ribonucleoproteins (RNPs) and can act on similar RNAs. Interestingly, loss of DCP5 expression results in a slight decrease of three endogenous mRNAs that exhibit NMD-inducing features, including an upstream ORF, a long 3′-UTR and a premature stop codon. A similar decrease was observed in *eif4g*, *rh12* and to some extent in *rh14*, indicating that these factors might influence the degradation of some NMD targets in wild-type plants. This effect might be linked to the fact that some of these components, DCP5 and RH12, localize in P-bodies, a central structure for the homeostasis of mRNAs.

The use of FAPS has recently allowed the purification of P-bodies from human epithelial cells. This study revealed that P-bodies are associated to translationally repressed mRNAs rather than to mRNAs undergoing decay as previously proposed (Hubstenberger *et al.*, 2017). In fact, there is accumulating evidence, both *in vitro* and *in vivo*, that mRNAs are not decayed within P-bodies but rather that RNA degradation occurs outside P-bodies in the cytoplasm (Horvathova *et al.*, 2017; Hubstenberger *et al.*, 2017; Schütz *et al.*, 2017; Tutucci *et al.*, 2018). The mRNA composition of P-bodies is enriched in transcripts coding for regulatory proteins, a clear difference with the mRNA composition of stress granules, in which most mRNAs accumulate without any functional distinctions (Khong *et al.*, 2017). As our understanding of the properties and functions of RNA granules and, in particular, P-bodies evolves, the functional relevance of the presence of UPF1 within P-bodies, together with translationally repressed mRNPs (Horvathova *et al.*, 2017; Hubstenberger *et al.*, 2017; Schütz *et al.*, 2017), becomes less obvious. The loss of DCP5 in the *dcp5* mutant does not lead to increased NMD target accumulation, suggesting that DCP5 is not a limiting factor for NMD target degradation, but may be involved in transitory storage or translation repression in P-bodies, with only marginal consequences on the steady-state level of NMD targets. A disruption of P-bodies was observed in *dcp5* (Xu and Chua, 2009), suggesting that P-bodies themselves are not required for NMD targets degradation in plants, as observed in Drosophila (Eulalio *et al.*, 2007).

In addition to its function in RNA surveillance, NMD clearly emerges as an important mechanism in the control of the expression of functional genes, important for plant immunity and stress response in plants (Drechsel *et al.*,

2013; Gloggnitzer *et al.*, 2014; Lorenzo *et al.*, 2017). Moreover, Arabidopsis UPF1 was shown to participate to the EIN2-induced translational repression of the EBF1 and EBF2 transcripts, key components of the signaling pathway of the phytohormone ethylene (Li *et al.*, 2015; Merchante *et al.*, 2015). The possible involvement of the translation repressors DCP5 and the DDX6 homologs in this process would be coherent with the association of UPF1 with these proteins, and remains to be tested. Translation repression of NMD targets was observed in mammals. This could be a conserved trend in plants, possibly through association of UPF1 with DCP5 and DDX6 homologs. Our work reveals a large set of proteins co-purifying with the Arabidopsis UPF1 and DCP5 proteins. These two factors showed perfect co-localization in P-bodies. The comparison of the proteomes associated with UPF1 and DCP5 revealed a strong overlap and allowed the identification of five previously unknown P-body components. This resource could be used to discover additional components of P-bodies and study their cellular function in plants. Such studies will be important to better understand the role of P-bodies in plants in controlling mRNA homeostasis, and the balance between RNA degradation, storage and translation repression.

EXPERIMENTAL PROCEDURES

Plant material and primers

Documents listing the plant lines and primers used in this study are provided in Tables S5 and S6.

Plasmids

The genomic sequence of the *Arabidopsis thaliana UPF1* genomic sequence including 1.6 kb of its upstream regulatory sequences was amplified from genomic DNA and cloned in pCTL235-RFP, pCTL235-GFP and pCTL235-Flag-HA vectors for C-terminal fusions of UPF1 with RFP, GFP or a double Flag-HA tag, respectively. For transient expression of GFP-fusion proteins, the genomic sequences of GRP7, PTB3, RH11, a putative TNTase and a putative endonuclease were amplified from genomic DNA and cloned into pB7WGF2. DCP5 was amplified from genomic DNA, RH6, RH8 and RH12 coding sequences were amplified from cDNA and all cloned in pUB10-GFP.

RNA extraction and real-time reverse transcriptase-polymerase chain reaction (RT-PCR)

Total RNA was extracted from 15-day-old plantlets with Tri-Reagent (Sigma, https://www.sigmaaldrich.com). Real-time RT-PCR analyses were conducted by reverse-transcription of RNA samples into cDNA using SuperScript IV reverse transcriptase (Invitrogen) with a mix of oligo(dT) and random hexamers. The cDNA was quantified using a SYBR Green qPCR kit (Roche, https://lifescience.roche.com) and gene-specific primers. PCR was performed using a LightCycler® 480 Real-Time PCR System (Roche) in 384-well optical reaction plates heated for 10 min at 95°C, followed by 45 cycles of denaturation for 15 sec at 95°C, annealing for 20 sec at 60°C, and elongation for 40 sec at 72°C. Transcript levels were normalized to those of *ACTIN2* (AT3G18780). Statistical analysis was performed using the Mann–Whitney test with three biological replicates each composed of two technical replicates ($P < 0.025$, $n = 3$).

Antibodies

PABP2, AGO1 and DCP1 antibodies were kindly provided by J.F. Laliberté (Dufresne *et al.*, 2008), N. Baumberger (Baumberger *et al.*, 2007) and M. Fauth (Weber *et al.*, 2008). Rabbit antisera were raised against immunogenic peptides identified on the AtUPF1 protein (QPNQSSQNPKHPYNG and GVDDEPQPVQPKYED) and affinity purified following the double X protocol of Eurogentec SA (https://secure.eurogentec.com).

Co-IP and mass spectrometry analysis

For each IP replicate, 0.3 g of flowers from stable Arabidopsis transformants expressing UPF1 and DCP5 fusion proteins was extracted in 1.5 ml of IP lysis buffer containing 50 mM NaCl using the μMACS GFP or HA isolation kit (MACS purification system, Miltenyi Biotech, https://www.miltenyibiotec.com). The lysates were cleared by centrifugation at 10 000 ***g***. IPs were performed using magnetic microparticles coated with monoclonal HA or GFP antibodies according to the manufacturer's instructions, except that sodium dodecyl sulfate was omitted from washing buffers. Co-IP experiments were carried out at least in triplicate for each bait.

Eluted proteins were digested with sequencing-grade trypsin (Promega, https://www.promega.com) and analyzed by nanoLC-MS/MS on a TripleTOF 5600 mass spectrometer (Sciex, https://sciex.com) as described previously (Chicher *et al.*, 2015) or on a QExactive+ mass spectrometer coupled to an EASY-nanoLC-1000 (Thermo-Fisher Scientific, USA, https://www.thermofisher.com). Data were searched against the TAIR 10 database with a decoy strategy. Peptides were identified with Mascot algorithm (version 2.5; Matrix Science, London, UK) and data were further imported into Proline 1.4 software (http://proline.profiproteomics.fr/). Proteins were validated on Mascot pretty rank equal to 1, and 1% FDR on both peptide spectrum matches (PSM score) and protein sets (Protein Set score). The total number of MS/MS fragmentation spectra was used to quantify each protein from at least three independent biological replicates. These spectral counts were analyzed by the msmsTest R package (R v1.10.11; Gregori *et al.*, 2013) using the implemented negative binominal model based on the solution provided by the edgeR package (v3.20.1; Robinson *et al.*, 2010). For each IP, spectral counts were scaled by total number of counts. For each identified protein, an adjusted *P*-value corrected by Benjamini–Hochberg was calculated, as well as a protein fold-change (FC). A supplemental filter was applied to exclude proteins that have a low *P*-value, and low number of counts and poor effect size (DEP column in Tables S1–S4). The results are presented in volcano plots that display log2 (fold-change) and −log10 (adjusted *P*-value) on *x*- and *y*-axes, respectively. Annotation enrichment was analyzed using DAVID Bioinformatics Resources 6.8 and the TAIR9 annotations (Huang *et al.*, 2008, 2009).

Transient expression and subcellular localization

Fluorescent fusion proteins were transiently expressed in *N. benthamiana* leaves as described in Garcia *et al.* (2014), and pictures taken 4 days after infiltration, or stably expressed in *A. thaliana*. Subcellular localizations were imaged in *N. benthamiana* leaves or in root epidermal cells of 7-day-old Arabidopsis plants using a LSM780 confocal microscope (Zeiss) with a 40 × objective. Co-localization analysis was performed with ImageJ as follows: foci in images were determined with a user-supervised local maxima detection method (script available on demand). Local intensities in channels visualizing GFP or RFP fusion proteins were measured for every detected focus and the reported values were

then charted in a (IGFP versus IRFP) scatter plot for further qualitative assessment of fluorescent spot content correlation.

ACCESSION NUMBERS

UPF1: AT5G47010; *DCP1*: AT1G08370; *DCP5*: AT1G26110; RNA helicases DDX6 family, *RH6*: AT2G45810; *RH8*: AT4G00660; *RH12*: AT3G61240; RNA helicases DDX3 family, *RH37: AT2G42520*; *RH52*: AT3G58570; *RH11*: AT3G58510; RNA helicases DDX17 family, *RH14: AT3G01540*; *RH40*: AT3G06480; *RH30*: AT5G63120; *RH46*: AT5G14610; putative endonuclease: AT2G15560; putative terminal nucleotidyltransferase (*TNTase*): AT3G61690; *PTB3*: AT1G43190; *GRP7*: AT2G21660.

ACKNOWLEDGEMENTS

This work was supported by the Centre National de la Recherche Scientifique (CNRS) and realized in the frame of the LABEX NetRNA (ANR-2010-LABX-36) from the French National Research Agency as part of the Investments for the Future Program; the mass spectrometer was co-funded by the University of Strasbourg, IdEx 'Equipement mi-lourd' 2015 and by the LabEx NetRNA. C.C. was supported by a PhD fellowship from the French Ministry of Research. The authors thank C. Bousquet-Antonelli and J.M. Deragon for pDCP1:DCP1:YFP, PAB2-RFP constructs and lines; R. Wagner's team for plant care; Heike Lange for critical reading of the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Detection of UPF1 in DCP5 and AGO1 IPs.

Figure S2. Line-by-line acquisition allows the precise imaging of dynamic foci.

Video S1. Time-lapse acquisition of the co-localization and dynamics of UPF1–RFP and DCP5–GFP.

Table S1 Proteins present in UPF1 IPs

Table S2 Proteins present in UPF1 IPs independently of RNA

Table S3 Proteins present in UPF1 IPs in an RNA-dependent manner

Table S4 Proteins present in DCP5 IPs

Table S5 Plant material used and generated in this study

Table S6 DNA oligonucleotides used in this study

REFERENCES

Arciga-Reyes, L., Wootton, L., Kieffer, M. and Davies, B. (2006) UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in Arabidopsis. *Plant J.*, **47**, 480–489.

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. *Mol. Biol. Cell*, **26**, 2579–2595.

Balistreri, G., Horvath, P., Schweingruber, C., Zünd, D., McInerney, G., Merits, A., Mühlemann, O., Azzalin, C. and Helenius, A. (2014) The host nonsense-mediated mRNA decay pathway restricts mammalian RNA virus replication. *Cell Host Microbe*, **16**, 403–411.

Baumberger, N., Tsai, C.-H., Lie, M., Havecker, E. and Baulcombe, D.C. (2007) The Polerovirus silencing suppressor P0 targets ARGONAUTE proteins for degradation. *Curr. Biol.*, **17**, 1609–1614.

Bhullar, D.S., Sheahan, M.B. and Rose, R.J. (2017) RNA processing body (P-body) dynamics in mesophyll protoplasts re-initiating cell division. *Protoplasma*, **254**, 1627–1637.

Bond, A.T., Mangus, D.A., He, F. and Jacobson, A. (2001) Absence of Dbp2p alters both nonsense-mediated mRNA decay and rRNA processing. *Mol. Cell. Biol.*, **21**, 7366–7379.

Borges, F. and Martienssen, R.A. (2015) The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.*, **16**, 727–741.

Brandmann, T., Fakim, H., Padamsi, Z., Youn, J.-Y., Gingras, A.-C., Fabian, M.R. and Jinek, M. (2018) Molecular architecture of LSM14 interactions involved in the assembly of mRNA silencing complexes. *EMBO J.*, **37**, e97869.

Chamieh, H., Ballut, L., Bonneau, F. and Hir, H. Le (2008) NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nat. Struct. Mol. Biol.*, **15**, 85–93.

Chicher, J., Simonetti, A., Kuhn, L., Schaeffer, L., Hammann, P., Eriani, G. and Martin, F. (2015) Purification of mRNA-programmed translation initiation complexes suitable for mass spectrometry analysis. *Proteomics*, **15**, 2417–2425.

Colombo, M., Karousis, E.D., Bourquin, J., Bruggmann, R. and Muhlemann, O. (2017) Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. *RNA*, **23**, 189–201.

Drechsel, G., Kahles, A., Kesarwani, A.K., Stauffer, E., Behr, J., Drewe, P., Rätsch, G. and Wachter, A. (2013) Nonsense-mediated decay of alternative precursor mRNA splicing variants is a major determinant of the Arabidopsis steady state transcriptome. *Plant Cell*, **25**, 3726–3742.

Dufresne, P.J., Ubalijoro, E., Fortin, M.G. and Laliberte, J.-F. (2008) Arabidopsis thaliana class II poly(A)-binding proteins are required for efficient multiplication of turnip mosaic virus. *J. Gen. Virol.*, **89**, 2339–2348.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O. and Jensen, T.H. (2009) SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.*, **16**, 49–55.

Elbarbary, R.A., Miyoshi, K., Hedaya, O., Myers, J.R. and Maquat, L.E. (2017) UPF1 helicase promotes TSN-mediated miRNA decay. *Genes Dev.*, **31**, 1483–1493.

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.

Flury, V., Restuccia, U., Bachi, A. and Mu, O. (2014) Characterization of phosphorylation- and RNA-dependent UPF1 interactors by quantitative proteomics. *J. Proteome Res.*, **13**, 3038–3053.

Garcia, D., Garcia, S. and Voinnet, O. (2014) Nonsense-mediated decay serves as a general viral restriction mechanism in plants. *Cell Host Microbe*, **16**, 391–402.

Gloggnitzer, J., Akimcheva, S., Srinivasan, A. *et al.* (2014) Nonsense-mediated mRNA decay modulates immune receptor levels to regulate plant antibacterial defense. *Cell Host Microbe*, **16**, 376–390.

Gregori, J., Sanchez, A. and Villanueva, J. (2013) msmsTests: LC-MS/MS Differential Expression Tests. R package version 1.10.11. https://doi.org/10.18129/B9.bioc.msmsTests

He, F. and Jacobson, A. (1995) Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen. *Genes Dev.*, **9**, 437–454.

He, F., Brown, A.H. and Jacobson, A. (1997) Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway. *Mol. Cell. Biol.*, **17**, 1580–1594.

Höck, J., Weinmann, L., Ender, C., Rüdel, S., Kremmer, E., Raabe, M., Urlaub, H. and Meister, G. (2007) Proteomic and functional analysis of Argonaute-containing mRNA–protein complexes in human cells. *EMBO Rep.*, **8**, 1052–1060.

Hori, K. and Watanabe, Y. (2005) UPF3 suppresses aberrant spliced mRNA in Arabidopsis. *Plant J.*, **43**, 530–540.

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.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.

Hubstenberger, A., Courel, M., Bénard, M. *et al.* (2017) P-body purification reveals the condensation of repressed mRNA regulons. *Mol. Cell*, **68**, 144–157.

Hug, N. and Cáceres, J.F. (2014) The RNA helicase DHX34 activates NMD by promoting a transition from the surveillance to the decay-inducing complex. *Cell Rep.*, **8**, 1845–1856.

Isken, O., Kim, Y.K., Hosoda, N., Mayeur, G.L., Hershey, J.W. and Maquat, L.E. (2008) Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. *Cell*, **133**, 314–327.

Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W. and Kulozik, A.E. (2008) Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.*, **27**, 736–747.

Janknecht, R. (2010) Multi-talented DEAD-box proteins and potential tumor promoters: p68 RNA helicase (DDX5) and its paralog, p72 RNA helicase (DDX17). *Am. J. Transl. Res.*, **2**, 223–234.

Karam, R., Lou, C., Kroeger, H., Huang, L., Lin, J.H. and Wilkinson, M.F. (2015) The unfolded protein response is shaped by the NMD pathway. *EMBO Rep.*, **16**, 599–609.

Kerényi, Z., Mérai, Z., Hiripi, L., Benkovics, A., Gyula, P., Lacomme, C., Barta, E., Nagy, F. and Silhavy, D. (2008) Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. *EMBO J.*, **27**, 1585–1595.

Kertész, S., Kerényi, Z., Mérai, Z., Bartos, I., Pálfy, T., Barta, E. and Silhavy, D. (2006) Both introns and long 3′-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. *Nucleic Acids Res.*, **34**, 6147–6157.

Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R. and Parker, R. (2017) The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Mol. Cell*, **68**, 808–820.

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.

Li, W., Ma, M., Feng, Y., Li, H., Wang, Y., Ma, Y., Li, M., An, F. and Guo, H. (2015) EIN2-directed translational regulation of ethylene signaling in Arabidopsis. *Cell*, **163**, 670–683.

Lindeboom, R.G.H., Supek, F. and Lehner, B. (2016) The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat. Genet.*, **48**, 1112–1118.

Lloyd, J.P. and Davies, B. (2013) SMG1 is an ancient nonsense-mediated mRNA decay effector. *Plant J.*, **76**, 800–810.

De Lorenzo, L., Sorenson, R., Bailey-Serres, J. and Hunt, A.G. (2017) Non-canonical alternative polyadenylation contributes to gene regulation in response to hypoxia. *Plant Cell*, **29**, 1262–1277.

Lykke-Andersen, J. (2002) Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol. Cell. Biol.*, **22**, 8114–8121.

Mérai, Z., Benkovics, A.H., Nyikó, T., Debreczeny, M., Hiripi, L., Kerényi, Z., Kondorosi, E. and Silhavy, D. (2013) The late steps of plant nonsense-mediated mRNA decay. *Plant J.*, **73**, 50–62.

Merchante, C., Brumos, J., Yun, J. *et al.* (2015) Gene-specific translation regulation mediated by the hormone-signaling molecule EIN2. *Cell*, **163**, 684–697.

Mino, T., Murakawa, Y., Fukao, A. *et al.* (2015) Regnase-1 and roquin regulate a common element in inflammatory mRNAs by spatiotemporally distinct mechanisms. *Cell*, **161**, 1058–1073.

Minshall, N., Kress, M., Weil, D. and Standart, N. (2009) Role of p54 RNA helicase activity and its C-terminal domain in translational repression, P-body localization and assembly. *Mol. Biol. Cell*, **20**, 2464–2472.

Mitchell, P. and Tollervey, D. (2003) An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3 → 5 degradation. *Mol. Cell*, **11**, 1405–1413.

Mocquet, V., Neusiedler, J., Rende, F. *et al.* (2012) The human T-lymphotropic virus type 1 tax protein inhibits nonsense-mediated mRNA decay by interacting with INT6/EIF3E and UPF1. *J. Virol.*, **86**, 7530–7543.

Morozov, I.Y., Jones, M.G., Gould, P.D., Crome, V., Wilson, J.B., Hall, A.J., 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.

Muhlrad, D. and Parker, R. (1994) Premature translational termination triggers mRNA decapping. *Nature*, **370**, 578–581.

Nasif, S., Contu, L. and Mühlemann, O. (2017) Beyond quality control: the role of nonsense-mediated mRNA decay (NMD) in regulating gene expression. *Semin. Cell Dev. Biol.*, **S1084–9521**, 30 342–30 347.

Poulsen, C., Vaucheret, H. and Brodersen, P. (2013) Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *Plant Cell*, **25**, 22–37.

Presnyak, V. and Coller, J. (2013) The DHH1/RCKp54 family of helicases: an ancient family of proteins that promote translational silencing. *Biochim. Biophys. Acta*, **1829**, 817–823.

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.

Rayson, S., Arciga-reyes, L., Wootton, L., Zabala, M.D.T., Truman, W., Grant, M. and Davies, B. (2012) A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in *Arabidopsis thaliana* NMD mutants. *PLoS ONE*, **7**, e31917.

Reichel, M., Liao, Y., Rettel, M. *et al.* (2016) In planta determination of the mRNA-binding proteome of Arabidopsis etiolated seedlings. *Plant Cell*, **28**, 2435–2452.

Riehs-Kearnan, N., Gloggnitzer, J., Dekrout, B., Jonak, C. and Riha, K. (2012) Aberrant growth and lethality of Arabidopsis deficient in nonsense-mediated RNA decay factors is caused by autoimmune-like response. *Nucleic Acids Res.*, **40**, 5615–5624.

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

Rodriguez, M.A., Watt, S., Bähler, J. and Russel, P. (2006) Upf1, an RNA helicase required for nonsense-mediated mRNA decay, modulates the transcriptional response to oxidative stress in fission yeast. *Mol. Cell. Biol.*, **26**, 6347–6356.

Schöning, J.C., Streitner, C., Meyer, I.M., Gao, Y. and Staiger, D. (2008) Reciprocal regulation of glycine-rich RNA-binding proteins via an interlocked feedback loop coupling alternative splicing to nonsense-mediated decay in Arabidopsis. *Nucleic Acids Res.*, **36**, 6977–6987.

Schütz, S., Nöldeke, 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.

Schweingruber, C., Soffientini, P., Ruepp, M.-D., Bachi, A. and Mühlemann, O. (2016) Identification of interactions in the NMD complex using proximity-dependent biotinylation (BioID). *PLoS ONE*, **11**, e0150239.

Serin, G., Gersappe, A., Black, J.D., Aronoff, R. and Maquat, L.E. (2001) Identification and characterization of human orthologues to *Saccharomyces cerevisiae* Upf2 protein and Upf3 protein (*Caenorhabditis elegans* SMG-4). *Mol. Cell. Biol.*, **21**, 209–223.

Shoemaker, C.J. and Green, R. (2012) Translation drives mRNA quality control. *Nat. Struct. Mol. Biol.*, **19**, 594–601.

Slevin, M.K., Meaux, S., Welch, J.D., Bigler, R., Miliani de Marval, P.L., Su, W., Rhoads, R.E., Prins, J.F. and Marzluff, W.F. (2014) Deep sequencing shows multiple oligouridylations are required for 3′ to 5′ degradation of histone mRNAs on polyribosomes. *Mol. Cell*, **53**, 1020–1030.

Soto-Rifo, R. and Ohlmann, T. (2013) The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism. *Wiley Interdiscip. Rev. RNA*, **8**, 369–385.

Stalder, L. and Muhlemann, O. (2009) Processing bodies are not required for mammalian nonsense-mediated mRNA decay. *RNA*, **15**, 1265–1273.

Stauffer, E., Westermann, A., Wagner, G. and Wachter, A. (2010) Polypyrimidine tract-binding protein homologues from Arabidopsis underlie regulatory circuits based on alternative splicing and downstream control. *Plant J.*, **64**, 243–255.

Tian, M., Yang, W., Zhang, J., Dang, H., Lu, X., Fu, C. and Miao, W. (2017) Nonsense-mediated mRNA decay in Tetrahymena is EJC independent and requires a protozoa-specific nuclease. *Nucleic Acids Res.*, **45**, 6848–6863.

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.

Unterholzner, L. and Izaurralde, E. (2004) SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol. Cell*, **16**, 587–596.

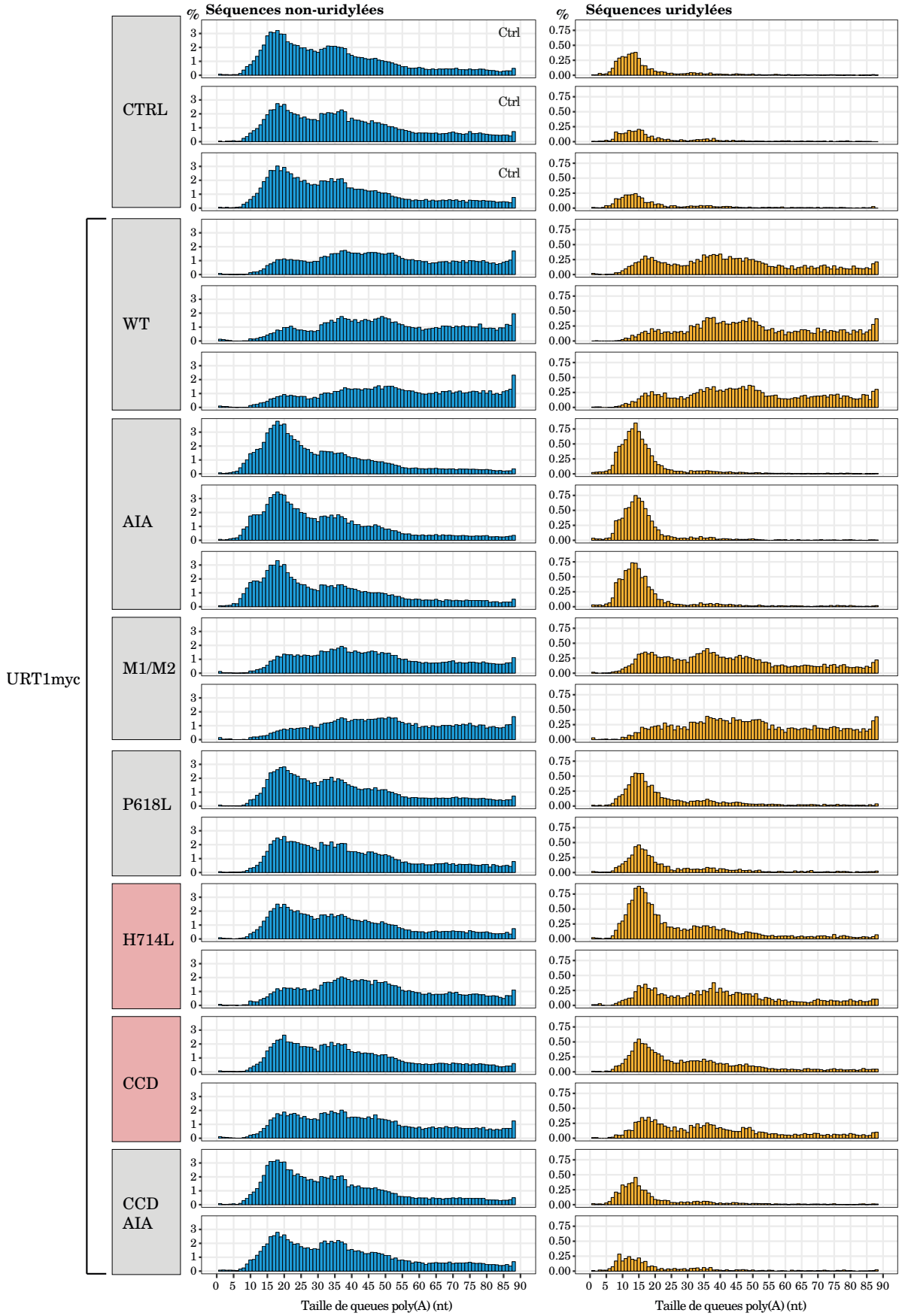
Weber, C., Nover, L. and Fauth, M. (2008) Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J.*, **56**, 517–530.

Weil, J.E. and Beemon, K.L. (2006) A 3′ UTR sequence stabilizes termination codons in the unspliced RNA of Rous sarcoma virus. *RNA*, **12**, 102–110.

Xu, J. and Chua, 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.

Yoine, M., Ohto, M.-A., Onai, K., Mita, S. and Nakamura, K. (2006) The lba1 mutation of UPF1 RNA helicase involved in nonsense-mediated mRNA decay causes pleiotropic phenotypic changes and altered sugar signalling in Arabidopsis. *Plant J.*, **47**, 49–62.

Zhang, Z., Hu, L. and Kong, X. (2013) MicroRNA or NMD: why have two RNA silencing systems? *J. Genet. Genomics*, **40**, 497–513.



Annexe 7 | **Analyse de la taille des queues poly(A) des transcrits GFP uridylés et non uridylés dans les patchs co-infiltrés avec les différentes versions de URT1myc.**
Diagrammes en barres représentant la distribution de la taille des queues poly(A) pour les transcrits GFP uridylés et non uridylés dans les différents patchs infiltrés avec P19, le rapporteur GFP et les différentes constructions de URT1myc. Le pourcentage de reads à chaque position a été normalisé par rapport au nombre total de reads correspondant à des tailles de queues poly(A) comprises entre 1 et 89A. l. Les échantillons URT1myc H714L et URT1myc CCD pour lesquels nous observons de la variabilité sont indiqués par un fond rouge.

Références

Abulfaraj, A.A., Mariappan, K., Bigeard, J., et al. (2018) The Arabidopsis homolog of human G3BP1 is a key regulator of stomatal and apoplastic immunity. *Life Science Alliance*, **1**, e201800046.

Agalarov, S.C., Sogorin, E.A., Shirokikh, N.E. and Spirin, A.S. (2011) Insight into the structural organization of the omega leader of TMV RNA: The role of various regions of the sequence in the formation of a compact structure of the omega RNA. *Biochemical and Biophysical Research Communications*, **404**, 250–253.

Aitken, C.E. and Lorsch, J.R. (2012) A mechanistic overview of translation initiation in eukaryotes. *Nature Structural & Molecular Biology*, **19**, 568–576.

Aizer, A., Kalo, A., Kafri, P., Shraga, A., Ben-Yishay, R., Jacob, A., Kinor, N. and Shav-Tal, Y. (2014) Quantifying mRNA targeting to P-bodies in living human cells reveals their dual role in mRNA decay and storage. *J Cell Sci*, **127**, 4443–4456.

Andrei, M.A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R. and Lührmann, R. (2005) A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA*, **11**, 717–727.

Aphasizhev, R. and Aphasizheva, I. (2014) Mitochondrial RNA editing in trypanosomes: small RNAs in control. *Biochimie*, **100**, 125–131.

Aphasizhev, R., Suematsu, T., Zhang, L. and Aphasizheva, I. (2016) Constructive edge of uridylation-induced RNA degradation. *RNA Biology*, **13**, 1078–1083.

Aphasizheva, I., Zhang, L., Wang, X., Kaake, R.M., Huang, L., Monti, S. and Aphasizhev, R. (2014) RNA binding and core complexes constitute the U-insertion/deletion editosome. *Mol. Cell. Biol.*, **34**, 4329–4342.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S. and Katada, T. (2001) Ski7p G protein interacts with the exosome and the Ski complex for 3′-to-5′ mRNA decay in yeast. *The EMBO Journal*, **20**, 4684–4693.

Arribas-Layton, M., Wu, D., Lykke-Andersen, J. and Song, H. (2013) Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1829**, 580–589.

Arribere, J.A., Doudna, J.A. and Gilbert, W.V. (2011) Reconsidering Movement of Eukaryotic mRNAs Between Polysomes and P-bodies. *Mol Cell*, **44**, 745–758.

Atkins, J., Boateng, S., Sorensen, T. and McGuffin, L. (2015) Disorder Prediction Methods, Their Applicability to Different Protein Targets and Their Usefulness for Guiding Experimental Studies. *International Journal of Molecular Sciences*, **16**, 19040–19054.

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. *Mol Biol Cell*, **26**, 2579–2595.

Aylett, C.H.S. and Ban, N. (2017) Eukaryotic aspects of translation initiation brought into focus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **372**, 20160186.

Baer, B.W. and Kornberg, R.D. (1980) Repeating structure of cytoplasmic poly(A)-

ribonucleoprotein. *Proc Natl Acad Sci U S A*, **77**, 1890–1892.

Banani, S.F., Lee, H.O., Hyman, A.A. and Rosen, M.K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology*, **18**, 285–298.

Barkoff, A., Ballantyne, S. and Wickens, M. (1998) Meiotic maturation in Xenopus requires polyadenylation of multiple mRNAs. *The EMBO Journal*, **17**, 3168–3175.

Basbouss-Serhal, I., Pateyron, S., Cochet, F., Leymarie, J. and Bailly, C. (2017) 5′ to 3′ mRNA Decay Contributes to the Regulation of Arabidopsis Seed Germination by Dormancy. *Plant Physiology*, **173**, 1709–1723.

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.

Bergamini, G., Preiss, T. and Hentze, M.W. (2000) Picornavirus IRESes and the poly(A) tail jointly promote cap-independent translation in a mammalian cell-free system. *RNA*, **6**, 1781–1790.

Bernstein, P., Peltz, S.W. and Ross, J. (1989) The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. *Molecular and Cellular Biology*, **9**, 659–670.

Bhullar, D.S., Sheahan, M.B. and Rose, R.J. (2017) RNA processing body (P-body) dynamics in mesophyll protoplasts re-initiating cell division. *Protoplasma*, **254**, 1627–1637.

Bianchin, C., Mauxion, F., Sentis, S., Séraphin, B. and Corbo, L. (2005) Conservation of the deadenylase activity of proteins of the Caf1 family in human. *RNA*, **11**, 487–494.

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. *Journal of Biological Chemistry*, **271**, 432–438.

Boeynaems, S., Alberti, S., Fawzi, N.L., et al. (2018) Protein Phase Separation: A New Phase in Cell Biology. *Trends in Cell Biology*, **28**, 420–435.

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.

Borman, A.M., Michel, Y.M. and Kean, K.M. (2000) Biochemical characterisation of cap–poly(A) synergy in rabbit reticulocyte lysates: the eIF4G–PABP interaction increases the functional affinity of eIF4E for the capped mRNA 5′-end. *Nucleic Acids Res*, **28**, 4068–4075.

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

Brandmann, T., Fakim, H., Padamsi, Z., Youn, J., Gingras, A., Fabian, M.R. and Jinek, M. (2018) Molecular architecture of LSM14 interactions involved in the assembly of mRNA silencing complexes. *The EMBO Journal*, **37**, e97869.

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., Huntzinger, E., Fauser, M. and Izaurralde, E. (2011) GW182 Proteins Directly Recruit Cytoplasmic Deadenylase Complexes to miRNA Targets. *Molecular Cell*, **44**, 120–133.

Brengues, M. (2005) Movement of Eukaryotic mRNAs Between Polysomes and Cytoplasmic Processing Bodies. *Science*, **310**, 486–489.

Browning, K.S. and Bailey-Serres, J. (2015) Mechanism of Cytoplasmic mRNA Translation. *The Arabidopsis Book*, **13**, e0176.

Bushell, M., Wood, W., Carpenter, G., Pain, V.M., Morley, S.J. and Clemens, M.J. (2001) Disruption of the Interaction of Mammalian Protein Synthesis Eukaryotic Initiation Factor 4B with the Poly(A)-binding Protein by Caspase- and Viral Protease-mediated Cleavages. *J. Biol. Chem.*, **276**, 23922–23928.

Caponigro, G. and Parker, R. (1995) Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes & Development*, **9**, 2421–2432.

Cerutti, H., Ma, X., Msanne, J. and Repas, T. (2011) RNA-Mediated Silencing in Algae: Biological Roles and Tools for Analysis of Gene Function. *Eukaryotic Cell*, **10**, 1164–1172.

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. *Molecular Cell*, **53**, 1044–1052.

Chang, H., Yeo, J., Kim, J., et al. (2018) Terminal Uridylyltransferases Execute Programmed Clearance of Maternal Transcriptome in Vertebrate Embryos. *Molecular Cell*, **70**, 72-82.e7.

Chantarachot, T. and Bailey-Serres, J. (2018) Polysomes, Stress Granules, and Processing Bodies: A Dynamic Triumvirate Controlling Cytoplasmic mRNA Fate and Function. *Plant Physiology*, **176**, 254–269.

Chapman, E.G., Costantino, D.A., Rabe, J.L., Moon, S.L., Wilusz, J., Nix, J.C. and Kieft, J.S. (2014) The Structural Basis of Pathogenic Subgenomic Flavivirus RNA (sfRNA) Production. *Science*, **344**, 307–310.

Charenton, C., Gaudon-Plesse, C., Fourati, Z., Taverniti, V., Back, R., Kolesnikova, O., Séraphin, B. and Graille, M. (2017) A unique surface on Pat1 C-terminal domain directly interacts with Dcp2 decapping enzyme and Xrn1 5′–3′ mRNA exonuclease in yeast. *Proceedings of the National Academy of Sciences*, **114**, E9493–E9501.

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 m7GDP and its Edc3 activator. *Nature Structural & Molecular Biology*, **23**, 982–986.

Charlesworth, A., Meijer, H.A. and Moor, C.H. de (2013) Specificity factors in cytoplasmic polyadenylation. *Wiley Interdisciplinary Reviews: RNA*, **4**, 437–461.

Chekanova, J.A., Gregory, B.D., Reverdatto, S.V., et al. (2007) Genome-Wide High-Resolution Mapping of Exosome Substrates Reveals Hidden Features in the Arabidopsis Transcriptome. *Cell*, **131**, 1340–1353.

Chekulaeva, M. and Landthaler, M. (2016) Eyes on Translation. *Molecular Cell*, **63**, 918–925.

Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R. and Filipowicz, W. (2011) miRNA repression involves GW182-mediated recruitment of CCR4–NOT through

conserved W-containing motifs. *Nature Structural & Molecular Biology*, **18**, 1218–1226.

Chen, C.-Y.A., Ezzeddine, N. and Shyu, A.-B. (2008) Messenger RNA Half-Life Measurements in Mammalian Cells. *Methods Enzymol*, **448**, 335–357.

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

Chen, Y., Boland, A., Kuzuoğlu-Öztürk, D., Bawankar, P., Loh, B., Chang, C.-T., Weichenrieder, O. and Izaurralde, E. (2014) A DDX6-CNOT1 Complex and W-Binding Pockets in CNOT9 Reveal Direct Links between miRNA Target Recognition and Silencing. *Molecular Cell*, **54**, 737–750.

Cheng, C.-P., Jaag, H.M., Jonczyk, M., Serviene, E. and Nagy, P.D. (2007) Expression of the Arabidopsis Xrn4p 5′–3′ exoribonuclease facilitates degradation of tombusvirus RNA and promotes rapid emergence of viral variants in plants. *Virology*, **368**, 238–248.

Cheng, S. and Gallie, D.R. (2007) eIF4G, eIFiso4G, and eIF4B Bind the Poly(A)-binding Protein through Overlapping Sites within the RNA Recognition Motif Domains. *J. Biol. Chem.*, **282**, 25247–25258.

Chiba, Y., Johnson, M.A., Lidder, P., Vogel, J.T., Erp, H. van 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. *The Plant Journal*, **0**. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/tpj.14022 [Accessed August 10, 2018].

Chlebowski, A., Lubas, M., Jensen, T.H. and Dziembowski, A. (2013) RNA decay machines: The exosome. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1829**, 552–560.

Chorghade, S., Seimetz, J., Emmons, R., Yang, J., Bresson, S.M., Lisio, M.D., Parise, G., Conrad, N.K. and Kalsotra, A. (2017) Poly(A) tail length regulates PABPC1 expression to tune translation in the heart. *eLife*, **6**. Available at: https://elifesciences.org/articles/24139 [Accessed June 25, 2018].

Chou, W.-L., Chung, Y.-L., Fang, J.-C. and Lu, C.-A. (2017) Novel interaction between CCR4 and CAF1 in rice CCR4–NOT deadenylase complex. *Plant Molecular Biology*, **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 Molecular Biology*, **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.

Christie, M., Boland, A., Huntzinger, E., Weichenrieder, O. and Izaurralde, E. (2013) Structure of the PAN3 Pseudokinase Reveals the Basis for Interactions with the PAN2

Deadenylase and the GW182 Proteins. *Molecular Cell*, **51**, 360–373.

Chu, C. and Rana, T.M. (2006) Translation Repression in Human Cells by MicroRNA-Induced Gene Silencing Requires RCK/p54. *PLOS Biology*, **4**, e210.

Chung, C.Z., Seidl, L.E., Mann, M.R. and Heinemann, I.U. (2017) Tipping the balance of RNA stability by 3′ editing of the transcriptome. *Biochimica et Biophysica Acta (BBA) - General Subjects*, **1861**, 2971–2979.

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.

Coller, J. and Parker, R. (2005) General Translational Repression by Activators of mRNA Decapping. *Cell*, **122**, 875–886.

Coller, J.M., Tucker, M., Sheth, U., Valencia-Sanchez, M.A. and Parker, R. (2001) The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA*, **7**, 1717–1727.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S. and Baulcombe, D.C. (2000) An RNA-Dependent RNA Polymerase Gene in Arabidopsis Is Required for Posttranscriptional Gene Silencing Mediated by a Transgene but Not by a Virus. *Cell*, **101**, 543–553.

Daugeron, M.-C., Mauxion, F. and Séraphin, B. (2001) The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Res*, **29**, 2448–2455.

Daughdrill, G.W., Narayanaswami, P., Gilmore, S.H., Belczyk, A. and Brown, C.J. (2007) Dynamic Behavior of an Intrinsically Unstructured Linker Domain Is Conserved in the Face of Negligible Amino Acid Sequence Conservation. *Journal of Molecular Evolution*, **65**, 277–288.

De Almeida, C., Scheer, H., Zuber, H. and Gagliardi, D. (2018) RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome: RNA uridylation. *Wiley Interdisciplinary Reviews: RNA*, **9**, e1440.

Decourty, L., Saveanu, C., Zemam, K., Hantraye, F., Frachon, E., Rousselle, J.-C., Fromont-Racine, M. and Jacquier, A. (2008) Linking functionally related genes by sensitive and quantitative characterization of genetic interaction profiles. *PNAS*, **105**, 5821–5826.

Delatte, B., Wang, F., Ngoc, L.V., et al. (2016) Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science*, **351**, 282–285.

Delis, C., Krokida, A., Tomatsidou, A., et al. (2015) AtHESPERIN: a novel regulator of circadian rhythms with poly(A)-degrading activity in plants. *RNA Biol*, **13**, 68–82.

Deller, M.C., Kong, L. and Rupp, B. (2016) Protein stability: a crystallographer's perspective. *Acta Crystallogr F Struct Biol Commun*, **72**, 72–95.

Deng, X., Chen, K., Luo, G.-Z., Weng, X., Ji, Q., Zhou, T. and He, C. (2015) Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids Res*, **43**, 6557–6567.

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.

Dever, T.E., Dinman, J.D. and Green, R. (2018) Translation Elongation and Recoding in

Eukaryotes. *Cold Spring Harbor Perspectives in Biology*, a032649.

Dhanraj, S., Gunja, S.M.R., Deveau, A.P., et al. (2015) Bone marrow failure and developmental delay caused by mutations in poly(A)-specific ribonuclease (PARN). *Journal of Medical Genetics*, **52**, 738–748.

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., et al. (2016) The dynamic N1 -methyladenosine methylome in eukaryotic messenger RNA. *Nature*, **530**, 441–446.

Dunckley, T., Tucker, M. and Parker, R. Two Related Proteins, Edc1p and Edc2p, Stimulate mRNA Decapping in Saccharomyces cerevisiae. , 11.

Eckmann, C.R., Rammelt, C. and Wahle, E. (2011) Control of poly(A) tail length: Control of poly(A) tail length. *Wiley Interdisciplinary Reviews: RNA*, **2**, 348–361.

Ernoult-Lange, M., Baconnais, S., Harper, M., et al. (2012) Multiple binding of repressed mRNAs by the P-body protein Rck/p54. *RNA*, **18**, 1702–1715.

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. *Molecular and Cellular Biology*, **27**, 3970–3981.

Evguenieva-Hackenberg, E., Hou, L., Glaeser, S. and Klug, G. (2014) Structure and function of the archaeal exosome: Structure and function of the archaeal exosome. *Wiley Interdisciplinary Reviews: RNA*, **5**, 623–635.

Eystathioy, T. (2003) The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. *RNA*, **9**, 1171–1173.

Ezzeddine, N., Chang, T.-C., Zhu, W., Yamashita, A., Chen, C.-Y.A., Zhong, Z., Yamashita, Y., Zheng, D. and Shyu, A.-B. (2007) Human TOB, an Antiproliferative Transcription Factor, Is a Poly(A)-Binding Protein-Dependent Positive Regulator of Cytoplasmic mRNA Deadenylation. *Molecular and Cellular Biology*, **27**, 7791–7801.

Fabian, M.R., Cieplak, M.K., Frank, F., et al. (2011) miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. *Nature Structural & Molecular Biology*, **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. *Nature Structural & Molecular Biology*, **24**, 658–665.

Fan, J., Yang, X., Wang, W., Wood, W.H., Becker, K.G. and Gorospe, M. (2002) Global analysis of stress-regulated mRNA turnover by using cDNA arrays. *Proc Natl Acad Sci U S A*, **99**, 10611–10616.

Ferraiuolo, M.A., Basak, S., Dostie, J., Murray, E.L., Schoenberg, D.R. and Sonenberg, N. (2005) A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. *J. Cell Biol.*, **170**, 913–924.

Fischer, N. and Weis, K. (2002) The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. *The EMBO Journal*, **21**, 2788–2797.

Flobinus, A., Chevigny, N., Charley, P.A., et al. (2018) Beet Necrotic Yellow Vein Virus Noncoding RNA Production Depends on a 5′→3′ Xrn Exoribonuclease Activity. *Viruses*, **10**, 137.

Fourati, Z. and Graille, M. (2014) Cytoplasmic mRNA Surveillance Pathways. In *Fungal RNA Biology*. Springer, Cham, pp. 195–216.

Fromm, S.A., Truffault, V., Kamenz, J., Braun, J.E., Hoffmann, N.A., Izaurralde, E. and Sprangers, R. (2012) The structural basis of Edc3- and Scd6-mediated activation of the Dcp1:Dcp2 mRNA decapping complex. *EMBO J*, **31**, 279–290.

Fu, Y., Dominissini, D., Rechavi, G. and He, C. (2014) Gene expression regulation mediated through reversible m^{6}A RNA methylation. *Nature Reviews Genetics*, **15**, 293–306.

Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J., Sonenberg, N., Thoma, C. and Fujiwara, T. (2014) MicroRNAs Trigger Dissociation of eIF4AI and eIF4AII from Target mRNAs in Humans. *Molecular Cell*, **56**, 79–89.

Furuichi, Y., LaFiandra, A. and Shatkin, A.J. (1977) 5′-Terminal structure and mRNA stability. *Nature*, **266**, 235–239.

Gallie, D.R. (1998) A tale of two termini: A functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. , 11.

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. (2002) The 5'-leader of tobacco mosaic virus promotes translation through enhanced recruitment of eIF4F. *Nucleic Acids Res.*, **30**, 3401–3411.

Gallie, D.R. and Liu, R. (2014) Phylogenetic analysis reveals dynamic evolution of the poly(A)-binding protein gene family in plants. *BMC Evolutionary Biology*, **14**. Available at: http://bmcevolbiol.biomedcentral.com/articles/10.1186/s12862-014-0238-4 [Accessed August 30, 2018].

Gallie, D.R., Lucas, W.J. and Walbot, V. (1989) Visualizing mRNA expression in plant protoplasts: factors influencing efficient mRNA uptake and translation. *The Plant Cell*, **1**, 301–311.

Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C. and Wilson, T.M. (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.

Geissler, R., Golbik, R.P. and Behrens, S.-E. (2012) The DEAD-box helicase DDX3 supports the assembly of functional 80S ribosomes. *Nucleic Acids Res*, **40**, 4998–5011.

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. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1829**, 571–579.

Goss, D.J. and Kleiman, F.E. (2013) Poly(A) binding proteins: are they all created equal?: Poly(A) binding proteins. *Wiley Interdisciplinary Reviews: RNA*, **4**, 167–179.

Gowrishankar, G., Winzen, R., Bollig, F., Ghebremedhin, B., Redich, N., Ritter, B., Resch, K., Kracht, M. and Holtmann, H. (2005) Inhibition of mRNA deadenylation and degradation

by ultraviolet light. *Biol. Chem.*, **386**, 1287–1293.

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

Greatrix, B.W. and Vuuren, H.J.J. van (2006) Expression of the HXT13, HXT15 and HXT17 genes in Saccharomyces cerevisiae and stabilization of the HXT1 gene transcript by sugar-induced osmotic stress. *Curr. Genet.*, **49**, 205–217.

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. *Developmental Cell*, **14**, 854–866.

Grigull, J., Mnaimneh, S., Pootoolal, J., Robinson, M.D. and Hughes, T.R. (2004) Genome-Wide Analysis of mRNA Stability Using Transcription Inhibitors and Microarrays Reveals Posttranscriptional Control of Ribosome Biogenesis Factors. *Mol. Cell. Biol.*, **24**, 5534–5547.

Grudzien-Nogalska, E. and Kiledjian, M. (2017) New insights into decapping enzymes and selective mRNA decay: mRNA Decapping Enzymes in mRNA Decay. *Wiley Interdisciplinary Reviews: RNA*, **8**, e1379.

Guan, Q., Zheng, W., Tang, S., Liu, X., Zinkel, R.A., Tsui, K.-W., Yandell, B.S. and Culbertson, M.R. (2006) Impact of Nonsense-Mediated mRNA Decay on the Global Expression Profile of Budding Yeast. *PLOS Genetics*, **2**, e203.

Haar, T. von der, Ball, P.D. and McCarthy, J.E.G. (2000) Stabilization of Eukaryotic Initiation Factor 4E Binding to the mRNA 5′-Cap by Domains of eIF4G. *J. Biol. Chem.*, **275**, 30551–30555.

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. *The Journal of Cell Biology*, **189**, 289–302.

Hagan, J.P., Piskounova, E. and Gregory, R.I. (2009) Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nature Structural & Molecular Biology*, **16**, 1021–1025.

Halbach, F., Reichelt, P., Rode, M. and Conti, E. (2013) The Yeast Ski Complex: Crystal Structure and RNA Channeling to the Exosome Complex. *Cell*, **154**, 814–826.

Harrison, P.M. (2017) fLPS: Fast discovery of compositional biases for the protein universe. *BMC Bioinformatics*, **18**. Available at: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-017-1906-3 [Accessed May 7, 2018].

He, F., Brown, A.H. and Jacobson, A. (1997) Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway. *Molecular and Cellular Biology*, **17**, 1580–1594.

He, F. and Jacobson, A. (2015a) Control of mRNA decapping by positive and negative regulatory elements in the Dcp2 C-terminal domain. *RNA*, **21**, 1633–1647.

He, F. and Jacobson, A. (2015b) Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. *Annual Review of Genetics*, **49**, 339–366.

He, F., Li, X., Spatrick, P., Casillo, R., Dong, S. and Jacobson, A. (2003) Genome-Wide Analysis of mRNAs Regulated by the Nonsense-Mediated and 5′ to 3′ mRNA Decay Pathways in Yeast. *Molecular Cell*, **12**, 1439–1452.

He, W. and Parker, R. The Yeast Cytoplasmic LsmI/Pat1p Complex Protects mRNA 3˜Termini From Partial Degradation. , 11.

Hellen, C.U.T. (2018) Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harbor Perspectives in Biology*, a032656.

Helm, M. and Motorin, Y. (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. *Nature Reviews Genetics*, **18**, 275–291.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J. and Kim, V.N. (2008) Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Molecular Cell*, **32**, 276–284.

Heo, I., Joo, C., Kim, Y.-K., Ha, M., Yoon, M.-J., Cho, J., Yeom, K.-H., Han, J. and Kim, V.N. (2009) TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. *Cell*, **138**, 696–708.

Heyer, E.E. and Moore, M.J. (2016) Redefining the Translational Status of 80S Monosomes. *Cell*, **164**, 757–769.

Hinnebusch, A.G. and Lorsch, J.R. (2012) The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges. *Cold Spring Harb Perspect Biol*, **4**, a011544.

Hir, H.L., Saulière, J. and Wang, Z. (2016) The exon junction complex as a node of post-transcriptional networks. *Nature Reviews Molecular Cell Biology*, **17**, 41–54.

Hirayama, T., Matsuura, T., Ushiyama, S., et al. (2013) A poly(A)-specific ribonuclease directly regulates the poly(A) status of mitochondrial mRNA in Arabidopsis. *Nature Communications*, **4**, 2247.

Hong, K.Y., Lee, S.H., Gu, S., Kim, E., An, S., Kwon, J., Lee, J.-B. and Jang, S.K. (2017) The bent conformation of poly(A)-binding protein induced by RNA-binding is required for its translational activation function. *RNA Biol*, **14**, 370–377.

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. *Molecular Cell*, **68**, 615-625.e9.

Houseley, J., LaCava, J. and Tollervey, D. (2006) RNA-quality control by the exosome. *Nature Reviews Molecular Cell Biology*, **7**, 529–539.

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.

Huang, J. and Yin, P. (2018) Structural Insights into N6-methyladenosine (m6A) Modification in the Transcriptome. *Genomics, Proteomics and Bioinformatics*, **16**, 85–98.

Hubstenberger, A., Courel, M., Bénard, M., et al. (2017) P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Molecular Cell*, **68**, 144-157.e5.

Hug, N., Longman, D. and Cáceres, J.F. (2016) Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res*, **44**, 1483–1495.

Imataka, H. (1998) A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *The EMBO Journal*, **17**, 7480–7489.

Inada, T. (2013) Quality control systems for aberrant mRNAs induced by aberrant translation elongation and termination. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1829**, 634–642.

Inada, T. (2017) The Ribosome as a Platform for mRNA and Nascent Polypeptide Quality Control. *Trends in Biochemical Sciences*, **42**, 5–15.

Ingolia, N.T. (2014) Ribosome profiling: new views of translation, from single codons to genome scale. *Nature Reviews Genetics*, **15**, 205–213.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M. and Weissman, J.S. (2012) The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature Protocols*, **7**, 1534–1550.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S. and Weissman, J.S. (2009) Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science*, **324**, 218–223.

Iwakawa, H. and Tomari, Y. (2013) Molecular Insights into microRNA-Mediated Translational Repression in Plants. *Molecular Cell*, **52**, 591–601.

Iwakawa, H. and Tomari, Y. (2015) The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends in Cell Biology*, **25**, 651–665.

Jaag, H.M. and Nagy, P.D. (2009) Silencing of Nicotiana benthamiana Xrn4p exoribonuclease promotes tombusvirus RNA accumulation and recombination. *Virology*, **386**, 344–352.

Jackson, R.J., Hellen, C.U.T. and Pestova, T.V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews Molecular Cell Biology*, **11**, 113–127.

Januszyk, K. and Lima, C.D. (2014) The eukaryotic RNA exosome. *Current Opinion in Structural Biology*, **24**, 132–140.

Jäschke, A., Höfer, K., Nübel, G. and Frindert, J. (2016) Cap-like structures in bacterial RNA and epitranscriptomic modification. *Current Opinion in Microbiology*, **30**, 44–49.

Jiang, S., Jiang, L., Yang, J., Peng, J., Lu, Y., Zheng, H., Lin, L., Chen, J. and Yan, F. (2018) Over-expression of Oryza sativa Xrn4 confers plant resistance to virus infection. *Gene*, **639**, 44–51.

Jiao, X., Doamekpor, S.K., Bird, J.G., Nickels, B.E., Tong, L., Hart, R.P. and Kiledjian, M. (2017) 5′ End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding. *Cell*, **168**, 1015-1027.e10.

Jinek, M., Coyle, S.M. and Doudna, J.A. (2011) Coupled 5′ Nucleotide Recognition and Processivity in Xrn1-Mediated mRNA Decay. *Molecular Cell*, **41**, 600–608.

Jones, C.I., Zabolotskaya, M.V. and Newbury, S.F. The 5′ → 3′ exoribonuclease XRN1/Pacman and its functions in cellular processes and development. *Wiley Interdisciplinary Reviews: RNA*, **3**, 455–468.

Kahvejian, A., Svitkin, Y.V., Sukarieh, R., M'Boutchou, M.-N. and Sonenberg, N. (2005) Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev.*, **19**, 104–113.

Kalisiak, K., Kuliński, T.M., Tomecki, R., Cysewski, D., Pietras, Z., Chlebowski, A., Kowalska, K. and Dziembowski, A. (2017) A short splicing isoform of HBS1L links the cytoplasmic exosome and SKI complexes in humans. *Nucleic Acids Res*, **45**, 2068–2080.

Karousis, E.D., Nasif, S. and Mühlemann, O. (2016) Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact: Nonsense-mediated mRNA decay. *Wiley Interdisciplinary Reviews: RNA*, **7**, 661–682.

Kastenmayer, J.P. and Green, P.J. (2000) Novel features of the XRN-family in Arabidopsis: Evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. *PNAS*, **97**, 13985–13990.

Kawa, D. and Testerink, C. (2017) Regulation of mRNA decay in plant responses to salt and osmotic stress. *Cell. Mol. Life Sci.*, **74**, 1165–1176.

Kedersha, N., Stoecklin, G., Ayodele, M., et al. (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *The Journal of Cell Biology*, **169**, 871–884.

Kessler, S.H. and Sachs, A.B. (1998) RNA Recognition Motif 2 of Yeast Pab1p Is Required for Its Functional Interaction with Eukaryotic Translation Initiation Factor 4G. *Mol. Cell. Biol.*, **18**, 51–57.

Kiledjian, M. (2018) Eukaryotic RNA 5′-End NAD+ Capping and DeNADding. *Trends in Cell Biology*, **28**, 454–464.

Ko, S., Kawasaki, I. and Shim, Y.-H. (2013) PAB-1, a Caenorhabditis elegans Poly(A)-Binding Protein, Regulates mRNA Metabolism in germline by Interacting with CGH-1 and CAR-1. *PLoS One*, **8**. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3868610/ [Accessed September 1, 2018].

Korneeva, N.L., Lamphear, B.J., Hennigan, F.L.C. and Rhoads, R.E. (2000) Mutually Cooperative Binding of Eukaryotic Translation Initiation Factor (eIF) 3 and eIF4A to Human eIF4G-1. *Journal of Biological Chemistry*, **275**, 41369–41376.

Kowalinski, E., Kögel, A., Ebert, J., Reichelt, P., Stegmann, E., Habermann, B. and Conti, E. (2016) Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. *Molecular Cell*, **63**, 125–134.

Kramer, S., Queiroz, R., Ellis, L., Hoheisel, J.D., Clayton, C. and Carrington, M. (2010) The RNA helicase DHH1 is central to the correct expression of many developmentally regulated mRNAs in trypanosomes. *Journal of Cell Science*, jcs.058511.

Kulkarni, M., Ozgur, S. and Stoecklin, G. (2010) On track with P-bodies. *Biochemical Society Transactions*, **38**, 242–251.

Kumakura, N., Otsuki, H., Tsuzuki, M., Takeda, A. and Watanabe, Y. (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**. Available at:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3820695/ [Accessed September 2, 2018].

Kurihara, Y., Matsui, A., Hanada, K., et al. (2009) Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in Arabidopsis. *PNAS*, **106**, 2453–2458.

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

Łabno, A., Tomecki, R. and Dziembowski, A. (2016) Cytoplasmic RNA decay pathways - Enzymes and mechanisms. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, **1863**, 3125–3147.

Ladomery, M., Wade, E. and Sommerville, J. (1997) Xp54, the Xenopus homologue of human RNA helicase p54, is an integral component of stored mRNP particles in oocytes. *Nucleic Acids Res*, **25**, 965–973.

Lai, M.-C., Lee, Y.-H.W., Tarn, W.-Y. and Walter, P. (2008) The DEAD-Box RNA Helicase DDX3 Associates with Export Messenger Ribonucleoproteins as well asTip-associated Protein and Participates in Translational Control. *MBoC*, **19**, 3847–3858.

Lamphear, B.J., Kirchweger, R., Skern, T. and Rhoads, R.E. (1995) Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *J. Biol. Chem.*, **270**, 21975–21983.

Lange, H. and Gagliardi, D. (2011) The exosome and 3'-5' RNA degradation in plants. *Adv. Exp. Med. Biol.*, **702**, 50–62.

Lange, H., Holec, S., Cognat, V., Pieuchot, L., Ret, M.L., 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 in Plant Science*, **14**, 497–504.

Lange, H., Zuber, H., Sement, F.M., et al. (2014) The RNA Helicases AtMTR4 and HEN2 Target Specific Subsets of Nuclear Transcripts for Degradation by the Nuclear Exosome in Arabidopsis thaliana X. Chen, ed. *PLoS Genetics*, **10**, e1004564.

Lawrence, R.J., Earley, K., Pontes, O., Silva, M., Chen, Z.J., Neves, N., Viegas, W. and Pikaard, C.S. (2004) A Concerted DNA Methylation/Histone Methylation Switch Regulates rRNA Gene Dosage Control and Nucleolar Dominance. *Molecular Cell*, **13**, 599–609.

Le, H., Tanguay, R.L., Balasta, M.L., Wei, C.-C., Browning, K.S., Metz, A.M., Goss, D.J. and Gallie, D.R. (1997) Translation Initiation Factors eIF-iso4G and eIF-4B Interact with the Poly(A)-binding Protein and Increase Its RNA Binding Activity. *J. Biol. Chem.*, **272**, 16247–16255.

Lee, C.-C., Lin, T.-L., Lin, J.-W., Han, Y.-T., Huang, Y.-T., Hsu, Y.-H. and Meng, M. (2016) Promotion of Bamboo Mosaic Virus Accumulation in Nicotiana benthamiana by 5′→3′ Exonuclease NbXRN4. *Front Microbiol*, **6**. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4702010/ [Accessed July 17, 2018].

Lee, C.-S., Dias, A.P., Jedrychowski, M., Patel, A.H., Hsu, J.L. and Reed, R. (2008) Human

DDX3 functions in translation and interacts with the translation initiation factor eIF3. *Nucleic Acids Res*, **36**, 4708–4718.

Lee, M., Kim, B. and Kim, V.N. (2014) Emerging Roles of RNA Modification: m 6 A and U-Tail. *Cell*, **158**, 980–987.

Lee, S.H., Oh, J., Park, J., Paek, K.Y., Rho, S., Jang, S.K. and Lee, J. (2014) Poly(A) RNA and Paip2 act as allosteric regulators of poly(A)-binding protein. *Nucleic Acids Research*, **42**, 2697–2707.

Li, W., Ma, M., Feng, Y., Li, H., Wang, Y., Ma, Y., Li, M., An, F. and Guo, H. (2015) EIN2-Directed Translational Regulation of Ethylene Signaling in Arabidopsis. *Cell*, **163**, 670–683.

Li, X., Ma, S. and Yi, C. (2016) Pseudouridine: the fifth RNA nucleotide with renewed interests. *Curr Opin Chem Biol*, **33**, 108–116.

Li, X., Peng, J. and Yi, C. (2017) Transcriptome-wide mapping of N1-methyladenosine methylome. *Methods in Molecular Biology*, **1562**, 245–255.

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

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., Kim, D., Lee, Y., et al. (2018) Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science*, eaam5794.

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. *Nature Structural & Molecular Biology*, **24**, 1057–1063.

Llácer, J.L., Hussain, T., Marler, L., Aitken, C.E., Thakur, A., Lorsch, J.R., Hinnebusch, A.G. and Ramakrishnan, V. (2015) Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. *Mol Cell*, **59**, 399–412.

Lowell, J.E., Rudner, D.Z. and Sachs, A.B. (1992) 3'-UTR-dependent deadenylation by the yeast poly(A) nuclease. *Genes & Development*, **6**, 2088–2099.

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. *The EMBO Journal*, **32**, 1855–1868.

Luo, G.-Z., MacQueen, A., Zheng, G., et al. (2014) Unique features of the m6A methylome in Arabidopsis thaliana. *Nat Commun*, **5**, 5630.

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.

Makino, D.L., Halbach, F. and Conti, E. (2013) The RNA exosome and proteasome: common principles of degradation control. *Nature Reviews Molecular Cell Biology*, **14**, 654–660.

Maldonado-Bonilla, L.D. (2014) Composition and function of P bodies in Arabidopsis thaliana.

Frontiers in Plant Science, **5**.

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. *The EMBO Journal*, **32**, 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 Biology*, 14.

Marguerat, S., Lawler, K., Brazma, A. and Bähler, J. (2014) Contributions of transcription and mRNA decay to gene expression dynamics of fission yeast in response to oxidative stress. *RNA Biol*, **11**, 702–714.

Marnef, A. and Standart, N. (2010) Pat1 proteins: a life in translation, translation repression and mRNA decay. *Biochemical Society Transactions*, **38**, 1602–1607.

Marshall, A.N., Han, J., Kim, M. and Hoof, A. van (2018) Conservation of mRNA quality control factor Ski7 and its diversification through changes in alternative splicing and gene duplication. *Proceedings of the National Academy of Sciences*, 201801997.

Marzluff, W.F. and Koreski, K.P. (2017) Birth and Death of Histone mRNAs. *Trends Genet.*, **33**, 745–759.

Mata, J. and Wise, J.A. (2017) 4-Thiouridine Labeling to Analyze mRNA Turnover in *Schizosaccharomyces pombe*. *Cold Spring Harbor Protocols*, **2017**, pdb.prot091645.

Mathys, H., Basquin, J., Ozgur, S., et al. (2014) Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in MicroRNA Repression. *Molecular Cell*, **54**, 751–765.

Mauer, J., Luo, X., Blanjoie, A., et al. (2017) Reversible methylation of m6Am in the 5′ cap controls mRNA stability. *Nature*, **541**, 371–375.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F. and Dietz, H.C. (2004) Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nature Genetics*, **36**, 1073–1078.

Merchante, C., Brumos, J., Yun, J., et al. (2015) Gene-Specific Translation Regulation Mediated by the Hormone-Signaling Molecule EIN2. *Cell*, **163**, 684–697.

Merret, R., Descombin, J., Juan, Y., Favory, J.-J., Carpentier, M.-C., Chaparro, C., Charng, Y., Deragon, J.-M. and Bousquet-Antonelli, C. (2013) XRN4 and LARP1 Are Required for a Heat-Triggered mRNA Decay Pathway Involved in Plant Acclimation and Survival during Thermal Stress. *Cell Reports*, **5**, 1279–1293.

Merret, R., Nagarajan, V.K., Carpentier, M.-C., et al. (2015) Heat-induced ribosome pausing triggers mRNA co-translational decay in Arabidopsis thaliana. *Nucleic Acids Res*, **43**, 4121–4132.

Mészáros, B., Dosztányi, Z., Magyar, C. and Simon, I. (2014) Bioinformatical Approaches to Unstructured/Disordered Proteins and Their Interactions. In *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes*. Springer Series in Bio-/Neuroinformatics. Springer, Berlin, Heidelberg, pp. 525–556. Available at:

https://link.springer.com/chapter/10.1007/978-3-642-28554-7_16 [Accessed May 9, 2018].

Metzstein, M.M. and Krasnow, M.A. (2006) Functions of the Nonsense-Mediated mRNA Decay Pathway in Drosophila Development. *PLOS Genetics*, **2**, e180.

Meyer, K.D. and Jaffrey, S.R. (2014) The dynamic epitranscriptome: N^6-methyladenosine and gene expression control. *Nature Reviews Molecular Cell Biology*, **15**, 313–326.

Moon, S.L., Blackinton, J.G., Anderson, J.R., Dozier, M.K., Dodd, B.J.T., Keene, J.D., Wilusz, C.J., Bradrick, S.S. and Wilusz, J. (2015) XRN1 Stalling in the 5' UTR of Hepatitis C Virus and Bovine Viral Diarrhea Virus Is Associated with Dysregulated Host mRNA Stability. *PLOS Pathogens*, **11**, e1004708.

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.

Morino, S., Imataka, H., Svitkin, Y.V., Pestova, T.V. and Sonenberg, N. (2000) Eukaryotic Translation Initiation Factor 4E (eIF4E) Binding Site and the Middle One-Third of eIF4GI Constitute the Core Domain for Cap-Dependent Translation, and the C-Terminal One-Third Functions as a Modulatory Region. *Mol Cell Biol*, **20**, 468–477.

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. *Molecular and Cellular Biology*, **32**, 2585–2595.

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. *Molecular and Cellular Biology*, **30**, 460–469.

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. *Nature Communications*, **9**, 1152.

Munoz-Tello, P., Gabus, C. and Thore, S. (2012) Functional Implications from the Cid1 Poly(U) Polymerase Crystal Structure. *Structure*, **20**, 977–986.

Munoz-Tello, P., Rajappa, L., Coquille, S. and Thore, S. (2015) Polyuridylation in Eukaryotes: A 3′-End Modification Regulating RNA Life. *BioMed Research International*, **2015**, 1–12.

Nakamura, A., Amikura, R., Hanyu, K. and Kobayashi, S. (2001) Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis. *Development*, **128**, 3233–3242.

Narsai, R., Howell, K.A., Millar, A.H., O'Toole, N., Small, I. and Whelan, J. (2007) Genome-Wide Analysis of mRNA Decay Rates and Their Determinants in Arabidopsis thaliana. *The Plant Cell*, **19**, 3418–3436.

Nasif, S., Contu, L. and Mühlemann, O. (2018) Beyond quality control: The role of nonsense-mediated mRNA decay (NMD) in regulating gene expression. *Seminars in Cell & Developmental Biology*, **75**, 78–87.

Nguyen, A.H., Matsui, A., Tanaka, M., et al. (2015) Loss of Arabidopsis 5'-3' Exoribonuclease AtXRN4 Function Enhances Heat Stress Tolerance of Plants Subjected to Severe Heat Stress.

Plant Cell Physiol., **56**, 1762–1772.

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. *The Plant Journal*, **44**, 972–984.

Nissan, T., Rajyaguru, P., She, M., Song, H. and Parker, R. (2010) Decapping Activators in Saccharomyces cerevisiae Act by Multiple Mechanisms. *Molecular Cell*, **39**, 773–783.

Norbury, C.J. (2013) Cytoplasmic RNA: a case of the tail wagging the dog. *Nature Reviews Molecular Cell Biology*, **14**, 643–653.

Oldfield, C.J. and Dunker, A.K. (2014) Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annual Review of Biochemistry*, **83**, 553–584.

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. *PNAS*, **103**, 13286–13293.

Ostareck, D.H., Vries, I.S.N. and Ostareck-Lederer, A. DDX6 and its orthologs as modulators of cellular and viral RNA expression. *Wiley Interdisciplinary Reviews: RNA*, **5**, 659–678.

Özeş, A.R., Feoktistova, K., Avanzino, B.C. and Fraser, C.S. (2011) Duplex Unwinding and ATPase Activities of the DEAD-Box Helicase eIF4A Are Coupled by eIF4G and eIF4B. *Journal of Molecular Biology*, **412**, 674–687.

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.

Panchy, N., Lehti-Shiu, M.D. and Shiu, S.-H. (2016) Evolution of gene duplication in plants. *Plant Physiology*, pp.00523.2016.

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. *Molecular Cell*, **62**, 462–471.

Parker, R. and Sheth, U. (2007) P Bodies and the Control of mRNA Translation and Degradation. *Molecular Cell*, **25**, 635–646.

Parsyan, A. ed. (2014) *Translation and Its Regulation in Cancer Biology and Medicine*, Dordrecht: Springer Netherlands. Available at: http://link.springer.com/10.1007/978-94-017-9078-9 [Accessed June 26, 2018].

Pelechano, V., Wei, W. and Steinmetz, L.M. (2015) Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell*, **161**, 1400–1412.

Pellegrini, O., Mathy, N., Condon, C. and Bénard, L. (2008) Chapter 9 In Vitro Assays of 5′ to 3′-Exoribonuclease Activity. In *Methods in Enzymology*. RNA Turnover in Eukaryotes: Nucleases, Pathways and Analysis of mRNA Decay. Academic Press, pp. 167–183. Available at: http://www.sciencedirect.com/science/article/pii/S0076687908026098 [Accessed July 16, 2018].

Pen, J.L., Jiang, H., Domenico, T.D., et al. (2017) Terminal uridylyltransferases target RNA

viruses as part of the innate immune system in animals. *bioRxiv*, 209114.

Peng, J., Yang, J., Yan, F., Lu, Y., Jiang, S., Lin, L., Zheng, H., Chen, H. and Chen, J. (2011) Silencing of NbXrn4 facilitates the systemic infection of Tobacco mosaic virus in Nicotiana benthamiana. *Virus Research*, **158**, 268–270.

Perea-Resa, C., Carrasco-López, C., Catalá, R., Turečková, V., Novak, O., Zhang, W., Sieburth, L., Jiménez-Gómez, J.M. and Salinas, J. (2016) The LSM1-7 Complex Differentially Regulates Arabidopsis Tolerance to Abiotic Stress Conditions by Promoting Selective mRNA Decapping. *The Plant Cell*, **28**, 505–520.

Pestova, T.V. and Kolupaeva, V.G. (2002) The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev.*, **16**, 2906–2922.

Piovesan, D., Walsh, I., Minervini, G. and Tosatto, S.C.E. (2017) FELLS: fast estimator of latent local structure. *Bioinformatics*, **33**, 1889–1891.

Piskounova, E., Polytarchou, C., Thornton, J.E., LaPierre, R.J., Pothoulakis, C., Hagan, J.P., Iliopoulos, D. and Gregory, R.I. (2011) Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. *Cell*, **147**, 1066–1079.

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. *The Plant Cell*, **18**, 3047–3057.

Pradhan, S.J., Nesler, K.R., Rosen, S.F., Kato, Y., Nakamura, A., Ramaswami, M. and Barbee, S.A. (2012) The conserved P body component HPat/Pat1 negatively regulates synaptic terminal growth at the larval Drosophila neuromuscular junction. *J Cell Sci*, **125**, 6105–6116.

Presnyak, V., Alhusaini, N., Chen, Y.-H., et al. (2015) Codon Optimality Is a Major Determinant of mRNA Stability. *Cell*, **160**, 1111–1124.

Ptushkina, M. (1998) Cooperative modulation by eIF4G of eIF4E-binding to the mRNA 5' cap in yeast involves a site partially shared by p20. *The EMBO Journal*, **17**, 4798–4808.

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.

Rajyaguru, P., She, M. and Parker, R. (2012) Scd6 Targets eIF4G to Repress Translation: RGG Motif Proteins as a Class of eIF4G-Binding Proteins. *Molecular 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.

Ramani, A.K., Nelson, A.C., Kapranov, P., Bell, I., Gingeras, T.R. and Fraser, A.G. (2009) High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective Caenorhabditis elegans. *Genome Biology*, **10**, R101.

Rayson, S., Arciga-Reyes, L., Wootton, L., Zabala, M.D.T., Truman, W., Graham, N., Grant, M. and Davies, B. (2012) A Role for Nonsense-Mediated mRNA Decay in Plants: Pathogen Responses Are Induced in Arabidopsis thaliana NMD Mutants. *PLOS ONE*, **7**, e31917.

Reimão-Pinto, M.M., Ignatova, V., Burkard, T.R., et al. (2015) Uridylation of RNA Hairpins by

Tailor Confines the Emergence of MicroRNAs in Drosophila. *Molecular Cell*, **59**, 203–216.

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. *The EMBO Journal*, **35**, 2417–2434.

Ren, G., Chen, X. and Yu, B. (2012) Uridylation of miRNAs by HEN1 SUPPRESSOR1 in Arabidopsis. *Current Biology*, **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. *PNAS*, **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.

Rissland, O.S. and Norbury, C.J. (2009) Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. *Nature Structural & Molecular Biology*, **16**, 616–623.

Rogers, G.W., Richter, N.J., Lima, W.F. and Merrick, W.C. (2001) Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J. Biol. Chem.*, **276**, 30914–30922.

Romero-Santacreu, L., Moreno, J., Pérez-Ortín, J.E. and Alepuz, P. (2009) Specific and global regulation of mRNA stability during osmotic stress in Saccharomyces cerevisiae. *RNA*, **15**, 1110–1120.

Roundtree, I.A., Evans, M.E., Pan, T. and He, C. (2017) Dynamic RNA Modifications in Gene Expression Regulation. *Cell*, **169**, 1187–1200.

Roundtree, I.A. and He, C. (2016) RNA epigenetics--chemical messages for posttranscriptional gene regulation. *Curr Opin Chem Biol*, **30**, 46–51.

Roux, M.E., Rasmussen, M.W., Palma, K., et al. (2015) The mRNA decay factor PAT1 functions in a pathway including MAP kinase 4 and immune receptor SUMM2. *The EMBO Journal*, **34**, 593–608.

Rouya, C., Siddiqui, N., Morita, M., Duchaine, T.F., Fabian, M.R. and Sonenberg, N. (2014) Human DDX6 effects miRNA-mediated gene silencing via direct binding to CNOT1. *RNA*, **20**, 1398–1409.

Roy, D. and Rajyaguru, P.I. (2018) Suppressor of clathrin deficiency (Scd6)-An emerging RGG-motif translation repressor. *Wiley Interdisciplinary Reviews: RNA*, e1479.

Rozen, F., Edery, I., Meerovitch, K., Dever, T.E., Merrick, W.C. and Sonenberg, N. (1990) Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Mol. Cell. Biol.*, **10**, 1134–1144.

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 Davis, R.W. (1989) The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell*, **58**, 857–867.

Safaee, N., Kozlov, G., Noronha, A.M., Xie, J., Wilds, C.J. and Gehring, K. (2012) Interdomain Allostery Promotes Assembly of the Poly(A) mRNA Complex with PABP and eIF4G. *Molecular Cell*, **48**, 375–386.

Saito, S., Hosoda, N. and Hoshino, S. (2013) The Hbs1-Dom34 Protein Complex Functions in Non-stop mRNA Decay in Mammalian Cells. *J. Biol. Chem.*, **288**, 17832–17843.

Salles, F., Lieberfarb, M., Wreden, C., Gergen, J. and Strickland, S. (1994) Coordinate initiation of Drosophila development by regulated polyadenylation of maternal messenger RNAs. *Science*, **266**, 1996–1999.

Sawazaki, R., Imai, S., Yokogawa, M., Hosoda, N., Hoshino, S., 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. *Scientific Reports*, **8**.

Scarpin, M.R., Sigaut, L., Temprana, S.G., Boccaccio, G.L., Pietrasanta, L.I. and Muschietti, J.P. (2017) Two Arabidopsis late pollen transcripts are detected in cytoplasmic granules. *Plant Direct*, **1**, e00012.

Schaefer, M., Kapoor, U. and Jantsch, M.F. (2017) Understanding RNA modifications: the promises and technological bottlenecks of the 'epitranscriptome.' *Open Biology*, **7**, 170077.

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

Scheer, H., Zuber, H., De Almeida, C. and Gagliardi, D. (2016) Uridylation Earmarks mRNAs for Degradation… and More. *Trends in Genetics*, **32**, 607–619.

Schneider, C. and Tollervey, D. (2013) Threading the barrel of the RNA exosome. *Trends in Biochemical Sciences*, **38**, 485–493.

Schütz, S., Nöldeke, 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. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1829**, 654–665.

Sement, F.M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J.-M., Bousquet-Antonelli, C., Lange, H. and Gagliardi, D. (2013) Uridylation prevents 3′ trimming of oligoadenylated mRNAs. *Nucleic Acids Res*, **41**, 7115–7127.

Sharif, H. and Conti, E. (2013) Architecture of the Lsm1-7-Pat1 Complex: A Conserved Assembly in Eukaryotic mRNA Turnover. *Cell Reports*, **5**, 283–291.

Sharif, H., Ozgur, S., Sharma, K., Basquin, C., Urlaub, H. and Conti, E. (2013) Structural analysis of the yeast Dhh1–Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. *Nucleic Acids Res*, **41**, 8377–8390.

Sharma, S., Poetz, F., Bruer, M., Ly-Hartig, T.B.N., Schott, J., Séraphin, B. and Stoecklin, G. (2016) Acetylation-Dependent Control of Global Poly(A) RNA Degradation by CBP/p300 and HDAC1/2. *Molecular Cell*, **63**, 927–938.

Shaul, O. (2015) Unique Aspects of Plant Nonsense-Mediated mRNA Decay. *Trends in Plant Science*, **20**, 767–779.

She, M., Decker, C.J., Chen, N., Tumati, S., Parker, R. and Song, H. (2006) Crystal structure and functional analysis of Dcp2p from Schizosaccharomyces pombe. *Nature Structural & Molecular Biology*, **13**, 63–70.

Sheth, U. (2003) Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies. *Science*, **300**, 805–808.

Shih, J.-W., Wang, W.-T., Tsai, T.-Y., Kuo, C.-Y., Li, H.-K. and Wu Lee, Y.-H. (2012) Critical roles of RNA helicase DDX3 and its interactions with eIF4E/PABP1 in stress granule assembly and stress response. *Biochemical Journal*, **441**, 119–129.

Shin, Y. and Brangwynne, C.P. (2017) Liquid phase condensation in cell physiology and disease. *Science*, **357**, eaaf4382.

Shoemaker, C.J. and Green, R. (2012) Translation drives mRNA quality control. *Nature Structural & Molecular Biology*, **19**, 594–601.

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. *Nature Communications*, **8**. Available at: http://www.nature.com/articles/s41467-017-02066-2 [Accessed June 20, 2018].

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*. *The Plant Journal*, **83**, 991–1004.

Simms, C.L., Thomas, E.N. and Zaher, H.S. (2017) Ribosome-based quality control of mRNA and nascent peptides: mRNA and protein quality control on the ribosome. *Wiley Interdisciplinary Reviews: RNA*, **8**, e1366.

Slayter, H.S., Warner, J.R., Rich, A. and Hall, C.E. (1963) The visualization of polyribosomal structure. *Journal of Molecular Biology*, **7**, 652-IN5.

Sloan, K.E. and Bohnsack, M.T. (2018) Unravelling the Mechanisms of RNA Helicase Regulation. *Trends in Biochemical Sciences*, **43**, 237–250.

Smillie, D.A. and Sommerville, J. (2002) RNA helicase p54 (DDX6) is a shuttling protein involved in nuclear assembly of stored mRNP particles. *Journal of Cell Science*, **115**, 395–407.

Song, M.-G. and Kiledjian, M. (2007) 3' Terminal oligo U-tract-mediated stimulation of decapping. *RNA*, **13**, 2356–2365.

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. *Proceedings of the National Academy of Sciences*, **115**, E1485–E1494.

Standart, N. and Weil, D. (2018) P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. *Trends in Genetics*, **34**, 612–626.

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**. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4773420/ [Accessed July 3, 2018].

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.

Sweet, T., Kovalak, C. and Coller, J. (2012) The DEAD-Box Protein Dhh1 Promotes Decapping by Slowing Ribosome Movement. *PLOS Biology*, **10**, e1001342.

Synowsky, S.A. and Heck, A.J.R. The yeast Ski complex is a hetero-tetramer. *Protein Science*, **17**, 119–125.

Szádeczky-Kardoss, I., Gál, L., Auber, A., Taller, J. and Silhavy, D. (2018) The No-go decay system degrades plant mRNAs that contain a long A-stretch in the coding region. *Plant Sci.*, **275**, 19–27.

Tanaka, K.J., Ogawa, K., Takagi, M., Imamoto, N., Matsumoto, K. and Tsujimoto, M. (2006) RAP55, a Cytoplasmic mRNP Component, Represses Translation in Xenopus Oocytes. *J. Biol. Chem.*, **281**, 40096–40106.

Tani, H., Imamachi, N., Salam, K.A., Mizutani, R., Ijiri, K., Irie, T., Yada, T., Suzuki, Y. and Akimitsu, N. (2012) Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. *RNA Biology*, **9**, 1370–1379.

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.

Tarun, S.Z. and Sachs, A.B. (1996) Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J*, **15**, 7168–7177.

Tarun, S.Z. and Sachs, A.B. (1997) Binding of eukaryotic translation initiation factor 4E (eIF4E) to eIF4G represses translation of uncapped mRNA. *Mol. Cell. Biol.*, **17**, 6876–6886.

Tarun, S.Z., Wells, S.E., Deardorff, J.A. and Sachs, A.B. (1997) Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *PNAS*, **94**, 9046–9051.

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., Zaessinger, S., Meyer, S., Simonelig, M. and Wahle, E. (2004) A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in Drosophila. *The EMBO Journal*, **23**, 2862–2871.

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., He, W., Mayes, A.E., Lennertz, P., Beggs, J.D. and Parker, R. (2000) Yeast Sm-like proteins function in mRNA decapping and decay. , **404**, 4.

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. *Molecular Cell*, **8**, 1075–1083.

Thore, S., Mauxion, F., Séraphin, B. and Suck, D. (2003) X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex. *EMBO Rep*, **4**, 1150–1155.

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

Tritschler, F., Braun, J.E., Eulalio, A., Truffault, V., Izaurralde, E. and Weichenrieder, O. (2009) Structural Basis for the Mutually Exclusive Anchoring of P Body Components EDC3 and Tral to the DEAD Box Protein DDX6/Me31B. *Molecular Cell*, **33**, 661–668.

Tritschler, F., Eulalio, A., Helms, S., Schmidt, S., Coles, M., Weichenrieder, O., Izaurralde, E. and Truffault, V. (2008) Similar Modes of Interaction Enable Trailer Hitch and EDC3 To Associate with DCP1 and Me31B in Distinct Protein Complexes. *Mol. Cell. Biol.*, **28**, 6695–6708.

Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrad, D. and Parker, R. (2002) Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. *EMBO J*, **21**, 1427–1436.

Tutucci, E., Livingston, N.M., Singer, R.H. and Wu, B. (2018) Imaging mRNA In Vivo, from Birth to Death. *Annual Review of Biophysics*, **47**, 85–106.

Uchida, N., Hoshino, S. and Katada, T. (2004) Identification of a Human Cytoplasmic Poly(A) Nuclease Complex Stimulated by Poly(A)-binding Protein. *Journal of Biological Chemistry*, **279**, 1383–1391.

Ustianenko, D., Hrossova, D., Potesil, D., et al. (2013) Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA*, **19**, 1632–1638.

Vacic, V., Uversky, V.N., Dunker, A.K. and Lonardi, S. (2007) Composition Profiler: a tool for discovery and visualization of amino acid composition differences. *BMC Bioinformatics*, **8**, 211.

Villa, N., Do, A., Hershey, J.W.B. and Fraser, C.S. (2013) Human Eukaryotic Initiation Factor 4G (eIF4G) Protein Binds to eIF3c, -d, and -e to Promote mRNA Recruitment to the Ribosome. *J Biol Chem*, **288**, 32932–32940.

Viswanathan, P., Ohn, T., Chiang, Y.-C., Chen, J. and Denis, C.L. (2004) Mouse CAF1 can function as a processive deadenylase/3'-5'-exonuclease in vitro but in yeast the deadenylase function of CAF1 is not required for mRNA poly(A) removal. *J. Biol. Chem.*, **279**, 23988–23995.

Vogel, C. and Marcotte, E.M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics*, **13**, 227–232.

Wahle, E. and Winkler, G.S. (2013) RNA decay machines: Deadenylation by the Ccr4–Not and Pan2–Pan3 complexes. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1829**, 561–570.

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

Wang, C., Schmich, F., Srivatsa, S., Weidner, J., Beerenwinkel, N. and Spang, A. Context-dependent deposition and regulation of mRNAs in P-bodies. *eLife*, **7**. Available at:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5752201/ [Accessed June 29, 2018].

Wang, H., Morita, M., Yang, X., et al. (2010) Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. *The EMBO Journal*, **29**, 2566–2576.

Wang, L., Lewis, M.S. and Johnson, A.W. (2005) Domain interactions within the Ski2/3/8 complex and between the Ski complex and Ski7p. *RNA*, **11**, 1291–1302.

Wang, L., Nam, Y., Lee, A.K., Yu, C., Roth, K., Chen, C., Ransey, E.M. and Sliz, P. (2017) LIN28 Zinc Knuckle Domain Is Required and Sufficient to Induce let-7 Oligouridylation. *Cell Reports*, **18**, 2664–2675.

Wang, X. and He, C. (2014) Dynamic RNA modifications in posttranscriptional regulation. *Mol. Cell*, **56**, 5–12.

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**. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4415790/ [Accessed August 31, 2018].

Wang, Y., Liu, C.L., Storey, J.D., Tibshirani, R.J., Herschlag, D. and Brown, P.O. (2002) Precision and functional specificity in mRNA decay. *PNAS*, **99**, 5860–5865.

Warkocki, Z., Krawczyk, P.S., Adamska, D., Bijata, K., Garcia-Perez, J.L. and Dziembowski, A. (2018) Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s. *Cell*. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0092867418309176 [Accessed August 24, 2018].

Webster, M.W., Chen, Y.-H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Coller, J. and Passmore, L.A. (2018) mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Molecular Cell*, **70**, 1089-1100.e8.

Wei, C.-C., Balasta, M.L., Ren, J. and Goss, D.J. (1998) Wheat Germ Poly(A) Binding Protein Enhances the Binding Affinity of Eukaryotic Initiation Factor 4F and (iso)4F for Cap Analogues †. *Biochemistry*, **37**, 1910–1916.

Weill, L., Belloc, E., Bava, F.-A. and Méndez, R. (2012) Translational control by changes in poly(A) tail length: recycling mRNAs. *Nature Structural & Molecular Biology*, **19**, 577–585.

Wells, S.E., Hillner, P.E., Vale, R.D. and Sachs, A.B. (1998) Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell*, **2**, 135–140.

Wilhelm, J.E., Buszczak, M. and Sayles, S. (2005) Efficient Protein Trafficking Requires Trailer Hitch, a Component of a Ribonucleoprotein Complex Localized to the ER in Drosophila. *Developmental Cell*, **9**, 675–685.

Wolf, J., Valkov, E., Allen, M.D., et al. (2014) Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. *The EMBO Journal*, **33**, 1514–1526.

Wormington, M., Searfoss, A.M. and Hurney, C.A. (1996) Overexpression of poly(A) binding protein prevents maturation-specific deadenylation and translational inactivation in Xenopus oocytes. *EMBO J*, **15**, 900–909.

Wright, P.E. and Dyson, H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein

structure-function paradigm. *Journal of Molecular Biology*, **293**, 321–331.

Xu, J. and Chua, N.-H. (2009) Arabidopsis Decapping 5 Is Required for mRNA Decapping, P-Body Formation, and Translational Repression during Postembryonic Development. *The Plant Cell*, **21**, 3270–3279.

Xu, J. and Chua, N.-H. (2011) Processing bodies and plant development. *Current Opinion in Plant Biology*, **14**, 88–93.

Xu, J., Yang, J.-Y., Niu, Q.-W. and Chua, N.-H. (2006) Arabidopsis DCP2, DCP1, and VARICOSE Form a Decapping Complex Required for Postembryonic Development. *The Plant Cell*, **18**, 3386–3398.

Yamashita, A., Chang, T.-C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.-Y.A. and Shyu, A.-B. (2005) Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nature Structural & Molecular Biology*, **12**, 1054–1063.

Yang, E., Schroeder, M., Magnasco, M. and Jr, J.E.D. Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes. , 11.

Yang, H., Song, Jianbo, Yue, L., Mo, X., Song, Jun and Mo, B. (2017) Identification and expression profiling of Oryza sativa nucleotidyl transferase protein (NTP) genes under various stress conditions. *Gene*, **628**, 93–102.

Yang, L., Wu, G. and Poethig, R.S. (2012) Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. *PNAS*, **109**, 315–320.

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. *Molecular Cell*, **70**, 1081-1088.e5.

You, C., Cui, J., Wang, H., Qi, X., Kuo, L.-Y., Ma, H., Gao, L., Mo, B. and Chen, X. (2017) Conservation and divergence of small RNA pathways and microRNAs in land plants. *Genome Biology*, **18**. Available at: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-017-1291-2 [Accessed May 6, 2018].

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 Co-translational RNA Decay in Arabidopsis. *The Plant Cell*, tpc.00456.2016.

Yu, Y., Jia, T. and Chen, X. (2017) The 'how' and 'where' of plant microRNAs. *New Phytologist*, **216**, 1002–1017.

Zanetti, M.E., Chang, I.-F., Gong, F., Galbraith, D.W. and Bailey-Serres, J. (2005) Immunopurification of Polyribosomal Complexes of Arabidopsis for Global Analysis of Gene Expression. *Plant Physiol*, **138**, 624–635.

Zaringhalam, M. and Papavasiliou, F.N. (2016) Pseudouridylation meets next-generation sequencing. *Methods*, **107**, 63–72.

Zekri, L., Kuzuoğlu-Öztürk, D. and Izaurralde, E. (2013) GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *The EMBO Journal*,

32, 1052–1065.

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 U S A*, **107**, 15981–15985.

Zhang, X., Zhu, Y., Wu, H. and Guo, H. (2016) Post-transcriptional gene silencing in plants: a double-edged sword. *Science China Life Sciences*, **59**, 271–276.

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B. and Chen, X. (2012) The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. *Curr. Biol.*, **22**, 689–694.

Zinder, J.C. and Lima, C.D. (2017) Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes & Development*, **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 Reports*, **14**, 2707–2717.

Zukin, R.S., Richter, J. and Bagni, C. (2009) Signals, synapses, and synthesis: how new proteins control plasticity. *Front. Neural Circuits*, **3**.].

Zuo, Y. and Deutscher, M.P. (2001) Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res*, **29**, 1017–1026.

Hélène SCHEER

Rôle complexe de l'uridylation dans le métabolisme des ARNm chez *Arabidopsis thaliana*

Résumé

La dégradation des ARNm participe à la régulation de l'expression des gènes, régulation essentielle au développement et à la survie des organismes. L'**uridylation** des ARNm favorise leur dégradation et est un processus conservé chez la plupart des eucaryotes.

Mes travaux de thèse ont porté sur la caractérisation de URT1, une TUTase responsable de l'uridylation des ARNm chez Arabidopsis. J'ai notamment établi une chaine d'interactions entre URT1 et des facteurs impliqués dans l'inhibition de la traduction. Ces interactions sont probablement conservées dans l'ensemble des plantes terrestres. De plus, je montre grâce à une méthode d'analyse par séquençage à haut-débit des extrémités 3' d'ARNm candidats que le niveau d'expression de URT1 ou de différentes versions mutées ou tronquées peut contrôler les profils de distributions de taille des queues poly(A) uridylées et non-uridylées. De manière très intéressante, la modification de ces profils de distributions est associée à la répression de l'expression d'un gène rapporteur, sans modification de l'accumulation des ARNm correspondants.

Mes résultats contribuent à une meilleure compréhension des rôles moléculaires de l'uridylation des ARNm chez Arabidopsis et ouvrent des perspectives nouvelles quant à la diversité des rôles de l'uridylation dans le métabolisme des ARNm.

Uridylation – Dégradation des ARNm – Déadénylation – Arabidopsis

Abstract

RNA degradation is an essential step of gene regulation and is critical for development and survival of organisms. Uridylation of mRNAs triggers their degradation and is conserved in most eukaryotes.

My PhD thesis is focused on the characterization of URT1, a TUTase responsible for the uridylation of mRNAs in Arabidopsis. I have identified a chain of interactions linking URT1 to factors involved in translation inhibition. Those interactions are likely conserved in all land plants. In addition, I show by 3' RACE-seq, a method using high throughput sequencing to characterize the 3' extremities of candidate mRNAs, that URT1 expression levels can control the distribution profiles of uridylated and non-uridylated poly(A) tails. Interestingly, the modification of these distribution profiles is associated to the repression of the expression of a reporter gene, without affecting mRNA steady-state levels.

My results contribute to a better understanding of the molecular roles of mRNA uridylation in Arabidopsis and open new perspectives related to the diversity of the roles of uridylation in the metabolism of mRNAs.

Uridylation – mRNA decay – Deadenylation – Arabidopsis