Study of the interactome of UPF1, a key factor of Nonsense-mediated decay in Arabidopsis thaliana

Clara Chicois

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

Institut de Biologie Moléculaire des Plantes – CNRS – UPR2357

THÈSE DE DOCTORAT

présentée par :

Clara CHICOIS

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

Discipline: Sciences de la Vie et de la Santé

Spécialité : Aspects Moléculaires et Cellulaires de la Biologie

Study of the interactome of UPF1, a key factor of Nonsense-Mediated Decay in *Arabidopsis thaliana*

Thèse soutenue le 31 janvier 2018 devant la Commission d'Examen :

Dr BOUSQUET-ANTONELLI Cécile	(CNRS, Université de Perpignan)	Rapporteur externe
Dr WACHTER Andreas	(Université de Tübingen)	Rapporteur externe
Dr PFEFFER Sébastien	(CNRS, Université de Strasbourg)	Examinateur interne
Dr GARCIA Damien	(CNRS, Université de Strasbourg)	Directeur de thèse

Acknowledgments / Remerciements

Firstly, I would like to thank the members of the jury Cécile Bousquet-Antonelli, Andreas Wachter and Sébastien Pfeffer for having accepted to evaluate my thesis work.

Je souhaite remercier mon directeur de thèse, Damien Garcia pour sa patience et sa présence au cours de ces trois années de thèse. Je remercie également Shahinez qui m'a beaucoup appris, pour son soutien indéfectible et sa bonne humeur.

Merci aux membres de mon ancienne équipe d'accueil, Marco, Thomas, Yerim, Fabrice et Pat ; ainsi que Lucie et Réjane pour leur travail fourni. Je remercie aussi l'équipe Dégradation des ARNs, notamment Caroline, Marlène et Hélène S. pour leur aide, leur disponibilité ainsi que pour nos moments partagés. Je remercie également Hélène Z., Heike et Gag pour leurs précieux conseils.

J'aimerais également remercier les différentes plateformes de l'IBMP, Malek et Sandrine de la plateforme de séquençage, Jérome à la plateforme d'imagerie, Michèle de la laverie mais également toute la plateforme de production de plantes, pour avoir pris soin de mes plantes pendant toutes ces années. Je remercie également les membres de l'IBMP qui m'ont aidé au cours de ma thèse, par leur mise à disposition de matériel et leurs conseils. Enfin, j'aimerais aussi remercier Laure pour son soutien moral et nos découvertes culinaires.

Évidemment, j'aimerais remercier ma famille et mes amis de m'avoir soutenue, épaulée et encouragée pendant toutes ces années.

Abbreviation list

AD	Activation domain
Ade	Adenine
BD	Binding domain
bp	Base pair
$CaCl_2$	Calcium chloride
Ci	Curie
cDNA	Complementary DNA
Ct	Cycle threshold
D	Aspartic acid
DAI	Days after infiltration
DCP	Decapping
DNA	Desoxyribonucleic Acid
dATP	Desoxyadenosine triphosphate
dCTP	Desoxycytidine triphosphate
dGTP	Desoxyguanine triphosphate
dNTP	Desoxynucleotide triphosphate
DTT	Dithiolthreitol
dTTP	Desoxythymidine triphosphate
EDTA	Ethylenediaminetetra-acetic acid
EGTA	ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
FC	Fold change
g	Centrifuge acceleration
GFP	Green Fluorescent Protein
h	Hour
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
IP	Immunoprecipitation
L	Liter
LB	Lysogeny Broth
LiAc	Lithium acetate
Leu	Leucine
mA	Milliampere
$MgSO_4$	Magnesium sulfate
min	Minute
mJ	Millijoules
mRNA	Messenger RNA
MS	Murashige & Skoog
N	Alanine
NaCl	Sodium chloride
ng	Nanogram

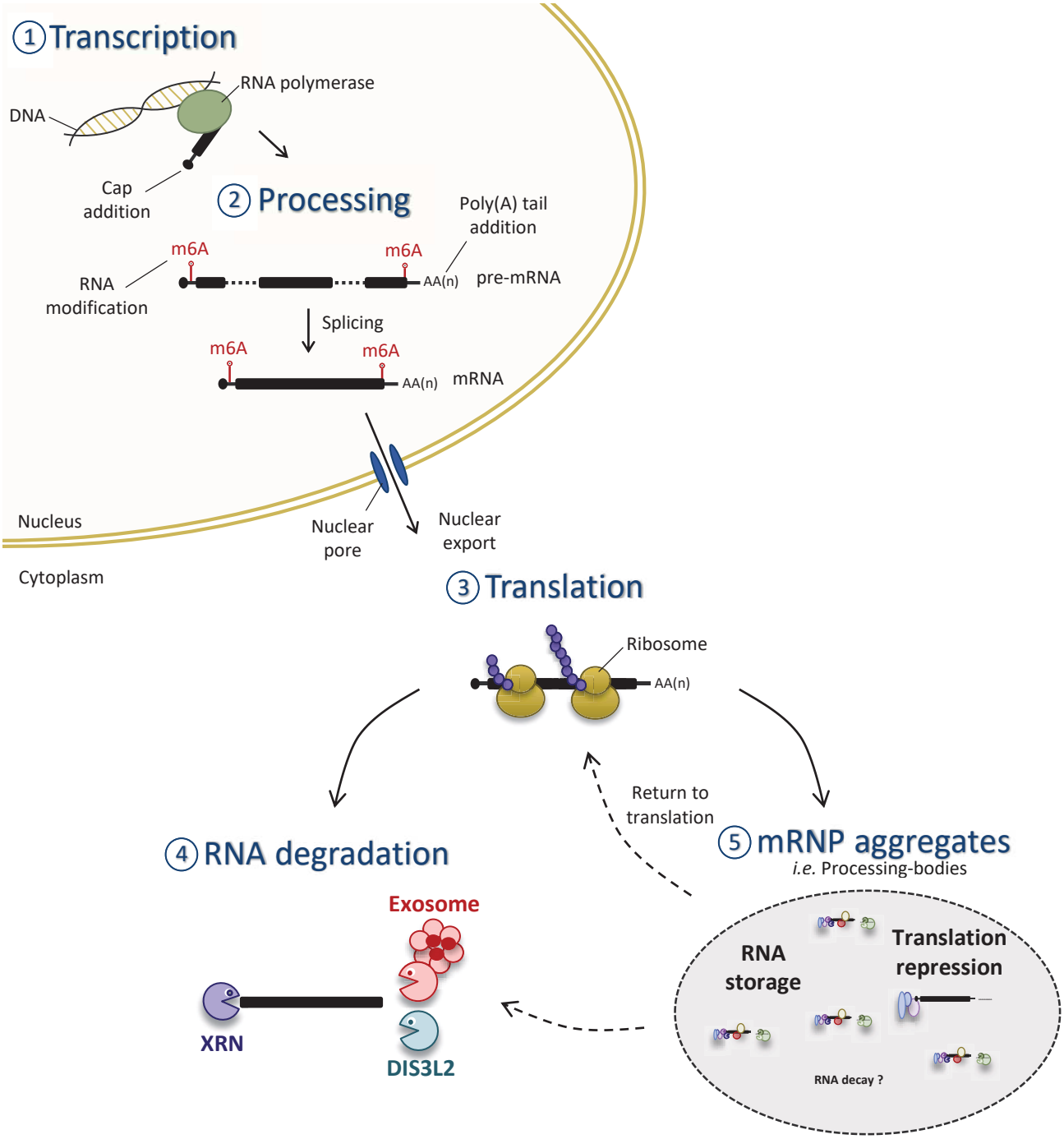
nm	Nanometer
NMD	Nonsense Mediated Decay
Nt	Nucleotide
ORF	Open Reading Frame
PEG	Polyethylene Glycol
PCR	Polymerase Chain Reaction
PVX	Potato Virus X
qPCR	Quantitative PCR
RFP	Red Fluorescent Protein
RNA	Ribonucleic Acid
rpm	Rotation per minute
RT	Reverse transcription
SDS	Sodium Dodecyl Sulfate
Sec	Second
SSC	Sodium saline citrate
TBS	Tris-buffered saline
TE	Tris EDTA
TCV	Turnip Crinkle Virus
Trp	Tryptophane
TuMV	Turnip Mosaic Virus
UCN	UFP1 Co-purified Nuclease
UPF1	Up frameshift
UV	Ultra-violet
V	Volt
v/v	Volume/volume
XRN	Exoribonuclease
w/v	Weight/volume

Table of contents

List of figures

List of tables

Introduction



Inspired from Kervestin & Jacobson, 2012

Figure 1: mRNA life cycle. The first step of the mRNA life cycle is the transcription of an mRNA from a DNA template. The second step of mRNA life cycle is the processing of the mRNA, *i.e.* the cap and poly(A)-tail additions and splicing, as well as nucleotide modification. The third step of the mRNA cycle is mRNA translation after its export in the cytoplasm. After being translated, an mRNA can either be degraded or can be stored in cytoplasmic granules called Processing-bodies where they undergo translation repression. Stored mRNA can return to translation or can be degraded.

Ribonucleic acid (RNA) molecules are present in all organisms and are essential for many different cellular functions. Different types of RNA exist: messenger RNAs (mRNAs), ribozymes, transfer RNAs (tRNAs), microRNAs (miRNAs), small nucleolar RNAs (snoRNAs) and many more. This diversity is correlated with the numerous functions supported by RNA molecules from translation, regulation of gene expression and RNA processing to central functions in cellular macromolecular complexes such as the ribosome. mRNAs are key RNA molecules that carry the genetic information present on DNA and represent the matrix used by the translation machinery for protein synthesis. The accumulation of mRNAs is tightly regulated at different levels of their life cycle such as transcription, processing, nuclear export, translation, storage and degradation (Figure 1). Transcription is the first step of mRNA cycle and occurs from a DNA template through the action of a protein complex containing RNA polymerase II. Transcription rate is different for each gene and depends on intrinsic properties of the corresponding promoter. Transcription represents the first layer of gene regulation, and both transcription factors and chromatin remodeling factors can modulate its efficiency. The second step of the mRNA life cycle is mRNA processing. This step includes cap and poly(A)-tail addition, splicing, as well as nucleoside modification, such as N6-methyladenosine (m6A) (Figure 1). mRNA processing represents an additional layer in the regulation of gene expression. Differences in mRNA processing can lead to the production of several mRNA isoforms, allowing a single gene to encode for different proteins, thereby increasing proteome diversity, and leading to the production of protein isoforms with the potential to exert different functions or localizations. The third major step of the mRNA life cycle is the translation of a mature mRNA into a functional protein. This crucial step is tightly regulated at different levels including translation initiation, elongation and termination. Additional factors can also act as modulators of translation by exerting either enhancement or repression of this process. When an mRNA stops being translated, it can have different fates: it can either be targeted for degradation or can be redirected to RNA granules to be stored in a non-translating state. mRNA degradation and mRNA storage participate to the regulation of gene expression, and they respectively exert an irreversible and reversible repressive action. These different steps of the mRNA cycle are tightly interconnected and can influence each other. For example, transcription and mRNA maturation are tightly linked as they often occur simultaneously. In addition, stored mRNAs can return to translation and translating transcripts, initially thought to be protected from the RNA degradation machinery, can actually undergo co-translational degradation.

In this introduction, I will first present general mRNA degradation pathways and their link relative to repression of mRNA translation. I will then focus on Nonsense-Mediated Decay (NMD), an essential surveillance pathway found in all eukaryotes that participates in the identification and degradation of RNAs carrying features that induce premature translation termination. Finally, I will focus on the knowledge currently available on Up-frameshift 1 (UPF1) protein, the key RNA helicase involved in NMD target identification and degradation.

I. General mRNA degradation

All mRNAs need to be degraded to maintain mRNA homeostasis, or to allow a fast response to environmental changes or developmental stimuli. This housekeeping RNA degradation exists in all organisms studied from bacteria to eukaryotes (Łabno et al., 2016). Although all mRNAs are subject to turnover, mRNAs are not all degraded at the same rate; mRNA half-lives range from minutes to several hours (Chen et al., 2008). In fact, mRNA degradation is tightly regulated and this regulation depends on intrinsic characteristics of a given mRNA. For example, traits such as the presence of motifs allowing protein binding, codon composition, mRNA processing and translation efficiency as well as many other features can modulate the dynamics of mRNA degradation by either favoring or inhibiting mRNA decay pathways (Yang et al., 2003; Presnyak et al., 2015).

In addition to its housekeeping functions, the mRNA decay machinery is also necessary to prevent aberrant forms of mRNAs from accumulating in the cell, in order to avoid potential deleterious effects (Shoemaker and Green, 2012). Deleterious effects can arise from an mRNA carrying a mutation introduced after an uncorrected replication mistake or exposure to mutagenic agents. These abnormal mRNAs could produce aberrant or truncated proteins with potential deleterious or dominant-negative effects, leading to cellular dysfunctions (Miller and Pearce, 2014). Likewise, aberrant non-functional mRNAs could have negative effects if they saturate important cellular machineries such as translation, RNA processing or if they enter the RNA silencing pathway, for example (Eulalio et al., 2007).

mRNAs are modified with a 5’ cap and a 3’ poly-adenosine tail (poly(A)-tail) protecting them from being degraded. As a consequence, mRNA decay usually begins with the removal of the poly(A) and cap structures, steps termed deadenylation and decapping respectively (Łabno et al., 2016). mRNA degradation can occur by (i) the digestion of sequential nucleotides from the 5’ extremity to the 3’ extremity (5’-3’ pathway), (ii) in the 3’ to 5’ direction (3’-5’ pathway) (iii) or in both direction, for example following an endonucleolytic cleavage (Figure 2). Some simple mutations of RNA degradation factors

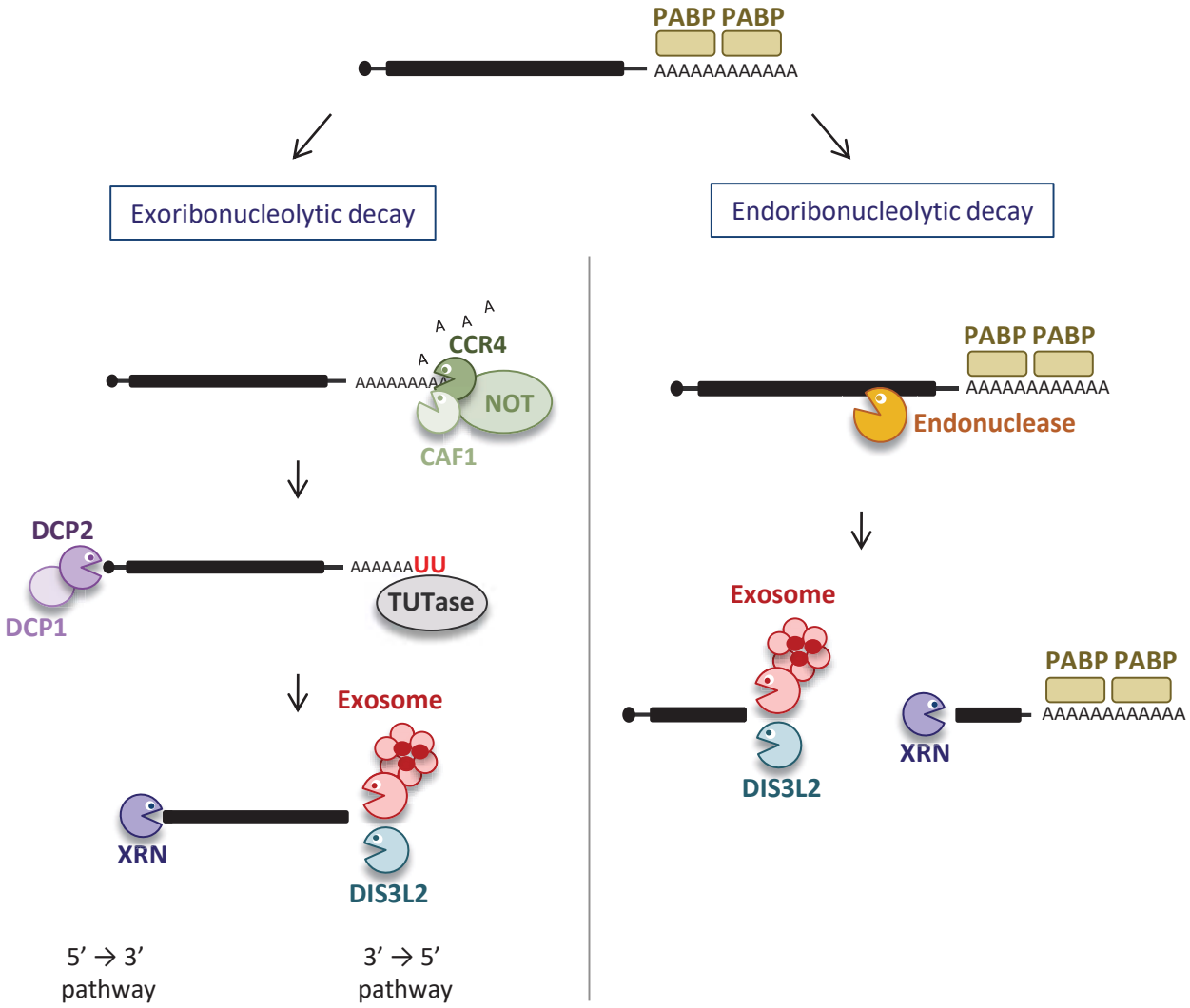


Figure 2: General mRNA decay. An mRNA targeted for decay can be subjected either to exoribonucleolytic or endoribonucleolytic decay. The exoribonucleolytic decay involves the DECAPPING 1 and DECAPPING 2 factors (DCP1 and DCP2) as well as the major deadenylation complex CARBON CATABOLITE REPRESSOR 4/NOT (CCR4/NOT). Deadenylated mRNA can be uridylated, which favors 5'-3 and 3'-5' mRNA decay. The removal of the cap induces the action of exoribonucleases XRN while the trimming of the poly(A)-tail recruits the exosome complex or DIS3L2 to further degrade mRNA. The endoribonucleolytic decay is initiated by an endonuclease cut, followed by exosome, DIS3L2 and XRN activities.

involving only one of these two pathways have mild or no effects on cell viability, while the mutation of both 5'-3' and 3'-5' pathways is lethal in both yeast and plants (Schneider et al., 2009; Zhang et al., 2015). These results demonstrate the functional redundancy between mRNA decay pathways (Houseley and Tollervey, 2009; Zhang et al., 2015). This feature highlights the essential function of bulk mRNA degradation. In addition to this housekeeping role, mRNA degradation can also be triggered on specific mRNA targets to regulate their stability and expression.

In the following section, I will first outline the mRNA degradation pathways responsible for the 3' to the 5' degradation including deadenylation of the poly(A)-tail, uridylation and action of 3' to 5' exoribonucleases. Next, I will describe the decay occurring from 5' to 3' starting with the removal of the cap and the subsequent decay by 5' to 3' exoribonucleases. Finally, I will describe the RNA granules called Processing-bodies (P-bodies) and their proposed roles in mRNA degradation and translational repression.

1. Degradation pathway from 3' to 5'

1.1. Deadenylation

The 3' extremity of mRNAs are protected by a poly(A)-tail, which is bound by members of the RNA binding protein family called POLY(A) BINDING PROTEIN (PABP). Deadenylation, the first step of mRNA degradation, corresponds to the shortening of the poly(A)-tail, and, results in the removal of the bound PABP factors (Figure 2). This initial step, which is considered in most cases as the main limiting step of mRNA degradation, is followed by the recruitment of other RNA degradation factors. Based on many observations, it has been proposed that different decay rates of individual mRNAs largely result from differences in the rate of deadenylation (Brogna et al., 2008), identifying deadenylation as the key step of mRNA degradation. The universality of this statement has to be moderated as deadenylation-independent RNA decay pathways have also been described, such as yeast deadenylation-independent decapping, histone mRNA decay, or the degradation of mRNA following endonucleolytic cleavage (Gera and Baker, 1998; Łabno et al., 2016; Marzluff and Koreski, 2017). As the poly(A)-tail and cap structures are necessary for mRNA translation, it has also been shown that deadenylation is connected to the translation status of the transcript: untranslated mRNAs are more exposed to deadenylation indicating that translation protects mRNAs from this process (Schwartz and Parker, 1999).

Three different complexes are responsible for the shortening of the poly(A)-tail. The POLY(A) NUCLEASE 2/POLY(A) NUCLEASE 3 (PAN2/PAN3) complex is thought to be responsible for initial poly(A) trimming to subsequently allow the CARBON CATABOLITE REPRESSOR 4/NOT (CCR4/NOT) complex, the major

deadenylation complex in eukaryotes, to complete deadenylation (Yamashita et al., 2005; Łabno et al., 2016). As no plant PAN2/PAN3 homolog has been identified, deadenylation in plants was proposed to be performed by the CCR4/NOT complex. Finally, the POLY(A) RIBONUCLEASE (PARN) deadenylase was proposed to regulate plant embryogenesis (Reverdatto et al., 2004; Łabno et al., 2016). PARN has also been shown to localize in mitochondria and to play a role in the regulation of mitochondrial mRNAs poly(A) tract (Hirayama et al., 2013).

Deadenylated mRNA 3' tails are able to bind the RNA binding protein SM-like LSM1-7 complex that recruits and activates the decapping complex (see part 2.1) or can be the substrate for the exosome machinery (Łabno et al., 2016) (see part 1.3). Finally, it has also been shown that the shortening of the poly(A) tail is a reversible process: mRNA being deadenylated can be re-adenylated and PABP proteins can re-associate to the poly(A) tail, increasing mRNA stability (Chen and Shyu, 2011).

1.2. mRNA uridylation

mRNA uridylation appears to be a common mechanism between *Schizosaccharomyces pombe*, *Aspergilus nidulans* and human cells to elicit 5'-3' and 3'-5' decay (Scheer et al., 2016; Morozov et al., 2010; Rissland and Norbury, 2009; Chang et al., 2014b). This terminal addition of a small number of uridines is proposed to facilitate the binding of the LSM1-7 complex, which has greater affinity for poly(U) tract over poly(A) tract (Chowdhury et al., 2007). IUridylated transcripts are then degraded by the exosome complex or by DIS3L2 in the 3'-5' direction (Mullen and Marzluff, 2008; Malecki et al., 2013). Interestingly, in *Arabidopsis thaliana*, it has been shown that mRNA uridylation by the URIDYLYLTRANSFERASE 1 (URT1) is responsible for the restoration of a poly(A) size allowing the binding of PABP (Zuber et al., 2016). This control of the extent of deadenylation by URT1 favors the 5'-3' polarity of degradation, which could be important for mRNAs still engaged on polysomes, for example (Zuber et al., 2016).

1.3. 3'-5' exoribonucleolytic decay

After the initial deadenylation step or after an additional uridylation step, mRNA degradation can occur from 3' to 5' either through the action of the cytoplasmic exosome or by DIS3L2/SOV.

The exosome is a very conserved multi-subunit complex: a core EXO9 comprising nine factors is associated with specific cofactors depending on its subcellular localization. The exosome carries an exonuclease activity either by one of its cofactors, such as Dis3p/DIS3 or RRP6 in yeast, mammals and plants (Kumakura et al., 2013; Chlebowski et al., 2013; Lange et al., 2008; Tomecki et al., 2010) or by one of its core subunits in plants, RRP41 (Sikorska et al., 2017, in press). It is known that exosome target specificity is conferred

by its cofactors, which are localized in the cytoplasm and/or the nucleus. The exosome is responsible for the cytosolic turnover of non-aberrant mRNAs, the maturation and decay of numerous non-coding RNAs, such as ribosomal RNA (rRNA) and small nuclear RNA (snoRNA), as well as the elimination of pervasive RNAs (Houseley and Tollervey, 2009; Schaeffer et al., 2009; Schneider et al., 2012; Chlebowski et al., 2013).

DIS3L2 in human and yeast and its plant homolog SOV were shown to be involved in the 3'-5' degradation of mRNA in the cytoplasm, independently on the exosome (Malecki et al., 2013; Lubas et al., 2013; Zhang et al., 2010). DIS3L2 preferentially degrades transcripts that are uridylated, as observed in *S. pombe* and in mouse embryonic stem cells (Malecki et al., 2013; Ustianenko et al., 2013).

2. Degradation pathway from 5' to 3'

2.1. Decapping

mRNAs are modified at their 5' end by a cap structure composed of a methylated guanosine (7-methylguanosine) required for translation and protecting transcripts from 5' to 3' degradation. Decapping (*i.e.* the removal of the cap) is usually stimulated after the initial step of deadenylation and irreversibly triggers mRNA degradation (Figure 2) (Łabno et al., 2016). The decapping core complex is composed of DECAPPING 1 (DCP1) and DECAPPING 2 (DCP2) proteins, which are highly conserved among eukaryotes. DCP2 is the catalytic subunit, whereas DCP1 interacts with DCP2 and stimulates its activity (Chang et al., 2014a). Several other proteins have been shown to enhance decapping, such as ENHANCER OF DECAPPING 3 (EDC3), SUPPRESSOR OF CLATHRIN DEFICIENCY (SCD6) (both are homologs to plant DECAPPING 5 (DCP5)), EDC4/Ge-1 (homolog to plant VARICOSE (VCS)), the LSM1-7 complex and the Dhh1/DDX6 homologs (Fromm et al., 2012; Nissan et al., 2010). Some of these activators, such as DCP1 and EDC4 directly stimulate DCP2 activity (Franks and Lykke-andersen, 2008). On the contrary, other activators, such as EDC3 and Dhh1 have been proposed to promote decapping as a secondary effect of their key function in translation inhibition (Franks and Lykke-andersen, 2008; Sweet et al., 2012).

2.2. 5' to 3' exoribonucleolytic decay

Following mRNA decapping, the action of a crucial family of 5' to 3' exoribonucleases (XRN), represents a major step of RNA decay (Figure 2). In mammals, XRN1 is known to play a key role in bulk mRNA decay (Łabno et al., 2016). In plants, several XRN proteins are described and localize to different compartments: XRN4 localizes in the cytoplasm, while there are indications for XRN2 and XRN3 localization in the nucleus (Kastenmayer and Green, 2000). XRN4 degrades all cytoplasmic mRNAs targeted for decay, but also cleavage products such as miRNA-directed cleaved transcripts for example (Souret et al., 2004). Both

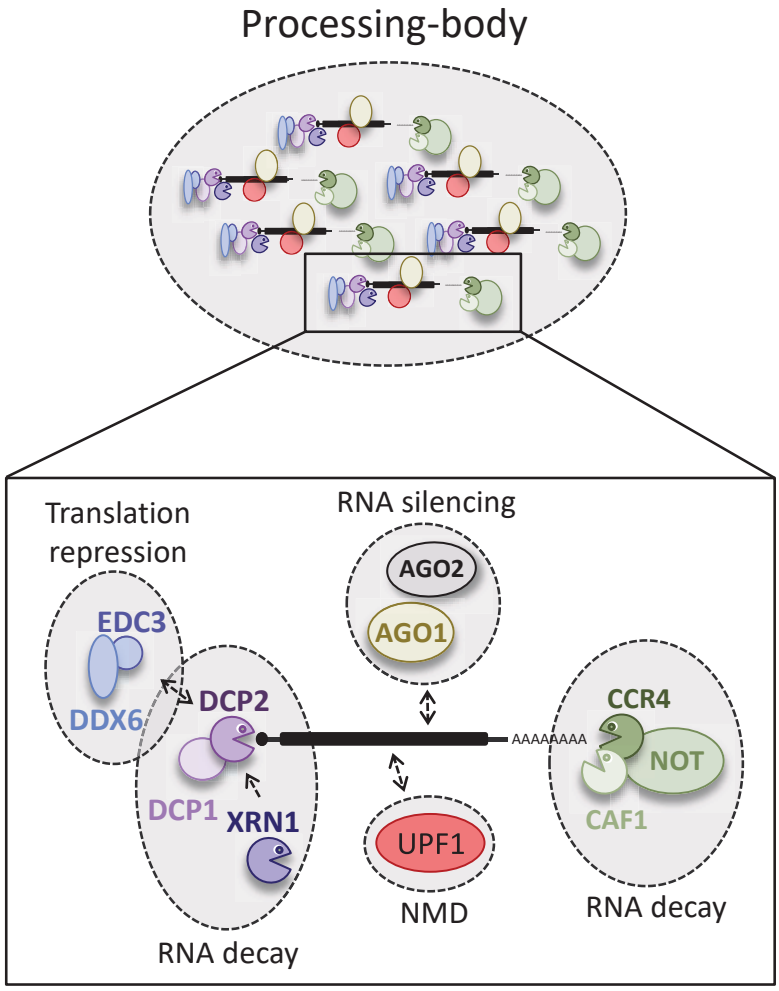


Figure 3: Mammals processing-bodies composition. Processing-bodies (P-bodies) are composed of aggregates of ribonucleoprotein (mRNP) complexes. These mRNP are composed of RNA decay factors such as decapping factors DCP1 and DCP2, exoribonuclease XRN1 and the CCR4/NOT deadenylation complex. DDX6 and EDC3 translation repressors are involved in P-bodies composition as well as the RNA silencing factors AGO1 and AGO2 and the Nonsense-Mediated Decay factor UPF1

XRN2 and XRN3 have a role in pre-ribosomal RNA processing (Zakrzewska-Placzek et al., 2010). XRN2, XRN3 and XRN4 proteins have been shown to act as endogenous RNA silencing suppressors in plants (Gazzani et al., 2004; Gy et al., 2007). Translating transcripts were first thought to be immune to XRN degradation, but recent studies have shown that mRNA decay can also occur co-translationally (Hu et al., 2009; Pelechano et al., 2015; Yu et al., 2016).

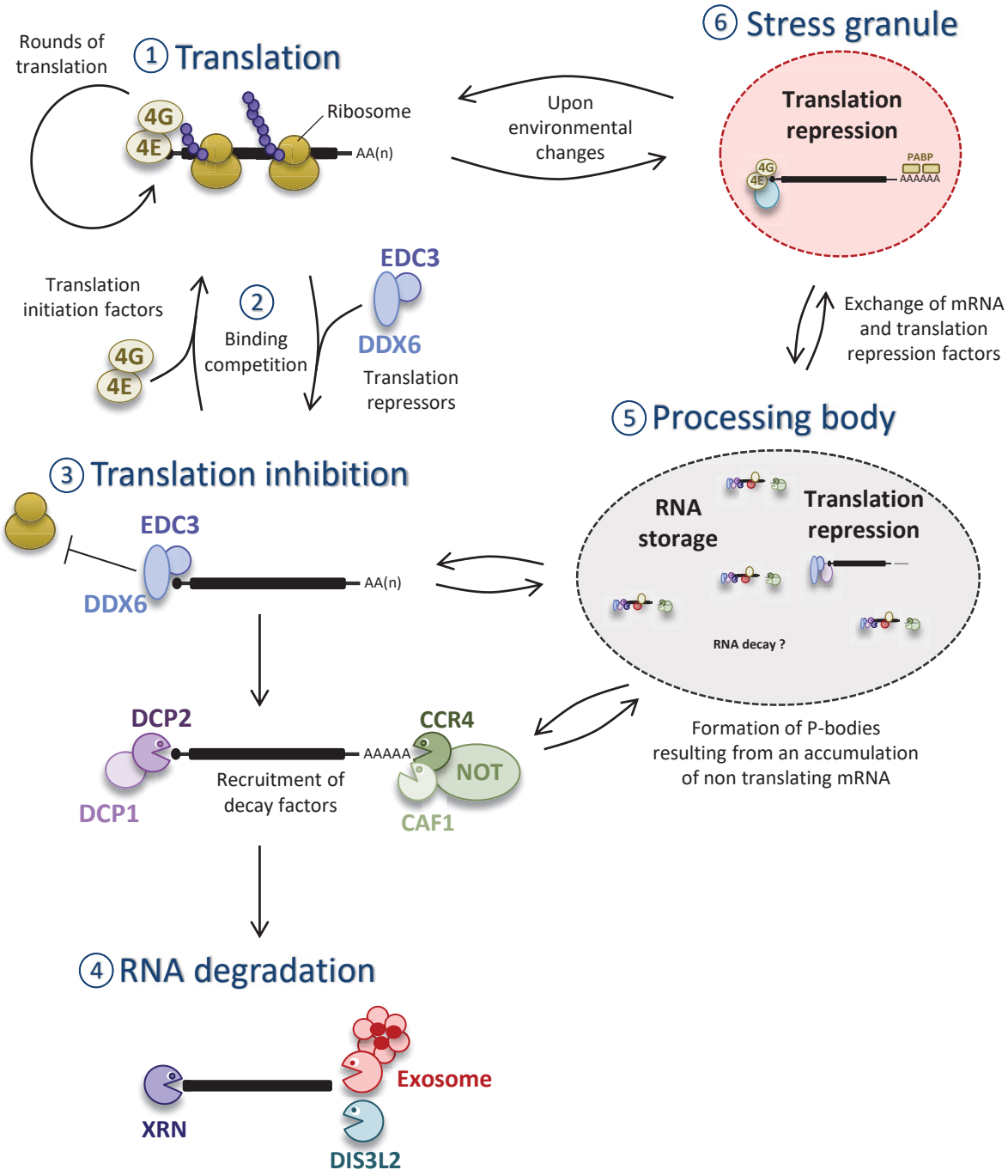
3. Endoribonucleolytic decay

Endoribonucleolytic decay is initiated by the internal cleavage of an mRNA by an endonuclease (Figure 2). An example of well characterized endonucleases is the family of Argonaute (AGO) proteins, which are involved in RNA silencing in mammals, plants and *Drosophila* (Meister, 2013). Rrp44/Dis3 is a PIN-domain endonuclease, which provides endoribonucleolytic activity to the exosome (Houseley and Tollervey, 2009). This activity was shown to have a role in RNA processing rather than mRNA decay (Schneider et al., 2009; Schaeffer et al., 2009). On the contrary, SMG6 is another PIN-domain endonuclease, involved in the specific decay of Nonsense-Mediated Decay targets in *Drosophila*, worms and mammals (Kervestin and Jacobson, 2012; Boehm et al., 2014; Glavan et al., 2006). 5' and 3' cleavage products resulting from endoribonucleolytic cleavage are then degraded by the exosome complex and by XRN1 respectively (Figure 2) (Łabno et al., 2016).

4. Processing-bodies, at the crossroad between mRNA decay, storage and translation

4.1. Processing-body composition

Processing-bodies (P-bodies) are small, dynamic and membrane-devoid structures present in the cytoplasm. These cytoplasmic foci are often detectable in unstressed cells, but are induced in stress conditions leading to translation inhibition (Teixeira et al., 2005). P-bodies contain messenger Ribonucleoproteins (mRNPs) formed by non-translating RNAs and many factors linked to RNA metabolism and decay. P-body components include decapping factors (DCP1, DCP2, EDC3/DCP5, Ge-1/VCS), deadenylases (PARN, CCR4), exoribonucleases (XRN1/XRN4), RNA silencing factors (AGO1, AGO2), proteins of the Nonsense-Mediated Decay (UPF1, SMG7) and many RNA binding proteins (Figure 3) (Sheth and Parker, 2003; Kedersha et al., 2005; Brogna et al., 2008; Maldonado-Bonilla, 2014). The targets of different RNA degradation pathways accumulate in P-bodies, such as miRNA-targeted transcripts, NMD targets or AU-rich elements-containing transcripts (Sheth and Parker, 2006 ; Kedersha, 2005 ; Liu, 2005). Because of this high local concentration of RNA decay factors and mRNA decay intermediates, P-bodies



Inspired from Decker et al., 2012

Figure 4: P-bodies at the crossroad of mRNA decay, storage and translation. During mRNA translation, a competition operates between translation initiation factors and translation repressors. Translation stops upon the binding of translation repressors DDX6 and EDC3. The mRNA is then either degraded or relocalized in processing bodies to repress its translation. mRNA within P-bodies can shift in and out of P-bodies and can be transferred to Stress Granules or can return to translation when translation is restored.

have been for many years proposed to be the major subcellular compartment involved in mRNA decay (Sheth and Parker, 2003; Cougot et al., 2004).

4.2. Processing-body functions

The observation that the inhibition of decapping or exoribonuclease activities results in the increase of P-body size and number in yeast was an argument suggesting that P-bodies are the site of decapping and 5' to 3' decay (Figure 4) (Sheth and Parker, 2003). Nevertheless, no proof of mRNA decay occurring within P-bodies has ever been described.

On the contrary, several reports have recently shown, using diversified methods, that P-bodies are not a site of mRNA decay. First, the demonstration of the existence of co-translational decay, has shown that mRNA degradation process can occur in other cytoplasmic compartments challenging the view that P-bodies would represent the only location of mRNA decay (Hu et al., 2009; Pelechano et al., 2015; Yu et al., 2016). In addition, imaging studies of single RNA molecule decay could show that RNA decay occurs dispersedly in the cytoplasm and not in P-bodies (Tutucci et al., 2017; Horvathova et al., 2017). Moreover, it has been shown that P-bodies are not mandatory for decay: bulk mRNA decay, NMD and RNA-mediated gene silencing pathways are functional in cells lacking P-bodies (Decker et al., 2007; Eulalio et al., 2007; Stalder and Mühlemann, 2009). Other functions were proposed for P-bodies: these mRNP aggregates could, for example, be generated following a temporarily high load on decay pathways (Stalder and Mühlemann, 2009). Indeed, this mechanism could prevent decay intermediates from accumulating in the cytoplasm, avoiding potential saturation of the translation machinery or the titration of limiting RNA binding proteins (Eulalio et al., 2007). Consistent with this idea, liquid phase transition has been suggested to define cellular granules assembly such as P-bodies (Weber and Brangwynne, 2012). Liquid phases are characterized by several multivalent low interactions between RNA molecules and proteins, such as arginine/glycine-rich RNA binding proteins for P-bodies (Banani et al., 2016; Alberti, 2017). A recent study showed that this organization is able to repress endoribonucleolytic cleavage and to reduce decapping activity of mRNA that are localized within P-bodies, suggesting that P-bodies are a site of mRNA storage, and not of decay (Schütz et al., 2017).

The function of P-bodies seems also to be linked to translational repression, since they are devoid of translation initiation factors and mRNAs present in P-bodies are in a non-translating state (Figure 4) (Teixeira et al., 2005). P-bodies also accumulate the known translation repressors Dhh1/DDX6, Scd6 and EDC3 (Parker and Sheth, 2007; Eulalio et al., 2007; Maldonado-Bonilla, 2014). Translation repressors are important for P-body formation in human and yeast as the lack of Dhh1/DDX6, or DCP5 in plants leads to

P-body formation defects (Coller and Parker, 2005; Ayache et al., 2015; Xu and Chua, 2009). Moreover, translation inhibition mediated by miRNAs also takes place in P-bodies: both miRNA and the targeted RNA are localized in P-bodies (Liu et al., 2005). In a recent report, Hubstenberger et al., 2017 showed using P-body purification experiments that mRNAs encoding regulatory processes are targeted to P-bodies, where they are translationally repressed (Figure 4) (Hubstenberger et al., 2017).

4.3. The central role of the P-body factor Dhh1/DDX6 in translation repression

Dhh1 is the founding member of the RNA helicase family homolog to the human DDX6 protein. It was first identified as a decapping activator in yeast as both mRNA decapping and P-body integrity are affected in a *dhh1* mutant (Coller and Parker, 2005). Several studies have shown that Dhh1/DDX6 is a key factor of translation repression and P-body formation (Franks and Lykke-andersen, 2008; Ayache et al., 2015). Strikingly, the overexpression of Dhh1 induces the loss of mRNA from polysomes and the accumulation of Dhh1 in P-bodies (Presnyak and Coller, 2013b). More recently, it has been proposed that the multimeric assembly of Dhh1 bound to mRNA is important for P-body formation (Mugler et al., 2016). It has also been shown that Dhh1 targets the majority of cytoplasmic mRNA and relocates those translationally repressed mRNA in P-bodies (Coller and Parker, 2005). The tethering of Xenopus Xp54, the Dhh1 homolog, leads to the specific translation repression of capped transcripts but not of transcripts harboring an Internal Ribosome Entry Site (IRES) (Cooke et al., 2010). This observation demonstrates the tight link between decapping and translation repression. Consistent with this idea, Dhh1/DDX6 have been shown to directly interact with another translation repressor, decapping activator and component of P-bodies, Edc3 in both yeast and mammals (Decker et al., 2007; Tritschler et al., 2009). It has been shown that the tethering of Dhh1 resulted in translation repression followed by mRNA decay in yeast (Carroll et al., 2007), but also in mammals through the interaction between DDX6 and the CCR4/NOT1 complex (Ozgur et al., 2015). Nevertheless, it was also shown that miRNA-mediated translation inhibition and decay occurs without detectable P-bodies and that location of miRNA in P-bodies is a consequence of translation repression mediated by DDX6 (Chu and Rana, 2006; Eulalio et al., 2007).

4.4. Other key factors involved in translation repression

In addition to the key role of Dhh1, many other factors have an important role in translation repression. Scd6, one of the two yeast homologs of plant DCP5, has been shown to repress translation upon its binding to the Carboxy-terminal (C-ter) domain of the translation initiator factor eIF4G, through a domain of Scd6 containing RGG motifs (Rajyaguru et al., 2012). This interaction is performed within an intact complex composed of the translation initiation factors eIF4E, eIF4A, eIF4G and PAB1, allowing a rapid return to

translation upon changing conditions and the dissociation of only one factor, Scd6 (Figure 4) (Rajyaguru and Parker, 2012).

Ded1, the yeast homolog of human DDX3 has been shown to form an mRNP complex stalled in translation. Ded1 also possesses an RGG domain and interacts with eIF4G (Rajyaguru et al., 2012). Ded1 and Scd6 could have the same mode of action while allowing the regulation of different subsets of mRNA targets through sequence-specific RNA binding (Rajyaguru et al., 2012). Ded1 localization is unclear; it has been shown to localize either in Stress Granules or in P-bodies, but it seems that Ded1 has a function in P-bodies as its depletion results in P-body defects (Beckham et al., 2008).

4.5. Processing-bodies are highly dynamic structures

P-bodies are mobile cytoplasmic foci, but the role of this movement within the cell is not understood. P-body movement is dependent on the cytoskeleton, microtubules in mammals and acto-myosin in *Arabidopsis,* in which a direct interaction between DCP1 and different myosins has been shown (Aizer et al., 2008; Loschi et al., 2009; Hamada et al., 2017; Steffens et al., 2014). P-body movement appears to be regulated and two types of movements have been observed: either fast and directional or slow and anisotropic (Steffens et al., 2014). While moving, P-bodies have also been observed to enter in contact with other RNA granules, including Stress Granules. Stress Granules are composed of mRNA in translation initiation arrest and share some components with P-bodies (Kedersha et al., 2005). P-body factors and mRNAs present in P-bodies can shift in and out or can be transferred to other compartments like Stress Granules (Figure 4) (Kedersha et al., 2005). Finally, in yeast, it has been shown that mRNAs accumulating in P-bodies when translation initiation is blocked can exit P-bodies and return to translation (Brengues et al., 2007). All these reports identify P-bodies as a key structure to fine-tune the balance between mRNA decay, storage and polysomal engagement.

II. Nonsense-mediated decay

The quality control of translation

One important function of mRNA decay is to remove aberrant RNA from the cell to prevent possible deleterious effects. Aberrant mRNA containing a defect could give rise to aberrant proteins with abnormal functions. Moreover, even aberrant RNAs that will not give rise to protein could also have negative effects for the cell: they could saturate important cellular processes such as translation, or titrate limiting wide spectra RNA binding proteins.

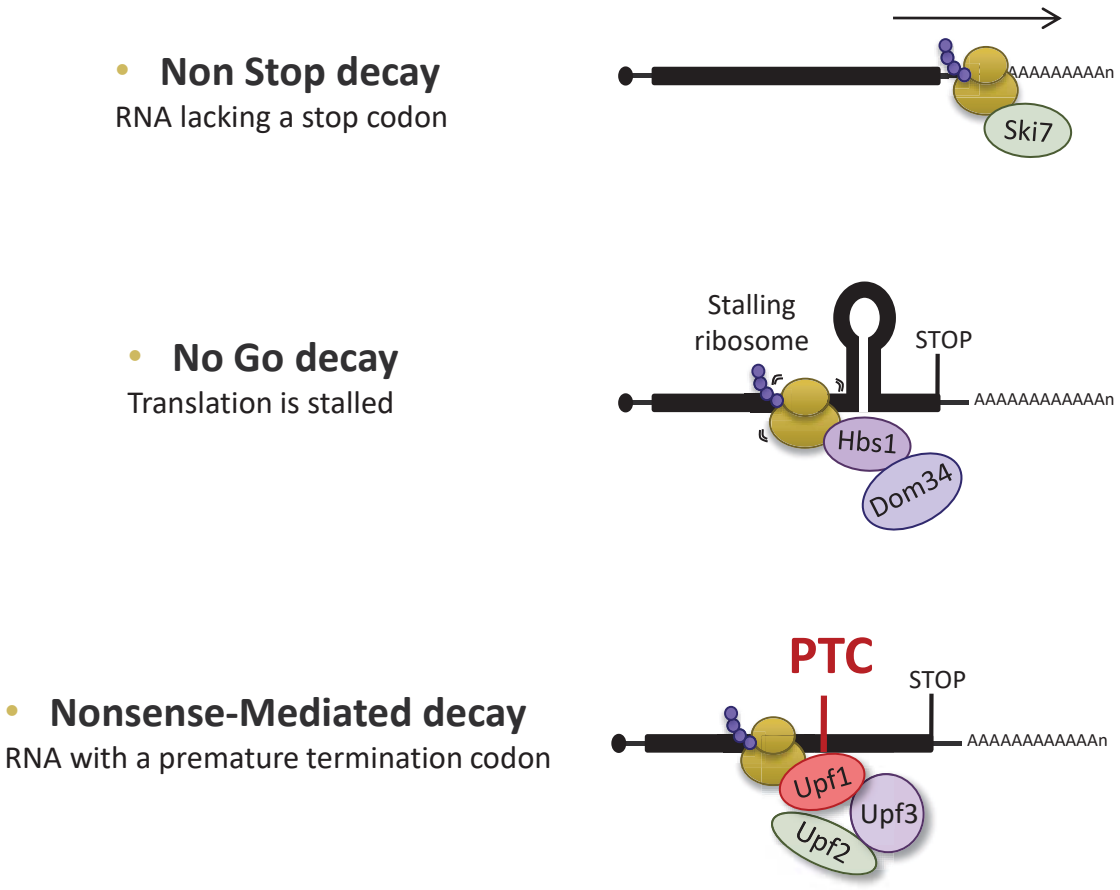


Figure 5: mRNA quality control of translation. Three different mRNA surveillance pathways recognize and accelerate the decay of transcripts showing aberrant translation. The Non-Stop Decay (NSD) recognizes mRNA lacking a termination codon, it results in ribosome translating the poly(A)-tail. Ski7 recognizes this feature and recruits the exosome for decay. The No-Go Decay (NGD) targets transcripts for which the ribosome is stalling due to an RNA secondary structure for example. The NGD involves the action of the Dom34 and Hsb1 factors to initiate mRNA decay. The Nonsense-Mediated Decay (NMD) targets mRNA possessing a premature termination codon (PTC). NMD target recognition is performed by the UPF complex which induces a cascade of reactions allowing its degradation.

There are three different mRNA surveillance pathways able to recognize and accelerate the decay of transcripts showing signs of aberrant translation. These surveillance pathways are Non-Stop Decay (NSD), No-Go Decay (NGD) and Nonsense-Mediated Decay (NMD) (Figure 5). The latter is the best characterized translation-mediated surveillance pathway. Non-Stop Decay targets RNAs lacking an in-frame termination codon which leads the ribosome to translate the poly(A) tail (Figure 5). No-Go Decay recognizes stalling translation, caused for example by a strong secondary structure such as a stable stem loop, pseudoknots or damaged RNA bases, that inhibit further ribosome translocation (Figure 5). The distinction between NSD and NGD has become blurred, as some targets of NSD are also targeted by the NGD pathway because the translation of the poly(A) tail into polylysine can also cause ribosome stalling (Shoemaker and Green, 2012). Moreover, NSD and NGD share some common mechanistic features: the Dom34 and Hsb1 factors interact with the ribosome and allow ribosome recycling in both NSD and NGD. NGD also involves an endoribonucleolytic cleavage triggered by an unknown nuclease (Shoemaker and Green, 2012). During yeast NSD, the Ski7 factor, which is structurally related to Hbs1, recruits the exosome to stimulate decay. A recent report showed that a splicing variant of HBS1 in human produce a Ski7 analogue acting as a molecular link with the exosome to induce mRNA decay (Kalisiak et al., 2017).

Nonsense-Mediated Decay allows the recognition and the subsequent decay of mRNAs containing a premature termination codon (PTC), *i.e.* a termination codon upstream of the normal termination codon. Those mRNAs could give rise to truncated proteins some of these with a dominant negative effect, altering cell fitness or inducing diseases (Chang and Kan, 1979; Kervestin and Jacobson, 2012; Miller and Pearce, 2014). NMD acts via the recruitment of factors stimulating deadenylation, decapping and exonucleolytic activities in yeast and mammals (Muhlrad and Parker, 1994; Mitchell and Tollervey, 2003; Lejeune et al., 2003). NMD target recognition occurs upon binding of the conserved NMD core complex, which is composed of the three Up Frameshift proteins UPF1, UPF2 and UPF3. In addition to mRNA decay, it has also been shown that NMD activation can lead to an initial step of translation repression in both yeast and mammals (Muhlrad and Parker, 1999; Isken et al., 2008). This step happens upon the interaction between UPF1 and the translation initiation factor eIF3 (Muhlrad and Parker, 1999; Isken et al., 2008). It was shown that apart from its mRNA surveillance role, NMD also regulates the expression of functional genes with specific NMD stimulating features.

In the following section I will focus on NMD functions, targets and mode of action.

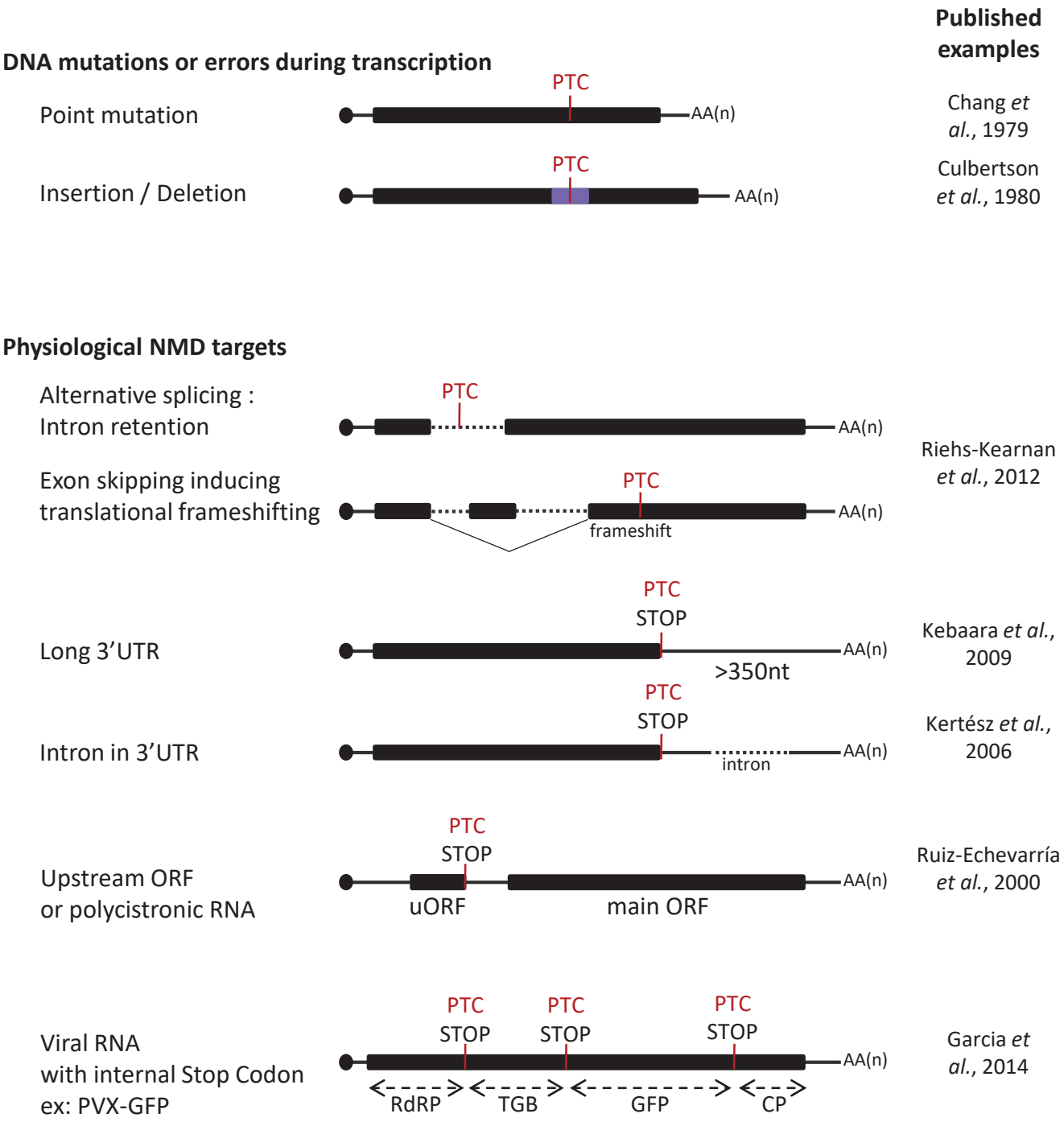


Figure 6: NMD-inducing features. NMD targets aberrant mRNA possessing a PTC following to a point mutation or insertion/deletion. NMD also targets physiological mRNA such as spliced isoforms, mRNA possessing a long 3'-UTR, a short upstream ORF, intron in its 3'-UTR or viral RNA for which normal termination codons are recognized as PTC.

1. Biological roles of NMD

NMD depletion results in lethality in different organisms such as plants, insects or mammals, indicating that this mRNA surveillance pathway is essential (Riehs-Kearnan et al., 2012; Rehwinkel et al., 2005; Medghalchi et al., 2001). Nevertheless, the lethality of NMD mutants could be due either to the accumulation of aberrant mRNA and the corresponding aberrant truncated proteins or to the deregulation of functional genes. In plants, NMD represses the expression of defense genes, resulting in a constitutive defense response in NMD mutants, which is in turn largely responsible for the lethality of the knock-out (KO) *upf1-3* mutant (Rayson et al., 2012; Riehs-Kearnan et al., 2012) (see part 3.2 for more details). A similar concept is true in mammals, in which NMD mutant lethality is due to the misregulation of only one transcript: the GADD45 transcript, its degradation by NMD is therefore essential for viability (Nelson et al., 2016). NMD is also involved in salt stress response in plants (Vexler et al., 2016; Drechsel et al., 2013) or hypoxic stress in mouse and plants (Gardner, 2008; Lorenzo et al., 2017). Finally, diverse developmental processes are influenced by NMD: examples go from brain development in mammals to the control of flowering time in plants (Bruno et al., 2011; Nasim et al., 2017).

Despite these crucial functions, the molecular mechanisms of NMD are not yet completely solved. Different models have been proposed to explain how NMD targets are recognized and degraded.

2. NMD-inducing features

NMD targets contain special features able to elicit their recognition by UPF1 and their subsequent degradation (Figure 6). These NMD-eliciting features include nonsense mutations, the presence of an intron in a 3'UTR, an upstream ORF (uORF) or a long 3'UTR (longer than 350 nucleotides (nt)) (Kebaara and Atkin, 2009; Hurt et al., 2013). In addition, unspliced transcripts, bicistronic transcripts, transcripts that underwent frameshifting and transcripts of pseudogenes can also be targeted by NMD (He et al., 2003). Apart from protein coding mRNA, NMD also regulates a wide range of long non-coding RNAs, such as snoRNA and antisense RNA (Kurihara et al., 2009; Colombo et al., 2017).

Interestingly, not all mRNA with potential NMD-inducing features are targeted by NMD. Endogenous as well as exogenous mechanisms have been developed to protect specific RNAs from NMD (Ruiz-Echevarria and Peltz, 2000; Fatscher et al., 2014) (see part 4.2 and 5.1.2 for more details). Moreover, UPF1 can also target mRNA devoid of any specific feature, indicating that this list is not sufficient to discriminate NMD-sensitive targets from NMD-immune targets. A recent study has shown that some NMD-sensitive transcripts do not possess any NMD-triggering features but tend to have a lower translation efficiency, a

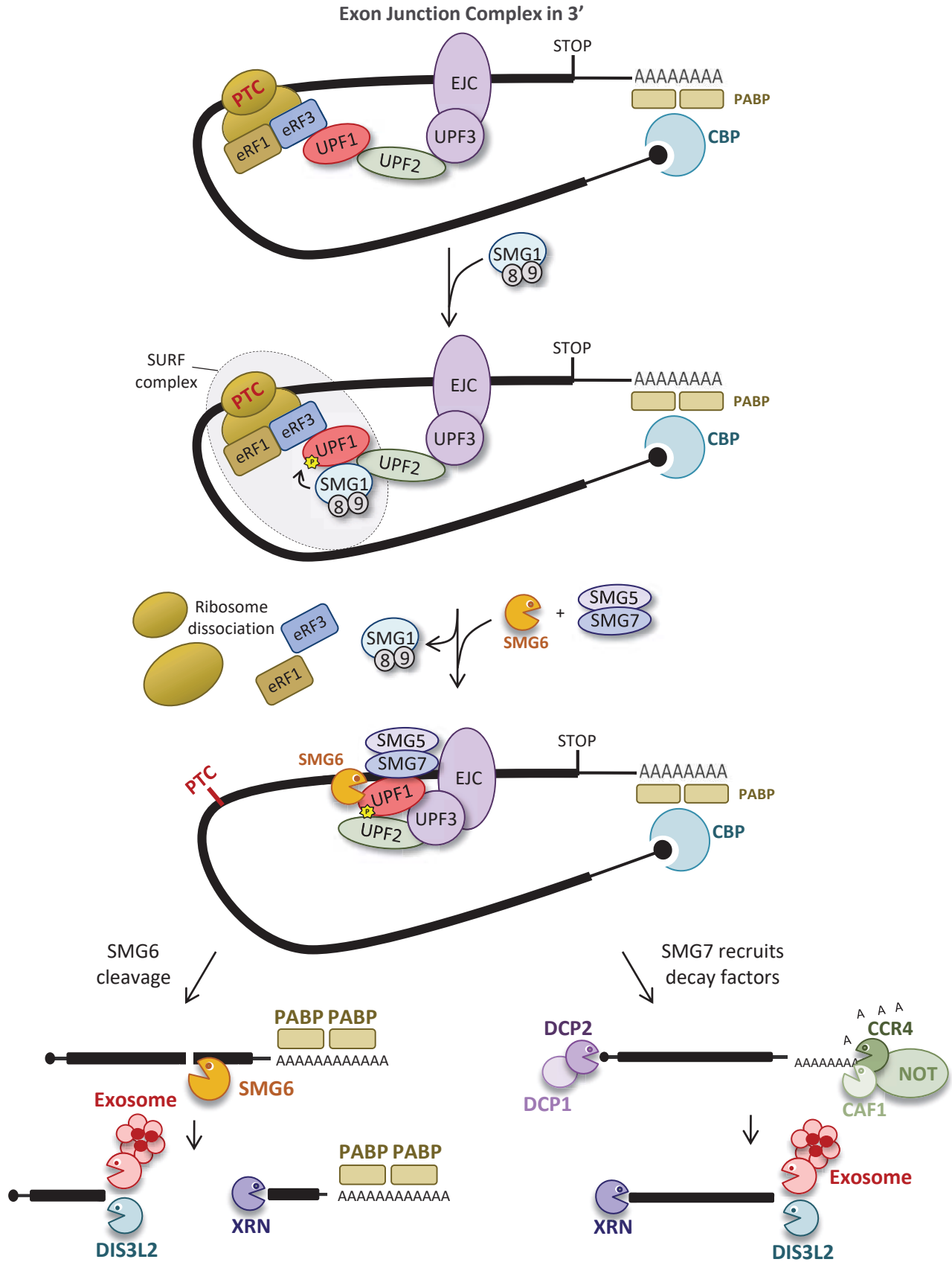


Figure 7: EJC-mediated NMD. The presence of the Exon Junction Complex (EJC) in 3' of the PTC allows the formation of the core NMD complex composed of the three UPF proteins. This complex recruits the SMG1 proteins, which phosphorylates UPF1. The formation of the SURF complex (SMG1-UPF-eRF1/3) allows conformation changes and ribosome dissociation. Phosphorylated UPF1 recruits either the endonuclease SMG6 or the SMG7-SMG5 complex. SMG6 is responsible for the cleavage of the mRNA target followed by exoribonuclease XRN and exosome decay. The SMG7-SMG5 complex recruits decapping and deadenylation factors to initiate decay of the NMD target. This mRNA is then completely degraded by XRN and exosome activities.

higher out-of-frame translation rate as well as lower average codon optimality (Celik et al., 2017a). Codon optimality has been shown to be a major determinant of mRNA stability: stable mRNAs have optimal codons while unstable mRNAs contain higher proportions of non-optimal codons, NMD could participate to this intrinsic instability (Presnyak et al., 2015). A model to reconcile these notions would be that low translation efficiency may lead to higher rate of translational errors, increasing the potentiality of premature translation termination and the induction of NMD (Celik et al., 2017b).

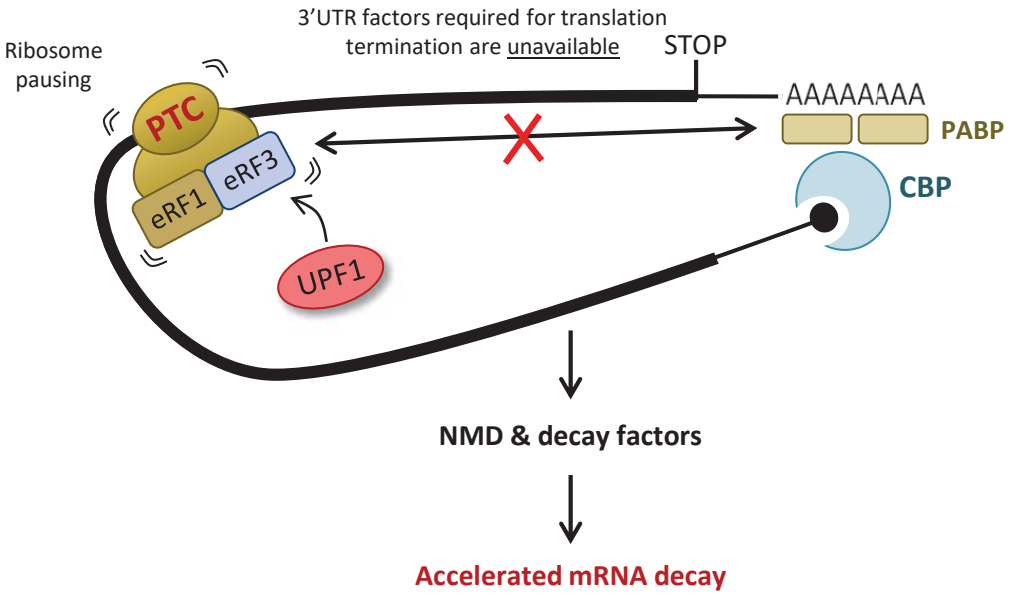
3. Mechanistic models of NMD

Several models to describe the NMD mechanisms have been proposed throughout the years, each one having its limits. I will first present the two major models: (i) the EJC-mediated NMD involving the Exon Junction Complex (EJC), which is the most documented NMD mechanism in mammals, and (ii) the faux 3'-UTR model, which is the predominant NMD model in yeast and *C. elegans*. In plants, both EJC-mediated NMD and the faux 3'-UTR models have been described (Kertész et al., 2006). In addition, I will present a more recent model, the ribosome release model, which was proposed to reconcile data conflicting with other NMD models (Brogna et al., 2016).

3.1. EJC-mediated NMD

EJC-mediated NMD, also known as the intron-based NMD, was first identified in mammals, where it is the NMD mechanism that has been investigated in most detail (Figure 7). The EJC is a complex present on exon-exon junctions after splicing. It is composed of multiple factors, among which the core factors Y14, Barentsz, Magoh and the eukaryotic initiation factor eIF4A-III. The core NMD factor UPF3 is also part of the EJC complex and constitutes a bridge between UPF1 and the EJC, thus allowing EJC-mediated NMD targets degradation (Figure 7) (Chamieh et al., 2008). The formation of the complex between the three UPF proteins allows the recruitment of SMG1 and its associated factors SMG8 and SMG9. Their binding on the UPF complex bound on the terminating ribosome constitutes the SURF (SMG1-UPF1-eRF3) complex, which enhances the SMG1-mediated phosphorylation of UPF1 necessary to recruit SMG6 and/or SMG7-SMG5 factors (Figure 7) (Kashima et al., 2006). Two redundant decay pathways are involved in NMD in human cells, one involving the SMG5-SMG7 complex and the other involving SMG6-mediated cleavage of the NMD target (Metze et al., 2013; Colombo et al., 2017). UPF1 interacts with SMG7 to initiate decapping-dependent RNA degradation (Mühlemann and Lykke-Andersen, 2010), but also to recruit the CCR4/NOT deadenylation complex in human cells (Loh et al., 2013). It has been shown that the artificial binding of SMG7 on any part of an mRNA leads to the degradation of this mRNA independently of UPF1, identifying SMG7 as the bridge between NMD and decay factors (Mérai et al., 2013; Unterholzner and

① Faux 3'-UTR model



② Long 3'-UTR model

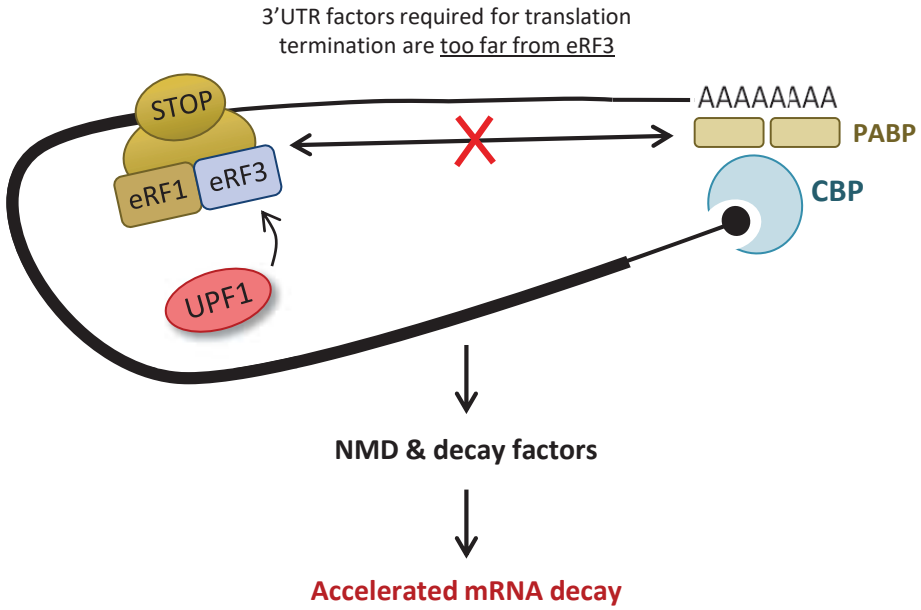


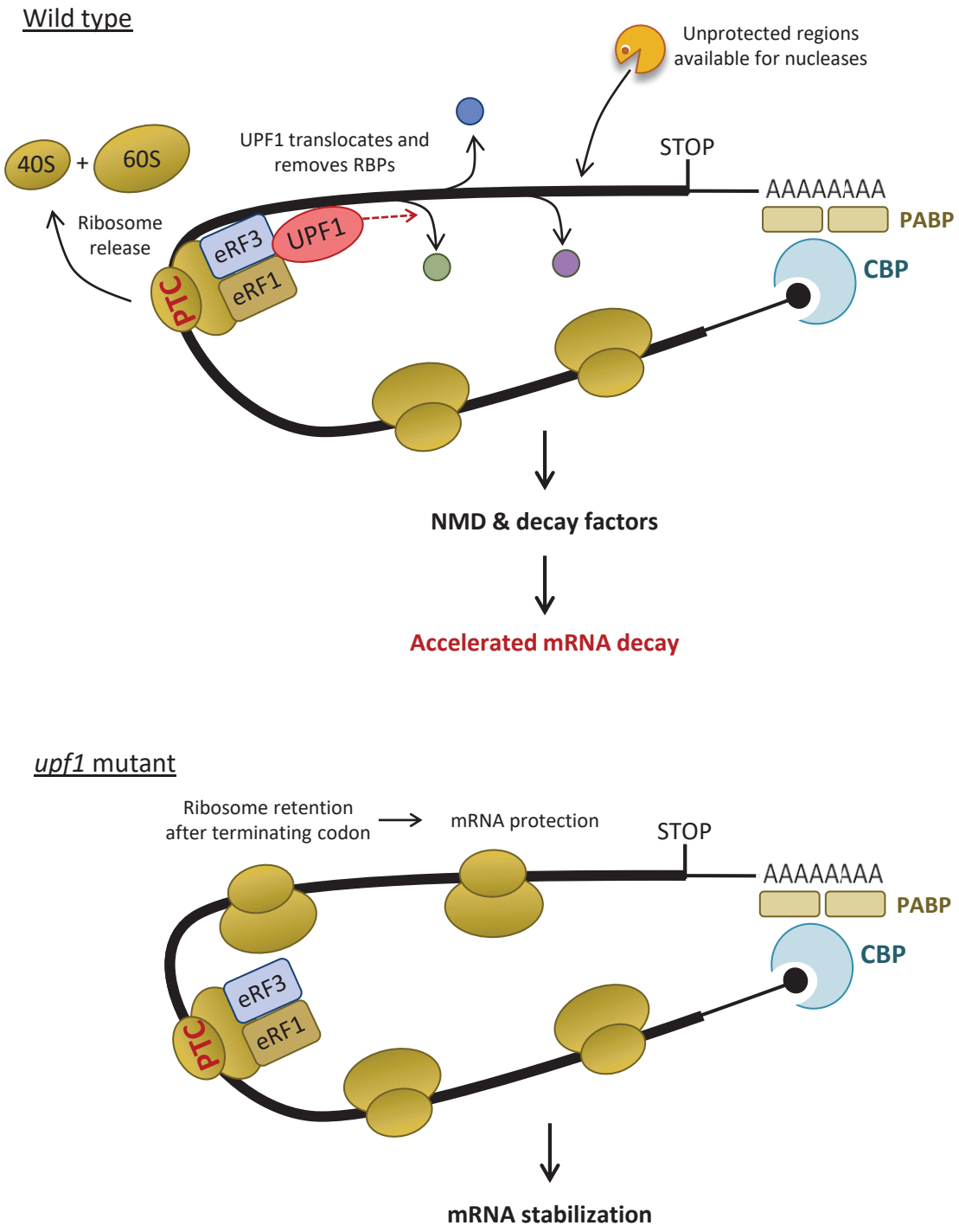
Figure 8: The NMD 3'-UTR model. The faux 3'-UTR model postulates that a non-favorable environment for translation termination, such as ribosome pausing and the absence of 3'UTR factors required for proper translation termination, induces the interaction between the release factor eRF3 and UPF1. This interaction induces a cascade of reactions involving NMD and decay factors to degrade the mRNA target. The long 3'-UTR model resemble to the faux 3'-UTR model as a long 3'-UTR prevents the association between PABP and eRF3.

Izaurralde, 2004). The second decay pathway in vertebrate and insect NMD involves SMG6, a PIN-domain endonuclease, which is responsible for the cleavage of all major classes of NMD substrates (Figure 7) (Boehm et al., 2014).

It has been shown that NMD is activated if a PTC is localized at least 50 nt upstream of an exon-exon junction (Nagy and Maquat, 1998). This mechanism exists also in plants where EJC-mediated NMD is position-dependent: an intron located 28 nt downstream of a stop codon fails to destabilize the mRNA, whereas the same intron located 99 nt downstream of the stop codon elicits NMD (Kertész et al., 2006). As a result, the length between the EJC and the stop codon plays an important role in the definition of transcript sensitivity to NMD. Plants orthologs of the core EJC proteins are conserved, such as Y14, Magoh, eIF4A-III and Barentsz. These are essential for the degradation of transcripts containing an intron downstream of the stop codon, but not for the degradation of transcripts containing a long 3'-UTR (Nyikó et al., 2013). In contrast to animals, the artificial binding of EJC components on mRNA in plants is not able to induce NMD-mediated degradation (Gehring et al., 2005; Kertész et al., 2006). Despite the clear enhancement of NMD by the EJC factors, several lines of evidence show that the EJC is not involved in all NMD events. Indeed, even in mammals some NMD targets do not possess any intron and are also degraded by NMD (Eberle et al., 2008; Singh et al., 2008). Moreover, no EJC factors are present in yeast where NMD is functional, while in *Drosophila*, EJC components are present but do not seem to be involved in NMD (Gatfield and Izaurralde, 2004). As a result, it is more likely that the EJC is an enhancer of NMD rather than a clear requisite for NMD degradation.

3.2. The faux 3'-UTR model

Another model proposed for NMD is the faux 3'-UTR model (Figure 8). It is well known that translation termination at a PTC is sub-optimal compared to normal translation termination, as ribosomes pause at a PTC yield toeprints, whereas ribosomes at normal stop codon do not (Peixeiro et al., 2012). In the faux 3'-UTR model, the lack of termination regulatory factors, that are normally present on a 3'-UTR and allow translation termination, leads to ribosome pausing and aberrant translation termination (Figure 8) (Amrani et al., 2004). It has also been shown that either the mRNP structure or some downstream elements (DSE) after the PTC are required to allow normal translation termination (Amrani et al., 2004). PABP has been proposed to be such a termination regulatory factor as the interaction between PABP and eukaryotic release factor 3 (eRF3), which is associated to the terminating ribosome, is necessary to allow normal translation termination (Ivanov et al., 2008). Nevertheless, if the interaction between the PABP and eRF3 is prevented by factors, complexes or by a long distance between them, translation termination



Inspired from Brogna et al., 2012

Figure 9: The ribosome release model. The ribosome release model postulates that the recognition of an NMD target by UPF1 is followed by the translocation of UPF1 on the mRNA and the removal of RNA Binding Protein (RBP) deposited on it. Unprotected regions are then available for digestion by nucleases. In a *upf1* mutant, ribosomes are not released after the PTC and accumulate on the target, resulting in mRNA protection from nucleases and mRNA stabilization.

is perturbed and NMD is activated (Figure 8). The faux 3'-UTR model is also applicable to EJC-containing NMD targets if complexes such as the EJC can prevent the interaction between termination regulatory factors and the terminating ribosome.

The faux 3'-UTR model is closely linked to another NMD model, the long 3'-UTR model, in which the length of the 3'-UTR promotes aberrant translation termination thus enhancing NMD (Figure 8) (Hogg and Goff, 2010). The long 3'-UTR model was initially proposed in both yeast and *C. elegans* but subsequent studies have shown that it is also valid in human cells, plants and *Drosophila* (Kertész et al., 2006; Hogg and Goff, 2010). It has been shown that rearrangement of the long 3'-UTR to bring the PABP closer to the PTC, for example via the introduction of intra molecular base pairing in the 3'-UTR results in NMD evasion (Eberle et al., 2008). These results indicate that the regulation of factors affecting the secondary structure of the 3'-UTR of an mRNA leads to the modulation of NMD-recognition, contributing to mRNA fate. A wide range of RNA binding proteins are known to bind 3'-UTRs (Yamashita and Takeuchi, 2017), the modulation of RNA binding protein expression could be a strategy for NMD-dependent fine-tuning of gene expression.

3.3. A new unifying NMD model: the ribosome release model

The exact mechanism of NMD is still not completely understood and a new NMD model, the ribosome release model, has been proposed to reconcile all the information gathered on NMD over the years (Brogna and Wen, 2009). The ribosome release model postulates that when a ribosome is stopped on a PTC, the downstream sequence is no longer protected by translocating ribosomes, the transcript becomes an NMD target and is digested by nucleases (Figure 9) (Brogna and Wen, 2009). It has been proposed that UPF1, thanks to its ATPase activity, is responsible for the clearance of RNA binding proteins bound to the transcript, which in turn allows digestion by nucleases (Figure 9) (Brogna et al., 2016). This hypothesis is comforted by a recent study using an *in vitro* system based on magnetic tweezers, that was able to test the biochemical properties of the UPF1 enzyme (Fiorini et al., 2015). They showed that UPF1 is able of both unwinding and slowly translocating along single and double stranded nucleic acids with a processivity of more than 10 kb, as well as to remodel the mRNP upon the removal of RNA bound proteins.

4. Special features of NMD in plants

It seems that plant NMD shares similarities with both invertebrate and vertebrates NMD, but it also possesses certain unique traits. Both the EJC-mediated NMD, extensively studied in mammals, and the long 3'UTR-mediated NMD, predominant in yeast, operate in plants (Kertész et al., 2006). Nevertheless, EJC components do not seem to be as important in plant as in mammals, since the tethering of EJC

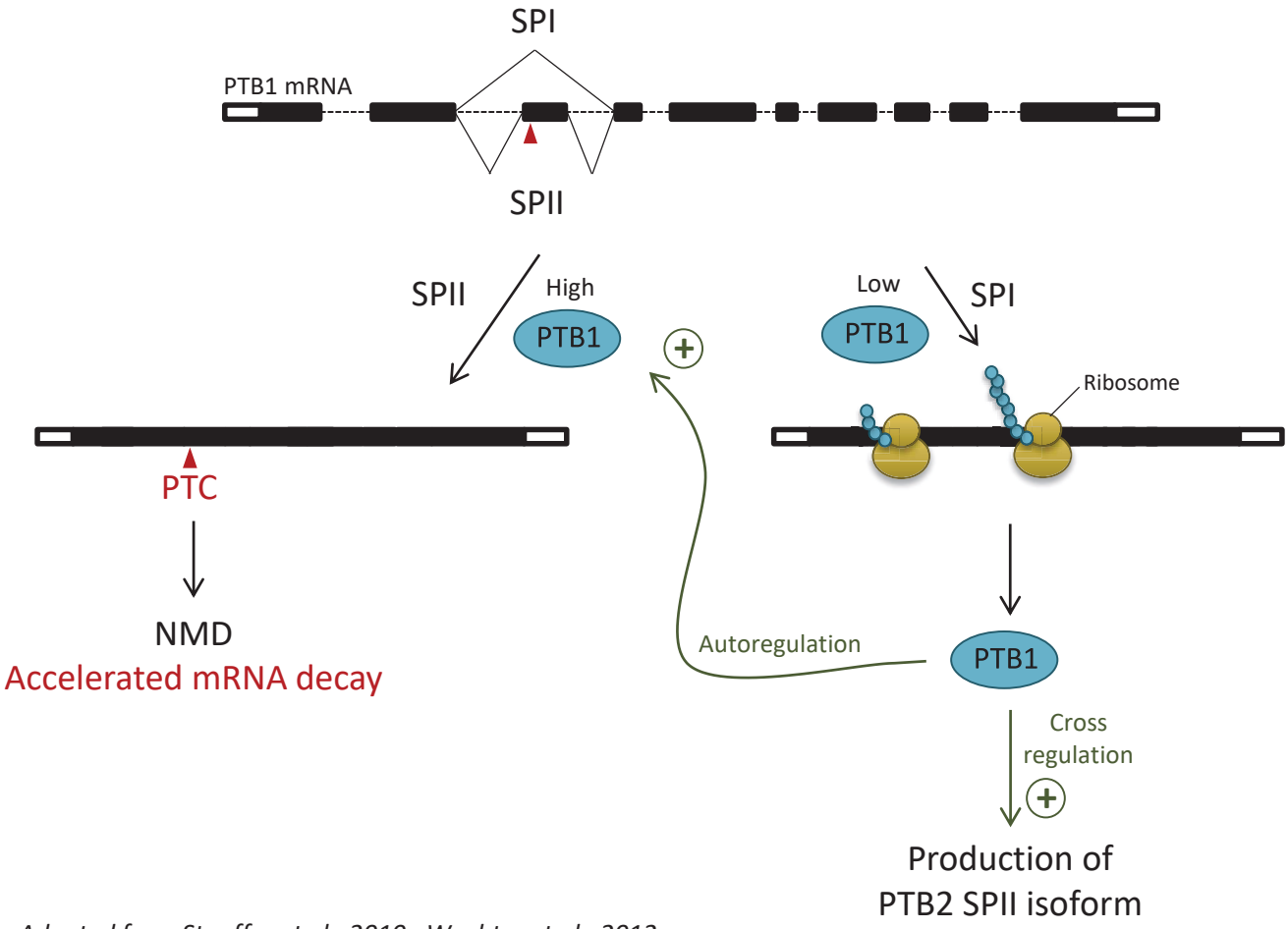
components on RNAs induces their degradation in mammals, but not in plants (Gehring et al., 2005; Kertész et al., 2006).

Some NMD factors have no homologs in *Arabidopsis*, including the SMG1 kinase and its associated proteins SMG9 and SMG8. However, SMG1 is present and active in NMD in other plant species, suggesting that *Arabidopsis thaliana* lost the SMG1 homolog within the last 5–10 million years (Lloyd and Davies, 2013). SMG1 is involved in UPF1 phosphorylation; this modification is necessary for UPF1 function in plants (Kerényi et al., 2013). The identification of *Arabidopsis* UPF1-phosphorylating protein would represent an important progress in understanding NMD in plants. Other NMD factors are absent in plants, including PNRC2, a protein that links NMD to the decapping machinery in mammals. The absolute necessity of PNRC2 in NMD has been discussed as it does not act on all NMD targets. It was proposed that PNRC2 is not specific to NMD but might be more generally linked to the decapping machinery (Loh et al., 2013). PNRC2 is also absent from yeast, and NMD operates in both yeast and plants, challenging the absolute requirement for PNRC2 in NMD (Loh et al., 2013). Finally, no endoribonucleolytic cleavage of NMD targets, and no homologs of the NMD-specific endonuclease SMG6 were identified in plants. More work is still needed to assess whether a similar activity is conserved in plants and, if so, to identify the possible nuclease that could substitute SMG6 in performing mRNA cleavage.

Finally, it has been observed that the only well described cytoplasmic 5'-3' exoribonuclease in plants – XRN4 – is not a limiting factor for NMD. Indeed, *Arabidopsis xrn4* mutant does not deregulate NMD targets accumulation, and XRN4 inactivation in *Nicotiana benthamiana* doesn't perturb NMD targets regulation (Mérai et al., 2013). The RNA decay pathways influencing NMD targets accumulation remain to be identified in plants. Factors responsible for NMD target degradation in plants, as well as their mode of action, have not yet been elucidated and may differ from other models.

5. NMD in the fine-tuning of gene expression

Regulation of gene expression is a very important process that allows the fine-tuning of protein production at different levels, in order to face changing conditions such as biotic or abiotic stresses or developmental changes for example. In addition to its role in the mRNA surveillance pathway and the decay of PTC-containing mRNA, NMD has also been shown to fine-tune the expression of functional genes. Indeed, it has been observed that the expression of up to 10 % of cellular transcripts is upregulated in NMD mutants in yeast and mammals (Kervestin and Jacobson, 2012).



Adapted from Stauffer et al., 2010 ; Wachter et al., 2012

Figure 10: Regulation of gene expression by AS-NMD: the example of plant PTB. Two isoforms of PTB mRNA are produced after splicing (SPI and SPII). Upon high PTB1 levels, PTB1 binds on PTB1 mRNA splicing sites and favors the production of SPII isoform. This isoform possesses a PTC and induces NMD recognition and accelerated decay. Upon low PTB1 level, the SPI isoform is produced and translated into PTB1 protein. PTB1 autoregulates its own expression as well as cross-regulates the expression of the PTB2 protein with the same strategy.

5.1. Alternative splicing coupled-NMD

Alternative splicing is a mean for the cell to produce different isoforms of mRNAs and therefore proteins, from a single transcription event. It participates in the proteome diversity expansion and exerts important functions in several biological processes, such as the regulation of circadian rhythms (Sanchez et al., 2011). Alternative splicing is tightly regulated by NMD, as it has been observed in plants but also in mammals that 30% to 40% of alternatively spliced transcripts possess putative NMD inducing features, such as a premature termination codon (Kalyna et al., 2012; McGlincy and Smith, 2008). A mechanism called AS-NMD (Alternative Splicing coupled to NMD) corresponds to the selective degradation of one alternatively spliced isoform containing a PTC, while the isoform lacking a PTC is not affected.

One example of proteins regulated by AS-NMD are those belonging to the POLYPYRIMIDINE TRACT BINDING (PTB) family (Figure 10) (Stauffer et al., 2010). PTB are heterogeneous nuclear ribonucleoproteins (hnRNPs) splicing factors able to repress splice site selection upon their binding to sequences known as splicing silencers. PTB proteins are able to exert feedback regulation on the alternative splicing of their own mRNA in mammals and plants (Figure 10) (Wollerton et al., 2004; Wachter et al., 2012). In plants, mRNAs of the three PTB homologs are subjected to AS-NMD: two isoforms are generated for each homolog, one isoform with a PTC inducing NMD while the other AS isoform being immune to NMD and encoding a full-length protein (Figure 10) (Stauffer et al., 2010). In addition, the regulation of splicing by plant PTB1 and PTB2 has also been shown to have a major role in AS of genes important for the regulation of seed germination and for the control of flowering time (Rühl et al., 2012).

Similar feedback-loop regulation mechanism linked to AS-NMD are not restricted to PTBs. AS-NMD is probably widely used to autoregulate the expression of alternative splicing factors, as exemplified by the GRP7/8 or SR proteins (Schöning et al., 2008; Palusa and Reddy, 2010). In addition to alternative splicing factors, the expression of a large group of genes encoding RNA- and DNA-binding proteins, as well as chromatin modifiers appeared to be under AS–NMD control in mammals (Hamid and Makeyev, 2014). In plants, transcription factors, RNA processing factors and stress response genes are also targeted by AS-NMD (Kalyna et al., 2012), suggesting that AS-NMD represents an important and widespread mechanism to fine-tune gene expression.

5.2. Sequence-dependent NMD inhibition

Modulation of transcript fate via NMD participates to the regulation of gene expression, but in fact many transcripts with potential NMD-inducing features are immune to this process. For example, a wide range of transcripts contain a long 3'-UTR, which are supposed to be targeted by NMD, yet only a fraction of

these are actually targeted by NMD, suggesting the existence of NMD-evasion mechanisms. Factors responsible for NMD-evasion could either inhibit UPF1 binding on the target mRNA or promote its dissociation via UPF1 ATPase activity (Lee et al., 2015). Apart from long 3'-UTR transcripts, the presence of uORF on a transcript is also an NMD-inducing feature, yet many of such transcripts somehow evade NMD. As an example, in yeast, some uORF-containing transcripts harbor a sequence called a stabilizer element (STE) that interacts with the protein Pub1, promoting NMD-evasion (Ruiz-Echevarria and Peltz, 2000). Only few examples of trans-factors able to bind mRNAs and promote NMD-evasion have been identified. Such a mechanism has also been proposed in the case of PTBP1 in mammals both for Rous Sarcoma Virus NMD-evasion and the regulation of endogenous mRNAs (Fatscher et al., 2014) (see part 5.1.2 for more details). It now clearly appears that the regulation of factors binding 3'-UTR mRNA or affecting its secondary structure can influence NMD activation, thus contributing to mRNA fate (Eberle et al., 2008). A wide range of RNA binding proteins are known to bind 3'-UTRs (Yamashita and Takeuchi, 2017), providing huge opportunities for this strategy to fine-tune gene expression.

6. NMD functions in host-pathogen interactions

6.1. NMD and viruses

6.1.1. NMD in host-virus interactions

Due to the size constrains on their genome, viruses commonly express multiple viral protein from a single RNA by producing polycistronic viral RNAs. As a result, many viruses express viral RNAs containing potential NMD-eliciting features, including internal termination codons, uORFs and long 3'-UTRs. NMD was shown to play a role in intrinsic antiviral immunity in plants and mammals, and to restrict the accumulation of several RNA viruses (Garcia et al., 2014; Balistreri et al., 2014). In plants, the key NMD components UPF1, UPF3 and SMG7 have been shown to act as restriction factors of the Potato Virus X (PVX). NMD targeting of the PVX viral RNAs has been found to be dependent on internal termination codons and long 3'-UTRs (Garcia et al., 2014). This intrinsic viral restriction mechanism is not limited to plants as the depletion of UPF1, SMG7 or SMG5 leads to higher Semliki Forest Virus (SFV) titers in mammals (Balistreri et al., 2014). In this case, it has been proposed that UPF1 counteracts the replication and amplification of viral RNAs at an early step of infection occurring just after viral entry (Balistreri et al., 2014).

The role of NMD in the defense against viral pathogens entails the possibility that viruses could have developed strategies to counteract NMD. Some families of retroviruses infecting mammals have evolved various NMD-evading strategies to cope with this viral restriction mechanism.

6.1.2. Retroviral strategies to evade NMD

Some viruses have evolved molecular mechanisms to evade host general RNA decay pathways (Narayanan and Makino, 2013; Hogg, 2016). The best example is the acquisition of a cap structure mimicking the endogenous mRNA cap to prevent viral RNA from being degraded by host decay machinery (Narayanan and Makino, 2013). In the following section, I will focus on strategies specifically dedicated to alleviate viral targeting by NMD.

The Rous Sarcoma Virus (RSV) is a retrovirus, and causal agent of sarcoma in infected chicken. This retrovirus contains a 3'-UTR sequence named RSE for RSV Stability Element, responsible for NMD-evasion (Weil and Beemon, 2006). The experimental removal of RSE leads to enhanced viral RNA degradation in a UPF1- and NMD-dependent manner (Weil and Beemon, 2006). Alignment of 20 RSE sequences from different avian retroviruses showed that both the structure and the sequence of RSE are well conserved, indicating that evolution selected efficient NMD-evasion mechanism as a beneficial trait for the virus (Weil et al., 2009). Polypyrimidine Tract Binding Protein 1 (PTBP1) was shown to bind to RSE and to repress NMD. The artificial binding of PTBP1 alone on an NMD-sensitive mRNA is able to repress NMD (Ge et al., 2016). Several mechanisms of NMD evasion upon PTBP1 binding have been proposed. The first hypothesis is that PTBP1 could prevent UPF1 binding to the 3'-UTR thanks to 3'-UTR overloading by several PTBP1 molecules (Ge et al., 2016). As PTBP1 possesses several RNA binding sites, it might also restructure the mRNA 3'-UTR, thus bringing the PABP closer to the terminating ribosome and allowing NMD suppression (Fatscher and Gehring, 2016). Finally, PTBP1 could also enhance UPF1 ATPase-dependent dissociation from the mRNP to prevent UPF core complex formation (Lee et al., 2015). Transcripts down-regulated in PTBP1-depleted cells have particularly long 3'-UTRs, which is a major trigger of NMD. Genome-wide investigations revealed that PTBP1 also protect a large pool of endogenous mRNAs from being turned-over by NMD (Ge et al., 2016).

Another mechanism of NMD-evasion was found to be deployed by the Human T-lymphotropic Virus 1 (HTLV1), another retrovirus infecting humans. HTLV1 is affected by NMD: its RNA undergoes ribosome frameshifting and alternative splicing, which leads to the production of NMD-sensitive isoforms (Nakano et al., 2013). UPF1-depleted cells show enhanced virus accumulation indicating that UPF1 restricts viral accumulation. HTLV1 expresses two viral RNA binding proteins, Rex and Tax, exhibiting NMD-suppressing

action (Nakano et al., 2013; Mocquet et al., 2012). The exact mechanism of NMD-inhibition by Rex is not yet understood, but Tax was proposed to inhibit NMD through its interaction with UPF1 and sequestration of UPF1 away from eIF3E (Mocquet et al., 2012). This UPF1 sequestration also affects endogenous NMD targets, which were found to accumulate to higher levels upon HTLV1 infection (Mocquet et al., 2012). NMD inhibition thus appears to be an effective viral strategy to reduce the impact of this intrinsic viral restriction pathway.

6.2. NMD orchestrates plant antibacterial defense

NMD also contributes to the immune response in *Arabidopsis*, in which it controls innate immunity. High levels of salicylic acid (SA) and strong morphological phenotypes were observed in *Arabidopsis* NMD mutants (Yoine et al., 2006; Jeong et al., 2011). NMD mutants also show a constitutive activation of plant immune responses, shown by the resistance to bacterial infection (Jeong et al., 2011; Riehs-Kearnan et al., 2012; Rayson et al., 2012). It was later shown that plant NMD regulates the immune response to bacterial infection by controlling the turnover of several mRNA coding for immune receptors (Gloggnitzer et al., 2014). *Arabidopsis* infections by *Pseudomonas syringae* causes NMD impairment as part of the host-programmed immune response, leading to the stabilization of TIR domain-containing, Nucleotide-binding, Leucine-rich repeat (TNL) immune receptors potentiating plant immunity (Gloggnitzer et al., 2014). This NMD dampening is abolished in *upf1*, *upf3* and *smg7* mutants resulting in exacerbated resistance to bacterial infection. Interestingly, NMD efficiency also declines during cell differentiation or upon stress exposure in yeast and mammals (Bruno et al., 2011; Gardner, 2008); NMD dampening could be used as a widespread mechanism to regulate gene expression.

III. The Up Frameshift 1 (UPF1) protein

UPF1 is a key component and one of the most studied factor of NMD. It is involved in the discrimination between NMD targets and non-NMD targets and activates the subsequent recruitment of decay factors to promote the degradation of NMD targets. The UPF1 interactome in yeast and mammals has been extensively studied: UPF1 interacts directly or indirectly with a wide range of proteins among which RNA decay factors, RNA binding proteins and RNA silencing factors for example (Lejeune et al., 2003; Flury et al., 2014; Schweingruber et al., 2016; Isken et al., 2008). By contrast, the UPF1 interactome in plants is poorly documented. A subcellular localization of UPF1 in P-bodies has been observed in different species (Sheth and Parker, 2006; Stalder and Mühlemann, 2009; Unterholzner and Izaurralde, 2004; Brogna et al., 2008). The exact function of UPF1 shuttling in and out P-bodies is not yet completely understood, its study

Table 1 : Current knowledge on UPF1 interactome. Principal known interactors are listed here. Proteins in bold letters are RNA-independent partners of UPF1, while proteins underlined are direct interactors of UPF1

Cellular Pathway	Proteins	Conservation	References
NMD factors			
• Core NMD factors	**UPF2**, **UPF3**		He, 1997 ; Chamieh, 2008 ; Nyiko, 2013
• Other NMD factors	**SMG5** ; SMG7 ; SMG6		Franks, 2010
	SMG1		Yamashita 2001 ; Isken, 2008 ; Schweingruber, 2016 ; Grimson, 2004
• EJC components	MAGOH ; CASC3 ; eIF4A3		Schweingruber, 2016 ; Flury, 2014
Decay factors			
• Decapping factors/ translation repressors	DCP1 ; DCP2 ; EDC4 EDC3 ; DDX6		Schweingruber, 2016 ; Fenger-Grøn, 2005 ; Swisher, 2011
• Exoribonucleases XRN	XRN1		Isken, 2008 ; Franks, 2010 ; Lejeune, 2003
• Exosome	RRP4 ; RRP41		Isken, 2008 ; Lejeune, 2003
Other RNA binding proteins			
• RNA helicases	DDX3		Schweingruber, 2016 ; Flury, 2014
• RNA silencing factors	**AGO1** ; **AGO2**		Flury, 2014 ; Jin, 2009
• Others	PABP		Schweingruber, 2016 ; Flury, 2014 ; Franks, 2010
	STAU1 ; STAU2		Schweingruber, 2016 ; Flury, 2014 ; Kim, 2005
Translation factors			
• Translation initiation	**eIF3** ; eIF4G		Schweingruber,2016 ; Isken, 2008 ; Flury, 2014
• Translation release factors	**eRF1** ; **eRF3**		Czaplinski, 1998 ; Isken, 2008

could give more insight into the role of both UFP1 and P-bodies (Sheth and Parker, 2006; Stalder and Mühlemann, 2009). Moreover, in some species, UPF1 is involved in NMD-independent decay pathways, such as the degradation of histone mRNA, the Staufen-Mediated Decay and the Regnase-Mediated Decay in mammals (Kaygun and Marzluff, 2005; Kim et al., 2005a; Mino et al., 2015). The exact mechanisms of NMD targets recognition by UPF1, and the molecular events leading to UPF1-mediated degradation are still debated.

1. The UPF1 interactome

The study of the UPF1 interactome has played an important role in understanding mechanistic aspects of UPF1 action, both in the NMD pathway and in non-NMD pathways. Several studies have identified UPF1 partners through immunoprecipitation and mass spectrometry analysis, identifying a wide range of RNA binding proteins as UPF1 partners (Flury et al., 2014; Isken et al., 2008; Schweingruber et al., 2016). Other studies have been able to identify direct interaction between UPF1 and its partners (He et al., 1997; Swisher and Parker, 2011; Mocquet et al., 2012). The principal UPF1 partners identified in different organisms are listed in Table 1.

UPF1 interacts directly with its main protein partner UPF2; this interaction is mediated by UPF1 cystein and histidine-rich (CH) domain (He et al., 1997; Kerényi et al., 2008). UPF2 also mediates the interaction between UPF1 and UPF3, thus allowing the formation of the NMD core complex (He et al., 1997; Serin et al., 2001). In mammals, another interaction important for NMD is the association of UPF1 with the 14-3-3 domain-containing factors SMG6 and SMG7. These interactions are dependent on UPF1 phosphorylation on both its N- and C-terminal parts (Chakrabarti et al., 2014; Flury et al., 2014). Some EJC components, such as eIF4AIII, also co-purify with UPF1. This interaction is independent on RNA but dependent on UPF1 phosphorylation (Flury et al., 2014). On the contrary, SMG1, the kinase responsible for UPF1 phosphorylation, and eRF3 are both associated with hypo-phosphorylated UPF1 (Isken et al., 2008).

Another group of UPF1 partners are the RNA decay factors. UPF1 has been shown to interact with different RNA decay factors to mediate the accelerated degradation of NMD targets. It was shown that the phosphorylated version of UPF1 has higher affinity for decay factors DCP1, XRN1, RRP4 (Isken et al., 2008). Moreover, the interaction between UPF1 and the decapping factors DCP2 and PNRC2 is direct in human cells and both UPF1 N- and C-terminal extremities are required for this interaction (Loh et al., 2013; Fenger-Grøn et al., 2005). In contrast, Dcp2 interaction with Upf1 in yeast is mediated by another

protein, Edc3 (Swisher and Parker, 2011). Independently of NMD, UPF1 has been shown to interact directly with factors involved in other RNA decay pathways such as the Staufen 1 protein (STAU1) which participates in Staufen-Mediated Decay, and the NYN-domain endonuclease Regnase-1, which is involved in Regnase-Mediated Decay (Kim et al., 2005a; Mino et al., 2015). Finally, UPF1 has been shown to interact with Argonaute RNA silencing factors AGO1 and AGO2 in an RNA-independent manner in human cells. Furthermore, it has been proposed that UPF1 RNA helicase activity could facilitate the binding of the RNA-Induced Silencing Complex (RISC) (Jin et al., 2009).

As UPF1 is a RNA binding protein, during immunoprecipitation, it co-purifies with a wide range of RNA binding proteins and its interactions are either RNA-dependent or RNA-independent. An example of an RNA-dependent interaction is that of UPF1 with PABP family proteins (Schell et al., 2003; Flury et al., 2014; Schweingruber et al., 2016). Moreover, as NMD is a process dependent on translation, the co-purification of UPF1 with some ribosomal proteins, translation initiation, elongation or termination factors have also been reported (Flury et al., 2014; Schweingruber et al., 2016). For example, eIF3 associates with UPF1 and this interaction is enhanced with phospho-UPF1. In contrast, eRF1 and eRF3 are preferentially bound to hypo-phosphorylated UPF1 (Isken et al., 2008; Ivanov et al., 2008). Interestingly, UPF1 association with ribosomal proteins is probably not only be due to the experimental co-purification of the translation machinery, as the interaction between UPF1 and eRF3 or with the ribosome factor Rps26 was shown to be a direct protein-protein interaction, indicating a possible role of UPF1 in the dissociation of the translation complex on PTC (Min et al., 2013).

Among the RNA binding proteins, several members of the RNA helicase family also co-purify with UPF1. The DDX1 and DDX3X RNA helicases were shown to co-purify with UPF1 independently of RNA, while the association with other RNA helicases such as DDX15 and DDX36 was RNA dependent (Flury et al., 2014) (Flury, 2014). In addition to UPF1, DDX3X also interacts with both UPF2 and SMG5 as shown by the BioID strategy recently used to specifically identify proteins that are located in close proximity of these NMD factors (Schweingruber et al., 2016), indicating a possible link between DDX3X and NMD in mammals. Moreover, DDX3X and DDX1 have been shown to interact with both HIV-1 mRNP and UPF1, suggesting a possible role of UPF1 and these RNA helicases in HIV-1 RNA export from the nucleus (Ajamian et al., 2008). Finally, UPF1 directly interacts with the MOV10 RNA helicase (Gregersen et al., 2014), which also plays a role in HIV-1 accumulation (Serquiña et al., 2013).

Other RNA binding proteins can be associated with UPF1, potentially because they are bound to the same RNA. This is the case of several regulatory RNA binding factors such as the Pumilio proteins, PTBP1 and

other hnRNP or splicing factors, as well as the RNA m6A methylation reader YTHDF2 (Flury et al., 2014). By contrast, PTBP3 (also called Rod1) interacts with UPF1 after benzonase treatment, suggesting a protein-mediated interaction (Brazão et al., 2012). The Stem Loop Binding Protein (SLBP), which has an important role in replication-dependent histone mRNA maturation and co-purifies with UPF1 in both direct and reverse IP, but the RNA dependency of this interaction has not been tested (Kaygun and Marzluff, 2005).

The list of UPF1 interactors will likely still evolve with the technical improvements of the immunoprecipitation methods, as already observed for BioID (Schweingruber et al., 2016). In plants, the UPF1 interactome has not yet been studied in detail and the extent of conserved and plant-specific UPF1 interactors remains to be determined. In addition to expanding the knowledge on plant NMD, plant-specific UPF1 partner identification could lead to the characterization of novel UPF1-dependent RNA degradation pathways not linked to NMD, as observed in mammals (Kim et al., 2005b; Mino et al., 2015).

2. UPF1 localization

The UPF1 protein is localized in the cytoplasm in yeast, but relocalizes to P-bodies in strains lacking either decay factors or NMD factors such as DCP1, DCP2, XRN1, UPF2 and UPF3, indicating that UPF1 and its targets are only transiently trafficking through P-bodies (Sheth and Parker, 2006). The same situation is observed in *Drosophila*, where NMD factors and substrates accumulate in P-bodies upon mutations of NMD or decay factors such as DCP1, DCP2 or XRN1 (Brogna et al., 2008). In human cells, UPF1 is localized in the cytoplasm and in P-bodies (Stalder and Mühlemann, 2009), but strongly accumulates in P-bodies upon SMG7 overexpression (Unterholzner and Izaurralde, 2004) or upon treatment with the NMD inhibitor NMDI1 (Durand et al., 2007). Despite these clear localizations, the role of P-bodies in NMD is not well established, as NMD still occurs in cells lacking detectable P-bodies in *Drosophila* and mammals (Eulalio et al., 2007; Stalder and Mühlemann, 2009).

In plants, UPF1 localization in relation to P-bodies remains elusive. UPF1 was first localized to P-bodies in transient assays with or without SMG7 overexpression (Mérai et al., 2013). Nevertheless, this localization has been questioned by another report suggesting that UPF1 is localized at 50% to adjacent cytoplasmic foci, which are not P-bodies (Moreno et al., 2013).

3. UPF1-dependent but NMD-independent pathways

Apart from its important role in NMD, UPF1 is also involved in other cellular processes linked to the degradation of subsets of transcripts, such as the decay of human histone mRNA, Staufen-Mediated Decay

(SMD), Regnase-Mediated Decay (RMD) and Glucocorticoid receptor-Mediated Decay (GMD). These pathways have in common a targeting strategy, based on the recognition of an RNA stem-loop structure, followed by UPF1 recruitment, which participates to the stimulation of RNA decay. In addition, UPF1 has also been recently shown to have a role in Tudor-staphylococcal/micrococcal-like nuclease (TNS)-mediated miRNA degradation.

3.1. UPF1-mediated decay of histone mRNA

Replication-dependent histone mRNAs are the only non-polyadenylated mRNAs in metazoans (Kaygun and Marzluff, 2005). Instead of the classical poly(A) tail, histone mRNAs possess a stem-loop structure in their 3'-UTR which is bound by the Stem-Loop Binding Protein (SLBP). UPF1 has been shown to co-purify with SLBP in both direct and reverse co-immunoprecipitation experiments (Kaygun and Marzluff, 2005). Histone mRNA degradation is initiated by the addition of untemplated uridines, that promotes the binding of the Lsm1-7 complex interacting with the 3'-5' exonuclease Eri1, supported by Upf1 (Mullen and Marzluff, 2008). Mutations abolishing either UPF1 helicase activity or ATP affinity resulted in inhibition of histone mRNA degradation. By contrast, mutations affecting UPF2 did not affect histone mRNA accumulation, demonstrating that this process is UPF1-dependent but NMD-independent (Kaygun and Marzluff, 2005).

3.2. Staufen-Mediated Decay

Staufen-Mediated Decay (SMD) is a mammal-specific RNA decay pathway that depends on the double-stranded RNA-binding protein Staufen 1 (STAU1). STAU1 can recognize double stranded structures formed in the 3'-UTR of its targets and recruits UPF1 to allow subsequent target degradation (Kim et al., 2005a). STAU1 is involved in the turnover of mRNA coding for proteins controlling cell metabolism, cell division or cell cycle, underlying the physiological importance of SMD (Park and Maquat, 2013). Tethering of STAU1 on the coding sequence or on the 3'-UTR of a transcript causes its degradation. UPF1 participates in SMD as it interacts directly with STAU1, and is required for the degradation of STAU1 targets (Kim et al., 2005a). This degradation is independent of UPF2, UPF3 or EJC factors, showing that SMD is an NMD-independent process. It was also described that STAU1 competes with UPF2 for binding to UPF1. Conversely, STAU1 appears to play no role in NMD, as NMD targets are not affected upon STAU1 depletion (Kim et al., 2005a).

3.3. Regnase-Mediated Decay

The Regnase-1 protein is a NYN-domain endonuclease required for the regulation of a subset of mRNAs including inflammation-related mRNAs preventing immune disorders in mouse (Matsushita et al., 2009). Translation and the presence of UPF1 were shown to be necessary for this regulation (Mino et al., 2015).

In human cells, the Regnase-1 homolog MCPIP1, when overexpressed, suppresses miRNA biogenesis by inducing the cleavage of pre-miRNA (Suzuki et al., 2011). Regnase-1 has been shown to directly interact with UPF1 in a translation-dependent manner (Mino et al., 2015). Regnase-1 localization is not clear: Regnase-1 homologs have been shown both to be P-body components and to be absent from P-bodies (Suzuki et al., 2011; Mino et al., 2015; Mao et al., 2017). Several lines of evidence show that Regnase-1 targets mRNAs with a stem loop, usually in their 3'-UTR, that has to be located further than 20 nt away from the stop codon (Mao et al., 2017).

3.4. Glucocorticoid Receptor-Mediated Decay

Glucocorticoid Receptor (GR) bind a subset of mRNAs involved in different pathways, such as immune response, and elicit their rapid decay. This process will be referred to as GMD for Glucocorticoid receptor-Mediated Decay. This decay pathway requires UPF1 and PNRC2, the latter acting as a bridge between NMD and the decapping machinery (Park et al., 2016). GR were proposed to recruit first PNRC2 upon their binding to mRNA targets, which will then recruit both UPF1 and the decapping factor DCP1 to elicit accelerated decay of the transcript. Apart from UPF1 and PNRC2 that are involved in other decay pathways, GMD requires its own set of factors, such as endoribonuclease HRSP12 and RNA binding protein YBX1. NMD factors UPF2, SMG1, SMG5, SMG7 and EJC factors are dispensable for GMD, indicating that this process is independent from NMD (Park et al., 2016). In contrast to SMD and RMD, GMD is a translation-independent process.

3.5. Tudor-staphylococcal/micrococcal-like nuclease-mediated miRNA decay

Tudor-staphylococcal/micrococcal-like nuclease (TSN) in an endonuclease constitutive of the RNA-Induced Silencing Complex (RISC) and is involved in the decay of functional miRNA (Caudy et al., 2003; Elbarbary et al., 2017). It has been recently shown that UPF1 is required for TSN-mediated miRNA decay. UPF1 was proposed to remodel mRNAs bound by RISC, therefore dissociating the miRNA from its target, and allowing miRNA degradation involving TSN (Elbarbary et al., 2017). Consistent with this model, UPF1 was identified as an RNA-independent interactor of AGO1 and AGO2 in human cells (Jin et al., 2009), but the rational for this interaction remained unclear as UPF1 knockdown did not affect miRNA targeted reporter constructs (Höck et al., 2007).

In conclusion, UPF1 is a multifunctional protein involved in several cellular RNA decay pathways. These include both the well conserved NMD pathway and some secondary decay pathways as SMD, RMD and GMD. Each secondary decay pathway acts on discrete subsets of RNA targets with specific cellular functions. UPF1 is required for viability and for the regulation of a wide number of functional transcripts. Nevertheless, some aspects of the mode of action of UPF1 in both NMD and other decay pathways remain obscure. How NMD targets are discriminated from non-NMD targets by UPF1, as well as the precise hierarchical recruitment of decay factors upon UPF1 activation remain in many cases elusive. This situation is particularly true in plants, as the knowledge of the factors associated with this pathway is very limited; novel strategies need to be developed to uncover the mechanistic details of NMD.

IV. Thesis objectives

Nonsense-Mediated Decay is a conserved mechanism among eukaryotes allowing the recognition and subsequent decay of mRNAs harboring specific NMD-inducing features. This mechanism is necessary for cellular fitness and the regulation of a wide number of protein-coding transcripts involved in mechanisms such as stress response or RNA processing. UPF1, the major factor in NMD, is conserved in all eukaryotes. UPF1 deficiency is lethal in *Drosophila*, mammals and plants, illustrating the essential function of this protein. NMD pathway is less characterized in plants compared to mammals and yeast. In plants, many NMD factors are not conserved: some homologs of NMD proteins necessary for UPF1 phosphorylation or for the subsequent decay of NMD targets are not present in plants. These observations raise the question of how NMD and UPF1 function in the plant kingdom.

In this context, my thesis objectives were the following:

- Identify new partners of UPF1 in *Arabidopsis*
- Determine how UPF1 partners participate in the regulation of NMD targets
- Study UPF1 partners subcellular localization
- Identify factors potentially involved in UPF1-dependent NMD-independent pathways

To address these questions, we designed an approach to study the UPF1 interactome using an unbiased method coupling immunoprecipitation of UPF1 and detection of UPF1 partners by mass spectrometry. This approach allowed the identification of 245 UPF1-associated factors in plants, among which RNA decay and RNA processing factors as well as translation repressors. We tested NMD targets accumulation in 22 mutants affected in some identified partners, and several mutants showed reduced NMD targets accumulation. Among these mutants we focused on the translation repressor DCP5. We identified that UPF1 shares around 50% of common protein partners with DCP5 suggesting their participation to similar mRNPs. We identified that both UPF1 and DCP5 interact directly with RNA helicases homologs of the DDX6 family, a known translation repressor in mammals. We showed that all these proteins co-localize in P-bodies, we thus propose a model involving DCP5 and DDX6 homologs in the protection of NMD targets from degradation via their re-localization into P-bodies.

Following this study, we focused on an endonuclease co-purifying with UPF1, that we named UCN, for UPF1 Co-purified Nuclease. A reverse study by mass spectrometry identified protein partners of UCN, validating its association with UPF1 and identifying new partners, such as components of the decapping complex, DCP1 and VCS. The use of yeast two-hybrid experiments identified a direct interaction between

UCN and DCP1. In co-localization studies, we validated that UCN is a novel component of plant P-bodies. Finally, the functional analysis of UCN function suggests that its expression could influence hormone signaling, cell death and defense responses. The study of this novel UPF1 partner and P-body component could identify a novel RNA degradation pathway in plants.

This work provides novel insights into NMD targets regulation in plants, in direct connection to the role of P-bodies in the equilibrium between mRNA translation, storage and decay. It also identifies several novel components of this important cytosolic structure, which will be very useful to better understand the functions and cellular processes occurring in P-bodies in plants.

Results

Chapter I: Interaction between UPF1 and translation repressors identifies novel regulatory networks in plant NMD

Manuscript in preparation

Interaction between UPF1 and translation repressors identifies novel regulatory networks in plant NMD

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

[1] Institut de Biologie Moléculaire des Plantes (IBMP), Centre National de la Recherche Scientifique, University of Strasbourg, 67000 Strasbourg, France.

[2] Plateforme Protéomique Strasbourg-Esplanade, Centre National de la Recherche Scientifique, University of Strasbourg, 67000 Strasbourg, France.

* To whom correspondence should be addressed. Email: damien.garcia@ibmp-cnrs.unistra.fr

Abstract

Nonsense-Mediated Decay (NMD) is a RNA decay pathway involved in RNA quality control and in the fine-tuning of gene expression. The RNA helicase UPF1 is essential for this process through its interaction with other NMD and RNA decay factors. We report here the identification of a large set of proteins that co-purify with *Arabidopsis* UPF1. Together with UPF1, several of these factors have a dual localization in the cytosol and in P-bodies. Most of them also co-purify with DCP5, a known translation repressor present in P-bodies. P-bodies are formed by the aggregation of translationally repressed mRNPs. In line with this idea, we show that both UPF1 and DCP5 directly interact with plant homologs of the RNA helicases DDX6/Dhh1, that are known translation repressors in human and yeast, respectively. Interestingly, NMD targets that show increased accumulation upon UPF1 knockdown are decreased in *dcp5* mutant. We propose a model in which the recruitment of translation repressors by UPF1 leads to the protection rather than degradation of NMD targets. This study identifies a novel UPF1 centered protein-protein interaction network linked to P-bodies and associated with the first NMD antagonist function in plants.

Introduction

Eukaryotes possess three translation-dependent quality control pathways to reduce the accumulation of mRNAs showing aberrant ribosomal progression: No-Go decay (NGD), Non-Stop Decay (NSD) 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). These PTCs can result from mutations, transcription errors or inefficient splicing, but can also happen in alternatively spliced transcript isoforms, in mRNAs harboring long 3'-UTRs or upstream open reading frames (uORF). 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 UPF2 is one of the key steps, which triggers the recruitment of UPF3 leading to 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). Interestingly, despite its discovery more than twenty years ago, the exact mechanisms by which a nonsense codon is recognized as premature, and the individual roles of the UPF proteins, are still poorly understood. One of the initial functions proposed for this pathway was to avoid the accumulation of potentially deleterious truncated proteins but NMD is now recognized to be a genuine post-transcriptional regulation pathway for mRNAs encoding functional proteins (Nasif et al., 2017). In mammals, NMD controls endoplasmic reticulum stress response by limiting the expression of IRE1a and repress 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, avoiding the fitness cost of defense activation and modulating the output of bacterial infection (Rayson et al., 2012; Riehs-Kearnan et al., 2012; Gloggnitzer et al., 2014). Interestingly, many interactions between the NMD pathway and viral infections have been observed across kingdoms. NMD is involved in the restriction of RNA virus infection in plants and mammals (Garcia et al., 2014; Balistreri et al., 2014). Retroviral proteins and RNA structures in the human T-lymphotropic virus type 1 (HTMLV1) and the Rous Sarcoma Virus (RSV) respectively, act as suppressors of NMD in mammals (Mocquet et al., 2012; Weil and Beemon, 2006). NMD suppressing motifs also exist on cellular mRNAs and their inhibiting action depends on the binding of RNA binding proteins with NMD inhibiting properties, such as the Polypyrimidine Tract Binding Protein 1 in mammals and the Pumilio 1 in yeast (Ruiz-Echevarria and Peltz, 2000; Ge et al., 2016). Of note, many mRNAs containing putative NMD inducing features are not regulated by NMD, suggesting the broad action of NMD inhibition mechanisms. Thus, the identification and study of these mechanisms are a prerequisite

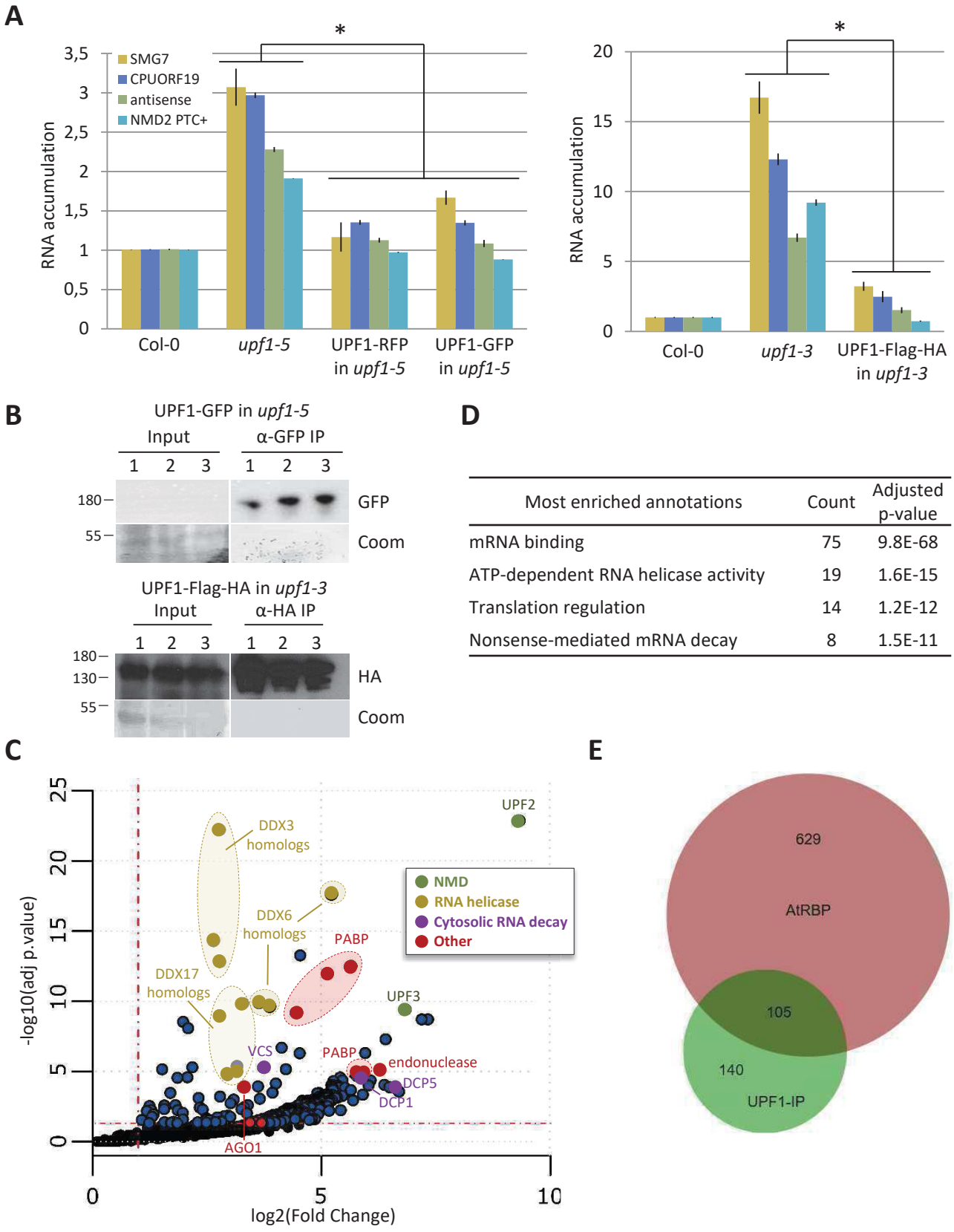
to fully understand how regulation by NMD operates and is modulated. UPF1 is a major player in NMD, its activity induces a cascade of events leading to accelerated mRNA decay. Decay induction was shown to involve the action of deadenylases, decapping factors and exoribonucleases in mammals and yeast (Muhlrad and Parker, 1994; Mitchell and Tollervey, 2003; Lejeune et al., 2003). In contrast, these events are still poorly documented in plants.

In this work, we identified a large set of proteins associated with *Arabidopsis* UPF1 using an immuno-affinity approach coupled to mass spectrometry. A subcellular localization study validated the co-localization of UPF1 with selected partners in the cytosol and/or P-bodies, identifying novel P-body components. We analyzed the link between UPF1 and the translation repressor DCP5 and identified that UPF1 shares half of its partners with DCP5. We show that mutations in DCP5 and UPF1 have an opposite effect on the accumulation of model NMD targets. Finally, we uncovered that both UPF1 and DCP5 directly interact with members of the plant DDX6 translation repressors family. Those results led us to propose a mechanistic model integrating the role of DCP5 and DDX6-like factors in NMD targets translation inhibition in P-bodies. This model is coherent with a recent report identifying a major function of P-bodies in translation inhibition rather than RNA decay and could have general implications for NMD in other organisms (Hubstenberger et al., 2017).

Results

Complementation of *upf1* mutations by UPF1-tagged versions

Early studies of NMD in plants were based on the analysis of mutants in homologs of UPF1 and UPF3. These mutants showed defects regarding flowering time, response to abiotic stresses and germination (Hori and Watanabe, 2005; Yoine et al., 2006). Major advances in our understanding of plant NMD was also achieved using NMD reporter genes in transient expression assays (Kerényi et al., 2008; Kertész et al., 2006; Mérai et al., 2013). Nevertheless, attempts to discover plant UPF1 interacting proteins by the combination of yeast two-hybrid screens and transient expression strategies could only validate previously suspected candidate genes (Kerényi et al., 2008). These approaches demonstrated that the core UPF1-UPF2-UPF3 complex is conserved in plants, but did not allow de novo identification of UPF1 interacting partners. In order to address this issue, we designed RFP, GFP or a double Flag-HA tagged versions of UPF1 driven by its endogenous promoter. Those constructs were introduced into either null (*upf1-3*) or hypomorphic (*upf1-5*) *upf1* mutant backgrounds and could restore the regulation of four



Most enriched annotations	Count	Adjusted p-value
mRNA binding	75	9.8E-68
ATP-dependent RNA helicase activity	19	1.6E-15
Translation regulation	14	1.2E-12
Nonsense-mediated mRNA decay	8	1.5E-11

Figure 11. Analysis of the plant UPF1 interactome. (A) qPCR analysis of NMD targets in *upf1* mutants and lines complemented with pUPF1:UPF1-GFP, pUPF1:UPF1-RFP and pUPF1:UPF1-FHA. Values are normalized to *ACTIN2* presented as mean of FC and arbitrary set to 1 in Col-0 ± SEM. Asterisk represents significant differences, *p<0.025*; n=2. (B) Western blot analysis of UPF1-GFP and UPF1-FHA IPs. Six replicates used for mass spectrometry analysis are shown. (C) Most enriched Gene Ontology and Uniprot terms among UPF1 partners. This analysis was done with a total of 245 input proteins using DAVID Bioinformatics Resources 6.8. (D) Volcano plot representing the proteins enriched in UPF1 IPs. Adjusted p-value is presented on the y axis, Fold Change (FC) in x axis. The red dotted lines represent the significance thresholds of *p<0.05* and log2(FC)<1. (E) Venn diagram showing the overlap between UPF1-IPs and the At-RBP proteome identified in Reichel et al., 2016. IP: immunoprecipitation.

distinct NMD targets with different NMD inducing features, demonstrating that they encode functional UPF1 fusion proteins (Figure 11A).

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

Plants stably expressing UPF1-GFP and UPF1-Flag-HA were used to perform immunoprecipitations (IPs) with antibodies directed against the GFP and HA tags, respectively. We observed a strong UPF1 enrichment in the IP fractions as shown on western blots (Figure 11B), indicating that we produced adequate material to purify UPF1-enriched fractions. We used mass spectrometry to analyze these fractions (Figure 11C and Table S1), and identified 245 proteins significantly enriched in UPF1 IPs compared to the control IPs with an adjusted p-value<0.05. The success of our approach was validated by the identification of UPF2 and UPF3, the only two proteins known to be part of the UPF core complex in plants, among the most enriched partners (Figure 11C, Table S1). A global annotation enrichment analysis of proteins associated with UPF1 revealed the presence of an important proportion of RNA binding proteins including many RNA helicases, in addition to proteins involved in translation regulation (Figure 11D). This global analysis is coherent with the RNA binding capacity of UPF1 and the fact that multiple RNA helicases, RNA binding proteins and proteins involved in translation interact with UPF1 in other organisms (Bond et al., 2001; Flury et al., 2014; Schweingruber et al., 2016). Among partners significantly enriched with UPF1, we noticed the decapping associated factors DCP1, DCP5 and VARICOSE (VCS), as well as the plant DDX6/Dhh1 homologs RH6, RH8 and RH12. DDX6/Dhh1 is an RNA helicase and translation repressor localized in P-bodies and required for their assembly in mammals and yeast (Ayache et al., 2015). DDX6 was shown to associate with the decapping complex as well as UPF1 in human cells but the rational of this association was not further investigated (Ayache et al., 2015). Of note, DCP5, its yeast homolog Scd6 and DDX6 have all been involved in translation repression (Rajyaguru et al., 2012; Minshall et al., 2009; Presnyak and Coller, 2013a). We also noticed the presence in UPF1 IPs of homologs of two other RNA helicase families, DDX3 and DDX17, the former has functions in stress granule assembly and translation inhibition and the latter in mRNA processing in mammals (Soto-Rifo and Ohlmann, 2013; Janknecht, 2010). As previously described in mammals, we observe that AGO1, a major RNA silencing effector in plants co-purified with UPF1 (Poulsen et al., 2013). Finally, UPF1 enriched fractions contained the eIF4G protein, known to influence NMD in mammals (Lejeune et al., 2003) and several members of the plant Poly(A) Binding Protein (PABP) family. In order to validate our mass spectrometry data, we performed co-IP tests of UPF1 with selected partners by either IP or reverse IP. This approach was applied

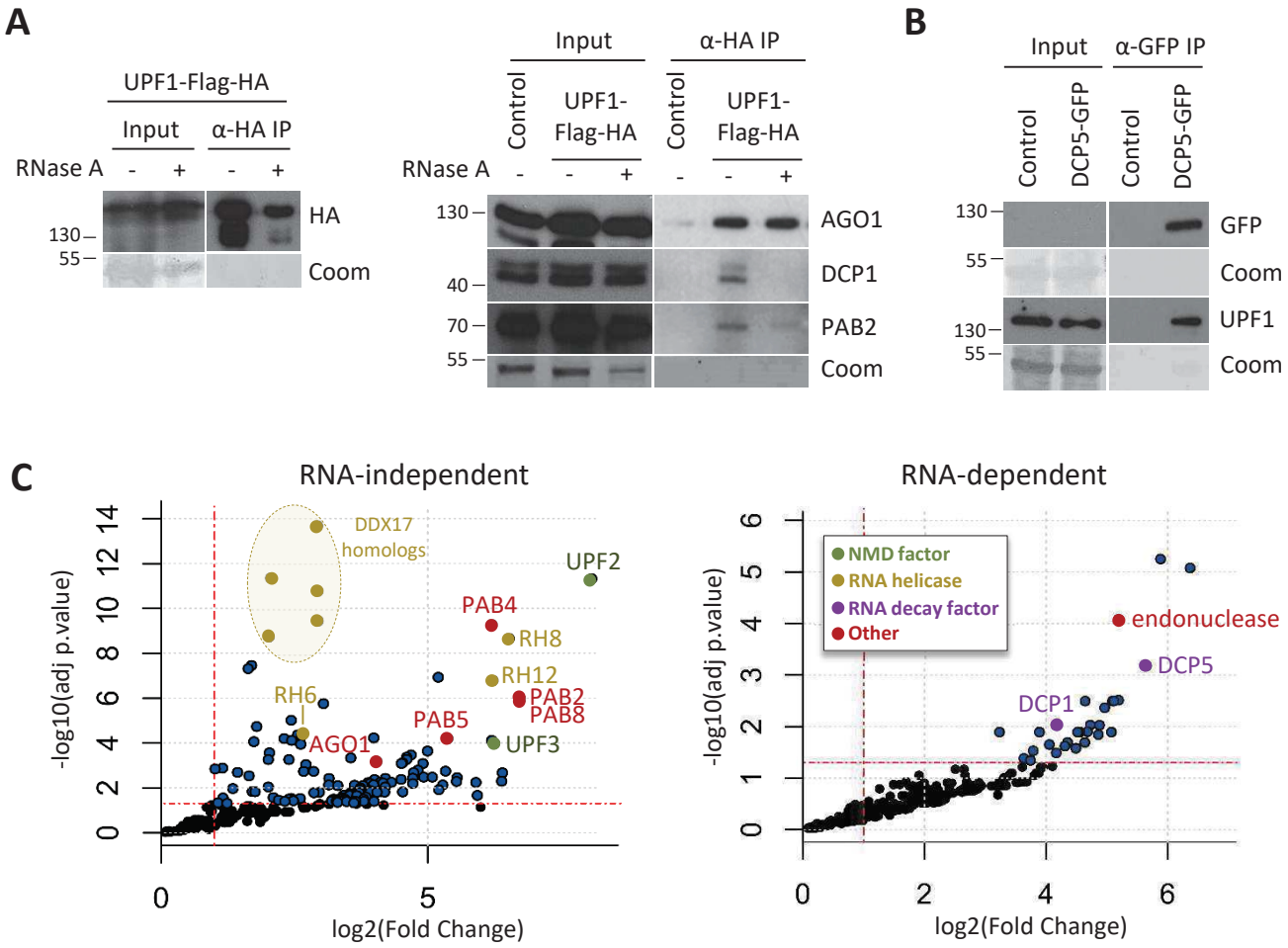


Figure 12. Analysis of RNA-dependent and RNA-independent UFP1 partners. (A) Volcano plot representing the proteins enriched in both RNaseA treated and non-treated UPF1-FHA IP (RNA independent) and proteins enriched in non-treated UPF1-FHA IP only (RNA dependent). Adjusted p-value is presented on the y axis, Fold Change (FC) in x axis. The red dotted lines represent the significance thresholds of *p<0.05* and log2(FC)<1. (B) Western blot analysis of the co-IPs of UPF1-FHA with selected UPF1 partners. (C) Western blot detection of UPF1 in DCP5-GFP IPs. UPF1, AGO1, DCP1 and PAB2 protein levels were analyzed using antibodies raised against endogenous proteins.

to proteins for which we could obtain specific antibodies or produce transgenic lines expressing tagged-versions of the corresponding factors. These experiments were performed for PAB2 and for proteins known to be present in P-bodies including AGO1, DCP1, and DCP5. We validated the co-purification of all these partners with UPF1 by either co-IPs or reverse co-IPs (Figure 12A-12B and Figure S1). These validations illustrate the striking similarities between the interactome of animal and plant UPF1, as homologs of PAB2, AGO1, DCP1 and DCP5 are also known to co-purify with UPF1 in animals.

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

Our strategy identified a set of 245 proteins significantly enriched in the UPF1-purified fraction. Overall, 43% of these proteins are also present in an experimentally validated *Arabidopsis* mRNA-binding proteome (At-RBP) determined by "mRNA interactome capture" (Reichel et al., 2016). This striking overlap suggests that their presence in the UPF1 IPs could be due to RNA-mediated interactions (Figure 11E and Table S2). To determine RNA-dependent and RNA-independent UPF1 partners, we performed IPs upon RNaseA treatment; proteins present in both RNaseA treated and non-treated IPs were considered as RNA independent partners, whereas proteins depleted from RNaseA treated IPs compared to non-treated IPs were considered as RNA dependent (Figure 12A and Table S3-S4). This analysis revealed the RNA dependency status for 107 proteins: 78 of them co-purified with UPF1 in a RNA-independent manner, whereas 29 were found to be RNA-dependent (Table S3-S4). The specific recovery of the core UPF1-UPF2-UPF3 complex in the RNA-independent experiments, while they were absent from the RNA-dependent pool, confirmed the efficiency of our approach. RNA-independent partners also included the RISC complex component AGO1, as well as the three DDX6 homologs RH6, RH8, RH12 and the PABP PAB2, 4, 5 and 8. As expected, more than 50% of the RNA-dependent partners had clear predicted functions in nucleic acid metabolism. They notably included DCP5, DCP1, four proteins containing a RRM domain and a putative endonuclease. The validation of RNA dependency status was achieved for AGO1 and DCP1 by co-IPs (Figure 12A). In contrast, for PABPs, which were present in the RNA-independent pool seem to be partially dependent on RNA (Figure 12A).

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

The extent of UPF1 localization within P-bodies in plants is still not completely clear as different UPF1 localization patterns were reported (Mérai et al., 2013; Kim et al., 2009; Moreno et al., 2013). These studies were exclusively based on transient expression studies and they all show some degree of co-localization with P-bodies. However, the most recent work also indicates that half of the UPF1 cytoplasmic

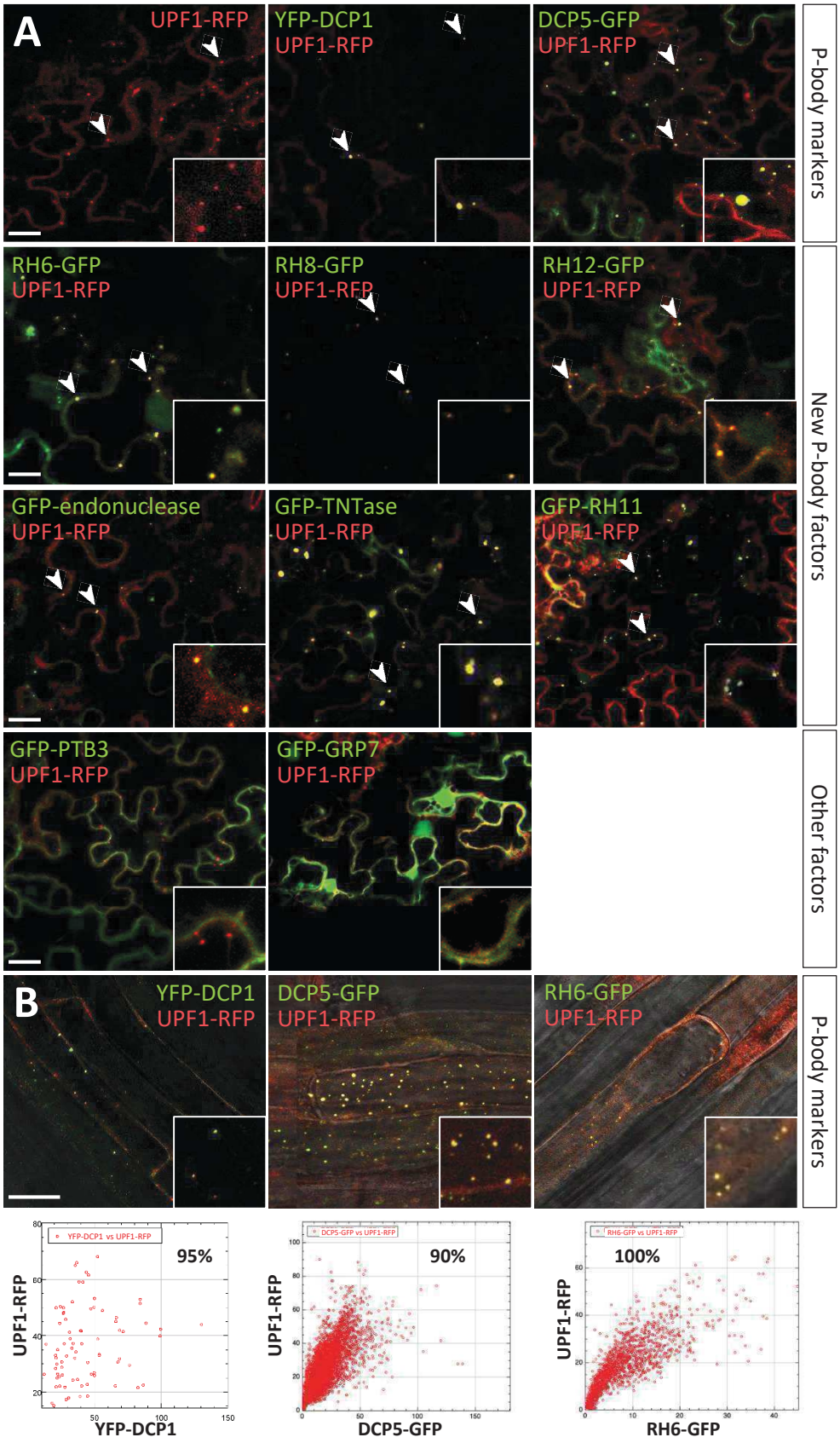


Figure 13. Subcellular localization of UPF1 and its partners. (A) Confocal microscopy images of UPF1-RFP and GFP-fusion of UPF1 partners transiently co-expressed in *N. benthamiana* leaves. Pictures taken at 4 days after infiltration. (B) Confocal microscopy images of epidermal root cells of stable Arabidopsis lines expressing both UPF1-RFP and the P-body markers: YFP-DCP1, DCP5-GFP or RH6-GFP. Dot plot representing the quantification of the signal in both GFP/YFP and RFP channels in x or y axis respectively are represented under each confocal image. Each plot was acquired following the analysis of the foci composition of 10 randomly chosen confocal sections. The colocalization percentage between P-body markers and UPF1-RFP is indicated for each analysis. Insets on the right corner show details of the cytosolic foci at higher magnification. Scale bars represent 10 µm.

foci are not P-bodies but adjacent cytoplasmic foci (Moreno et al., 2013). In order to clarify the subcellular localization of UPF1 and its potential co-localization with protein partners, we performed co-localization studies. We first used the full length genomic UPF1-RFP construct in transient expression assays and the results were then further refined by the quantification of co-localization signals using plants stably expressing UPF1-RFP and P-body markers. In transient expression assays, the UPF1-RFP fusion protein accumulated in the cytosol as well as in cytosolic foci, as previously described (Figure 13A). We first focused our co-localization assays on archetypal markers of P-bodies, DCP1 and DCP5, also found as UPF1 co-purified proteins. P-bodies are very dynamic (Figure S3) and fast P-body movement can have blurring effects, which alters the reliability of co-localization studies. We solved this problem by using specific detection settings, based on multitrack line-by-line acquisition instead of the commonly used frame-by-frame acquisition of the GFP and RFP signals (Figure S2A). Using these parameters, we identified a near-perfect co-localization of UPF1 with the P-body markers DCP1 and DCP5 both in transient assays and upon stable expression in *Arabidopsis* (Figure 13A-13B). This P-body localization was confirmed in time-lapse experiments monitoring the concomitant movement of UPF1 and the P-body marker DCP5 (Figure S3). We then quantified the co-localization of UPF1 with P-body components in stable lines, using an ImageJ macro to systematically identify cytoplasmic foci in our images and measure the extent of co-localization between foci in both GFP and RFP channels. The efficiency of this approach to discriminate between two distinct populations of cytoplasmic foci was confirmed using plants expressing the P-body marker DCP1 and the stress granule marker PAB2 (Figure S2B). This quantification demonstrated that UPF1 and the P-body markers DCP1 and DCP5 show near-perfect co-localization with between 90% and 95% co-localization respectively (Figure 13B). We also co-expressed UPF1 with RH6, RH8 and RH12, the three homologs of DDX6, a key component of P-bodies in mammals. Again, co-localization was perfect upon both transient and stable expression (Figures 13A-13B).

In addition, we tested the co-localization of UPF1 with five other factors that were enriched in UPF1 IPs: the Polypyrimidine Tract Binding factor PTB3, the Glycine Rich Protein GRP7, a putative endonuclease, a predicted terminal nucleotidyltransferase (TNTase) and an RNA helicase (RH11). All these proteins have known or putative RNA related functions. Of note, the alternative splicing of PTB3 and GRP7 mRNAs is regulated by a mechanism coupling PTB or GRP protein action with NMD (Stauffer et al., 2010; Schöning et al., 2008). For both RNA binding proteins PTB3 and GRP7, we noticed a co-localization with UPF1 in the cytosol only but no specific enrichment in cytosolic foci (Figure 13A). By contrast, for the endonuclease, the TNTase and RH11, we observed a co-localization in the cytosol as well as in cytoplasmic foci, previously identified as P-bodies (Figure 13A). In conclusion, two different co-localization patterns were identified

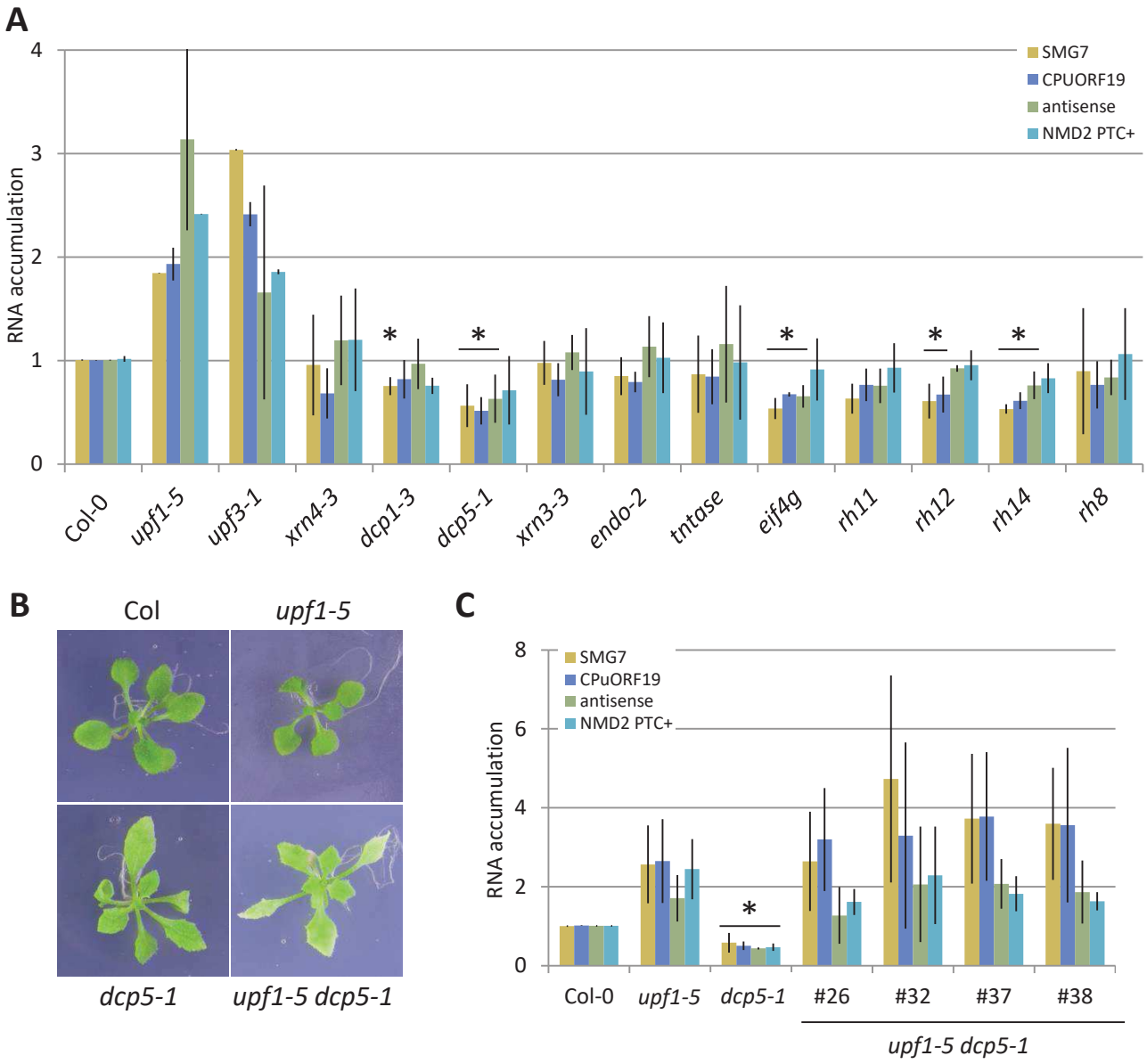


Figure 14. Functional analysis of mutants of UPF1 partners. (A) Real-time RT-PCR analysis of known NMD target transcripts in mutants in UPF1 partners. (B) Phenotype of the *upf1-5 dcp5-1* double mutant (20 days old seedlings). (C) Real-time RT-PCR analysis of known NMD target transcripts in *upf1-5 dcp5-1*. All Real-time RT-PCR results are normalized to *ACTIN2* and presented as mean of FC and arbitrary set to 1 in Col-0 ± SEM. Asterisk represents significant differences with the control Col-0, $p<0.025$; n=3.

here for UPF1 and associated factors: either a cytosolic-only pattern, or a dual localization in both the cytosol and P-bodies. These localizations may be linked to specific functions of these partners in these different subcellular compartments.

Mutations in UPF1 and DCP5 have opposite effects on NMD targets accumulation

To determine the influence of UPF1 partners on NMD, we tested the accumulation of NMD targets in several T-DNA mutants of UPF1 co-purifying factors. A total of 22 mutants were tested for the accumulation of four NMD targets with distinct NMD activating features corresponding to an upstream ORF, a long 3'-UTR with or without stimulating intron and a splicing variant containing a PTC. The goal was to identify mutations in either NMD activators or inhibitors, if the mutation leads to enhanced or decreased NMD target accumulation, respectively. To our surprise none of the mutation tested led to increased NMD target accumulation (Figure 14A and S4). This result suggests that none of the partners tested is a limiting factor for NMD. In contrast, we identified several mutations for which a significant decrease in NMD targets was observed, this was the case for *dcp5*, *rh14* and *eif4g* mutants. We focused on *dcp5*, a mutant showing robust molecular effect (Figure 14A-14C), and defective for DCP5, a validated protein partner of UPF1 localized in P-bodies. To test the genetic interaction between *upf1* and *dcp5*, we crossed the mutants and selected double mutants. The *dcp5 upf1* plants harbored stronger developmental defects compared to the two single mutants with thinner leaves, more serrations and a pale color (Figure 14B). The reduced NMD target accumulation in *dcp5* could either be due to increased UPF1 activity or to an indirect effect of *dcp5*. In order to discriminate between these two hypotheses, we tested the steady state levels of NMD targets in *dcp5 upf1*. If the reduction in *dcp5* is independent on UPF1, the reduction of NMD targets accumulation should be observed in both *dcp5* and *dcp5 upf1*. In contrast, if the reduction of NMD targets accumulation in *dcp5* is caused by an exacerbated action of UPF1, this reduction should be abolished in the double mutant. According to the latter hypothesis, we did not observe a reduction of NMD targets accumulation comparing *upf1* and *dcp5 upf1* (Figure 14C). This result suggests that the decrease of NMD target accumulation in *dcp5* is dependent on UPF1 function.

UPF1 and DCP5 are part of similar mRNPs and directly interact with DDX6 homologs

To better understand the molecular basis of the genetic and biochemical links between UPF1 and DCP5, we used IPs coupled to mass spectrometry to identify the network of DCP5 associated factors (Figure 15A-

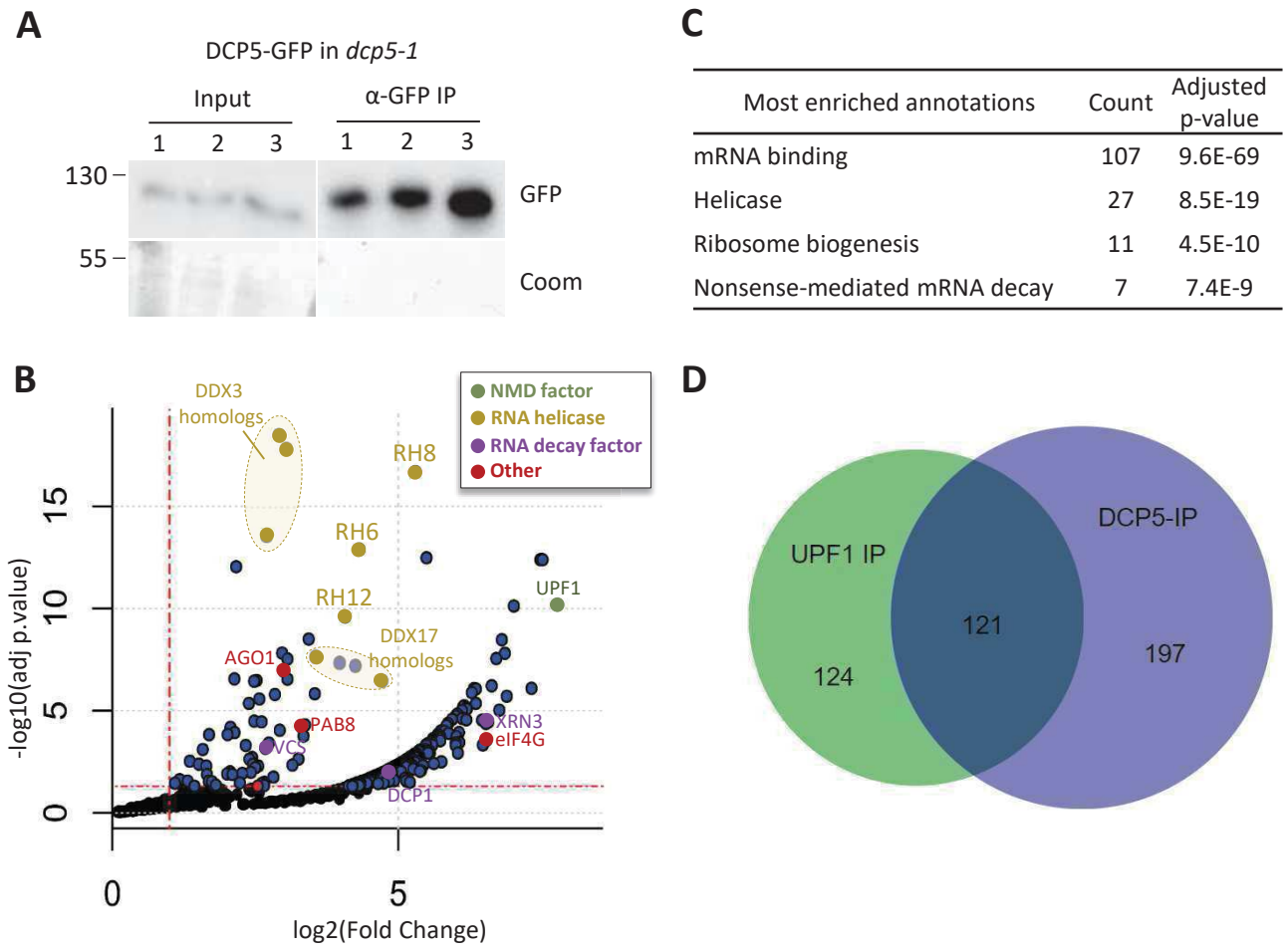


Figure 15. Overlap between DCP5 and UPF1 interactome. (A) Western blot analysis of DCP5-GFP fusion in DCP5-GFP IPs from *dcp5-1* complemented lines. The three replicates of IP fractions were used for mass spectrometry analysis. (B) Volcano plot representing the proteins enriched in DCP5 IPs Adjusted p-value is presented on the y axis, Fold Change (FC) in x axis. The red dotted lines represent the significance thresholds of *p<0,05* and log2(FC)<1 (C) Venn diagram showing the overlap between UPF1 and DCP5 co-purified partners. IP: immunoprecipitation.

15B). This approach identified 318 proteins significantly enriched in DCP5 IPs (Figure 15B, Table S5). A global annotation analysis revealed an enrichment in terms related to RNA binding, RNA helicases, as well as ribosome biogenesis and Nonsense-Mediated Decay (Figure 15C). As observed for UPF1, 40% of these proteins were present in the experimentally determined At-RBP proteome (Figure S5, Table S6) (Reichel et al., 2016). Strikingly, UPF1 is among the most significantly enriched proteins co-purifying with DCP5. Importantly, we identified that the overlap between UPF1 and DCP5 interactome covers 49% of UPF1 co-purified partners (Figure 15D, Table S7). This result indicated that UPF1 and DCP5 likely participate to similar mRNP complexes localized in P-bodies as observed by confocal microscopy (Figure 13A-13B). Among common partners of DCP5 and UPF1, we noticed multiple families of RNA helicase including the three DDX6 homologs among the most prominent partners considering both adjusted p-value and fold change (RH6, 8, 12, Figure 15B). DDX6 is a major component of P-bodies and is strictly required for their formation in human (Ayache et al., 2015), suggesting that the presence of UPF1 and DCP5 in P-bodies could be due to their concurrent interaction with homologs of this P-body organizing factors. The fact that DDX6 homologs are among the best-ranked UPF1 and DCP5 co-purified proteins licenses them as good candidates for direct UPF1 and DCP5 protein partners. To test this hypothesis, we used the yeast two-hybrid system and looked for possible direct interactions between UPF1, DCP5 and DDX6 homologs (Figure 16). As suggested by the RNA-dependent interaction, we observe no growth on selective media for co-transformation of UPF1 and DCP5, indicating that UPF1 does not directly interact with DCP5 (Figure 16). In contrast, significant yeast growth was observed on selective media when combining UPF1 with the RH8 protein, indicating a direct interaction. A similar result was obtained when combining DCP5 and RH6 (Figure 16). Therefore, both UPF1 and DCP5 are directly linked to DDX6 homologs.

Overall, our results reveal a network of interactions connecting UPF1 to the translation inhibitors, DCP5 and DDX6 homologs: DDX6 homologs directly contact UPF1 and mediate the interaction between UPF1 and DCP5. UPF1 together with DCP5 and DDX6 homologs are major constituents of P-bodies. In light of the main role of P-bodies in condensing translationally repressed mRNPs rather than corresponding to mRNA degradation sites (Hubstenberger et al., 2017; Horvathova et al., 2017; Schütz et al., 2017), we conclude that DCP5 and DDX6 recruitments by UPF1 protect NMD targets from degradation, explaining the downregulation of different types of NMD targets in *dcp5* mutants (Figure 17).

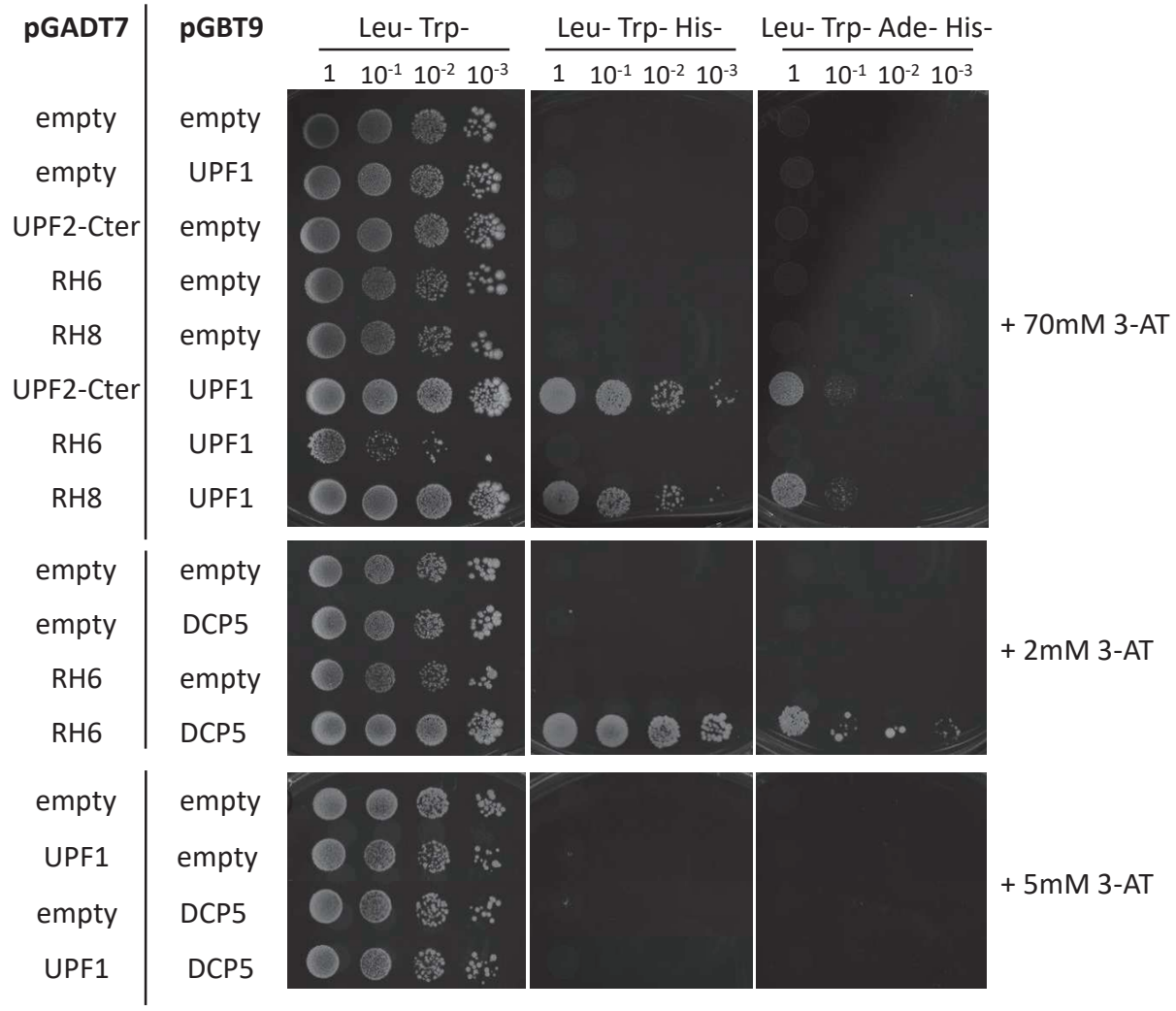


Figure 16. Yeast two-hybrid interaction between DDX6 homologs, UPF1 and DCP5. Specific growth on selective media for the UPF1-RH8 and DCP5-RH6 combination highlights the direct interaction between these factors.

Discussion

Conserved and specific hallmarks of the plant UPF1 interactome

Our unbiased IP coupled to mass spectrometry approach validates the conservation of the core NMD complex and provides a list of other putative direct UPF1 interacting factors in *Arabidopsis*. Among these factors we could demonstrate the direct interaction between UPF1 and homologs of the translation repressor DDX6. The interaction network of UPF1 with the translation inhibitors DCP5 and DDX6 homologs that are central components of P-bodies, combined with the observation that several NMD targets are downregulated in *dcp5* mutants, led us to propose a model involving DCP5 and DDX6 in NMD antagonist functions (Figure 17).

We provide here an extended overview of the UPF1 protein interaction network in *Arabidopsis*. Among other UPF1 partner, we validated UPF1 co-purification with DCP5, DCP1, AGO1 and PAB2. Interestingly UPF1 interactions with homologs of these partners were also observed in other species (Lejeune et al., 2003; Flury et al., 2014; Swisher and Parker, 2011; Höck et al., 2007). In human, co-purification of UPF1 with DCP2 was proposed to be a connection of NMD with downstream RNA degradation pathways as downregulating DCP2 decreased the degradation rate of NMD targets (Lejeune et al., 2003). Interestingly, in yeast direct interaction between Edc3 (DCP5 homolog) and Upf1 N-terminal domain was reported (Swisher and Parker, 2011). Loss of Edc3 led to heightened interaction between Upf1 and Upf2 supporting the conclusion that Edc3 and Upf2 independently interact at partially overlapping sites on Upf1. Among other UPF1-copurified partners, we found AGO1, a key RNA silencing component in plants (Borges and Martienssen, 2015). Interestingly, UPF1 was identified as an RNA-independent interactor of AGO1 and AGO2 in human cells, but the rational for this interaction remained unclear as UPF1 knockdown did not affect miRNA targeted reporter constructs (Höck et al., 2007). Similarly to these results, no specific overlap between miRNA and NMD targets were found by comparing miRNA and NMD targets in *Arabidopsis* at the genome-wide scale (Zhang et al., 2013). Nevertheless, this interaction could have other impacts on the miRNA pathway, as revealed by a recent report in human describing a function of UPF1 in inducing Tudor-SN-mediated miRNA decay (Elbarbary et al., 2017). Despite the fact that such pathway has not been described in plants, our identification of AGO1 as potential direct UPF1 partner encourages the verification of such a function in *Arabidopsis*. Regarding the link with the PABPs, they were found to exert an inhibitory effect on NMD in human in a position dependent manner, which is consistent with their presence in our UPF1 purifications (Ivanov et al., 2008).

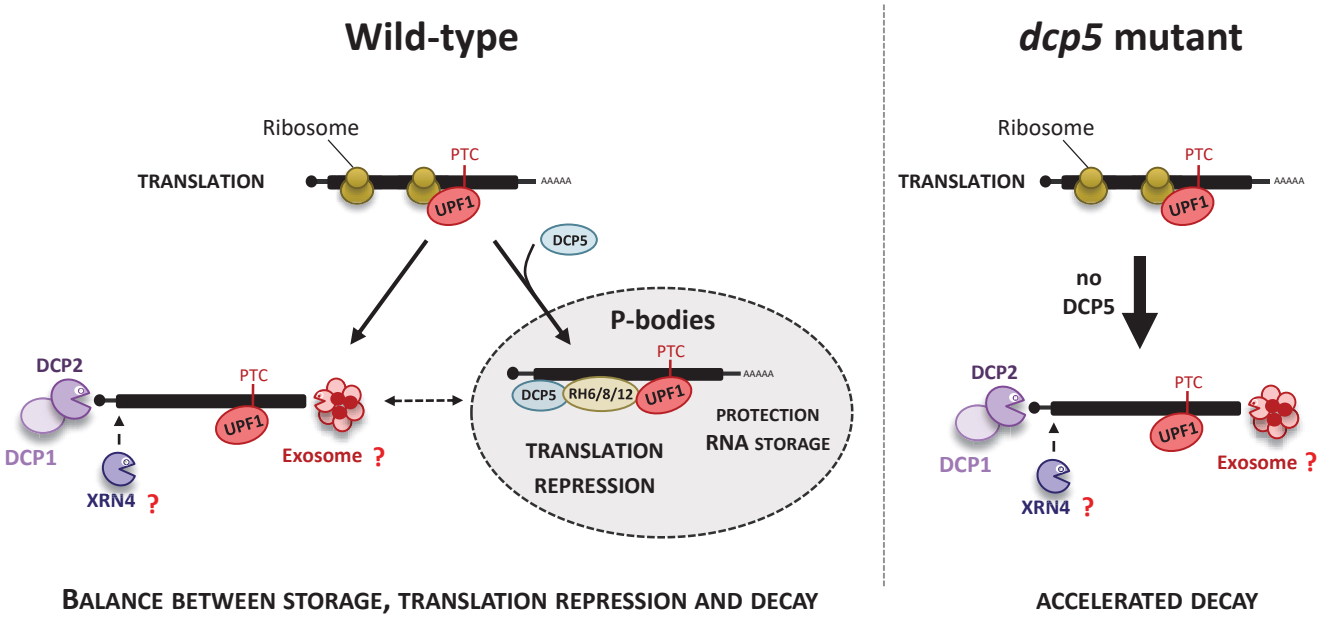


Figure 17. Integrated model of the link between DCP5 and UPF1 in NMD. We represented an archetypal NMD target containing a PTC. It is recognized by UPF1 during translation, stimulating accelerated decay possibly through exoribonucleases XRN4 and/or the exosome activities (left panel). Upon the binding of DCP5 and/or DDX6 homologs, a fraction of the transcripts is relocalized into P-bodies and escapes this accelerated decay possibly by being translationally repressed. This UPF1-DCP5-DDX6 molecular network influences the fate of the PTC containing transcript. Upon reduced DCP5 action the proportion this protection is alleviated leading to increased UPF1 dependent NMD target degradation (right panel).

Our immunoprecipitation experiments led to co-localization studies, which allowed the identification of three novel plant P-body components: the RNA helicases RH11, a predicted TNTase and a putative endonuclease. We also show that all three DDX6 homologs, RH6, RH8 and RH12, are also localized in P-bodies, confirming a prior result for RH12 (Bhullar et al., 2017). All these proteins were retrieved in both UPF1 and DCP5 IPs. Although not all proteins that co-purify with DCP5 and UPF1 are P-body components, this common protein set is valuable for de novo identification of P-body components. These proteins could play specific roles in RNA decay, storage or translation inhibition. The detailed analysis of some of the identified factors could provide new insights into P-body functions in plants. Overall, our experimental approach gives access to an enlarged UPF1 interacting network, providing an interesting resource for the identification of proteins involved in RNA metabolism, P-body function and NMD.

Genetic and proteomic links between UPF1 and DCP5

Our analysis of mutants in UPF1 interacting factors identified a significant decrease in NMD targets accumulation in the *dcp5* mutant (Figure 14A, Figure S4). Using *dcp5 upf1* double mutants we determined that this molecular phenotype depends on UPF1 action (Figure 14C). As DCP5 is a known translation repressor in plant (Xu and Chua, 2009), we propose that DCP5 can exert a protective effect on NMD targets, through translation inhibition and possible sequestration of NMD targets in P-bodies (Figure 17). In this model, the mutation in *dcp5* reduces this translation inhibition and favors the translation-dependent UPF1-mediated decay (Figure 17). We found that homologs of the DDX6/Dhh1 translational repressors, required for P-body assembly in mammals (Ayache et al., 2015; Rajyaguru et al., 2012; Presnyak and Coller, 2013b; Xu and Chua, 2009), directly interact with UPF1 and DCP5. We propose that the newly identified UPF1-DDX6-DCP5 interaction network influences the fate of NMD targets while they transit into P-bodies. UPF1 has been shown to trigger translational repression during NMD in mammals (Isken et al., 2008), UPF1 interaction with translation repressors suggests the possible conservation of this function in plants. An alternative and non-mutually exclusive explanation would be that DDX6-DCP5 competes with UFP2 for UPF1 binding. When DCP5 function is reduced in *dcp5* mutant, this competition would be in favor of the formation of the UPF1-UPF2-UPF3 complex, stimulating degradation instead of storage/translation repression. This equilibrium between UPF2 and DDX6-DCP5 binding could also be influenced by the disruption of P-bodies in *dcp5* (Xu and Chua, 2009). Interestingly, such a competition was observed in yeast, in which the DCP5 homolog Edc3, directly competes with Upf2 for binding the CH domain of Upf1 (Swisher and Parker, 2011).

In plant, 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). It would be very interesting to determine the possible involvement of the translation repressors DCP5 and the DDX6 in this process.

Central function of P-bodies in translation repression

We identified in this study the proteome associated with two components of plant P-bodies, UPF1 and DCP5. The use of fluorescence-activated particle sorting (FAPS), recently allowed the purification of P-bodies and identified the associated proteome and transcriptome from P-bodies in human epithelial cells. This study revealed that P-bodies are associated to translationally repressed mRNAs rather than mRNA undergoing decay as previously proposed (Hubstenberger et al., 2017). The model we favor here to explain the influence of *dcp5* on NMD targets is coherent with this proposed function of P-bodies. It is reinforced by our results showing direct interaction between UPF1 and the translational repressors of the DDX6 family. One argument against this model could be that P-bodies were shown to be dispensable for the regulation of an endogenous NMD target in Drosophila (Eulalio et al., 2007). Nevertheless, in this report the inactivation of the DCP5 homologs Edc3 and Tral, were not tested. In addition, this report only addresses the importance of detectable microscopic P-bodies. Whether processes taking place in P-bodies could also occur in submicroscopic aggregates or associated with diffuse protein complexes in the absence of detectable microscopic P-bodies remains to be determined. Similarly, the action of DCP5 on NMD target accumulation could occur both within and outside P-bodies. In *Arabidopsis*, *dcp5* mutant was shown to disrupt P-bodies formation (Xu and Chua, 2009), suggesting that NMD targets release from DCP5 and P-bodies could lead to the increased action of UPF1 as proposed in our model (Figure 17). Our study also revealed a direct interaction between UPF1 and DDX6 homologs in plants. In addition to its function in RNA surveillance, NMD clearly emerges as an important mechanism involved in the control of the expression of functional genes, important for plant immunity and stress response in plants (Gloggnitzer et al., 2014; Lorenzo et al., 2017; Drechsel et al., 2013). The P-body localization identified here would fit with such a regulatory function, as P-bodies were recently shown to preferentially contain translationally arrested transcripts coding for regulatory proteins (Hubstenberger et al., 2017). This is in clear opposition with the mRNA composition of stress granules, in which most mRNAs accumulate without any functional distinctions (Khong et al., 2017). Translation repression of NMD targets in P-bodies could allow a more flexible regulation of these regulatory transcripts, allowing their inactivation and storage to possibly be reengaged into the translation process in response to environmental changes.

SUPPLEMENTARY DATA

Supplementary Data are available following the manuscript.

ACKNOWLEDGEMENT

The authors thank C. Bousquet-Antonelli and JM. Deragon for sharing pDCL1:DCL1:YFP, PAB2-RFP constructs and transgenic lines; R. Wagner's team for plant care; members of the RNA degradation team and of the IBMP for critical reading of the manuscript.

FUNDING

This work was 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; French Ministry of Research [C.C.].

CONFLICT OF INTEREST

Conflict of interest statement. None declared.

Supplementary figures Chapter I

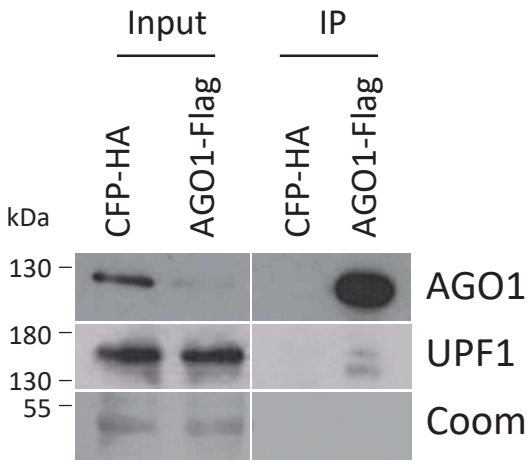


Figure S1. Reverse co-immunoprecipitation of AGO1. Western blot detection of UPF1 in AGO1-Flag IPs. AGO1 and UPF1 protein levels were analyzed using antibodies raised against endogenous proteins. IP: immunoprecipitation.

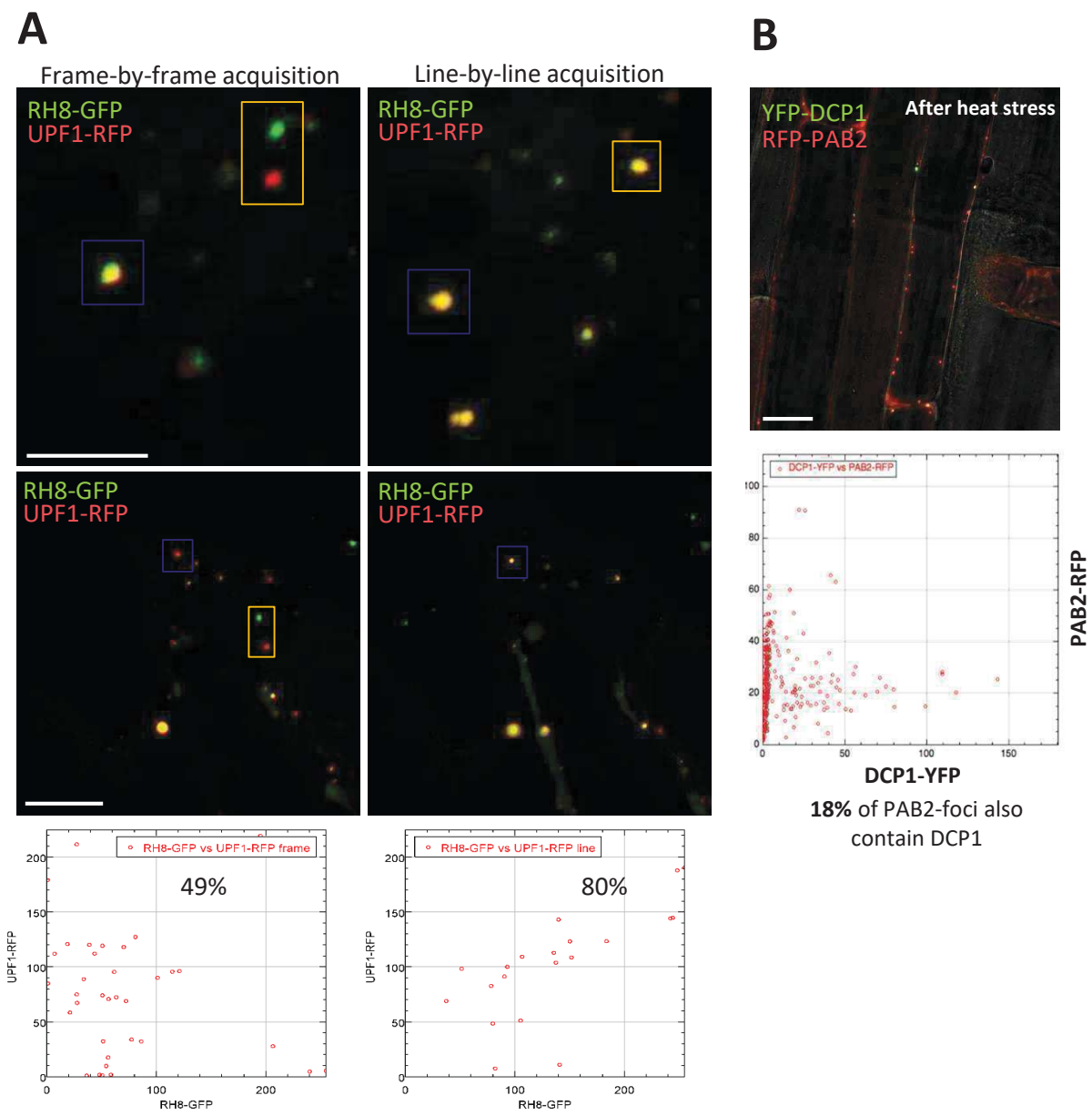


Figure S2. Line-by-line acquisition prevents imaging artefacts for fast moving objects. (A) Confocal microscopy images of UPF1-RFP and RH8-GFP fusions transiently co-expressed in *N. benthamiana* leaves. Two sets of the same confocal section are acquired sequentially using frame-by-frame acquisition and line-by-line acquisition to image the same foci (identical foci are highlighted by colored insets). Two distinct or docked foci are detected for frame-by-frame acquisition while line-by-line prevents this imaging artefact. Dot plot representing the quantification of the signal in both GFP and RFP channels in x or y axis respectively are represented under confocal images. Each plot was created following the analysis of the foci composition of the two presented identical confocal sections for frame-by-frame and line-by-line acquisition. The colocalization percentage between RH8-GFP and UPF1-RFP is indicated for each acquisition. Pictures taken at 4 days after infiltration. (B) Foci quantification is able to distinguish between two distinct populations of foci. Confocal microscopy images of epidermal root cells of stable *Arabidopsis* lines expressing both the stress granule marker PAB2-RFP and the P-body marker YFP-DCP1. Dot plot representing the quantification of the signal in both YFP and RFP channels in x or y axis respectively are represented under each confocal image. Each plot was acquired following the analysis of the foci composition of 10 randomly chosen confocal sections. The colocalization percentage between P-body (YFP-DCP1) and stress granules (RFP-PAB2) is indicated for each analysis. Scale bars represent 10 µm.

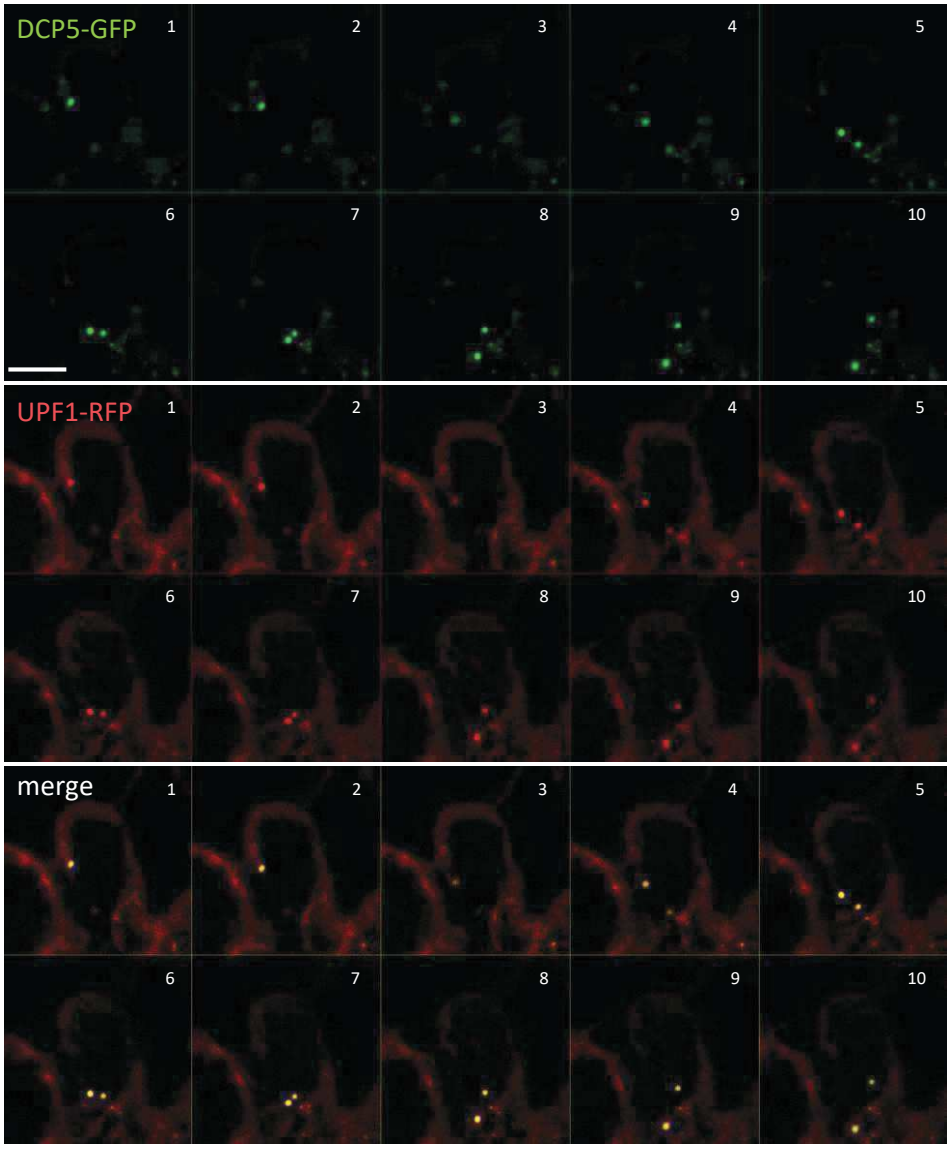


Figure S3. Time-lapse acquisition of the co-localization between UPF1-RFP and DCP5-GFP. Confocal microscopy time-lapse images of UPF1-RFP and DCP5-GFP fusions transiently co-expressed in *N. benthamiana* leaves. Pictures taken at 4 days after infiltration.

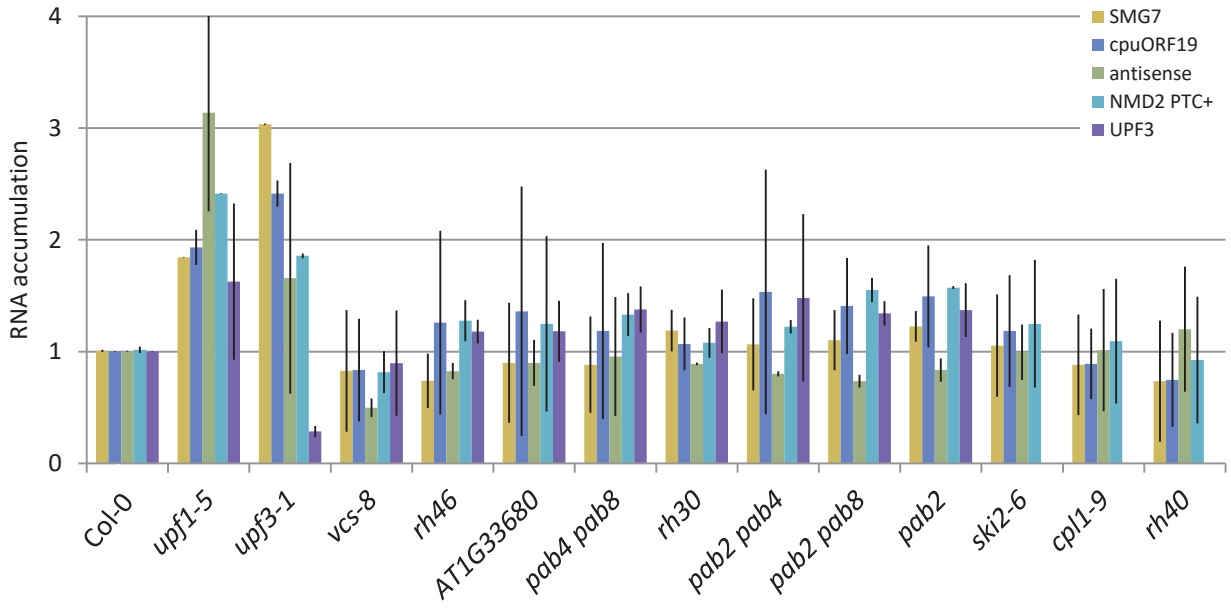


Figure S4. Functional analysis of mutants of UPF1 partners. (A) Real time RT-PCR analysis of known NMD target transcripts in mutants in UPF1 partners Real-time RT-PCR results are normalized to *ACTIN2* and presented as mean of FC and arbitrary set to 1 in Col-0 ± SEM. Asterisk represents significant differences with the control Col-0, *p<0.025*; n=3.

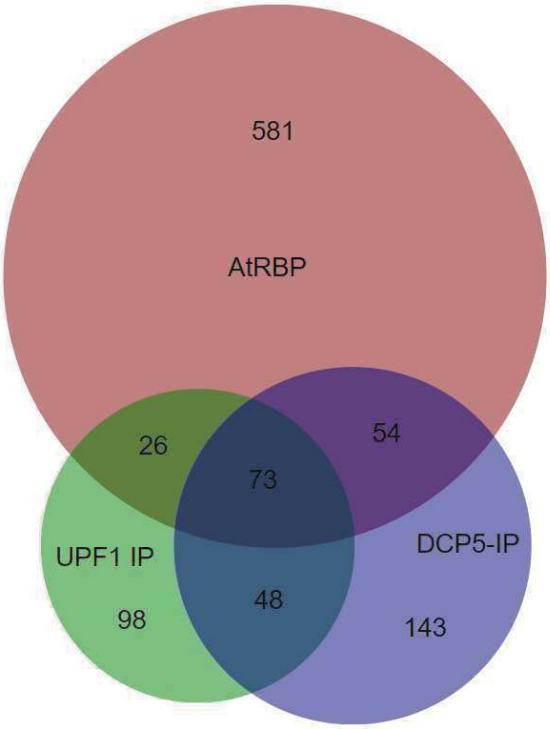


Figure S5. Comparison between UPF1 IPs, DCP5 IPs and the At-RBP proteome. Venn diagram showing the overlap between UPF1-IPs, DCP5-IPs and the At-RBP proteome identified in Reichel *et al.*, 2016.

Table S1. Proteins enriched in UPF1-IPs. In this table are indicated the informations used to create the volcano plot in Chapter I, Figure 11C (only proteins with an *adjp<0.08* are listed here). Mean of spectral counts for control or UPF1 IPs, log2(Fold Change), adjusted p-value (adjp) are detailed. DEP indicates whether the considered protein passed the statistical test (TRUE) or not (FALSE). Proteins in grey are not significantly enriched in UPF1-IPs.

AGI	Mean spectral count **Control**	Mean spectral count **UPF1**	LogFC	adjp	DEP	Description
AT5G47010	0	2230	13,75	6,8E-292	TRUE	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in
AT2G39260	0	89,6	9,339	1,36E-23	TRUE	"binding;RNA binding; FUNCTIONS IN: RNA binding, b
AT3G58510	6,3	95	2,766	5,59E-23	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT4G00660	0,3	36,4	5,233	1,94E-18	TRUE	"RNAhelicase-like 8 (RH8); FUNCTIONS IN: helicase
AT2G42520	6,7	89,1	2,641	4,16E-15	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT3G22330	0,5	29,9	4,543	5,42E-14	TRUE	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT3G58570	6,2	86,9	2,77	1,28E-13	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT1G49760	0,3	39,9	5,644	3,42E-13	TRUE	"poly(A) binding protein 8 (PAB8); FUNCTIONS IN: R
AT2G23350	0,3	31,5	5,136	1,09E-12	TRUE	"poly(A) binding protein 4 (PAB4); FUNCTIONS IN: R
AT2G45810	0,8	26,6	3,645	1,14E-10	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G01540	1,8	38,6	3,266	1,63E-10	TRUE	"DEAD box RNA helicase 1 (DRH1); FUNCTIONS IN: ATP
AT3G61240	0,8	30,4	3,863	2,09E-10	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT1G33980	0	17,6	6,846	3,89E-10	TRUE	"UPF3; FUNCTIONS IN: nucleotide binding; INVOLVED
AT4G34110	0,8	41,8	4,461	6,2E-10	TRUE	"poly(A) binding protein 2 (PAB2); FUNCTIONS IN: R
AT5G14610	1,8	29	2,777	1,15E-09	TRUE	"DEAD box RNA helicase family protein; FUNCTIONS I
AT3G10770	0	20,6	7,337	1,97E-09	TRUE	"Single-stranded nucleic acid binding R3H protein;
AT5G61020	0	19	7,201	1,97E-09	TRUE	"evolutionarily conserved C-terminal region 3 (ECT
AT3G13920	8,3	81,6	1,984	2,94E-09	TRUE	"eukaryotic translation initiation factor 4A1 (EIF
AT1G54270	8	84,5	2,092	8,38E-09	TRUE	"eif4a-2 (EIF4A-2); FUNCTIONS IN: ATP-dependent he
AT5G05100	0	13,9	6,411	5,16E-08	TRUE	"Single-stranded nucleic acid binding R3H protein;
AT5G40490	0,3	16,2	4,137	2,05E-07	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G04280	0	9,5	5,965	4,9E-07	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein w
AT3G13460	0,8	34,4	4,532	5,06E-07	TRUE	"evolutionarily conserved C-terminal region 2 (ECT
AT1G29250	1,8	28,9	3,144	3,9E-06	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G06850	2,5	27,6	2,184	5,07E-06	TRUE	"xyloglucan endotransglucosylase/hydrolase 4 (XTH4
AT3G13300	0,5	16,1	3,736	5,34E-06	TRUE	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT3G22310	0,2	10,4	4,23	6,99E-06	TRUE	"putative mitochondrial RNA helicase 1 (PMH1); CON
AT4G20360	6,3	42,6	1,525	7,16E-06	TRUE	"RAB GTPase homolog E1B (RABE1b); FUNCTIONS IN: GT
AT2G15560	0	9,4	6,287	8,03E-06	TRUE	"Putative endonuclease or glycosyl hydrolase; BEST
AT5G63120	0,8	17,6	3,139	9,57E-06	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT1G71770	0	8,1	5,814	1,28E-05	TRUE	"poly(A)-binding protein 5 (PAB5); FUNCTIONS IN: R
AT1G22760	0	9,1	5,896	1,49E-05	TRUE	"poly(A) binding protein 3 (PAB3); CONTAINS InterP
AT3G06480	1	17,1	2,944	1,57E-05	TRUE	"DEAD box RNA helicase family protein; FUNCTIONS I
AT1G08370	0	9,1	5,778	2,1E-05	TRUE	"decapping 1 (DCP1); CONTAINS InterPro DOMAIN/s: D
AT5G26742	6,8	78	2,076	2,63E-05	TRUE	"embryo defective 1138 (emb1138); FUNCTIONS IN: in
AT2G17870	0	7,4	5,418	2,72E-05	TRUE	"cold shock domain protein 3 (CSP3); FUNCTIONS IN:
AT5G53060	0	11,1	5,87	2,72E-05	TRUE	"RNA-binding KH domain-containing protein; FUNCTIO
AT4G34980	3,5	28,5	1,86	2,87E-05	TRUE	"subtilisin-like serine protease 2 (SLP2); FUNCTIO
AT1G43190	0	7,6	5,467	4,31E-05	TRUE	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT2G34160	0	9,5	6,082	4,42E-05	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G29200	0	6	5,268	0,00005	TRUE	"pumilio 1 (PUM1); FUNCTIONS IN: RNA binding, bind
AT1G54080	0,2	9,1	4,011	7,27E-05	TRUE	"oligouridylate-binding protein 1A (UBP1A); FUNCTI
AT1G14710	0	18,1	6,384	9,08E-05	TRUE	"hydroxyproline-rich glycoprotein family protein;
AT2G29140	0	5,9	5,225	9,48E-05	TRUE	"pumilio 3 (PUM3); FUNCTIONS IN: RNA binding, bind
AT4G35850	0	6	5,278	9,77E-05	TRUE	"Pentatricopeptide repeat (PPR) superfamily protei
AT1G07310	0	12,9	6,038	9,89E-05	TRUE	"Calcium-dependent lipid-binding (CaLB domain) fam
AT3G03950	0	6,9	5,312	0,000103	TRUE	"evolutionarily conserved C-terminal region 1 (ECT
AT1G18660	0	7,4	5,363	0,000104	TRUE	"zinc finger (C3HC4-type RING finger) family prote
AT3G15010	0	6,4	5,346	0,000106	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G38680	1	35,8	3,622	0,000118	TRUE	"glycine rich protein 2 (GRP2); FUNCTIONS IN: doub
AT4G10610	0	6,4	5,238	0,00013	TRUE	"CTC-interacting domain 12 (CID12); CONTAINS Inter
AT3G60240	0	19,9	6,494	0,000133	TRUE	"eukaryotic translation initiation factor 4G (EIF4
AT1G48410	1,5	34,6	3,295	0,000133	TRUE	"ARGONAUTE 1 (AGO1); FUNCTIONS IN: protein binding
AT3G51950	0	7,8	5,392	0,000138	TRUE	"Zinc finger (CCCH-type) family protein / RNA reco
AT1G70710	0	8,9	5,526	0,000145	TRUE	"glycosyl hydrolase 9B1 (GH9B1); FUNCTIONS IN: cel
AT1G26110	0	9,4	6,607	0,00015	TRUE	"decapping 5 (DCP5); CONTAINS InterPro DOMAIN/s: D
AT1G32790	0	6,5	5,255	0,000156	TRUE	"CTC-interacting domain 11 (CID11); FUNCTIONS IN:
AT5G44780	0	6	5,225	0,000174	TRUE	"unknown protein; BEST Arabidopsis thaliana protei
AT5G13530	0	6,9	5,265	0,000215	TRUE	"KEEP ON GOING (KEG); FUNCTIONS IN: ubiquitin-prot
AT1G75660	0	10	5,651	0,000217	TRUE	"5'-3' exoribonuclease 3 (XRN3); CONTAINS InterPro
AT3G12915	0	23,2	6,703	0,000278	TRUE	"Ribosomal protein S5/Elongation factor G/III/V fa
AT3G14100	0,2	8,9	3,937	0,000328	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G22010	7,7	50,6	1,704	0,000342	TRUE	"SKU5 similar 4 (sks4); FUNCTIONS IN: oxidoreduct
AT3G07010	0	5	4,957	0,000346	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT2G29190	0	4,6	4,915	0,000367	TRUE	"pumilio 2 (PUM2); FUNCTIONS IN: mRNA binding, RNA
AT5G56000	0	12,6	5,913	0,000453	TRUE	"HEAT SHOCK PROTEIN 81.4 (Hsp81.4); FUNCTIONS IN:
AT2G39800	0	5,9	5,074	0,000465	TRUE	"delta1-pyrroline-5-carboxylate synthase 1 (P5CS1)
AT2G27100	0	6,1	5,242	0,000468	TRUE	"SERRATE (SE); FUNCTIONS IN: DNA binding, sequence
AT1G64390	0	5,8	5,249	0,000482	TRUE	"glycosyl hydrolase 9C2 (GH9C2); FUNCTIONS IN: car
AT4G18480	0	6,9	5,172	0,000568	TRUE	"CHLI1; FUNCTIONS IN: magnesium chelatase activity
AT3G26420	0	5,9	5,462	0,000568	TRUE	"ATRZ-1A; CONTAINS InterPro DOMAIN/s: RNA recognit
AT4G25550	0,5	9,4	2,791	0,000621	TRUE	"Cleavage/polyadenylation specificity factor, 25kD
AT3G23990	0	5,9	5,019	0,000628	TRUE	"heat shock protein 60 (HSP60); FUNCTIONS IN: copp
AT1G23490	0	5,6	4,961	0,000635	TRUE	"ADP-ribosylation factor 1 (ARF1); FUNCTIONS IN: p
AT4G27060	0	6,8	5,149	0,000715	TRUE	"TORTIFOLIA 1 (TOR1); CONTAINS InterPro DOMAIN/s:
AT5G26360	0	8,2	5,358	0,000746	TRUE	"TCP-1/cpn60 chaperonin family protein; FUNCTIONS
AT3G04610	0	7,4	5,313	0,000753	TRUE	"flowering locus KH domain (FLK); FUNCTIONS IN: RN
AT1G76010	1,3	15,9	2,321	0,000785	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle

AT5G11170	0	5,5	4,919	0,00079	TRUE	"DEAD/DEAH box RNA helicase family protein ; FUNCT
AT5G48650	0	6,1	5,128	0,00079	TRUE	"Nuclear transport factor 2 (NTF2) family protein
AT3G07030	0	5,5	4,916	0,000879	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G11650	0,2	7,6	3,66	0,000891	TRUE	"RBP45B; FUNCTIONS IN: RNA binding; EXPRESSED IN:
AT5G35360	0	5,8	4,967	0,000953	TRUE	"acetyl Co-enzyme a carboxylase biotin carboxylase
AT1G27430	0	5,8	4,971	0,000965	TRUE	"GYF domain-containing protein; CONTAINS InterPro
AT5G24770	0	5,9	4,99	0,000969	TRUE	"vegetative storage protein 2 (VSP2); FUNCTIONS IN
AT4G27000	0,3	7,8	3,087	0,001024	TRUE	"ATRBP45C; CONTAINS InterPro DOMAIN/s: RNA recogni
AT4G03110	0	4,4	4,825	0,001024	TRUE	"RNA-binding protein-defense related 1 (RBP-DR1);
AT4G29040	0	7,8	5,243	0,001117	TRUE	"regulatory particle AAA-ATPase 2A (RPT2a); FUNCTI
AT5G08680	0	10,5	5,698	0,001157	TRUE	"ATP synthase alpha/beta family protein; FUNCTIONS
AT2G30950	0	4,9	4,776	0,001159	TRUE	"VARIEGATED 2 (VAR2); FUNCTIONS IN: metallopeptida
AT2G31810	0	5,5	4,905	0,001175	TRUE	"ACT domain-containing small subunit of acetolacta
AT2G26280	0	4,1	4,718	0,001175	TRUE	"CID7; FUNCTIONS IN: damaged DNA binding, protein
AT3G15000	0	5,1	5,186	0,001189	TRUE	"cobalt ion binding; FUNCTIONS IN: cobalt ion bind
AT5G37720	0	4,8	5,007	0,001213	TRUE	"ALWAYS EARLY 4 (ALY4); FUNCTIONS IN: nucleotide b
AT3G49390	0	4,5	4,78	0,001239	TRUE	"CTC-interacting domain 10 (CID10); FUNCTIONS IN:
AT5G54900	0	3,9	4,587	0,001245	TRUE	"RNA-binding protein 45A (RBP45A); FUNCTIONS IN: R
AT5G22060	0	4,1	4,626	0,001307	TRUE	"DNAJ homologue 2 (J2); FUNCTIONS IN: unfolded pro
AT2G42570	0	5,9	4,985	0,001317	TRUE	"TRICHOME BIREFRINGENCE-LIKE 39 (TBL39); INVOLVED
AT1G13190	0	4	4,684	0,001463	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT2G46780	0	3,8	5,096	0,001466	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G47700	0	5,9	4,998	0,00147	TRUE	"60S acidic ribosomal protein family; FUNCTIONS IN
AT5G55670	0	3,9	4,76	0,001571	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G13060	0	4,6	4,786	0,001599	TRUE	"evolutionarily conserved C-terminal region 5 (ECT
AT2G32700	0	7,5	5,167	0,001729	TRUE	"LEUNIG_homolog (LUH); CONTAINS InterPro DOMAIN/s:
AT1G33680	0	7,4	5,171	0,001792	TRUE	"KH domain-containing protein; FUNCTIONS IN: RNA b
AT4G21670	0	4,5	4,819	0,001838	TRUE	"C-terminal domain phosphatase-like 1 (CPL1); FUNC
AT3G13290	0,3	6,9	3,022	0,001851	TRUE	"varicose-related (VCR); FUNCTIONS IN: nucleotide
AT4G21710	0	5,4	4,858	0,001892	TRUE	"NRPB2; CONTAINS InterPro DOMAIN/s: DNA-directed R
AT1G21540	0	4,5	4,661	0,001946	TRUE	"AMP-dependent synthetase and ligase family protei
AT5G55070	0	4,6	4,675	0,002061	TRUE	"Dihydrolipoamide succinyltransferase; FUNCTIONS I
AT5G20020	0	5	4,8	0,002093	TRUE	"RAS-related GTP-binding nuclear protein 2 (RAN2);
AT3G02530	0	5,9	4,906	0,002167	TRUE	"TCP-1/cpn60 chaperonin family protein; FUNCTIONS
AT3G49490	0	3,9	4,881	0,002178	TRUE	"unknown protein; Has 722 Blast hits to 186 protei
AT5G58470	0	4,6	4,962	0,002178	TRUE	"TBP-associated factor 15B (TAF15b); FUNCTIONS IN:
AT5G16070	0	5,4	4,819	0,002178	TRUE	"TCP-1/cpn60 chaperonin family protein; FUNCTIONS
AT1G24510	0	5,9	4,915	0,002211	TRUE	"TCP-1/cpn60 chaperonin family protein; FUNCTIONS
AT2G32730	0	5,2	4,781	0,002358	TRUE	"26S proteasome regulatory complex, non-ATPase sub
AT5G43330	0	4,4	4,637	0,002358	TRUE	"Lactate/malate dehydrogenase family protein; FUNC
AT1G02800	0	5,2	4,788	0,002496	TRUE	"cellulase 2 (CEL2); FUNCTIONS IN: cellulase activ
AT2G33410	0,2	6,1	3,397	0,002521	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G01300	3	23,8	1,814	0,002555	TRUE	"Eukaryotic aspartyl protease family protein; FUNC
AT1G52730	0	3,6	4,492	0,002647	TRUE	"Transducin/WD40 repeat-like superfamily protein;
AT1G60780	0	5,2	4,768	0,002655	TRUE	"HAPLESS 13 (HAP13); INVOLVED IN: intracellular pr
AT5G44340	10,3	81,4	1,573	0,002677	TRUE	"tubulin beta chain 4 (TUB4); FUNCTIONS IN: struct
AT1G20010	10,3	62,4	1,248	0,002949	TRUE	"tubulin beta-5 chain (TUB5); FUNCTIONS IN: struct
AT1G43850	0	5,1	4,734	0,002967	TRUE	"seuss (SEU); BEST Arabidopsis thaliana protein ma
AT3G55410	0	7,4	5,136	0,002999	TRUE	"2-oxoglutarate dehydrogenase, E1 component; FUNCT
ATMG01190	0	5,9	4,978	0,0035	TRUE	#N/A
AT5G23540	0,2	5,6	3,318	0,003562	TRUE	"Mov34/MPN/PAD-1 family protein; INVOLVED IN: resp
AT1G27090	3	24,1	2,347	0,003733	TRUE	"glycine-rich protein; BEST Arabidopsis thaliana p
AT5G14170	0	5,4	4,776	0,003836	TRUE	"CHC1; CONTAINS InterPro DOMAIN/s: SWIB/MDM2 domai
AT2G28950	0	3,4	4,41	0,003907	TRUE	"expansin A6 (EXPA6); CONTAINS InterPro DOMAIN/s:
AT3G42170	0	4,6	4,603	0,003952	TRUE	"BED zinc finger ;hAT family dimerisation domain;
AT1G53750	0	4,8	4,632	0,003952	TRUE	"regulatory particle triple-A 1A (RPT1A); FUNCTION
AT3G15610	0	3,5	4,368	0,004204	TRUE	"Transducin/WD40 repeat-like superfamily protein;
AT4G38630	0	3	4,263	0,004438	TRUE	"regulatory particle non-ATPase 10 (RPN10); FUNCTI
AT4G26910	0	4,1	4,501	0,004564	TRUE	"Dihydrolipoamide succinyltransferase; FUNCTIONS I
AT3G05020	0	4	4,518	0,004737	TRUE	"acyl carrier protein 1 (ACP1); FUNCTIONS IN: acyl
AT5G26210	0	3,4	4,314	0,004757	TRUE	"alfin-like 4 (AL4); CONTAINS InterPro DOMAIN/s: Z
AT4G20890	10,2	81,4	1,57	0,004808	TRUE	"tubulin beta-9 chain (TUB9); FUNCTIONS IN: struct
AT1G26460	0,2	5,6	3,259	0,00488	TRUE	"Tetratricopeptide repeat (TPR)-like superfamily p
AT3G60960	0	3,1	4,296	0,005076	TRUE	"Tetratricopeptide repeat (TPR)-like superfamily p
AT1G06410	0	6,8	4,979	0,005096	TRUE	"trehalose-phosphatase/synthase 7 (TPS7); FUNCTION
AT3G56800	1	13	2,411	0,005291	TRUE	"calmodulin 3 (CAM3); FUNCTIONS IN: calcium ion bi
AT5G49460	4,8	28,8	1,286	0,005694	TRUE	"ATP citrate lyase subunit B 2 (ACLB-2); FUNCTIONS
AT2G21660	14	89	1,864	0,005793	TRUE	"cold, circadian rhythm, and rna binding 2 (CCR2);
AT4G08350	0	10,1	5,429	0,005836	TRUE	"global transcription factor group A2 (GTA2); FUNC
AT1G17370	0	3,1	4,254	0,005846	TRUE	"oligouridylate binding protein 1B (UBP1B); FUNCTI
AT4G34450	0	4,2	4,479	0,006162	TRUE	"coatomer gamma-2 subunit, putative / gamma-2 coat
AT3G12780	0	7,1	5,176	0,006307	TRUE	"phosphoglycerate kinase 1 (PGK1); FUNCTIONS IN: p
AT2G47250	0	3,8	4,362	0,006339	TRUE	"RNA helicase family protein; FUNCTIONS IN: in 7 f
AT3G62310	0	3,8	4,363	0,006386	TRUE	"RNA helicase family protein; FUNCTIONS IN: RNA he
AT1G36160	0	4	4,44	0,006473	TRUE	"acetyl-CoA carboxylase 1 (ACC1); FUNCTIONS IN: ac
AT3G54230	0	3,1	4,216	0,006663	TRUE	"suppressor of abi3-5 (SUA); FUNCTIONS IN: nucleot
AT5G62700	18,8	116,9	1,191	0,006663	TRUE	"tubulin beta chain 3 (TUB3); FUNCTIONS IN: struct
AT1G25490	0	3,6	4,371	0,006723	TRUE	"ROOTS CURL IN NPA (RCN1); FUNCTIONS IN: protein p
AT2G36870	0	3,1	4,2	0,006967	TRUE	"xyloglucan endotransglucosylase/hydrolase 32 (XTH
AT5G03690	0	3,1	4,201	0,007014	TRUE	"Aldolase superfamily protein; FUNCTIONS IN: fruct

AT3G15290	0	3	4,166	0,007103	TRUE	"3-hydroxyacyl-CoA dehydrogenase family protein; F
AT5G02740	0	3	4,16	0,007892	TRUE	"Ribosomal protein S24e family protein; FUNCTIONS
AT1G35160	0	4,5	4,633	0,007926	TRUE	"GF14 protein phi chain (GF14 PHI); FUNCTIONS IN:
AT5G43960	0	3,9	4,404	0,008633	TRUE	"Nuclear transport factor 2 (NTF2) family protein
AT1G62940	0	3,1	4,178	0,008778	TRUE	"acyl-CoA synthetase 5 (ACOS5); CONTAINS InterPro
AT3G25860	1,5	15,2	1,974	0,009091	TRUE	"LTA2; FUNCTIONS IN: dihydrolipoyllysine-residue a
AT3G61690	0	2,9	4,131	0,009202	TRUE	"nucleotidyltransferases; FUNCTIONS IN: nucleotidy
AT3G61600	0	4,1	4,539	0,009262	TRUE	"POZ/BTB containin G-protein 1 (POB1); CONTAINS In
AT3G09820	0	4,1	4,493	0,009427	TRUE	"adenosine kinase 1 (ADK1); FUNCTIONS IN: adenosin
AT2G38110	0,2	6,8	3,379	0,009567	TRUE	"glycerol-3-phosphate acyltransferase 6 (GPAT6); C
AT1G02080	0	5,2	4,668	0,01076	TRUE	"transcription regulators; FUNCTIONS IN: transcrip
AT4G02290	3,7	22,5	1,339	0,01127	TRUE	"glycosyl hydrolase 9B13 (GH9B13); FUNCTIONS IN: h
AT1G09430	0	2,8	4,054	0,01127	TRUE	"ATP-citrate lyase A-3 (ACLA-3); CONTAINS InterPro
AT3G60830	0	6	4,829	0,01143	TRUE	"actin-related protein 7 (ARP7); CONTAINS InterPro
AT5G20490	0	5,1	4,934	0,01182	TRUE	"XIK; FUNCTIONS IN: motor activity; INVOLVED IN: i
AT1G55500	0	2,5	4,217	0,01226	TRUE	"evolutionarily conserved C-terminal region 4 (ECT
AT5G49220	0	2,5	3,956	0,01244	TRUE	"Protein of unknown function (DUF789); CONTAINS In
AT2G34590	0	2,8	4,043	0,01282	TRUE	"Transketolase family protein; FUNCTIONS IN: pyruv
AT1G56190	0	4	4,469	0,01314	TRUE	"Phosphoglycerate kinase family protein; FUNCTIONS
AT2G46020	0	6	4,755	0,01372	TRUE	"BRAHMA (BRM); FUNCTIONS IN: helicase activity, tr
AT1G10170	0	2,4	3,929	0,0138	TRUE	"NF-X-like 1 (NFXL1); FUNCTIONS IN: zinc ion bindi
AT4G32551	0	4,5	4,46	0,01428	TRUE	"LEUNIG (LUG); CONTAINS InterPro DOMAIN/s: WD40 re
AT2G15430	0,2	4,4	2,969	0,01438	TRUE	"NRPB3; FUNCTIONS IN: DNA-directed RNA polymerase
AT3G63460	0	2,9	4,046	0,0144	TRUE	"transducin family protein / WD-40 repeat family p
AT1G06430	0,2	4,5	2,961	0,01444	TRUE	"FTSH protease 8 (FTSH8); FUNCTIONS IN: metallopep
AT3G13930	0	3,1	4,118	0,01496	TRUE	"Dihydrolipoamide acetyltransferase, long form pro
AT3G50370	0	4,4	4,66	0,01546	TRUE	"unknown protein; FUNCTIONS IN: molecular_function
AT4G11380	0,3	8,2	2,882	0,0156	TRUE	"Adaptin family protein; FUNCTIONS IN: protein tra
AT1G17720	0	2,4	3,867	0,01633	TRUE	"ATB BETA; CONTAINS InterPro DOMAIN/s: WD40 repeat
AT2G38610	0	2,5	4,038	0,01735	TRUE	"RNA-binding KH domain-containing protein; FUNCTIO
AT4G38780	0,2	4,8	3,016	0,0176	TRUE	"Pre-mRNA-processing-splicing factor; FUNCTIONS IN
AT2G29550	13,5	92,9	1,303	0,0176	TRUE	"tubulin beta-7 chain (TUB7); FUNCTIONS IN: struct
AT5G01410	0	3	4,089	0,01834	TRUE	"REDUCED SUGAR RESPONSE 4 (RSR4); FUNCTIONS IN: pr
AT1G04410	0,3	7	2,733	0,01899	TRUE	"Lactate/malate dehydrogenase family protein; FUNC
AT3G25800	0	2,5	3,932	0,02015	TRUE	"protein phosphatase 2A subunit A2 (PP2AA2); FUNC
AT1G20200	0	4,5	4,437	0,02045	TRUE	"EMBRYO DEFECTIVE 2719 (EMB2719); FUNCTIONS IN: en
AT1G79530	0	2,5	3,897	0,02054	TRUE	"glyceraldehyde-3-phosphate dehydrogenase of plast
AT3G57890	0	2,5	3,935	0,02098	TRUE	"Tubulin binding cofactor C domain-containing prot
AT5G18410	0	2,4	3,934	0,02122	TRUE	"PIROGI 121 (PIR121); CONTAINS InterPro DOMAIN/s:
AT5G46290	0	2,5	3,866	0,02137	TRUE	"3-ketoacyl-acyl carrier protein synthase I (KAS I
AT3G61650	0	2,2	3,801	0,02141	TRUE	"gamma-tubulin (TUBG1); FUNCTIONS IN: structural m
AT4G02840	0	2,4	3,832	0,02142	TRUE	"Small nuclear ribonucleoprotein family protein; F
AT5G51750	0,7	11,1	2,491	0,02223	TRUE	"subtilase 1.3 (SBT1.3); FUNCTIONS IN: identical p
AT5G03300	0	2,8	4,007	0,02242	TRUE	"adenosine kinase 2 (ADK2); FUNCTIONS IN: adenosin
AT5G48900	0	2	3,725	0,02352	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT1G70830	0	2,8	3,939	0,02358	TRUE	"MLP-like protein 28 (MLP28); FUNCTIONS IN: molecu
AT3G07590	0	2,2	3,776	0,02358	TRUE	"Small nuclear ribonucleoprotein family protein; F
AT3G19130	0	2,6	3,967	0,02448	TRUE	"RNA-binding protein 47B (RBP47B); FUNCTIONS IN: R
AT3G27690	0	3,9	4,433	0,02471	TRUE	"photosystem II light harvesting complex gene 2.3
AT3G60980	0	2,1	3,759	0,02485	TRUE	"Tetratricopeptide repeat (TPR)-like superfamily p
AT5G22780	0	6,9	4,867	0,025	TRUE	"Adaptor protein complex AP-2, alpha subunit; FUNC
AT5G02530	0,7	10,8	2,498	0,0255	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G22770	0	6,5	4,8	0,02574	TRUE	"alpha-adaptin (alpha-ADR); FUNCTIONS IN: protein
AT3G44110	0,7	7	1,981	0,02645	TRUE	"DNAJ homologue 3 (ATJ3); FUNCTIONS IN: unfolded p
AT3G08770	0	3,1	4,157	0,02759	TRUE	"lipid transfer protein 6 (LTP6); FUNCTIONS IN: li
AT3G60750	0	5,1	4,746	0,02844	TRUE	"Transketolase; FUNCTIONS IN: catalytic activity,
AT1G04680	0	2,5	3,864	0,02865	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT3G06650	4,8	24,2	1,08	0,02865	TRUE	"ATP-citrate lyase B-1 (ACLB-1); FUNCTIONS IN: ATP
AT3G59770	0	3,8	4,209	0,02876	TRUE	"SUPPRESSOR OF ACTIN 9 (SAC9); FUNCTIONS IN: inosi
AT2G21060	0,5	5,6	2,556	0,02902	TRUE	"glycine-rich protein 2B (GRP2B); FUNCTIONS IN: DN
AT3G15680	0	1,9	3,633	0,0292	FALSE	"Ran BP2/NZF zinc finger-like superfamily protein;
AT4G36960	0	2,2	3,729	0,03082	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G55320	0	2,2	3,731	0,03089	TRUE	"acyl-activating enzyme 18 (AAE18); FUNCTIONS IN:
AT1G47490	0	2	3,77	0,03201	TRUE	"RNA-binding protein 47C (RBP47C); CONTAINS InterP
AT2G41040	0	2	3,618	0,03257	TRUE	"S-adenosyl-L-methionine-dependent methyltransfera
AT1G22410	0,2	4,4	2,821	0,03664	TRUE	"Class-II DAHP synthetase family protein; FUNCTION
AT3G58500	0	2,5	3,807	0,03751	TRUE	"protein phosphatase 2A-4 (PP2A-4); CONTAINS Inter
AT2G36530	0	2,5	3,85	0,0382	TRUE	"LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 2
AT5G54160	0	1,9	3,534	0,03894	FALSE	"O-methyltransferase 1 (OMT1); FUNCTIONS IN: myric
AT5G64960	0	3,6	4,12	0,03909	TRUE	"cyclin dependent kinase group C2 (CDKC2); FUNCTIO
AT2G28790	1,7	14,5	1,772	0,03957	TRUE	"Pathogenesis-related thaumatin superfamily protei
AT5G35630	0	1,9	3,538	0,0401	FALSE	"glutamine synthetase 2 (GS2); FUNCTIONS IN: gluta
AT5G20000	0,3	7,2	2,691	0,04059	TRUE	"AAA-type ATPase family protein; FUNCTIONS IN: hyd
AT5G55280	0	1,9	3,572	0,04215	FALSE	"homolog of bacterial cytokinesis Z-ring protein F
AT5G53330	0	2,1	3,758	0,04215	TRUE	"Ubiquitin-associated/translation elongation facto
AT4G39260	4,3	23,2	1,575	0,04221	TRUE	"cold, circadian rhythm, and RNA binding 1 (CCR1);
AT2G34810	0,3	4,4	2,303	0,04265	TRUE	"FAD-binding Berberine family protein; FUNCTIONS I
AT4G35890	0	2,9	3,999	0,04317	TRUE	"winged-helix DNA-binding transcription factor fam
AT1G47128	0	1,8	3,448	0,04452	FALSE	"responsive to dehydration 21 (RD21); FUNCTIONS IN
AT4G29010	0,8	17,2	2,492	0,04452	TRUE	"ABNORMAL INFLORESCENCE MERISTEM (AIM1); FUNCTIONS

AT5G67630	0	2	3,557	0,04578	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT1G20620	6	50,2	1,715	0,04617	TRUE	"catalase 3 (CAT3); FUNCTIONS IN: catalase activit
AT2G43970	0	2,9	3,871	0,04649	TRUE	"RNA-binding protein; FUNCTIONS IN: RNA binding, n
AT3G07810	0	1,6	3,698	0,04693	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G28470	1	9,5	1,726	0,04694	TRUE	"26S proteasome regulatory subunit S2 1B (RPN1B);
AT3G12400	0	2	3,614	0,04694	TRUE	"ELC; CONTAINS InterPro DOMAIN/s: Ubiquitin-conjug
AT3G25150	0,3	8,9	3,438	0,04798	TRUE	"Nuclear transport factor 2 (NTF2) family protein
AT5G65750	0	2	3,532	0,04857	TRUE	"2-oxoglutarate dehydrogenase, E1 component; FUNCT
AT3G61820	1,7	9,5	1,24	0,04857	TRUE	"Eukaryotic aspartyl protease family protein; FUNC
AT1G12310	0,8	8,1	1,857	0,04857	TRUE	"Calcium-binding EF-hand family protein; FUNCTIONS
AT2G36580	0	2	3,543	0,0493	TRUE	"Pyruvate kinase family protein; FUNCTIONS IN: pyr
AT3G52990	0	2	3,543	0,0493	TRUE	"Pyruvate kinase family protein; FUNCTIONS IN: pyr
AT4G31770	0	2,5	3,726	0,0493	TRUE	"debranching enzyme 1 (DBR1); CONTAINS InterPro DO
AT1G03457	0	1,6	3,404	0,05227	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G54110	0	2	3,531	0,05231	FALSE	"plant uncoupling mitochondrial protein 1 (PUMP1);
AT1G53240	0	2,4	3,801	0,05286	FALSE	"Lactate/malate dehydrogenase family protein; FUNC
AT1G54490	0	1,8	3,48	0,05486	FALSE	"exoribonuclease 4 (XRN4); CONTAINS InterPro DOMAI
AT1G30630	0,2	2,9	2,466	0,05515	FALSE	"Coatomer epsilon subunit; FUNCTIONS IN: protein t
AT3G21100	0	2,5	3,703	0,05627	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G51980	0	2,2	3,621	0,05633	FALSE	"Transducin/WD40 repeat-like superfamily protein;
AT3G42790	0	1,9	3,459	0,05647	FALSE	"alfin-like 3 (AL3); CONTAINS InterPro DOMAIN/s: Z
AT1G07360	0	3,9	4,165	0,05686	FALSE	"CCCH-type zinc fingerfamily protein with RNA-bind
AT1G12000	0	1,6	3,342	0,05809	FALSE	"Phosphofructokinase family protein; FUNCTIONS IN:
AT1G13320	0	2	3,544	0,05984	FALSE	"protein phosphatase 2A subunit A3 (PP2AA3); FUNC
AT5G25780	0	4,1	4,2	0,06033	FALSE	"eukaryotic translation initiation factor 3B-2 (EI
AT3G09100	0	2,4	3,645	0,06033	FALSE	"mRNA capping enzyme family protein; FUNCTIONS IN:
AT1G23860	0	1,5	3,324	0,06059	FALSE	"RS-containing zinc finger protein 21 (RSZP21); FU
AT3G19760	1,7	12,8	1,529	0,06224	FALSE	"eukaryotic initiation factor 4A-III (EIF4A-III);
AT1G11860	0	1,8	3,384	0,06289	FALSE	"Glycine cleavage T-protein family; FUNCTIONS IN:
AT5G51550	0	1,5	3,269	0,06319	FALSE	"EXORDIUM like 3 (EXL3); CONTAINS InterPro DOMAIN/
AT4G27440	0	1,5	3,279	0,06319	FALSE	"protochlorophyllide oxidoreductase B (PORB); FUNC
AT1G54580	0	1,5	3,287	0,06319	FALSE	"acyl carrier protein 2 (ACP2); FUNCTIONS IN: acyl
AT1G13730	0	3,5	4,5	0,06386	FALSE	"Nuclear transport factor 2 (NTF2) family protein
AT3G22270	0	2,2	3,575	0,0642	FALSE	"Topoisomerase II-associated protein PAT1; BEST Ar
AT1G62640	0	1,9	3,435	0,06436	FALSE	"3-ketoacyl-acyl carrier protein synthase III (KAS
AT1G10200	0	1,6	3,311	0,06513	FALSE	"WLIM1; FUNCTIONS IN: sequence-specific DNA bindin
AT4G23460	0,3	4,6	2,194	0,06655	FALSE	"Adaptin family protein; FUNCTIONS IN: protein tra
AT4G34430	0	2,6	3,885	0,0671	FALSE	"CHB3; FUNCTIONS IN: DNA binding, sequence-specifi
AT1G17580	0	3,2	3,942	0,06802	FALSE	"myosin 1 (MYA1); FUNCTIONS IN: motor activity; IN
AT4G28440	1,7	8,8	1,275	0,06817	FALSE	"Nucleic acid-binding, OB-fold-like protein; CONTA
AT5G53620	0	1,5	3,305	0,06854	FALSE	"unknown protein; INVOLVED IN: biological_process
AT2G27880	0	1,8	3,357	0,0691	FALSE	"ARGONAUTE 5 (AGO5); FUNCTIONS IN: nucleic acid bi
AT5G12440	0	1,8	3,443	0,0691	FALSE	"CCCH-type zinc fingerfamily protein with RNA-bind
AT1G30680	0,2	3,4	2,516	0,06967	FALSE	"toprim domain-containing protein; FUNCTIONS IN: D
AT3G06850	0	1,8	3,361	0,07088	FALSE	"BCE2; FUNCTIONS IN: acetyltransferase activity, d
AT2G27720	0	1,5	3,271	0,07201	FALSE	"60S acidic ribosomal protein family; FUNCTIONS IN
AT4G34660	0	2	3,447	0,07271	FALSE	"SH3 domain-containing protein; CONTAINS InterPro
AT3G15020	0	1,8	3,436	0,07384	FALSE	"Lactate/malate dehydrogenase family protein; FUNC
AT1G78260	0	1,4	3,193	0,07493	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G24820	0,3	4,4	2,175	0,07583	FALSE	"26S proteasome, regulatory subunit Rpn7;Proteasom
AT5G20160	0,3	3,8	2,004	0,07636	FALSE	"Ribosomal protein L7Ae/L30e/S12e/Gadd45 family pr
AT1G53650	0	1,4	3,164	0,07699	FALSE	"CTC-interacting domain 8 (CID8); CONTAINS InterPr
AT1G13930	0	2,6	3,905	0,07721	FALSE	"FUNCTIONS IN: molecular_function unknown; INVOLVE
AT1G31280	0	3,5	3,987	0,07772	FALSE	"argonaute 2 (AGO2); FUNCTIONS IN: nucleic acid bi
AT2G32080	0	1,2	3,227	0,07848	FALSE	"purin-rich alpha 1 (PUR ALPHA-1); FUNCTIONS IN: n
AT5G46630	0	2,1	3,545	0,0785	FALSE	"Clathrin adaptor complexes medium subunit family
AT5G23860	11	92,5	1,475	0,0785	FALSE	"tubulin beta 8 (TUB8); FUNCTIONS IN: protein bind
AT3G12130	0	2	3,532	0,07982	FALSE	"KH domain-containing protein / zinc finger (CCCH
AT1G18450	0,3	6,9	2,555	0,0799	FALSE	"actin-related protein 4 (ARP4); FUNCTIONS IN: str

Table S2. Common protein set between UPF1-IPs and At-RBP. Comparison of proteins enriched in UPF1-IPs (this study) with *Arabidopsis* RNA Binding Proteins (At-RBP) from Reichel *et al.*, 2016.

AT1G02080	"transcription regulators; FUNCTIONS IN: transcrip
AT1G10170	"NF-X-like 1 (NFXL1); FUNCTIONS IN: zinc ion bindi
AT1G11650	"RBP45B; FUNCTIONS IN: RNA binding; EXPRESSED IN:
AT1G13190	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G14710	"hydroxyproline-rich glycoprotein family protein;
AT1G17370	"oligouridylate binding protein 1B (UBP1B); FUNCTI
AT1G20010	"tubulin beta-5 chain (TUB5); FUNCTIONS IN: struct
AT1G20620	"catalase 3 (CAT3); FUNCTIONS IN: catalase activit
AT1G26110	"decapping 5 (DCP5); CONTAINS InterPro DOMAIN/s: D
AT1G27090	"glycine-rich protein; BEST Arabidopsis thaliana p
AT1G27430	"GYF domain-containing protein; CONTAINS InterPro
AT1G29250	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G32790	"CTC-interacting domain 11 (CID11); FUNCTIONS IN:
AT1G33680	"KH domain-containing protein; FUNCTIONS IN: RNA b
AT1G33980	"UPF3; FUNCTIONS IN: nucleotide binding; INVOLVED
AT1G36160	"acetyl-CoA carboxylase 1 (ACC1); FUNCTIONS IN: ac
AT1G43190	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT1G47490	"RNA-binding protein 47C (RBP47C); CONTAINS InterP
AT1G48410	"ARGONAUTE 1 (AGO1); FUNCTIONS IN: protein binding
AT1G49760	"poly(A) binding protein 8 (PAB8); FUNCTIONS IN: R
AT1G54080	"oligouridylate-binding protein 1A (UBP1A); FUNCTI
AT1G54270	"eif4a-2 (EIF4A-2); FUNCTIONS IN: ATP-dependent he
AT1G55500	"evolutionarily conserved C-terminal region 4 (ECT
AT1G75660	"5'-3' exoribonuclease 3 (XRN3); CONTAINS InterPro
AT1G76010	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G06850	"xyloglucan endotransglucosylase/hydrolase 4 (XTH4
AT2G17870	"cold shock domain protein 3 (CSP3); FUNCTIONS IN:
AT2G21060	"glycine-rich protein 2B (GRP2B); FUNCTIONS IN: DN
AT2G21660	"cold, circadian rhythm, and rna binding 2 (CCR2);
AT2G23350	"poly(A) binding protein 4 (PAB4); FUNCTIONS IN: R
AT2G26280	"CID7; FUNCTIONS IN: damaged DNA binding, protein
AT2G27100	"SERRATE (SE); FUNCTIONS IN: DNA binding, sequence
AT2G29140	"pumilio 3 (PUM3); FUNCTIONS IN: RNA binding, bind
AT2G29190	"pumilio 2 (PUM2); FUNCTIONS IN: mRNA binding, RNA
AT2G29200	"pumilio 1 (PUM1); FUNCTIONS IN: RNA binding, bind
AT2G33410	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT2G34160	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G36530	"LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 2
AT2G38610	"RNA-binding KH domain-containing protein; FUNCTIO
AT2G39260	"binding;RNA binding; FUNCTIONS IN: RNA binding, b
AT2G42520	"P-loop containing nucleoside triphosphate hydrola
AT2G43970	"RNA-binding protein; FUNCTIONS IN: RNA binding, n
AT2G45810	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT2G46780	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G01540	"DEAD box RNA helicase 1 (DRH1); FUNCTIONS IN: ATP
AT3G03950	"evolutionarily conserved C-terminal region 1 (ECT
AT3G04610	"flowering locus KH domain (FLK); FUNCTIONS IN: RN
AT3G06480	"DEAD box RNA helicase family protein; FUNCTIONS I
AT3G07030	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G10770	"Single-stranded nucleic acid binding R3H protein;
AT3G12915	"Ribosomal protein S5/Elongation factor G/III/V fa
AT3G13060	"evolutionarily conserved C-terminal region 5 (ECT
AT3G13300	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT3G13460	"evolutionarily conserved C-terminal region 2 (ECT
AT3G13920	"eukaryotic translation initiation factor 4A1 (EIF
AT3G14100	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G15000	"cobalt ion binding; FUNCTIONS IN: cobalt ion bind
AT3G15010	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G15610	"Transducin/WD40 repeat-like superfamily protein;
AT3G19130	"RNA-binding protein 47B (RBP47B); FUNCTIONS IN: R
AT3G22330	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT3G25150	"Nuclear transport factor 2 (NTF2) family protein
AT3G26420	"ATRZ-1A; CONTAINS InterPro DOMAIN/s: RNA recognit
AT3G44110	"DNAJ homologue 3 (ATJ3); FUNCTIONS IN: unfolded p
AT3G49390	"CTC-interacting domain 10 (CID10); FUNCTIONS IN:
AT3G50370	"unknown protein; FUNCTIONS IN: molecular_function
AT3G51950	"Zinc finger (CCCH-type) family protein / RNA reco
AT3G58510	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G58570	"P-loop containing nucleoside triphosphate hydrola
AT3G60240	"eukaryotic translation initiation factor 4G (EIF4
AT3G60750	"Transketolase; FUNCTIONS IN: catalytic activity,
AT3G61240	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G61690	"nucleotidyltransferases; FUNCTIONS IN: nucleotidy
AT4G00660	"RNAhelicase-like 8 (RH8); FUNCTIONS IN: helicase
AT4G03110	"RNA-binding protein-defense related 1 (RBP-DR1);
AT4G10610	"CTC-interacting domain 12 (CID12); CONTAINS Inter
AT4G20360	"RAB GTPase homolog E1B (RABE1b); FUNCTIONS IN: GT
AT4G20890	"tubulin beta-9 chain (TUB9); FUNCTIONS IN: struct
AT4G21710	"NRPB2; CONTAINS InterPro DOMAIN/s: DNA-directed R
AT4G25550	"Cleavage/polyadenylation specificity factor, 25kD
AT4G27000	"ATRBP45C; CONTAINS InterPro DOMAIN/s: RNA recogni
AT4G34110	"poly(A) binding protein 2 (PAB2); FUNCTIONS IN: R
AT4G34450	"coatomer gamma-2 subunit, putative / gamma-2 coat
AT4G35890	"winged-helix DNA-binding transcription factor fam
AT4G36960	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G38680	"glycine rich protein 2 (GRP2); FUNCTIONS IN: doub
AT4G38780	"Pre-mRNA-processing-splicing factor; FUNCTIONS IN
AT4G39260	"cold, circadian rhythm, and RNA binding 1 (CCR1);
AT5G02530	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G04280	"RNA-binding (RRM/RBD/RNP motifs) family protein w
AT5G05100	"Single-stranded nucleic acid binding R3H protein;
AT5G11170	"DEAD/DEAH box RNA helicase family protein ; FUNCT
AT5G20490	"XIK; FUNCTIONS IN: motor activity; INVOLVED IN: i
AT5G26742	"embryo defective 1138 (emb1138); FUNCTIONS IN: in
AT5G37720	"ALWAYS EARLY 4 (ALY4); FUNCTIONS IN: nucleotide b
AT5G40490	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G43960	"Nuclear transport factor 2 (NTF2) family protein
AT5G47010	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in
AT5G48650	"Nuclear transport factor 2 (NTF2) family protein
AT5G49220	"Protein of unknown function (DUF789); CONTAINS In
AT5G54900	"RNA-binding protein 45A (RBP45A); FUNCTIONS IN: R
AT5G55670	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G58470	"TBP-associated factor 15B (TAF15b); FUNCTIONS IN:
AT5G61020	"evolutionarily conserved C-terminal region 3 (ECT
AT5G63120	"P-loop containing nucleoside triphosphate hydrola

Table S3. RNA –independent partners of UPF1. Analysis of UPF1 partners enriched in both RNase-treated and non-treated IPs. In this table are indicated the informations used to create the volcano plot in Figure 12A (only proteins with an *adjp<0.08* are listed here). Mean of spectral counts for control or UPF1 IPs, log2(Fold Change), adjusted p-value (adjp) are detailed. DEP indicates whether the considered protein passed the statistical test (TRUE) or not (FALSE). Proteins in grey are not significative.

AGI	Mean spectral count **Control**	Mean spectral count **RNase-treated and non-treated**	LogFC	adjp	DEP	Description
AT5G47010	0	1166	13,55	1,93E-69	TRUE	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in
AT2G42520	10,3	43,4	2,906	2,04E-14	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT5G02500	23,3	53,4	2,055	4,03E-12	TRUE	"heat shock cognate protein 70-1 (HSC70-1); FUNCTI
AT2G39260	0	25,6	8,075	4,77E-12	TRUE	"binding;RNA binding; FUNCTIONS IN: RNA binding, b
AT3G58570	9,7	40,6	2,913	1,53E-11	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT3G58510	9	38,6	2,918	3,12E-10	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT2G23350	0	7,4	6,201	5,42E-10	TRUE	"poly(A) binding protein 4 (PAB4); FUNCTIONS IN: R
AT3G09440	17	38,3	2,019	1,56E-09	TRUE	"Heat shock protein 70 (Hsp 70) family protein; FU
AT4G00660	0	9,3	6,527	2,22E-09	TRUE	"RNAhelicase-like 8 (RH8); FUNCTIONS IN: helicase
AT1G13440	10,7	19,9	1,696	3,51E-08	TRUE	"glyceraldehyde-3-phosphate dehydrogenase C2 (GAPC
AT3G04120	11	19,6	1,633	4,88E-08	TRUE	"glyceraldehyde-3-phosphate dehydrogenase C subuni
AT3G13460	0,3	11	5,197	1,18E-07	TRUE	"evolutionarily conserved C-terminal region 2 (ECT
AT3G61240	0	7,1	6,216	1,64E-07	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT1G49760	0	11,1	6,724	1,02E-06	TRUE	"poly(A) binding protein 8 (PAB8); FUNCTIONS IN: R
AT4G34110	0	11,9	6,724	1,21E-06	TRUE	"poly(A) binding protein 2 (PAB2); FUNCTIONS IN: R
AT3G01540	2,7	13,3	3,046	1,77E-06	TRUE	"DEAD box RNA helicase 1 (DRH1); FUNCTIONS IN: ATP
AT4G22010	10,7	31,1	2,44	9,69E-06	TRUE	"SKU5 similar 4 (sks4); FUNCTIONS IN: oxidoreduct
AT5G02490	19,7	37,1	1,793	1,85E-05	TRUE	"Heat shock protein 70 (Hsp 70) family protein; FU
AT5G14610	2,7	9,3	2,561	4,38E-05	TRUE	"DEAD box RNA helicase family protein; FUNCTIONS I
AT1G54270	8	27	2,635	4,69E-05	TRUE	"eif4a-2 (EIF4A-2); FUNCTIONS IN: ATP-dependent he
AT3G44110	0,3	4,6	3,988	5,94E-05	TRUE	"DNAJ homologue 3 (ATJ3); FUNCTIONS IN: unfolded p
AT1G71770	0	4	5,372	6,34E-05	TRUE	"poly(A)-binding protein 5 (PAB5); FUNCTIONS IN: R
AT5G26742	7,3	22	2,307	6,36E-05	TRUE	"embryo defective 1138 (emb1138); FUNCTIONS IN: in
AT5G61020	0	8	6,199	7,68E-05	TRUE	"evolutionarily conserved C-terminal region 3 (ECT
AT2G21660	16,7	49,3	2,448	7,69E-05	TRUE	"cold, circadian rhythm, and rna binding 2 (CCR2);
AT1G56410	19,7	35,6	1,74	8,72E-05	TRUE	"EARLY-RESPONSIVE TO DEHYDRATION 2 (ERD2); FUNCTIO
AT1G33980	0	6,9	6,251	0,000105	TRUE	"UPF3; FUNCTIONS IN: nucleotide binding; INVOLVED
AT3G13920	8,3	28,1	2,618	0,000114	TRUE	"eukaryotic translation initiation factor 4A1 (EIF
AT5G22060	0	3,1	4,997	0,000229	TRUE	"DNAJ homologue 2 (J2); FUNCTIONS IN: unfolded pro
AT3G06650	4	9,6	2,032	0,000274	TRUE	"ATP-citrate lyase B-1 (ACLB-1); FUNCTIONS IN: ATP
AT4G24280	0	2,6	4,688	0,00034	TRUE	"chloroplast heat shock protein 70-1 (cpHsc70-1);
AT5G49910	0	2,4	4,606	0,00043	TRUE	"chloroplast heat shock protein 70-2 (CPHSC70-2EAT
AT5G40490	0,7	5,6	3,557	0,000461	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G32790	0	2,4	4,606	0,000541	TRUE	"CTC-interacting domain 11 (CID11); FUNCTIONS IN:
AT1G20620	3,7	11,7	2,397	0,000548	TRUE	"catalase 3 (CAT3); FUNCTIONS IN: catalase activit
AT5G49460	4	9,4	2,01	0,000548	TRUE	"ATP citrate lyase subunit B 2 (ACLB-2); FUNCTIONS
AT1G48410	1	12	4,041	0,000666	TRUE	"ARGONAUTE 1 (AGO1); FUNCTIONS IN: protein binding
AT4G10610	0	2,3	4,53	0,000761	TRUE	"CTC-interacting domain 12 (CID12); CONTAINS Inter
AT3G23990	0	2,3	4,509	0,000814	TRUE	"heat shock protein 60 (HSP60); FUNCTIONS IN: copp
AT2G33210	0	2,3	4,509	0,000814	TRUE	"heat shock protein 60-2 (HSP60-2); FUNCTIONS IN:
AT4G25550	0	3,1	4,911	0,000823	TRUE	"Cleavage/polyadenylation specificity factor, 25kD
AT3G19760	1	4,3	2,716	0,000909	TRUE	"eukaryotic initiation factor 4A-III (EIF4A-III);
AT5G42020	14,7	18,6	1,153	0,001309	TRUE	"BIP2; FUNCTIONS IN: ATP binding; INVOLVED IN: pro
AT5G28540	14,7	18,6	1,153	0,001309	TRUE	"BIP1; FUNCTIONS IN: ATP binding; INVOLVED IN: pro
AT4G21280	23,7	26,4	1,013	0,001432	TRUE	"photosystem II subunit QA (PSBQA); FUNCTIONS IN:
AT3G03950	0	3	4,817	0,001623	TRUE	"evolutionarily conserved C-terminal region 1 (ECT
AT3G06480	1,3	6,3	2,861	0,001835	TRUE	"DEAD box RNA helicase family protein; FUNCTIONS I
AT5G63120	1,3	5,6	2,674	0,00185	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT1G16030	11,7	17,4	1,435	0,00204	TRUE	"heat shock protein 70B (Hsp70b); FUNCTIONS IN: AT
AT2G34160	0	2,9	4,736	0,002092	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT4G20360	7	12,9	1,649	0,002909	TRUE	"RAB GTPase homolog E1B (RABE1b); FUNCTIONS IN: GT
AT2G45810	0	3,7	5,551	0,003493	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT5G44780	0	2,7	4,72	0,003701	TRUE	"unknown protein; BEST Arabidopsis thaliana protei
AT2G39800	0	1,9	4,205	0,003776	FALSE	"delta1-pyrroline-5-carboxylate synthase 1 (P5CS1)
AT2G06850	1,3	7,7	3,149	0,003787	TRUE	"xyloglucan endotransglucosylase/hydrolase 4 (XTH4
AT4G18480	0	1,6	4,001	0,004268	FALSE	"CHLI1; FUNCTIONS IN: magnesium chelatase activity
AT1G75780	0	7,7	6,388	0,005038	TRUE	"tubulin beta-1 chain (TUB1); FUNCTIONS IN: struct
AT3G23450	0	4,6	5,54	0,005175	TRUE	"unknown protein; FUNCTIONS IN: molecular_function
AT3G07030	0	1,7	4,114	0,0055	FALSE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G18660	0	2	4,32	0,005628	TRUE	"zinc finger (C3HC4-type RING finger) family prote
AT5G23540	0	1,4	3,883	0,005818	FALSE	"Mov34/MPN/PAD-1 family protein; INVOLVED IN: resp
AT4G37910	0	1,9	4,161	0,006052	FALSE	"mitochondrial heat shock protein 70-1 (mtHsc70-1)
AT5G55670	0	1,4	3,922	0,00635	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G03960	0	1,6	3,983	0,007271	FALSE	"TCP-1/cpn60 chaperonin family protein; FUNCTIONS
AT4G29010	0,7	4,6	3,145	0,007388	TRUE	"ABNORMAL INFLORESCENCE MERISTEM (AIM1); FUNCTIONS
AT4G14301	0	3,7	5,282	0,007417	TRUE	"unknown protein; FUNCTIONS IN: molecular_function
AT3G22330	1	9,1	3,628	0,007607	TRUE	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT1G51380	0	1,3	3,785	0,007895	FALSE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO

AT1G51380	0	1,3	3,785	0,007895	FALSE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G22310	0,3	3	3,313	0,007937	TRUE	"putative mitochondrial RNA helicase 1 (PMH1); CON
AT1G78570	0	1,9	4,202	0,00909	FALSE	"rhamnose biosynthesis 1 (RHM1); CONTAINS InterPro
AT1G76010	2	5	2,01	0,009121	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G15000	0	1,3	3,738	0,01009	FALSE	"cobalt ion binding; FUNCTIONS IN: cobalt ion bind
AT1G20010	13	19,4	1,346	0,01072	TRUE	"tubulin beta-5 chain (TUB5); FUNCTIONS IN: struct
AT3G14100	0	1,3	3,726	0,01228	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G60240	0	1,6	3,917	0,01299	FALSE	"eukaryotic translation initiation factor 4G (EIF4
AT2G29550	13	25,7	1,711	0,01456	TRUE	"tubulin beta-7 chain (TUB7); FUNCTIONS IN: struct
AT5G04280	0	1,9	4,123	0,01487	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein w
AT1G01300	3,7	7,3	1,762	0,01487	TRUE	"Eukaryotic aspartyl protease family protein; FUNC
AT2G27100	0	1,7	4,035	0,01506	FALSE	"SERRATE (SE); FUNCTIONS IN: DNA binding, sequence
AT2G42570	0	1,6	3,925	0,01506	FALSE	"TRICHOME BIREFRINGENCE-LIKE 39 (TBL39); INVOLVED
AT1G26630	0	1,1	3,582	0,01533	FALSE	"FUMONISIN B1-RESISTANT12 (FBR12); FUNCTIONS IN: t
AT1G75660	0	2	4,193	0,01539	TRUE	"5'-3' exoribonuclease 3 (XRN3); CONTAINS InterPro
AT4G27060	0	1,7	4,048	0,01565	FALSE	"TORTIFOLIA 1 (TOR1); CONTAINS InterPro DOMAIN/s:
ATCG00500	0,3	2,3	2,967	0,01565	TRUE	"acetyl-CoA carboxylase carboxyl transferase subun
AT3G13060	0	1,3	3,71	0,01628	FALSE	"evolutionarily conserved C-terminal region 5 (ECT
AT1G22410	0,3	2,1	2,867	0,01656	TRUE	"Class-II DAHP synthetase family protein; FUNCTION
AT1G53750	0	1,1	3,576	0,01672	FALSE	"regulatory particle triple-A 1A (RPT1A); FUNCTION
AT3G12130	0	1,1	3,547	0,0172	FALSE	"KH domain-containing protein / zinc finger (CCCH
AT1G51570	0	1,1	3,555	0,02044	FALSE	"Calcium-dependent lipid-binding (CaLB domain) pla
AT4G35850	0	1,1	3,559	0,02056	FALSE	"Pentatricopeptide repeat (PPR) superfamily protei
AT1G53650	0	1	3,419	0,02166	FALSE	"CTC-interacting domain 8 (CID8); CONTAINS InterPr
AT1G55970	0	5,4	5,931	0,02166	TRUE	"histone acetyltransferase of the CBP family 4 (HA
AT5G13300	0	1,1	3,538	0,0218	FALSE	"SCARFACE (SFC); FUNCTIONS IN: ARF GTPase activato
AT2G28950	0	1,3	3,669	0,02213	FALSE	"expansin A6 (EXPA6); CONTAINS InterPro DOMAIN/s:
AT2G33845	0	1,6	3,84	0,02586	FALSE	"Nucleic acid-binding, OB-fold-like protein; FUNCT
AT2G46780	0	1	3,387	0,027	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G56010	0	1	3,387	0,027	FALSE	"heat shock protein 81-3 (HSP81-3); FUNCTIONS IN:
AT4G39260	4,7	11,9	2,104	0,027	TRUE	"cold, circadian rhythm, and RNA binding 1 (CCR1);
AT3G20820	10,3	13	1,161	0,02814	TRUE	"Leucine-rich repeat (LRR) family protein; INVOLVE
AT1G45000	0,7	2,6	2,44	0,02934	TRUE	"AAA-type ATPase family protein; FUNCTIONS IN: hyd
AT4G39980	0,3	1,7	2,609	0,02975	FALSE	"3-deoxy-D-arabino-heptulosonate 7-phosphate synth
AT2G43680	0	1	3,355	0,03159	FALSE	"IQ-domain 14 (IQD14); FUNCTIONS IN: calmodulin bi
AT1G33680	0	1	3,37	0,03159	FALSE	"KH domain-containing protein; FUNCTIONS IN: RNA b
AT1G62820	0	1,6	3,796	0,03159	FALSE	"Calcium-binding EF-hand family protein; FUNCTIONS
AT5G64040	0	0,9	3,237	0,03159	FALSE	"PSAN; FUNCTIONS IN: calmodulin binding; INVOLVED
AT2G02100	0,7	2,6	2,394	0,03159	TRUE	"low-molecular-weight cysteine-rich 69 (LCR69); FU
AT4G38680	0,3	3,6	3,451	0,03235	TRUE	"glycine rich protein 2 (GRP2); FUNCTIONS IN: doub
AT5G02530	0	1,1	3,461	0,03439	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT2G36250	0	0,9	3,206	0,03521	FALSE	"FTSZ2-1; FUNCTIONS IN: protein binding, structura
AT4G13940	0,7	2,1	2,203	0,03526	TRUE	"MATERNAL EFFECT EMBRYO ARREST 58 (MEE58); FUNCTIO
AT4G33510	0,3	1,6	2,482	0,03742	FALSE	"3-deoxy-d-arabino-heptulosonate 7-phosphate synth
AT5G53060	0	1	3,337	0,03789	FALSE	"RNA-binding KH domain-containing protein; FUNCTIO
AT3G14790	0	0,9	3,202	0,03822	FALSE	"rhamnose biosynthesis 3 (RHM3); CONTAINS InterPro
AT1G14710	0	1,3	3,606	0,03903	FALSE	"hydroxyproline-rich glycoprotein family protein;
AT1G34430	0	0,9	3,288	0,03903	FALSE	"embryo defective 3003 (EMB3003); FUNCTIONS IN: di
AT5G20000	0,7	2,4	2,321	0,03985	TRUE	"AAA-type ATPase family protein; FUNCTIONS IN: hyd
AT5G37720	0	1,9	3,992	0,03996	FALSE	"ALWAYS EARLY 4 (ALY4); FUNCTIONS IN: nucleotide b
AT1G13190	0	1,6	3,802	0,03999	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT2G33410	0	1,1	3,479	0,04076	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G25150	0	1	3,308	0,04421	FALSE	"Nuclear transport factor 2 (NTF2) family protein
AT1G54410	5	6	1,065	0,04635	TRUE	"dehydrin family protein; FUNCTIONS IN: molecular_
AT1G24510	0	1,1	3,616	0,04669	FALSE	"TCP-1/cpn60 chaperonin family protein; FUNCTIONS
AT4G34980	5	7	1,237	0,04903	TRUE	"subtilisin-like serine protease 2 (SLP2); FUNCTIO
AT5G04430	0	0,7	2,977	0,05025	FALSE	"BINDING TO TOMV RNA 1S (SHORT FORM) (BTR1S); FUNC
AT2G32700	0	0,9	3,149	0,05082	FALSE	"LEUNIG_homolog (LUH); CONTAINS InterPro DOMAIN/s:
AT2G38040	1,3	8,1	3,665	0,05196	FALSE	"acetyl Co-enzyme a carboxylase carboxyltransferas
AT5G14040	0	0,7	2,988	0,05321	FALSE	"phosphate transporter 3;1 (PHT3;1); FUNCTIONS IN:
AT3G52750	0	0,7	2,988	0,05321	FALSE	"FTSZ2-2; FUNCTIONS IN: structural molecule activi
AT5G54770	0	0,7	2,975	0,05331	FALSE	"THI1; FUNCTIONS IN: protein homodimerization acti
AT3G20250	2,3	14,1	3,857	0,0537	FALSE	"pumilio 5 (PUM5); FUNCTIONS IN: RNA binding, bind
AT1G26570	0	0,9	3,152	0,05402	FALSE	"UDP-glucose dehydrogenase 1 (UGD1); FUNCTIONS IN:
AT3G29360	0	0,9	3,152	0,05402	FALSE	"UDP-glucose 6-dehydrogenase family protein; FUNCT
AT1G80270	0,3	3,7	4,174	0,05901	FALSE	"PENTATRICOPEPTIDE REPEAT 596 (PPR596); FUNCTIONS
AT3G46520	6,3	6,9	0,9415	0,05901	FALSE	"actin-12 (ACT12); FUNCTIONS IN: structural consti
AT5G59370	6,3	6,9	0,9415	0,05901	FALSE	"actin 4 (ACT4); FUNCTIONS IN: structural constitu
AT3G07010	0	0,7	2,93	0,05977	FALSE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT1G08520	0	0,7	2,94	0,05977	FALSE	"ALBINA 1 (ALB1); FUNCTIONS IN: magnesium chelatas
AT1G55170	0	0,7	2,947	0,05977	FALSE	"unknown protein; BEST Arabidopsis thaliana protei
AT1G56070	9	11,6	1,129	0,06214	FALSE	"LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 1
AT5G16780	0,3	1,6	2,656	0,06321	FALSE	"DEFECTIVELY ORGANIZED TRIBUTARIES 2 (DOT2); FUNC
AT5G44340	10,7	20,9	1,658	0,06346	FALSE	"tubulin beta chain 4 (TUB4); FUNCTIONS IN: struct
AT5G12250	10	18,4	1,61	0,06485	FALSE	"beta-6 tubulin (TUB6); FUNCTIONS IN: structural c
AT4G20890	10,7	20,3	1,622	0,06602	FALSE	"tubulin beta-9 chain (TUB9); FUNCTIONS IN: struct
AT1G43850	0	0,7	2,898	0,06894	FALSE	"seuss (SEU); BEST Arabidopsis thaliana protein ma
AT3G05530	1	2,4	1,853	0,07104	FALSE	"regulatory particle triple-A ATPase 5A (RPT5A); C
AT1G29250	3,3	9,6	2,107	0,07171	FALSE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G55040	0	4,1	5,994	0,07211	FALSE	"zinc finger (Ran-binding) family protein; FUNCTIO
AT5G09810	9	9,1	0,8528	0,07293	FALSE	"actin 7 (ACT7); FUNCTIONS IN: protein binding, st

AT2G21390	2,3	4,3	1,561	0,07394	FALSE	"Coatomer, alpha subunit; FUNCTIONS IN: structural
AT5G06600	0	0,6	2,729	0,07584	FALSE	"ubiquitin-specific protease 12 (UBP12); FUNCTIONS
AT3G11910	0	0,6	2,729	0,07584	FALSE	"ubiquitin-specific protease 13 (UBP13); FUNCTIONS
AT1G06950	0	0,6	2,724	0,07584	FALSE	"translocon at the inner envelope membrane of chlo
AT2G02130	0	0,6	2,723	0,07584	FALSE	"low-molecular-weight cysteine-rich 68 (LCR68); FU
AT3G51950	0	0,6	2,723	0,07584	FALSE	"Zinc finger (CCCH-type) family protein / RNA reco
AT1G11860	0	0,6	2,722	0,07584	FALSE	"Glycine cleavage T-protein family; FUNCTIONS IN:
AT1G26460	0	0,6	2,722	0,07584	FALSE	"Tetratricopeptide repeat (TPR)-like superfamily p
AT3G05060	0	0,6	2,717	0,0759	FALSE	"NOP56-like pre RNA processing ribonucleoprotein;
AT3G49490	0	0,6	2,713	0,07614	FALSE	"unknown protein; Has 722 Blast hits to 186 protei
AT3G11830	0	0,6	2,712	0,07614	FALSE	"TCP-1/cpn60 chaperonin family protein; FUNCTIONS
AT5G12440	0	0,6	2,708	0,07619	FALSE	"CCCH-type zinc fingerfamily protein with RNA-bind
AT1G27090	4,7	10,3	1,74	0,07619	FALSE	"glycine-rich protein; BEST Arabidopsis thaliana p
AT5G53330	0	0,6	2,704	0,07633	FALSE	"Ubiquitin-associated/translation elongation facto
AT2G18330	0	0,6	2,698	0,07681	FALSE	"AAA-type ATPase family protein; FUNCTIONS IN: nuc
AT3G53460	0	0,6	2,693	0,07695	FALSE	"chloroplast RNA-binding protein 29 (CP29); FUNCTI
AT5G22040	0	0,6	2,693	0,07695	FALSE	"unknown protein; Has 30201 Blast hits to 17322 pr
AT3G06860	0	0,6	2,679	0,07828	FALSE	"multifunctional protein 2 (MFP2); FUNCTIONS IN: e
AT3G03920	0	0,6	2,674	0,07856	FALSE	"H/ACA ribonucleoprotein complex, subunit Gar1/Naf
AT3G60830	0	0,6	2,672	0,0786	FALSE	"actin-related protein 7 (ARP7); CONTAINS InterPro

Table S4. RNA –dependent partners of UPF1. Analysis of UPF1 partners enriched in non treated IPs compared to RNase-treated IPs. In this table are indicated the informations used to create the volcano plot in Figure 12A (only proteins with an *adjp<0.08* are listed here). Mean of spectral counts for control or UPF1 IPs, log2(Fold Change), adjusted p-value (adjp) are detailed. DEP indicates whether the considered protein passed the statistical test (TRUE) or not (FALSE). Proteins in grey are not significative.

AGI	Mean spectral count **Non-treated**	Mean spectral count **RNase-treated**	LogFC	adjp	DEP	Description
AT3G10770	0,1	15,3	5,884	5,607E-06	TRUE	"Single-stranded nucleic acid binding R3H protein;
AT3G54230	0	9,7	6,371	8,404E-06	TRUE	"suppressor of abi3-5 (SUA); FUNCTIONS IN: nucleot
AT2G15560	0,4	20,7	5,193	0,00008387	TRUE	"Putative endonuclease or glycosyl hydrolase; BEST
AT1G26110	0	6	5,65	0,0006666	TRUE	"decapping 5 (DCP5); CONTAINS InterPro DOMAIN/s: D
AT5G04280	0	4,3	5,194	0,00307	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein w
AT2G27100	0	4	5,095	0,0032	TRUE	"SERRATE (SE); FUNCTIONS IN: DNA binding, sequence
AT2G38110	0,1	6,7	4,645	0,0032	TRUE	"glycerol-3-phosphate acyltransferase 6 (GPAT6); C
AT4G27060	0	4	5,117	0,0032	TRUE	"TORTIFOLIA 1 (TOR1); CONTAINS InterPro DOMAIN/s:
AT2G42570	0	3,7	4,966	0,004314	TRUE	"TRICHOME BIREFRINGENCE-LIKE 39 (TBL39); INVOLVED
AT1G08370	0,3	7	4,169	0,009339	TRUE	"decapping 1 (DCP1); CONTAINS InterPro DOMAIN/s: D
AT3G13060	0	3	4,729	0,009339	TRUE	"evolutionarily conserved C-terminal region 5 (ECT
AT2G33845	0	3,7	4,88	0,009398	TRUE	"Nucleic acid-binding, OB-fold-like protein; FUNCT
AT2G28950	0	3	4,683	0,01256	TRUE	"expansin A6 (EXPA6); CONTAINS InterPro DOMAIN/s:
AT5G37720	0	4,3	5,077	0,0128	TRUE	"ALWAYS EARLY 4 (ALY4); FUNCTIONS IN: nucleotide b
AT4G35850	0	2,7	4,559	0,0128	TRUE	"Pentatricopeptide repeat (PPR) superfamily protei
AT1G51570	0	2,7	4,549	0,0128	TRUE	"Calcium-dependent lipid-binding (CaLB domain) pla
AT5G13300	0	2,7	4,521	0,0128	TRUE	"SCARFACE (SFC); FUNCTIONS IN: ARF GTPase activato
AT5G51750	0,7	7,3	3,236	0,0128	TRUE	"subtilase 1.3 (SBT1.3); FUNCTIONS IN: identical p
AT5G05100	0,4	8,7	3,893	0,0128	TRUE	"Single-stranded nucleic acid binding R3H protein;
AT1G13190	0	3,7	4,862	0,01431	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT2G46780	0	2,3	4,356	0,02028	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G14710	0	3	4,639	0,02028	TRUE	"hydroxyproline-rich glycoprotein family protein;
AT1G78570	0,1	4	4,055	0,02236	TRUE	"rhamnose biosynthesis 1 (RHM1); CONTAINS InterPro
AT2G43680	0	2,3	4,312	0,02391	TRUE	"IQ-domain 14 (IQD14); FUNCTIONS IN: calmodulin bi
AT2G33410	0	2,7	4,494	0,02635	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G53070	0,9	12,7	3,791	0,02948	TRUE	"Legume lectin family protein; FUNCTIONS IN: carbo
AT2G36250	0	2	4,169	0,03243	TRUE	"FTSZ2-1; FUNCTIONS IN: protein binding, structura
AT1G27430	0,3	5	3,638	0,04151	TRUE	"GYF domain-containing protein; CONTAINS InterPro
AT1G07310	0,4	7,3	3,745	0,04536	TRUE	"Calcium-dependent lipid-binding (CaLB domain) fam
AT2G32700	0	2	4,105	0,05885	FALSE	"LEUNIG_homolog (LUH); CONTAINS InterPro DOMAIN/s:
AT5G14040	0	1,7	3,953	0,06116	FALSE	"phosphate transporter 3;1 (PHT3;1); FUNCTIONS IN:
AT3G52750	0	1,7	3,953	0,06116	FALSE	"FTSZ2-2; FUNCTIONS IN: structural molecule activi
AT1G29670	0,1	3	3,612	0,06116	FALSE	"GDSL-like Lipase/Acylhydrolase superfamily protei
AT5G54770	0	1,7	3,913	0,06127	FALSE	"THI1; FUNCTIONS IN: protein homodimerization acti
AT4G32760	1	6	2,652	0,06734	FALSE	"ENTH/VHS/GAT family protein; FUNCTIONS IN: protei
AT3G61820	2	10,7	2,51	0,06734	FALSE	"Eukaryotic aspartyl protease family protein; FUNC
AT1G18660	0,3	4	3,441	0,07983	FALSE	"zinc finger (C3HC4-type RING finger) family prote
AT3G63400	0,4	4,7	3,426	0,07983	FALSE	"Cyclophilin-like peptidyl-prolyl cis-trans isomer
AT1G55170	0	1,7	3,883	0,07983	FALSE	"unknown protein; BEST Arabidopsis thaliana protei
AT1G08520	0	1,7	3,87	0,07983	FALSE	"ALBINA 1 (ALB1); FUNCTIONS IN: magnesium chelatas
AT3G07010	0	1,7	3,84	0,07986	FALSE	"Pectin lyase-like superfamily protein; FUNCTIONS

Table S5. Proteins enriched in DCP5-IP. In this table are indicated the informations used to create the volcano plot in ChapterI, Figure 15B (only proteins with an *adjp<0.08* are listed here). Mean of spectral counts for control or DCP5 IPs, log2(Fold Change), adjusted p-value (adjp) are detailed. DEP indicates whether the considered protein passed the statistical test (TRUE) or not (FALSE). Proteins in grey are not significantly enriched in DCP5-IPs.

AGI	Mean spectral count **Control**	Mean spectral count **DCP5**	LogFC	adjp	DEP	Description
AT1G26110	0	291,3	9,749	9,89E-31	TRUE	"decapping 5 (DCP5); CONTAINS InterPro DOMAIN/s: D
AT2G42520	1,7	139	2,912	3,23E-19	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT3G58570	1,6	138,7	3,033	1,51E-18	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT3G22330	0	75	7,785	1,1E-17	TRUE	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT4G00660	0,4	191,7	5,3	2,05E-17	TRUE	"RNAhelicase-like 8 (RH8); FUNCTIONS IN: helicase
AT3G58510	2	140	2,704	2,44E-14	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT2G45810	0,9	192,3	4,319	1,29E-13	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT1G29250	0,1	81,7	5,494	3,3E-13	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT5G40490	0	60,7	7,486	4,11E-13	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G14830	0	62,3	7,524	4,11E-13	TRUE	"DYNAMIN-like 1C (DL1C); FUNCTIONS IN: GTP binding
AT5G26742	3	144,3	2,164	9,05E-13	TRUE	"embryo defective 1138 (emb1138); FUNCTIONS IN: in
AT5G47010	0	73,3	7,786	6,55E-11	TRUE	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in
AT5G61020	0	44	7,022	7,59E-11	TRUE	"evolutionarily conserved C-terminal region 3 (ECT
AT3G61240	0,9	163	4,07	2,52E-10	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT5G51750	0,7	82,7	3,434	3,13E-09	TRUE	"subtilase 1.3 (SBT1.3); FUNCTIONS IN: identical p
AT2G34160	0	38	6,812	3,37E-09	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G10290	1	83,7	2,967	1,52E-08	TRUE	"dynamin-like protein 6 (ADL6); FUNCTIONS IN: GTPa
AT5G04280	0	39,3	6,865	1,55E-08	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein w
AT3G01540	0,4	57,7	3,579	2,59E-08	TRUE	"DEAD box RNA helicase 1 (DRH1); FUNCTIONS IN: ATP
AT1G43190	0	35,3	6,714	2,8E-08	TRUE	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT3G13460	0,6	53	3,062	2,95E-08	TRUE	"evolutionarily conserved C-terminal region 2 (ECT
AT4G28440	0,3	53,7	3,986	4,91E-08	TRUE	"Nucleic acid-binding, OB-fold-like protein; CONTA
AT1G76010	0,3	61,3	4,253	6,71E-08	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G48410	1	88,7	2,999	1,16E-07	TRUE	"ARGONAUTE 1 (AGO1); FUNCTIONS IN: protein binding
AT3G62250	1,6	73,3	2,134	2,78E-07	TRUE	"ubiquitin 5 (UBQ5); FUNCTIONS IN: structural cons
AT3G60190	0,6	53,3	3,062	2,84E-07	TRUE	"DYNAMIN-like 1E (DL1E); FUNCTIONS IN: GTP binding
AT5G63120	0,1	48,3	4,705	3,24E-07	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT1G59610	1,3	78,7	2,52	3,29E-07	TRUE	"dynamin-like 3 (DL3); FUNCTIONS IN: protein bindi
AT3G57150	1	61	2,478	3,6E-07	TRUE	"homologue of NAP57 (NAP57); FUNCTIONS IN: pseudou
AT3G22310	0	34	6,635	6,06E-07	TRUE	"putative mitochondrial RNA helicase 1 (PMH1); CON
AT2G27100	0	57	7,331	8,04E-07	TRUE	"SERRATE (SE); FUNCTIONS IN: DNA binding, sequence
AT5G63420	0	27	6,324	8,31E-07	TRUE	"embryo defective 2746 (emb2746); FUNCTIONS IN: hy
AT2G27880	0	27,7	6,358	8,64E-07	TRUE	"ARGONAUTE 5 (AGO5); FUNCTIONS IN: nucleic acid bi
AT1G33430	0	26,3	6,287	1,29E-06	TRUE	"Galactosyltransferase family protein; FUNCTIONS I
AT3G06480	0,3	38,7	3,539	1,48E-06	TRUE	"DEAD box RNA helicase family protein; FUNCTIONS I
AT4G34980	1,1	91,3	2,842	1,61E-06	TRUE	"subtilisin-like serine protease 2 (SLP2); FUNCTIO
AT1G31280	0	41,7	6,898	1,96E-06	TRUE	"argonaute 2 (AGO2); FUNCTIONS IN: nucleic acid bi
AT5G62390	1,1	73	2,569	2,67E-06	TRUE	"BCL-2-associated athanogene 7 (BAG7); FUNCTIONS I
AT1G70710	0	27,3	6,357	4,38E-06	TRUE	"glycosyl hydrolase 9B1 (GH9B1); FUNCTIONS IN: cel
AT4G38680	0,9	49	2,388	4,39E-06	TRUE	"glycine rich protein 2 (GRP2); FUNCTIONS IN: doub
AT3G13060	0	24	6,152	5,72E-06	TRUE	"evolutionarily conserved C-terminal region 5 (ECT
AT3G07030	0	23,7	6,138	6,39E-06	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G17870	0	24,3	6,185	7,77E-06	TRUE	"cold shock domain protein 3 (CSP3); FUNCTIONS IN:
AT5G44780	0	24,7	6,199	8,17E-06	TRUE	"unknown protein; BEST Arabidopsis thaliana protei
AT3G13860	0	35,7	6,768	9,38E-06	TRUE	"heat shock protein 60-3A (HSP60-3A); FUNCTIONS IN
AT5G04430	0	23,3	6,114	9,7E-06	TRUE	"BINDING TO TOMV RNA 1S (SHORT FORM) (BTR1S); FUNC
AT2G34810	0	22	6,035	1,47E-05	TRUE	"FAD-binding Berberine family protein; FUNCTIONS I
AT2G26280	0	21,7	6,007	1,84E-05	TRUE	"CID7; FUNCTIONS IN: damaged DNA binding, protein
AT5G53060	0	23,7	6,12	2,36E-05	TRUE	"RNA-binding KH domain-containing protein; FUNCTIO
AT5G65260	0	20,7	5,942	2,65E-05	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G48650	0	25	6,191	2,79E-05	TRUE	"Nuclear transport factor 2 (NTF2) family protein
AT1G64390	0	29,3	6,466	2,91E-05	TRUE	"glycosyl hydrolase 9C2 (GH9C2); FUNCTIONS IN: car
AT1G01300	1	61,3	2,481	3,25E-05	TRUE	"Eukaryotic aspartyl protease family protein; FUNC
AT4G21670	0	33,3	6,571	3,4E-05	TRUE	"C-terminal domain phosphatase-like 1 (CPL1); FUNC
AT1G20220	0,6	38,7	2,608	3,52E-05	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G75660	0	33,3	6,573	3,59E-05	TRUE	"5'-3' exoribonuclease 3 (XRN3); CONTAINS InterPro
AT5G20490	0	32,3	6,535	4,19E-05	TRUE	"XIK; FUNCTIONS IN: motor activity; INVOLVED IN: i
AT2G42570	0	19,7	5,872	4,19E-05	TRUE	"TRICHOME BIREFRINGENCE-LIKE 39 (TBL39); INVOLVED
AT3G25150	0,3	34,3	3,369	5,01E-05	TRUE	"Nuclear transport factor 2 (NTF2) family protein
AT3G03950	0	19,3	5,846	5,36E-05	TRUE	"evolutionarily conserved C-terminal region 1 (ECT
AT1G49760	0,3	32,7	3,302	6,08E-05	TRUE	"poly(A) binding protein 8 (PAB8); FUNCTIONS IN: R
AT5G52470	1	45,3	2,066	6,56E-05	TRUE	"fibrillarin 1 (FIB1); FUNCTIONS IN: snoRNA bindin
AT5G14610	0,4	36,3	2,926	9,05E-05	TRUE	"DEAD box RNA helicase family protein; FUNCTIONS I
AT5G03740	0	18	5,748	9,73E-05	TRUE	"histone deacetylase 2C (HD2C); FUNCTIONS IN: hist
AT1G16720	0	18,7	5,806	9,99E-05	TRUE	"high chlorophyll fluorescence phenotype 173 (HCF1
AT4G25630	0,9	41	2,136	0,000117	TRUE	"fibrillarin 2 (FIB2); FUNCTIONS IN: snoRNA bindin
AT4G27060	0	23	6,062	0,000117	TRUE	"TORTIFOLIA 1 (TOR1); CONTAINS InterPro DOMAIN/s:
AT1G17580	0	20	5,879	0,000129	TRUE	"myosin 1 (MYA1); FUNCTIONS IN: motor activity; IN

AT2G33845	0	17,3	5,694	0,000144	TRUE	"Nucleic acid-binding, OB-fold-like protein; FUNCT
AT4G02290	1	44	2,015	0,000146	TRUE	"glycosyl hydrolase 9B13 (GH9B13); FUNCTIONS IN: h
AT5G42080	2,6	88	1,691	0,000149	TRUE	"dynamin-like protein (DL1); FUNCTIONS IN: protein
AT2G44590	0	17,3	5,695	0,000164	TRUE	"DYNAMIN-like 1D (DL1D); FUNCTIONS IN: GTP binding
AT1G27090	0,7	79,3	3,343	0,000175	TRUE	"glycine-rich protein; BEST Arabidopsis thaliana p
AT5G58470	0	18	5,737	0,000177	TRUE	"TBP-associated factor 15B (TAF15b); FUNCTIONS IN:
AT2G44710	0	22,3	6,017	0,000187	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G37720	0	18	5,733	0,000221	TRUE	"ALWAYS EARLY 4 (ALY4); FUNCTIONS IN: nucleotide b
AT2G29190	0	16,3	5,609	0,000233	TRUE	"pumilio 2 (PUM2); FUNCTIONS IN: mRNA binding, RNA
AT3G60240	0	33,3	6,541	0,00028	TRUE	"eukaryotic translation initiation factor 4G (EIF4
AT1G48480	0	16,3	5,605	0,000288	TRUE	"receptor-like kinase 1 (RKL1); FUNCTIONS IN: prot
AT4G24780	0	17,7	5,736	0,000292	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT4G36020	0	16	5,585	0,000332	TRUE	"cold shock domain protein 1 (CSDP1); CONTAINS Int
AT4G26450	0	23,3	6,068	0,000355	TRUE	"unknown protein; Has 614 Blast hits to 492 protei
AT1G10170	0	22,3	6,005	0,000376	TRUE	"NF-X-like 1 (NFXL1); FUNCTIONS IN: zinc ion bindi
AT3G13300	0,4	31,3	2,737	0,000384	TRUE	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT2G29140	0	15,3	5,521	0,000409	TRUE	"pumilio 3 (PUM3); FUNCTIONS IN: RNA binding, bind
AT2G36870	0	15,3	5,522	0,000417	TRUE	"xyloglucan endotransglucosylase/hydrolase 32 (XTH
AT3G13224	0	16	5,577	0,000429	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G55150	0	15	5,489	0,000462	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT4G08350	0	32	6,479	0,000472	TRUE	"global transcription factor group A2 (GTA2); FUNC
AT1G04680	0	16	5,594	0,000485	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT1G14710	0	22,3	6,005	0,000495	TRUE	"hydroxyproline-rich glycoprotein family protein;
AT5G43810	0	15	5,492	0,0005	TRUE	"ZWILLE (ZLL); CONTAINS InterPro DOMAIN/s: Domain
AT1G72970	0,6	31,7	2,332	0,0005	TRUE	"HOTHEAD (HTH); CONTAINS InterPro DOMAIN/s: Glucos
AT3G06810	0,4	31	2,674	0,000679	TRUE	"IBA-RESPONSE 3 (IBR3); FUNCTIONS IN: acyl-CoA deh
AT3G26420	0	15	5,477	0,00078	TRUE	"ATRZ-1A; CONTAINS InterPro DOMAIN/s: RNA recognit
AT4G39260	2,1	66,7	1,581	0,000791	TRUE	"cold, circadian rhythm, and RNA binding 1 (CCR1);
AT5G55670	0	14	5,39	0,000829	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G02813	0	15,3	5,541	0,000915	TRUE	"Protein of unknown function, DUF538; INVOLVED IN:
AT5G13590	0	14	5,387	0,000938	TRUE	"unknown protein; Has 150 Blast hits to 121 protei
AT4G24270	0	13,7	5,36	0,000952	TRUE	"EMBRYO DEFECTIVE 140 (EMB140); FUNCTIONS IN: RNA
AT2G29200	0	13,7	5,358	0,000952	TRUE	"pumilio 1 (PUM1); FUNCTIONS IN: RNA binding, bind
AT4G21710	0	16	5,554	0,001158	TRUE	"NRPB2; CONTAINS InterPro DOMAIN/s: DNA-directed R
AT1G13020	0	23,3	6,042	0,001165	TRUE	"eukaryotic initiation factor 4B2 (EIF4B2); FUNCTI
AT2G43970	0	15,7	5,523	0,001177	TRUE	"RNA-binding protein; FUNCTIONS IN: RNA binding, n
AT1G69250	0	13,3	5,319	0,001238	TRUE	"Nuclear transport factor 2 (NTF2) family protein
AT3G07810	0	13	5,287	0,001322	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
ATCG00800	0	12,7	5,254	0,001576	TRUE	"RESISTANCE TO PSEUDOMONAS SYRINGAE 3 (RPS3); FUNC
AT2G32080	0	12,7	5,253	0,001576	TRUE	"purin-rich alpha 1 (PUR ALPHA-1); FUNCTIONS IN: n
AT1G54490	0	12,7	5,247	0,001728	TRUE	"exoribonuclease 4 (XRN4); CONTAINS InterPro DOMAI
AT4G31770	0	14	5,371	0,001752	TRUE	"debranching enzyme 1 (DBR1); CONTAINS InterPro DO
AT5G48900	0	13	5,293	0,001799	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT3G51950	0	12,3	5,214	0,001799	TRUE	"Zinc finger (CCCH-type) family protein / RNA reco
AT3G07010	0	12,7	5,259	0,001799	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT5G49360	0	12,3	5,212	0,001799	TRUE	"beta-xylosidase 1 (BXL1); FUNCTIONS IN: alpha-N-a
AT4G39680	0	13,7	5,345	0,00188	TRUE	"SAP domain-containing protein; FUNCTIONS IN: DNA
AT5G67240	0	13	5,295	0,001893	TRUE	"small RNA degrading nuclease 3 (SDN3); FUNCTIONS
AT5G64960	0	13,3	5,306	0,001918	TRUE	"cyclin dependent kinase group C2 (CDKC2); FUNCTIO
AT2G21060	0,4	26	2,444	0,001957	TRUE	"glycine-rich protein 2B (GRP2B); FUNCTIONS IN: DN
AT3G15000	0	12,3	5,209	0,00203	TRUE	"cobalt ion binding; FUNCTIONS IN: cobalt ion bind
AT5G57110	0	12	5,179	0,00214	TRUE	"autoinhibited Ca2+ -ATPase, isoform 8 (ACA8); FUN
AT5G15270	0	12	5,176	0,002151	TRUE	"RNA-binding KH domain-containing protein; FUNCTIO
AT3G54230	0	12,7	5,238	0,002265	TRUE	"suppressor of abi3-5 (SUA); FUNCTIONS IN: nucleot
AT3G12800	0,1	17,3	3,254	0,002336	TRUE	"short-chain dehydrogenase-reductase B (SDRB); FUN
AT3G23990	2,9	113,3	2,396	0,002362	TRUE	"heat shock protein 60 (HSP60); FUNCTIONS IN: copp
AT3G15010	0	11,7	5,134	0,002535	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G52530	0	11,7	5,134	0,002561	TRUE	"dentin sialophosphoprotein-related; Has 931 Blast
AT3G06980	0	11,7	5,136	0,002623	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT4G26630	0	13,7	5,333	0,002849	TRUE	"DEK domain-containing chromatin associated protei
AT3G25920	0	12,3	5,231	0,002955	TRUE	"ribosomal protein L15 (RPL15); FUNCTIONS IN: stru
AT2G06850	2,1	59,3	1,354	0,002957	TRUE	"xyloglucan endotransglucosylase/hydrolase 4 (XTH4
AT1G32790	0	11,3	5,092	0,003116	TRUE	"CTC-interacting domain 11 (CID11); FUNCTIONS IN:
AT5G22860	0	17,3	5,755	0,003161	TRUE	"Serine carboxypeptidase S28 family protein; FUNCT
AT1G55500	0	11,3	5,094	0,00317	TRUE	"evolutionarily conserved C-terminal region 4 (ECT
AT3G17840	0	12	5,163	0,00317	TRUE	"receptor-like kinase 902 (RLK902); FUNCTIONS IN:
AT5G12940	0	11,3	5,102	0,003197	TRUE	"Leucine-rich repeat (LRR) family protein; INVOLVE
AT2G28950	0	11,3	5,095	0,003292	TRUE	"expansin A6 (EXPA6); CONTAINS InterPro DOMAIN/s:
AT3G61690	0	13,3	5,293	0,003408	TRUE	"nucleotidyltransferases; FUNCTIONS IN: nucleotidy
AT2G15560	0	11	5,055	0,003443	TRUE	"Putative endonuclease or glycosyl hydrolase; BEST
AT1G71770	0	12,3	5,198	0,003583	TRUE	"poly(A)-binding protein 5 (PAB5); FUNCTIONS IN: R
AT4G08950	0	11	5,058	0,003583	TRUE	"EXORDIUM (EXO); FUNCTIONS IN: molecular_function
AT5G61970	0	11,3	5,107	0,003657	TRUE	"signal recognition particle-related / SRP-related
AT2G22400	0	12	5,156	0,003709	TRUE	"S-adenosyl-L-methionine-dependent methyltransfera
AT3G04620	0	11,3	5,102	0,003768	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G50370	0	17,3	5,622	0,003792	TRUE	"unknown protein; FUNCTIONS IN: molecular_function
AT5G10270	0	11,7	5,118	0,004033	TRUE	"cyclin-dependent kinase C;1 (CDKC;1); FUNCTIONS I
AT1G56110	1,3	39,7	1,513	0,004111	TRUE	"homolog of nucleolar protein NOP56 (NOP56); LOCAT
AT2G23350	0,4	26,7	2,46	0,004123	TRUE	"poly(A) binding protein 4 (PAB4); FUNCTIONS IN: R
AT1G47490	0	10,7	5,012	0,004123	TRUE	"RNA-binding protein 47C (RBP47C); CONTAINS InterP

AT3G21100	0	10,7	5,008	0,004377	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G27120	0,1	16	3,149	0,004555	TRUE	"NOP56-like pre RNA processing ribonucleoprotein;
AT1G30350	0	12	5,178	0,004615	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT1G50920	0	10,3	4,966	0,004791	TRUE	"Nucleolar GTP-binding protein; FUNCTIONS IN: GTP
AT3G04610	0	15	5,433	0,004792	TRUE	"flowering locus KH domain (FLK); FUNCTIONS IN: RN
AT5G55660	0	11,3	5,077	0,004792	TRUE	"DEK domain-containing chromatin associated protei
AT1G13190	0	13	5,247	0,004821	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G01080	0	10,3	4,97	0,00488	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G02830	0	10,3	4,97	0,00488	TRUE	"zinc finger protein 1 (ZFN1); CONTAINS InterPro D
AT3G05060	1	34	1,641	0,005059	TRUE	"NOP56-like pre RNA processing ribonucleoprotein;
AT1G30680	0	10,3	4,962	0,005069	TRUE	"toprim domain-containing protein; FUNCTIONS IN: D
AT4G29010	0,4	28,7	2,549	0,005075	TRUE	"ABNORMAL INFLORESCENCE MERISTEM (AIM1); FUNCTIONS
AT3G60830	0	10,7	4,999	0,00517	TRUE	"actin-related protein 7 (ARP7); CONTAINS InterPro
AT4G11050	0	11,7	5,164	0,005279	TRUE	"glycosyl hydrolase 9C3 (GH9C3); FUNCTIONS IN: car
AT1G08370	0	10	4,922	0,005583	TRUE	"decapping 1 (DCP1); CONTAINS InterPro DOMAIN/s: D
AT1G54560	0	13	5,264	0,005617	TRUE	"XIE; CONTAINS InterPro DOMAIN/s: Dil domain (Inte
AT2G35920	0	10,3	4,957	0,005625	TRUE	"RNA helicase family protein; FUNCTIONS IN: helica
AT5G54900	0	10	4,921	0,005625	TRUE	"RNA-binding protein 45A (RBP45A); FUNCTIONS IN: R
AT2G31810	0	12	5,144	0,005672	TRUE	"ACT domain-containing small subunit of acetolacta
AT4G31580	0	10,3	4,978	0,005672	TRUE	"serine/arginine-rich 22 (SRZ-22); FUNCTIONS IN: p
AT3G01780	0	10	4,918	0,005708	TRUE	"TPLATE; FUNCTIONS IN: binding; INVOLVED IN: polle
AT3G52150	0	10	4,922	0,005774	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G13290	0,3	18,7	2,501	0,006127	TRUE	"varicose-related (VCR); FUNCTIONS IN: nucleotide
AT4G09040	0	10	4,93	0,006376	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G26650	0	10,3	4,951	0,006435	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G08610	0	10,3	4,952	0,006543	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT5G11350	0	9,7	4,871	0,006707	TRUE	"DNAse I-like superfamily protein; CONTAINS InterP
AT3G28500	0	9,7	4,869	0,006793	TRUE	"60S acidic ribosomal protein family; FUNCTIONS IN
AT1G47128	0	10,3	4,987	0,0068	TRUE	"responsive to dehydration 21 (RD21); FUNCTIONS IN
AT4G33200	0	13,3	5,265	0,00716	TRUE	"XI-I; FUNCTIONS IN: motor activity; INVOLVED IN:
AT1G48650	0	9,3	4,825	0,007663	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT5G24710	0	9,3	4,825	0,007663	TRUE	"Transducin/WD40 repeat-like superfamily protein;
AT1G13730	0	9,3	4,819	0,008324	TRUE	"Nuclear transport factor 2 (NTF2) family protein
AT5G64270	0	9,3	4,828	0,00833	TRUE	"splicing factor, putative; FUNCTIONS IN: binding;
AT5G22060	0	10	4,901	0,00833	TRUE	"DNAJ homologue 2 (J2); FUNCTIONS IN: unfolded pro
AT1G22760	0	14,7	5,398	0,008375	TRUE	"poly(A) binding protein 3 (PAB3); CONTAINS InterP
AT3G12410	0	10,3	4,998	0,008557	TRUE	"Polynucleotidyl transferase, ribonuclease H-like
AT1G14840	0	14,7	5,378	0,009039	TRUE	"microtubule-associated proteins 70-4 (MAP70-4); F
AT2G47250	0	9	4,778	0,009289	TRUE	"RNA helicase family protein; FUNCTIONS IN: in 7 f
AT5G54600	0	9	4,776	0,009399	TRUE	"Translation protein SH3-like family protein; FUNC
AT1G47500	0	9	4,78	0,00943	TRUE	"RNA-binding protein 47C' (RBP47C'); CONTAINS Inte
AT1G33680	0	11,3	5,049	0,009475	TRUE	"KH domain-containing protein; FUNCTIONS IN: RNA b
AT3G49240	0	9,7	4,854	0,009552	TRUE	"embryo defective 1796 (emb1796); FUNCTIONS IN: AT
AT4G34110	0,7	31	2,057	0,009825	TRUE	"poly(A) binding protein 2 (PAB2); FUNCTIONS IN: R
AT5G51120	0	9	4,767	0,01004	TRUE	"polyadenylate-binding protein 1 (PABN1); CONTAINS
AT2G13540	0	10	4,893	0,01006	TRUE	"ABA HYPERSENSITIVE 1 (ABH1); CONTAINS InterPro DO
AT2G43680	0	9,7	4,85	0,01036	TRUE	"IQ-domain 14 (IQD14); FUNCTIONS IN: calmodulin bi
AT1G75930	0,4	25	2,422	0,01036	TRUE	"extracellular lipase 6 (EXL6); CONTAINS InterPro
AT1G75560	0	8,7	4,724	0,0107	TRUE	"zinc knuckle (CCHC-type) family protein; FUNCTION
AT5G55230	0	8,7	4,724	0,0107	TRUE	"microtubule-associated proteins 65-1 (MAP65-1); F
AT2G38610	0	8,7	4,719	0,0108	TRUE	"RNA-binding KH domain-containing protein; FUNCTIO
AT4G32460	0	8,7	4,725	0,01097	TRUE	"FUNCTIONS IN: molecular_function unknown; INVOLVE
AT1G07310	0	9,7	4,857	0,01097	TRUE	"Calcium-dependent lipid-binding (CaLB domain) fam
AT3G52500	0	8,7	4,716	0,01148	TRUE	"Eukaryotic aspartyl protease family protein; FUNC
AT4G25550	0,4	23,3	2,271	0,01168	TRUE	"Cleavage/polyadenylation specificity factor, 25kD
AT2G33210	2,7	83,7	2,007	0,01235	TRUE	"heat shock protein 60-2 (HSP60-2); FUNCTIONS IN:
AT4G29510	0	8,3	4,67	0,01267	TRUE	"arginine methyltransferase 11 (PRMT11); FUNCTIONS
AT5G47690	0	8,3	4,665	0,01283	TRUE	"binding; FUNCTIONS IN: binding; INVOLVED IN: biol
AT1G31850	0	8,3	4,663	0,013	TRUE	"S-adenosyl-L-methionine-dependent methyltransfera
AT3G61820	0,4	20,7	2,1	0,01315	TRUE	"Eukaryotic aspartyl protease family protein; FUNC
AT2G38040	1,1	37,3	1,654	0,01315	TRUE	"acetyl Co-enzyme a carboxylase carboxyltransferas
AT1G08730	0	10	4,893	0,01335	TRUE	"XIC; CONTAINS InterPro DOMAIN/s: Dil domain (Inte
AT5G14320	0	8,3	4,676	0,01344	TRUE	"Ribosomal protein S13/S18 family; FUNCTIONS IN: s
AT1G49670	0	8,3	4,665	0,01428	TRUE	"NQR; FUNCTIONS IN: oxidoreductase activity, bindi
AT5G21150	0	8,7	4,702	0,01446	TRUE	"ARGONAUTE 9 (AGO9); FUNCTIONS IN: siRNA binding;
AT3G49490	0	8,3	4,682	0,01446	TRUE	"unknown protein; Has 722 Blast hits to 186 protei
AT3G11964	0	8,3	4,677	0,01446	TRUE	"RNA binding;RNA binding; FUNCTIONS IN: RNA bindin
AT2G39990	0	8,7	4,705	0,01483	TRUE	"eukaryotic translation initiation factor 2 (EIF2)
AT4G03110	0	8,7	4,705	0,01483	TRUE	"RNA-binding protein-defense related 1 (RBP-DR1);
AT5G67500	0	8	4,608	0,01489	TRUE	"voltage dependent anion channel 2 (VDAC2); FUNCTI
AT1G24260	0	8	4,619	0,01598	TRUE	"SEPALLATA3 (SEP3); FUNCTIONS IN: protein binding,
AT5G35970	0	8	4,619	0,01598	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT2G32000	0	8	4,606	0,0163	TRUE	"DNA topoisomerase, type IA, core; FUNCTIONS IN: D
AT4G10610	0	8	4,606	0,0163	TRUE	"CTC-interacting domain 12 (CID12); CONTAINS Inter
AT3G13750	0	9	4,74	0,01668	TRUE	"beta galactosidase 1 (BGAL1); FUNCTIONS IN: beta-
AT5G26670	0	7,7	4,554	0,01751	TRUE	"Pectinacetylesterase family protein; FUNCTIONS IN
AT4G35890	0	9,3	4,781	0,01756	TRUE	"winged-helix DNA-binding transcription factor fam
AT3G12860	0,1	12,7	2,816	0,01759	TRUE	"NOP56-like pre RNA processing ribonucleoprotein;
AT1G02800	0	7,7	4,555	0,01776	TRUE	"cellulase 2 (CEL2); FUNCTIONS IN: cellulase activ
AT1G01100	0	7,7	4,549	0,01776	TRUE	"60S acidic ribosomal protein family; FUNCTIONS IN

AT5G65350	0	7,7	4,555	0,01866	TRUE	"histone 3 11 (HTR11); FUNCTIONS IN: DNA binding;
AT3G54540	0	8	4,637	0,01989	TRUE	"general control non-repressible 4 (GCN4); CONTAIN
AT2G46020	0	12,3	5,12	0,01989	TRUE	"BRAHMA (BRM); FUNCTIONS IN: helicase activity, tr
AT3G29310	0	9	4,763	0,01989	TRUE	"calmodulin-binding protein-related; FUNCTIONS IN:
AT3G13990	0	9	4,73	0,02074	TRUE	"Kinase-related protein of unknown function (DUF12
AT5G18550	0	7,3	4,489	0,02085	TRUE	"Zinc finger C-x8-C-x5-C-x3-H type family protein;
AT5G55920	0	7,3	4,493	0,02085	TRUE	"OLIGOCELLULA 2 (OLI2); FUNCTIONS IN: S-adenosylme
AT2G37230	0	7,7	4,572	0,02147	TRUE	"Tetratricopeptide repeat (TPR)-like superfamily p
AT1G07910	0	7,3	4,49	0,0217	TRUE	"RNAligase (RNL); CONTAINS InterPro DOMAIN/s: tRNA
AT3G53420	0	7,3	4,496	0,0217	TRUE	"plasma membrane intrinsic protein 2A (PIP2A); FUN
AT4G13710	0	7,3	4,497	0,02298	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT2G18960	1,6	38	1,176	0,02362	TRUE	"H(+)-ATPase 1 (HA1); FUNCTIONS IN: protein bindin
AT4G09150	0	8	4,58	0,02362	TRUE	"T-complex protein 11; FUNCTIONS IN: phosphopantet
AT2G37990	0	7,3	4,479	0,02399	TRUE	"ribosome biogenesis regulatory protein (RRS1) fam
AT5G22040	0	7	4,429	0,02472	TRUE	"unknown protein; Has 30201 Blast hits to 17322 pr
AT1G80270	0	7,3	4,482	0,02473	TRUE	"PENTATRICOPEPTIDE REPEAT 596 (PPR596); FUNCTIONS
AT1G63640	0	7,3	4,498	0,02575	TRUE	"P-loop nucleoside triphosphate hydrolases superfa
AT3G50590	0	7	4,42	0,02614	TRUE	"Transducin/WD40 repeat-like superfamily protein;
AT3G54670	0	7	4,42	0,02614	TRUE	"TITAN8 (TTN8); FUNCTIONS IN: transporter activity
AT5G38600	0	10,7	4,93	0,02624	TRUE	"Proline-rich spliceosome-associated (PSP) family
AT3G49390	0	7	4,421	0,02624	TRUE	"CTC-interacting domain 10 (CID10); FUNCTIONS IN:
AT2G40660	0	7	4,421	0,02647	TRUE	"Nucleic acid-binding, OB-fold-like protein; FUNCT
AT2G44530	0	7	4,433	0,02647	TRUE	"Phosphoribosyltransferase family protein; FUNCTIO
AT5G20950	0	7,3	4,477	0,02647	TRUE	"Glycosyl hydrolase family protein; FUNCTIONS IN:
AT2G30800	0	7	4,438	0,02647	TRUE	"helicase in vascular tissue and tapetum (HVT1); C
AT5G63260	0	7	4,418	0,02656	TRUE	"Zinc finger C-x8-C-x5-C-x3-H type family protein;
AT5G21160	0	7,7	4,538	0,02665	TRUE	"LA RNA-binding protein; FUNCTIONS IN: molecular_f
AT2G28290	0	12,7	5,14	0,02692	TRUE	"SPLAYED (SYD); CONTAINS InterPro DOMAIN/s: DEAD-l
AT3G19960	0	10	4,84	0,02789	TRUE	"myosin 1 (ATM1); FUNCTIONS IN: motor activity; IN
AT1G80410	0,9	27,7	1,683	0,02789	TRUE	"EMBRYO DEFECTIVE 2753 (EMB2753); FUNCTIONS IN: bi
AT4G30190	1,1	30,7	1,315	0,02794	TRUE	"H(+)-ATPase 2 (HA2); FUNCTIONS IN: ATPase activit
AT5G43960	0	6,7	4,358	0,02854	TRUE	"Nuclear transport factor 2 (NTF2) family protein
AT5G03280	0	7,7	4,517	0,02855	TRUE	"ETHYLENE INSENSITIVE 2 (EIN2); FUNCTIONS IN: tran
AT1G30360	0	6,7	4,358	0,02864	TRUE	"early-responsive to dehydration 4 (ERD4); INVOLVE
AT3G14750	0	6,7	4,358	0,02864	TRUE	"unknown protein; BEST Arabidopsis thaliana protei
AT1G10590	0	6,7	4,364	0,02864	TRUE	"Nucleic acid-binding, OB-fold-like protein; INVOL
AT5G02530	0,6	21	1,74	0,02967	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G22770	0	7,3	4,473	0,02995	TRUE	"alpha-adaptin (alpha-ADR); FUNCTIONS IN: protein
AT5G51550	0	7	4,447	0,03042	TRUE	"EXORDIUM like 3 (EXL3); CONTAINS InterPro DOMAIN/
AT2G40610	0	6,7	4,356	0,03042	TRUE	"expansin A8 (EXPA8); INVOLVED IN: plant-type cell
AT2G27170	0	7	4,408	0,03053	TRUE	"TITAN7 (TTN7); FUNCTIONS IN: ATP binding; INVOLVE
AT1G68060	0	13,3	5,204	0,03141	TRUE	"microtubule-associated proteins 70-1 (MAP70-1); F
AT3G50670	0	6,7	4,352	0,03141	TRUE	"U1 small nuclear ribonucleoprotein-70K (U1-70K);
AT3G06400	0	7,7	4,518	0,03149	TRUE	"chromatin-remodeling protein 11 (CHR11); FUNCTION
AT3G03920	0,1	11,3	2,657	0,03162	TRUE	"H/ACA ribonucleoprotein complex, subunit Gar1/Naf
AT1G70180	0	6,7	4,345	0,03354	TRUE	"Sterile alpha motif (SAM) domain-containing prote
AT5G59950	0,3	15,3	2,229	0,03362	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G30460	0	7	4,399	0,03362	TRUE	"cleavage and polyadenylation specificity factor 3
AT3G07050	0	7	4,399	0,03362	TRUE	"GTP-binding family protein; FUNCTIONS IN: GTP bin
AT5G53620	0	7	4,399	0,03362	TRUE	"unknown protein; INVOLVED IN: biological_process
AT2G07690	0	6,7	4,343	0,03362	TRUE	"MINICHROMOSOME MAINTENANCE 5 (MCM5); FUNCTIONS IN
AT2G19170	0	6,3	4,287	0,03362	TRUE	"subtilisin-like serine protease 3 (SLP3); FUNCTIO
AT3G15590	0	6,3	4,287	0,03362	TRUE	"Tetratricopeptide repeat (TPR)-like superfamily p
AT3G59770	0	6,7	4,344	0,0339	TRUE	"SUPPRESSOR OF ACTIN 9 (SAC9); FUNCTIONS IN: inosi
AT5G63140	0	6,3	4,297	0,03448	TRUE	"purple acid phosphatase 29 (PAP29); FUNCTIONS IN:
AT2G01750	0	9,3	4,743	0,03449	TRUE	"microtubule-associated proteins 70-3 (MAP70-3); F
AT3G44750	0	6,3	4,284	0,03459	TRUE	"histone deacetylase 3 (HDA3); FUNCTIONS IN: histo
AT3G12610	0	6,3	4,296	0,03475	TRUE	"DNA-DAMAGE REPAIR/TOLERATION 100 (DRT100); CONTAI
AT3G19130	0	6,3	4,288	0,03486	TRUE	"RNA-binding protein 47B (RBP47B); FUNCTIONS IN: R
AT5G55550	0	6,3	4,283	0,03495	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G22010	2,6	60,7	1,094	0,03506	TRUE	"SKU5 similar 4 (sks4); FUNCTIONS IN: oxidoreduct
AT4G36960	0	7,3	4,45	0,0354	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G28880	0	6,7	4,338	0,03564	TRUE	"casein kinase I-like 3 (ckl3); FUNCTIONS IN: prot
AT2G38110	0,1	11	2,615	0,03592	TRUE	"glycerol-3-phosphate acyltransferase 6 (GPAT6); C
AT1G61730	0	6,3	4,291	0,03642	TRUE	"DNA-binding storekeeper protein-related transcrip
AT5G43900	0	8,7	4,648	0,03657	TRUE	"myosin 2 (MYA2); CONTAINS InterPro DOMAIN/s: Dil
AT5G23400	0	6,3	4,277	0,03701	TRUE	"Leucine-rich repeat (LRR) family protein; INVOLVE
AT1G79350	0	6,3	4,284	0,03856	TRUE	"embryo defective 1135 (EMB1135); FUNCTIONS IN: DN
AT1G79280	0	6,7	4,332	0,03949	TRUE	"nuclear pore anchor (NUA); FUNCTIONS IN: molecula
AT1G67750	0	6	4,217	0,03982	TRUE	"Pectate lyase family protein; FUNCTIONS IN: pecta
AT3G23700	0	6	4,215	0,0399	TRUE	"Nucleic acid-binding proteins superfamily; FUNCTI
AT3G62600	0	6	4,219	0,04031	TRUE	"ATERDJ3B; FUNCTIONS IN: unfolded protein binding,
AT5G40200	0	6	4,222	0,04059	TRUE	"DegP protease 9 (DegP9); FUNCTIONS IN: serine-typ
AT3G61130	0	6	4,211	0,04069	TRUE	"galacturonosyltransferase 1 (GAUT1); CONTAINS Int
AT4G34660	0	6,3	4,268	0,0421	TRUE	"SH3 domain-containing protein; CONTAINS InterPro
AT3G44110	0,4	16,7	1,81	0,04385	TRUE	"DNAJ homologue 3 (ATJ3); FUNCTIONS IN: unfolded p
AT5G26710	0	6	4,22	0,04484	TRUE	"Glutamyl/glutaminyl-tRNA synthetase, class Ic; FU
AT1G18150	0	6	4,208	0,04507	TRUE	"ATMPK8; FUNCTIONS IN: MAP kinase activity; INVOLV
AT3G14100	0,1	11,3	2,655	0,04538	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G16210	0	6,3	4,316	0,04593	TRUE	"enoyl-CoA hydratase/isomerase A (ECHIA); FUNCTION

AT2G46780	0	6	4,2	0,04617	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G43700	0	6	4,197	0,04745	TRUE	"VIRE2-interacting protein 1 (VIP1); FUNCTIONS IN:
AT1G12770	0	5,7	4,138	0,04762	TRUE	"embryo defective 1586 (EMB1586); FUNCTIONS IN: RN
AT2G35410	0	5,7	4,136	0,04771	TRUE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G12440	0	5,7	4,136	0,04771	TRUE	"CCCH-type zinc fingerfamily protein with RNA-bind
AT1G19880	0	5,7	4,135	0,04814	TRUE	"Regulator of chromosome condensation (RCC1) famil
AT5G41460	0	5,7	4,143	0,04852	TRUE	"Protein of unknown function (DUF604); CONTAINS In
AT5G18590	0	5,7	4,133	0,04894	TRUE	"Galactose oxidase/kelch repeat superfamily protei
AT5G62670	0,9	24,7	1,434	0,04943	TRUE	"H(+)-ATPase 11 (HA11); FUNCTIONS IN: ATPase activ
AT3G58660	0	6	4,197	0,04991	TRUE	"Ribosomal protein L1p/L10e family; FUNCTIONS IN:
AT2G24590	0,1	10,3	2,542	0,05	FALSE	"RNA recognition motif and CCHC-type zinc finger d
AT1G10270	0	6,3	4,266	0,05079	FALSE	"glutamine-rich protein 23 (GRP23); FUNCTIONS IN:
AT1G75940	2,9	72,3	1,222	0,05115	FALSE	"ATA27; FUNCTIONS IN: cation binding, hydrolase ac
AT5G19010	0	5,7	4,125	0,0522	FALSE	"mitogen-activated protein kinase 16 (MPK16); FUNC
AT5G63810	0	5,7	4,125	0,0522	FALSE	"beta-galactosidase 10 (BGAL10); FUNCTIONS IN: cat
AT4G01037	0	6	4,215	0,0522	FALSE	"what's this factor? (WTF1); CONTAINS InterPro DOM
AT1G65440	0	7,3	4,42	0,05261	FALSE	"global transcription factor group B1 (GTB1); FUNC
AT4G39000	0	5,7	4,159	0,05377	FALSE	"glycosyl hydrolase 9B17 (GH9B17); FUNCTIONS IN: h
AT4G28990	0	5,3	4,058	0,05681	FALSE	"RNA-binding protein-related; FUNCTIONS IN: zinc i
AT1G04990	0	5,3	4,059	0,05686	FALSE	"Zinc finger C-x8-C-x5-C-x3-H type family protein;
AT4G24290	0	5,3	4,056	0,05718	FALSE	"MAC/Perforin domain-containing protein; FUNCTIONS
AT4G12600	0	5,3	4,053	0,05728	FALSE	"Ribosomal protein L7Ae/L30e/S12e/Gadd45 family pr
AT1G76810	0	7	4,369	0,05754	FALSE	"eukaryotic translation initiation factor 2 (eIF-2
AT5G10350	0	5,7	4,115	0,0578	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G24850	0	5,7	4,117	0,05801	FALSE	"cryptochrome 3 (CRY3); FUNCTIONS IN: FMN binding,
AT4G25340	0	5,3	4,048	0,05876	FALSE	"FK506 BINDING PROTEIN 53 (FKBP53); FUNCTIONS IN:
AT1G22015	0	5,3	4,053	0,05882	FALSE	"DD46; FUNCTIONS IN: transferase activity, transfe
AT5G46630	0	5,7	4,125	0,05912	FALSE	"Clathrin adaptor complexes medium subunit family
AT5G18620	0	6	4,186	0,05916	FALSE	"chromatin remodeling factor17 (CHR17); FUNCTIONS
AT5G13300	0	5,3	4,046	0,05942	FALSE	"SCARFACE (SFC); FUNCTIONS IN: ARF GTPase activato
AT5G51980	0	5,3	4,048	0,05992	FALSE	"Transducin/WD40 repeat-like superfamily protein;
AT4G23400	0	5,3	4,071	0,06212	FALSE	"plasma membrane intrinsic protein 1;5 (PIP1;5); F
AT1G11650	0,1	9,7	2,437	0,06226	FALSE	"RBP45B; FUNCTIONS IN: RNA binding; EXPRESSED IN:
AT5G58040	0	6	4,171	0,06345	FALSE	"homolog of yeast FIP1 [V] (FIP1[V]); FUNCTIONS IN
AT1G62390	0,1	10,3	2,524	0,06711	FALSE	"Octicosapeptide/Phox/Bem1p (PB1) domain-containin
AT2G39800	0	5	3,966	0,06711	FALSE	"delta1-pyrroline-5-carboxylate synthase 1 (P5CS1)
AT3G60245	0	5	3,966	0,06711	FALSE	"Zinc-binding ribosomal protein family protein; FU
AT5G14520	0	5	3,966	0,06711	FALSE	"pescadillo-related; FUNCTIONS IN: transcription c
AT1G27430	0	5,3	4,034	0,06714	FALSE	"GYF domain-containing protein; CONTAINS InterPro
AT3G09100	0	5,3	4,034	0,06714	FALSE	"mRNA capping enzyme family protein; FUNCTIONS IN:
AT3G18600	0	6,3	4,233	0,06739	FALSE	"P-loop containing nucleoside triphosphate hydrola
AT1G60650	0	5	3,973	0,06758	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein w
AT5G42390	0	6,3	4,361	0,06769	FALSE	"Insulinase (Peptidase family M16) family protein;
AT1G17370	0	5	3,963	0,06769	FALSE	"oligouridylate binding protein 1B (UBP1B); FUNCTI
AT3G59990	0	5	3,963	0,06769	FALSE	"methionine aminopeptidase 2B (MAP2B); FUNCTIONS I
AT2G39670	0	5	3,965	0,0678	FALSE	"Radical SAM superfamily protein; FUNCTIONS IN: ir
AT1G54080	0,1	9,7	2,449	0,0678	FALSE	"oligouridylate-binding protein 1A (UBP1A); FUNCTI
AT5G16750	0	5	3,962	0,0682	FALSE	"TORMOZEMBRYO DEFECTIVE (TOZ); CONTAINS InterPro D
AT1G24764	0	7,7	4,453	0,06857	FALSE	"microtubule-associated proteins 70-2 (MAP70-2); F
AT1G69840	0	5	3,968	0,06916	FALSE	"SPFH/Band 7/PHB domain-containing membrane-associ
AT1G14410	0	5	3,98	0,07013	FALSE	"WHIRLY 1 (WHY1); FUNCTIONS IN: DNA binding, telom
AT1G74970	0	5	3,98	0,07013	FALSE	"ribosomal protein S9 (RPS9); FUNCTIONS IN: struct
AT2G33410	0,1	9,7	2,437	0,07129	FALSE	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G14170	0	6,3	4,218	0,07214	FALSE	"CHC1; CONTAINS InterPro DOMAIN/s: SWIB/MDM2 domai
AT4G24680	0,1	9,3	2,386	0,07218	FALSE	"MODIFIER OF snc1 (MOS1); FUNCTIONS IN: molecular_
AT4G39200	0,3	12,3	1,919	0,07337	FALSE	"Ribosomal protein S25 family protein; FUNCTIONS I
AT5G04290	0	7	4,336	0,07375	FALSE	"kow domain-containing transcription factor 1 (KTF
AT3G62310	0,1	9,3	2,401	0,07414	FALSE	"RNA helicase family protein; FUNCTIONS IN: RNA he
AT5G07510	0	5	3,957	0,07414	FALSE	"glycine-rich protein 14 (GRP14); CONTAINS InterPr
AT4G16340	0	4,7	3,875	0,07837	FALSE	"SPIKE1 (SPK1); FUNCTIONS IN: GTPase binding, GTP
AT3G57660	0	4,7	3,873	0,07861	FALSE	"nuclear RNA polymerase A1 (NRPA1); FUNCTIONS IN:
AT4G30440	0	4,7	3,873	0,07861	FALSE	"UDP-D-glucuronate 4-epimerase 1 (GAE1); FUNCTIONS
AT3G54340	0	4,7	3,878	0,07918	FALSE	"APETALA 3 (AP3); CONTAINS InterPro DOMAIN/s: Tran
AT1G01510	0	4,7	3,871	0,07918	FALSE	"ANGUSTIFOLIA (AN); CONTAINS InterPro DOMAIN/s: D-
AT1G47980	0	4,7	3,881	0,07918	FALSE	"unknown protein; FUNCTIONS IN: molecular_function
AT4G23490	0	4,7	3,881	0,07918	FALSE	"Protein of unknown function (DUF604); CONTAINS In
AT1G78630	0	4,7	3,873	0,07923	FALSE	"embryo defective 1473 (emb1473); FUNCTIONS IN: st
AT5G17170	0	4,7	3,874	0,07923	FALSE	"enhancer of sos3-1 (ENH1); FUNCTIONS IN: electron

Table S6. Common protein set between DCP5-IPs and At-RBP. Comparison of proteins enriched in DCP5-IPs with *Arabidopsis* RNA Binding Proteins (At-RBP) from Reichel *et al.*, 2016.

AT1G01080	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G07910	"RNAligase (RNL); CONTAINS InterPro DOMAIN/s: tRNA
AT1G10170	"NF-X-like 1 (NFXL1); FUNCTIONS IN: zinc ion bindi
AT1G10290	"dynamin-like protein 6 (ADL6); FUNCTIONS IN: GTPa
AT1G10590	"Nucleic acid-binding, OB-fold-like protein; INVOL
AT1G12770	"embryo defective 1586 (EMB1586); FUNCTIONS IN: RN
AT1G13020	"eukaryotic initiation factor 4B2 (EIF4B2); FUNCTI
AT1G13190	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G13730	"Nuclear transport factor 2 (NTF2) family protein
AT1G14710	"hydroxyproline-rich glycoprotein family protein;
AT1G20220	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G26110	"decapping 5 (DCP5); CONTAINS InterPro DOMAIN/s: D
AT1G27090	"glycine-rich protein; BEST Arabidopsis thaliana p
AT1G29250	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G30460	"cleavage and polyadenylation specificity factor 3
AT1G31280	"argonaute 2 (AGO2); FUNCTIONS IN: nucleic acid bi
AT1G32790	"CTC-interacting domain 11 (CID11); FUNCTIONS IN:
AT1G33680	"KH domain-containing protein; FUNCTIONS IN: RNA b
AT1G43190	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT1G47490	"RNA-binding protein 47C (RBP47C); CONTAINS InterP
AT1G47500	"RNA-binding protein 47C' (RBP47C'); CONTAINS Inte
AT1G48410	"ARGONAUTE 1 (AGO1); FUNCTIONS IN: protein binding
AT1G49760	"poly(A) binding protein 8 (PAB8); FUNCTIONS IN: R
AT1G54490	"exoribonuclease 4 (XRN4); CONTAINS InterPro DOMAI
AT1G55500	"evolutionarily conserved C-terminal region 4 (ECT
AT1G59610	"dynamin-like 3 (DL3); FUNCTIONS IN: protein bindi
AT1G69250	"Nuclear transport factor 2 (NTF2) family protein
AT1G70180	"Sterile alpha motif (SAM) domain-containing prote
AT1G75560	"zinc knuckle (CCHC-type) family protein; FUNCTION
AT1G75660	"5'-3' exoribonuclease 3 (XRN3); CONTAINS InterPro
AT1G76010	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G80410	"EMBRYO DEFECTIVE 2753 (EMB2753); FUNCTIONS IN: bi
AT2G06850	"xyloglucan endotransglucosylase/hydrolase 4 (XTH4
AT2G17870	"cold shock domain protein 3 (CSP3); FUNCTIONS IN:
AT2G21060	"glycine-rich protein 2B (GRP2B); FUNCTIONS IN: DN
AT2G23350	"poly(A) binding protein 4 (PAB4); FUNCTIONS IN: R
AT2G26280	"CID7; FUNCTIONS IN: damaged DNA binding, protein
AT2G27100	"SERRATE (SE); FUNCTIONS IN: DNA binding, sequence
AT2G29140	"pumilio 3 (PUM3); FUNCTIONS IN: RNA binding, bind
AT2G29190	"pumilio 2 (PUM2); FUNCTIONS IN: mRNA binding, RNA
AT2G29200	"pumilio 1 (PUM1); FUNCTIONS IN: RNA binding, bind
AT2G32000	"DNA topoisomerase, type IA, core; FUNCTIONS IN: D
AT2G32080	"purin-rich alpha 1 (PUR ALPHA-1); FUNCTIONS IN: n
AT2G33845	"Nucleic acid-binding, OB-fold-like protein; FUNCT
AT2G34160	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G35410	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT2G35920	"RNA helicase family protein; FUNCTIONS IN: helica
AT2G37230	"Tetratricopeptide repeat (TPR)-like superfamily p
AT2G38610	"RNA-binding KH domain-containing protein; FUNCTIO
AT2G42520	"P-loop containing nucleoside triphosphate hydrola
AT2G43970	"RNA-binding protein; FUNCTIONS IN: RNA binding, n
AT2G45810	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT2G46780	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G01540	"DEAD box RNA helicase 1 (DRH1); FUNCTIONS IN: ATP
AT3G02830	"zinc finger protein 1 (ZFN1); CONTAINS InterPro D
AT3G03950	"evolutionarily conserved C-terminal region 1 (ECT
AT3G04610	"flowering locus KH domain (FLK); FUNCTIONS IN: RN
AT3G06480	"DEAD box RNA helicase family protein; FUNCTIONS I
AT3G06980	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G07030	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G07810	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G13060	"evolutionarily conserved C-terminal region 5 (ECT
AT3G13224	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G13300	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT3G13460	"evolutionarily conserved C-terminal region 2 (ECT
AT3G13990	"Kinase-related protein of unknown function (DUF12
AT3G14100	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G15000	"cobalt ion binding; FUNCTIONS IN: cobalt ion bind
AT3G15010	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G15590	"Tetratricopeptide repeat (TPR)-like superfamily p
AT3G19130	"RNA-binding protein 47B (RBP47B); FUNCTIONS IN: R
AT3G21100	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G22330	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT3G23700	"Nucleic acid-binding proteins superfamily; FUNCTI
AT3G25150	"Nuclear transport factor 2 (NTF2) family protein
AT3G25920	"ribosomal protein L15 (RPL15); FUNCTIONS IN: stru
AT3G26420	"ATRZ-1A; CONTAINS InterPro DOMAIN/s: RNA recognit
AT3G44110	"DNAJ homologue 3 (ATJ3); FUNCTIONS IN: unfolded p
AT3G49390	"CTC-interacting domain 10 (CID10); FUNCTIONS IN:
AT3G50370	"unknown protein; FUNCTIONS IN: molecular_function
AT3G50670	"U1 small nuclear ribonucleoprotein-70K (U1-70K);
AT3G51950	"Zinc finger (CCCH-type) family protein / RNA reco
AT3G52150	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G53420	"plasma membrane intrinsic protein 2A (PIP2A); FUN
AT3G54540	"general control non-repressible 4 (GCN4); CONTAIN
AT3G57150	"homologue of NAP57 (NAP57); FUNCTIONS IN: pseudou
AT3G58510	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G58570	"P-loop containing nucleoside triphosphate hydrola
AT3G60240	"eukaryotic translation initiation factor 4G (EIF4
AT3G61240	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G61690	"nucleotidyltransferases; FUNCTIONS IN: nucleotidy
AT3G62250	"ubiquitin 5 (UBQ5); FUNCTIONS IN: structural cons
AT4G00660	"RNAhelicase-like 8 (RH8); FUNCTIONS IN: helicase
AT4G03110	"RNA-binding protein-defense related 1 (RBP-DR1);
AT4G09040	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G10610	"CTC-interacting domain 12 (CID12); CONTAINS Inter
AT4G21710	"NRPB2; CONTAINS InterPro DOMAIN/s: DNA-directed R
AT4G25550	"Cleavage/polyadenylation specificity factor, 25kD
AT4G26650	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G28440	"Nucleic acid-binding, OB-fold-like protein; CONTA
AT4G28880	"casein kinase I-like 3 (ckl3); FUNCTIONS IN: prot
AT4G30190	"H(+)-ATPase 2 (HA2); FUNCTIONS IN: ATPase activit
AT4G31580	"serine/arginine-rich 22 (SRZ-22); FUNCTIONS IN: p
AT4G34110	"poly(A) binding protein 2 (PAB2); FUNCTIONS IN: R
AT4G35890	"winged-helix DNA-binding transcription factor fam
AT4G36020	"cold shock domain protein 1 (CSDP1); CONTAINS Int
AT4G36960	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G38680	"glycine rich protein 2 (GRP2); FUNCTIONS IN: doub
AT4G39260	"cold, circadian rhythm, and RNA binding 1 (CCR1);
AT5G02530	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G03280	"ETHYLENE INSENSITIVE 2 (EIN2); FUNCTIONS IN: tran
AT5G03740	"histone deacetylase 2C (HD2C); FUNCTIONS IN: hist
AT5G04280	"RNA-binding (RRM/RBD/RNP motifs) family protein w
AT5G04430	"BINDING TO TOMV RNA 1S (SHORT FORM) (BTR1S); FUNC
AT5G11350	"DNAse I-like superfamily protein; CONTAINS InterP
AT5G12440	"CCCH-type zinc fingerfamily protein with RNA-bind
AT5G15270	"RNA-binding KH domain-containing protein; FUNCTIO
AT5G18550	"Zinc finger C-x8-C-x5-C-x3-H type family protein;
AT5G18590	"Galactose oxidase/kelch repeat superfamily protei
AT5G20490	"XIK; FUNCTIONS IN: motor activity; INVOLVED IN: i
AT5G20950	"Glycosyl hydrolase family protein; FUNCTIONS IN:
AT5G21160	"LA RNA-binding protein; FUNCTIONS IN: molecular_f
AT5G26742	"embryo defective 1138 (emb1138); FUNCTIONS IN: in
AT5G37720	"ALWAYS EARLY 4 (ALY4); FUNCTIONS IN: nucleotide b
AT5G38600	"Proline-rich spliceosome-associated (PSP) family
AT5G40490	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G43960	"Nuclear transport factor 2 (NTF2) family protein
AT5G47010	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in
AT5G48650	"Nuclear transport factor 2 (NTF2) family protein
AT5G49360	"beta-xylosidase 1 (BXL1); FUNCTIONS IN: alpha-N-a
AT5G51120	"polyadenylate-binding protein 1 (PABN1); CONTAINS
AT5G53620	"unknown protein; INVOLVED IN: biological_process
AT5G54900	"RNA-binding protein 45A (RBP45A); FUNCTIONS IN: R
AT5G55550	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G55670	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G58470	"TBP-associated factor 15B (TAF15b); FUNCTIONS IN:
AT5G59950	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G61020	"evolutionarily conserved C-terminal region 3 (ECT
AT5G63120	"P-loop containing nucleoside triphosphate hydrola
AT5G63260	"Zinc finger C-x8-C-x5-C-x3-H type family protein;
AT5G64270	"splicing factor, putative; FUNCTIONS IN: binding;
AT5G65260	"RNA-binding (RRM/RBD/RNP motifs) family protein;

CHAPTER 2: Study of a new Processing-body component: the UCN endonuclease

Study of a new Processing-body component: the UCN endonuclease

I. Abstract

The analysis of the UPF1 proteome identified six new Processing-body components, among which four DEAD-box RNA helicases, three homologues of DDX6 (RH6, RH8, RH12) and RH11, a TNTase and a putative endonuclease. P-bodies are known to accumulate mRNAs and RNA decay factors, the study of these proteins thus represent a good opportunity to gain novel insights into RNA decay pathways in plants. This chapter will focus on the study of the putative endonuclease. As it was identified in our UPF1-enriched fractions, we named it UCN for UPF1-Copurified Nuclease. In this chapter, I will present the characterization of UCN, which contains a predicted PIN-like nuclease domain of the NYN family (Nedd4-BP1, YacP Nucleases) and two predicted RNA binding domains of the OST-HTH family (Oskar-TDRD5/TDRD7-HTH). UCN-associated factors were identified by mass spectrometry: we found a very robust and reproducible co-purification of UCN with the decapping factors and P-body components DCP1 and VCS, in addition to the previously documented association of UCN with UPF1. Yeast two-hybrid experiments demonstrated that UCN and DCP1 do not only co-purify but in fact directly interact, whereas UCN and UPF1 interaction is likely mediated by RNA. UCN P-body localization was also validated upon its co-localization with the P-body markers DCP1, DCP5 and RH8. We identified insertion mutants in UCN and analyzed endogenous NMD targets accumulation in these mutant backgrounds. We engineered *Arabidopsis* lines expressing wild type and NYN-domain catalytic mutant GFP fusion forms in the corresponding mutant backgrounds. Using a transient expression system, we monitored the impact of UCN expression on the accumulation of NMD targets. The results suggest that UCN expression preferentially perturb co-expressed transcripts with an extended 3'-UTR, and can modulate the accumulation of the NMD sensitive viruses PVX and TCV. UCN expression in transient assay or UCN ectopic expression in *Arabidopsis* lines show ethylene-related phenotypes, such as necrosis, dark induced senescence and ethylene insensitivity. This study identifies UCN as a potential new RNA decay factor, which could influence important biological processes including hormone signaling and defense responses in plants.

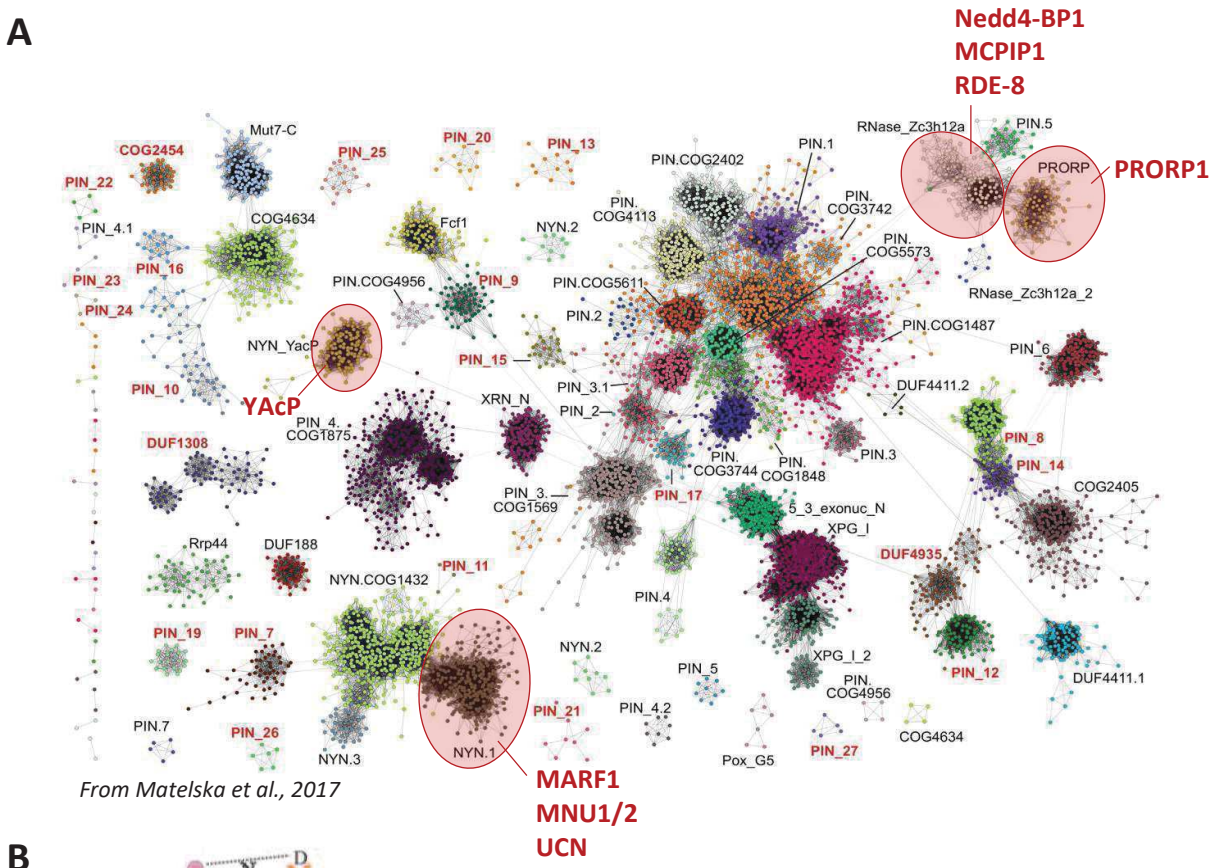


Figure 18. PIN-like domain endonucleases. (A) Sequence similarity network of the PIN-like domains. Represents the diversity of PIN-like proteins and their sub-families proposed by Matelska *et al.*, 2017. PIN-like domain proteins mentioned in this study are highlighted in red. (B) Topology diagram representative of the PIN/NYN domain as in Anantharaman *et al.,* 2006. The four conserved aspartic acid residues (D) are indicated. The pink circle represents Mg^{2+} ion. (C) Domain composition of the different NYN-domain discussed in this study.

II. Introduction

1. NYN-domain endonucleases

The NYN-domain (Nedd4-BP1, YacP nucleases) is a metallonuclease domain typified by the Nedd4-binding protein 1 and the bacterial YacP-like proteins. NYN-domain-containing proteins belong to one of the largest and most diverse nuclease superfamilies, the PIN (PilT N-terminal) domain-like superfamily. A comprehensive classification of the PIN domain-like superfamily, including NYN-domains was recently established to help the prediction of biological functions of family members of unknown functions (Figure 18A) (Matelska et al., 2017). NYN domain proteins are present in all three kingdoms of life: Eukarya, Bacteria and Archaea (Anantharaman and Aravind, 2006). In total, 5402 sequences of 2861 species are documented on the NYN-domain family in Pfam database (PF01936) (Figure S6). A vast majority of sequences are found in the bacteria kingdom, and NYN-domain endonucleases are prominently found in plant species but are also present in metazoan or fungi (Figure S6). NYN-domains underwent functional diversification and fusion to additional RNA binding domains in eukaryotes and can be traced back to LUCA (Last Universal Common Ancestor), implying several losses in eukaryotes and prokaryotes (Habacher and Ciosk, 2017; Anantharaman and Aravind, 2006). The NYN-domain, referred to as PF01936 on Pfam database comprises the NYN.1, NYN.2 and NYN.3 domains detailed in Matelska et al., 2017.

NYN-domains are closely related to two superfamilies: the PIN (PilT N-terminal) and the FLAP 5'-3' nuclease domain families. They share a common protein fold with the PIN and FLAP domains composed of five strands and four helices (Figure 18B) (Anantharaman and Aravind, 2006). NYN domains share four key aspartic acid residues (D) with PIN domains that are necessary for their nuclease activity and Mg^{2+} metal ion binding (Figure 18B (Anantharaman and Aravind, 2006). It was first thought that NYN domains were able to bind only a single metal ion, as PIN domains do, but it was later shown that the NYN metallonuclease domain can actually bind two metal ions at the active site (Howard et al., 2012).

NYN-domains are usually associated with RNA-binding domains (Figure 18C), such as PPR (Pentatricopeptide Repeat) domains, CCCH zinc finger domains or OST-HTH domains (Oskar-TDRD5/TDRD7 HTH). This last domain (also called LOTUS for Limkain, Oskar, and Tudor containing proteins 5 and 7) is found in the germline-specific proteins TDRD5/7 that have a role in piwi-mediated transposon silencing in vertebrates, but also in the Oskar protein which has a role in development in *Drosophila* (Anantharaman et al., 2010). OST-HTH domains are found in bacteria and eukaryotes, and are predicted to have an helix-turn-helix conformation allowing their binding to dsRNA or folded RNA structures

(Anantharaman et al., 2010). Recently, it was shown that some LOTUS domains could be involved in protein-protein interaction (Jeske et al., 2017). It is the case of the LOTUS domain of Oskar in *Drosophila*, which was shown to bind and stimulate the DEAD-box helicase activity of the Vasa protein, allowing its proper localization to pole plasm (Jeske et al., 2017). This domain was found in proteins known to nucleate or organize in structurally related ribonucleoprotein RNP complexes (arthropod polar granules and vertebrate nuage complexes) and was proposed to have a key role in the assembly and localization of these structures (Anantharaman et al., 2010).

2. NYN-domain endonucleases are involved in diverse cellular processes

This section compiles the information gathered on a few NYN-domain proteins functionally described in different model organisms. No common role has been described for NYN-domain endonucleases, it appears that NYN-domain endonucleases are involved in diverse cellular processes in eukaryotes such as RNA silencing, RNA maturation or the regulation of the stability of a subset of mRNA.

2.1. Regnase-1/MCPIP1

Regnase-1 is a NYN-domain endonuclease involved in the decay of a subset of transcripts in mouse including inflammation-related mRNAs thereby preventing immune disorders in mammals (Figure 18C) (Matsushita et al., 2009; Habacher and Ciosk, 2017). The recognition of targets by Regnase-1 or its human homolog MCPIP1 is mediated by a stem loop structure (Suzuki et al., 2011; Mino et al., 2015). Importantly, Regnase-Mediated Decay (RMD) was shown to be a translation- and UPF1-dependent process; a direct interaction between UPF1 and Regnase-1 has also been described (Mino et al., 2015). MCPIP1 is a suppressor of miRNA activity and biogenesis and acts as an antagonist of Dicer upon the MCPIP1-mediated cleavage of pre-miRNA, but also targets viral RNA and possesses broad-spectrum antiviral properties (Suzuki et al., 2011; Lin et al., 2017). Regnase-1/MCPIP1 subcellular localization is still unclear as some reports described MCPIP1 as a P-body component, while other studies showed that the Regnase-1 protein is devoid from P-bodies and Stress Granules and rather co-localizes with ribosomes (Suzuki et al., 2011; Mino et al., 2015). Regnase-1/MCPIP1 carries a CCCH-type Zinc Finger motif, which has been suggested to increase RNA affinity in the vicinity of the catalytic domain (Yokogawa et al., 2016).

2.2. RDE-8

The NYN domain protein RNAi Deficiency 8 (RDE-8) was identified in a genetic screen for worms with an RNAi deficiency (Rde) phenotype (Figure 18C) (Tsai et al., 2015). The ectopic expression of RDE-8 was able to rescue the Rde phenotype while catalytic mutants did not, demonstrating the importance of RDE-8

catalytic activity. In this study, Tsai and colleagues showed that the Argonaute RDE-1 protein recruits RDE-8 and allows the cleavage and uridylation of RNAi targets. RDE-8 also promotes RdRP activity and thus ensures RNAi signal amplification independently on Argonaute intrinsic cleavage activity (Tsai et al., 2015).

2.3. The PRORP family

The Proteinaceous RNase P (PRORP) proteins correspond to a family of NYN-domain endonucleases responsible for the essential 5‘ maturation of tRNA precursors (Holzmann et al., 2008; Gobert et al., 2010). PRORP proteins are found exclusively in eukaryotes (Lechner et al., 2017). Their function is restricted to mitochondria in metazoan, but they represent the only enzyme responsible for RNase P activity in plants (Holzmann et al., 2008; Gutmann et al., 2012). These proteins are composed of a catalytic NYN domain at their C-terminal region and of an RNA binding domain composed of Pentatricopeptide Repeats (PPR) at their N-terminal extremity (Figure 18C). The PPR domain is involved in PRORP substrate binding specificity (Howard et al., 2012).

2.4. MNU1 and MNU2

Recently, two novel NYN-endonucleases have been described in *Arabidopsis* mitochondria: the mitochondrial nucleases MNU1 and MNU2. MNU1 and MNU2 share common domains with UCN endonuclease: both NYN and OST-HTH domains are found in MNU1 and MNU2 (Figure 18C) (Stoll and Binder, 2016). MNU1 and MNU2 have been proposed to function in the 5’ processing of plant mitochondrial transcripts.

2.5. MARF1/Limkain B

Meiosis regulator and mRNA stability factor 1 (MARF1), also called Limkain B (LMKB) is a NYN-domain endonuclease involved in the control of meiosis and retrotransposon surveillance in oocytes in human (Su et al., 2012). MARF1 has also been shown to regulate the expression of a subset of specific transcripts including IFI44L, a component of the type I interferon and to be important for neuron differentiation (Bloch et al., 2014; Kanemitsu et al., 2017). MARF1 was originally localized to peroxisomes upon its partial colocalization with peroxisome markers, ABCD3 and PXF (Dunster et al., 2005). Nevertheless, later work reported that MARF1 localizes to P-bodies (Bloch et al., 2014). The identification of the direct interaction between MARF1 and the decapping activator Ge-1, a well-known P-body factor, is in favor of this second described localization of MARF1 in P-bodies in human (Bloch et al., 2014). MARF1 contains a NYN-domain at its N-terminal extremity, two RRM motifs in its center part and eight OST-HTH domains at its C-terminal extremity, a domain organization, which is related to plant UCN endonuclease (Figure 18C).

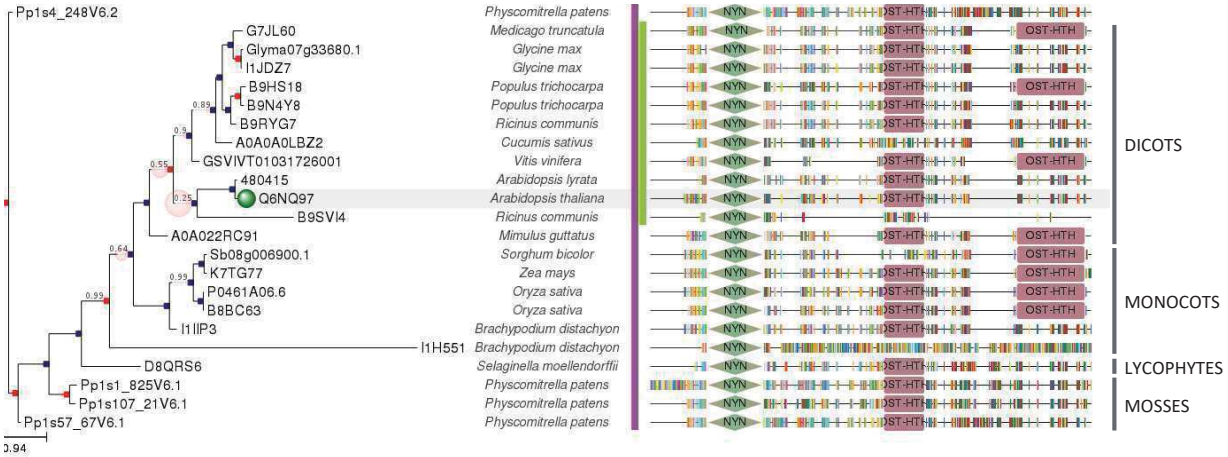


Figure 19. Evolutionary conservation of UCN endonuclease. Phylogenetic tree corresponding to the conservation of UCN endonuclease from mosses to dicots. Domain composition are represented in the right panel. The target sequence, UCN endonuclease (Q6NQ97) is indicated with a green circle. This phylogenetic tree was constructed with PhylomeDB.

Overall, NYN-domains have been identified in several organisms including animals, bacteria and plants, and characterized NYN-domains endonucleases appear to have a role in mRNA stability or RNA processing. NYN-domain proteins are particularly diverse in plants, but the function of many of these proteins remains to be discovered. This is the case of the UCN endonuclease, which was identified in upon UPF1 immunoprecipitation. In the following section, I will first describe the UCN interactome study performed by IPs coupled to mass spectrometry and yeast two-hybrid approaches, then, I will precisely detail the subcellular localization of UCN using confocal microscopy. Finally, I will describe the preliminary functional characterization of the UCN endonuclease with the analysis of *ucn* mutant lines, lines ectopically expressing UCN and transient expressions of UCN. These results provide novel insights in the possible functions and cellular roles of this hitherto unknown NYN-domain endonuclease.

III.Results

1. *In silico* analysis of the UCN endonuclease

As the UCN endonuclease (AT2G15560) was completely uncharacterized at the beginning of this study, we first performed an *in silico* analysis of UCN. The localization of UCN was predicted to be cytosolic in SUBA (http://suba.live/), and nuclear and peroxisomal according to GOterms. However, caution has to be taken, since these localizations were not curated; UCN localization remains to be experimentally confirmed. Using TileViz (http://jsp.weigelworld.org/tileviz/tileviz.jsp) and AtGenExpress Visualization Tool (http://jsp.weigelworld.org/expviz/expviz.jsp), tilling array experiments indicate a mild expression of UCN in every part of the plant, with an increased expression in flowers. No significant change in expression is observed upon abiotic stress exposure, pathogens infection (only fungi or bacteria were tested) or different light conditions. Interestingly, using these resources and bibliographic search, we identified that UCN expression is increased in the gibberellin mutant *ga1-5* as well as in plants overexpressing EIN3, the master regulator of the ethylene signaling pathway (Peng et al., 2014). These variations in UCN expression suggest a potential link between UCN expression and hormone biosynthesis and action. Using the PhylomeDB tool, we identified that UCN is conserved in plants from mosses to dicots (Figure 19).

A Protein Blast against the human genome identified MARF1/LimkainB as the closest homolog of UCN with 42% similarity and 27% identity. The NYN-domain of UCN can then be categorized as a NYN.1 domain, as for MARF1 and MNU1/2 (Matelska et al., 2017).

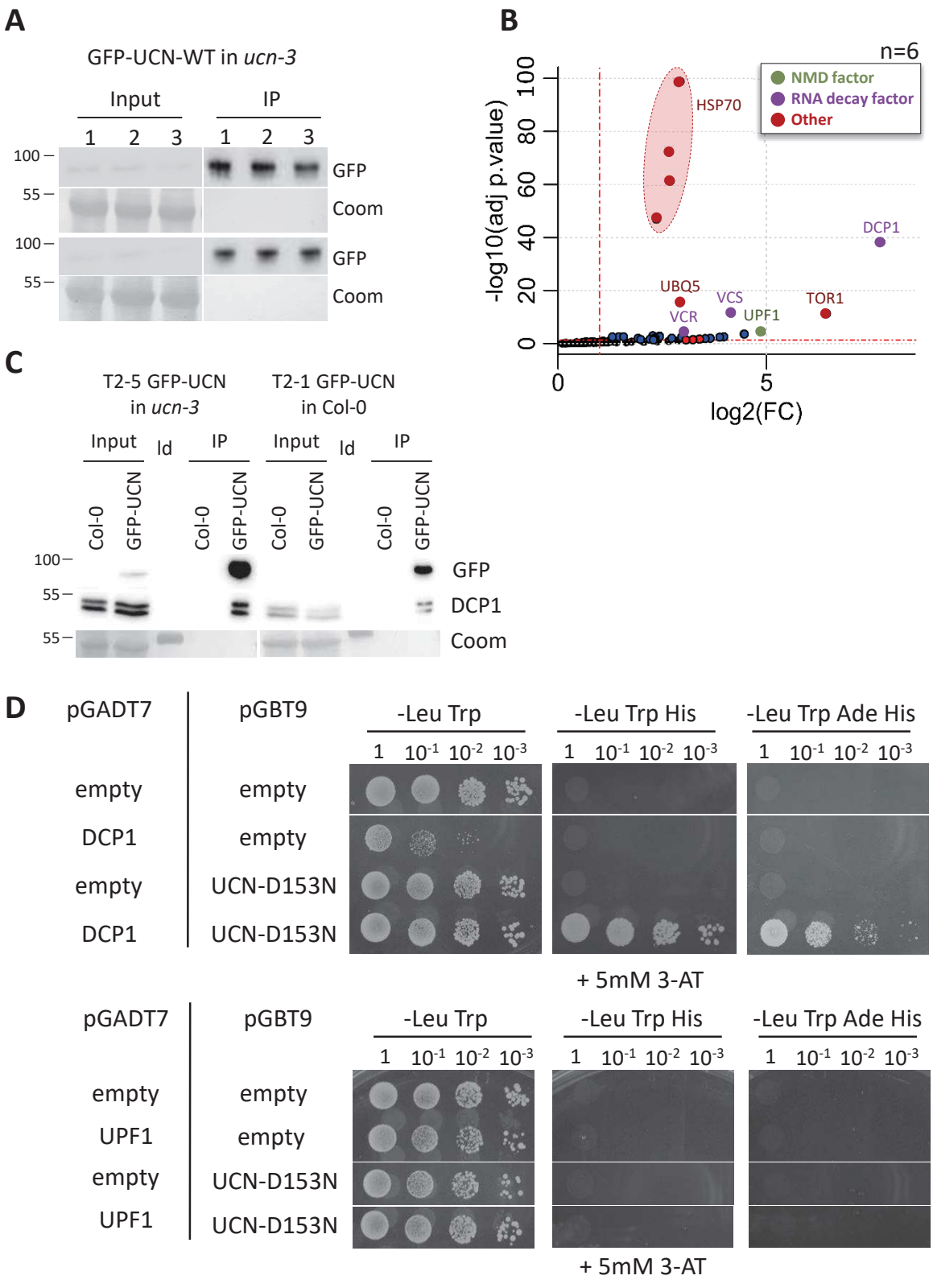


Figure 20. Analysis of UCN interactome by mass spectrometry and yeast two-hybrid. (A) Western blot validation of GFP-UCN-WT IPs. The six replicates used for mass spectrometry analysis are shown. (B) Volcano plot representing the proteins enriched in UCN IPs. Adjusted p-value is presented on the y axis, Fold Change (FC) in x axis. The red dotted lines represent the significativity thresholds of *p<0.05* and log2(FC)<1. (C) Western blot analysis showing the co-IPs of GFP-UCN with DCP1. IP: immunoprecipitation. (D) Yeast two-hybrid experiments showing specific growth on selective media for the UCN-D153N/DCP1 combination highlights the direct interaction between UCN-D153N and DCP1 (higher panel). No growth is observed on selective media for the UCN-D153N/UPF1 combination (lower panel).

2. Analysis of UCN interactome

2.1. Immunoprecipitation of the UCN endonuclease

In order to better understand the role of the UCN endonuclease, we looked for UCN-associated factors using an approach coupling IPs with mass spectrometry partners detection. For this purpose, *Arabidopsis* lines expressing a tagged-version of UCN were produced (GFP-UCN-WT, see part 4.2.2. for more details). An IP of UCN was performed using antibodies directed against the GFP tag and the enrichment of GFP-UCN in the IP fraction was validated by western blot analysis (Figure 20A). UCN partners enriched in IP fraction were analyzed by mass spectrometry. A total of 33 proteins were identified as specifically enriched in the UCN-purified fraction (Figure 20B, Table S8). This analysis first validated the association between UPF1 and UCN as observed in the reverse experiment in Chapter I, Figure 11B. This study of associated partners also revealed a strong association between UCN and the decapping activators DCP1 and VARICOSE (VCS), suggesting a possible role of UCN relative to decapping. Finally, several HEAT SHOCK PROTEIN 70 (HSP70) isoforms as well as several TUBULIN proteins (TUB) or factors associated to TUB such as TORTIFOLIA 1 were retrieved in UCN immunoprecipitations (Buschmann et al., 2004). The list of UCN partners was compared with UPF1 partners, DCP5 partners and the At-RBP list described in Reichel, 2016 (Figure S7). We observe that approximatively half of UCN partners also co-purifiy with UPF1 and DCP5.

The possible need of HSP70 for UCN protein folding, the relationship between UCN and the cytoskeleton, and the link between UCN and the decapping factors are very interesting tracks to follow to better understand UCN action.

2.2. Validations of UCN immunoprecipitation

In order to validate some candidate interactions highlighted by UCN IPs, we first performed a co-immunoprecipitation analysis coupled to western blotting using antibodies directed against DCP1 and UPF1. In a second time, we tested the direct interaction between UCN and these two candidates by yeast two-hybrid experiments.

2.2.1. Co-immunoprecipitation of UCN with DCP1 and UPF1

Co-IP followed by western blot analysis were performed using GFP-UCN-WT expressing lines in order to validate some UCN partners for which antibodies were available, *i.e.* UPF1 and DCP1. The immunoprecipitation of DCP1 with GFP-UCN-WT was validated by western blot using antibodies directed against DCP1 (Figure 20C). The presence of UPF1 protein was also tested but no signal was retrieved, indicating a weak association between the two proteins. This is consistent with the identified RNA-

dependent association between UCN and UPF1 described in Chapter I, Figure 12A. This analysis confirmed the association between UCN and the key decapping activator DCP1.

2.2.2. Analysis of UCN direct interaction by yeast two-hybrid

The prominent accumulation of decapping activators in UCN-enriched fractions suggests a functional link between UCN function and the decapping complex. While the association between UCN and UPF1 is likely mediated by RNA (see Chapter I), the strong enrichment of decapping activators in UCN IPs prompted us to look for possible direct interaction. To reach this objective, we used the yeast two-hybrid technique, which would allow to discriminate if the association between UCN, DCP1 and UPF1 are direct or indirect.

In a first set of experiments, the results were obtained using the catalytic mutant form of UCN (UCN-D153N; see part 3.3.1. for more details), because the transformation of UCN wild-type form in yeast was not successful. Yeast strain co-transformed with a vector containing DCP1 sequence fused to GAL4 activation domain (pGADT7-DCP1) and another vector containing UCN-D153N sequence fused to GAL4 binding domain (pGBT9-UCN-D153N) showed growth on selective medium (Figure 20D). These results indicate that the interaction between DCP1 and UCN-D153N is direct. During the time of my thesis, I was able to successfully transform the UCN-D153N version in yeast but not the WT version. A single mutation in UCN catalytic domain is not likely sufficient to promote this interaction, it suggests that UCN-WT probably also interacts with DCP1. Novel constructs of UCN-WT are currently being tested to confirm the results obtained with the UCN-D153N mutant version.

In addition, the direct interaction between UCN and UPF1 has been tested by yeast two-hybrid. As expected, no growth was observed for the co-transformation of pGADT7-UPF1 and pGBT9-UCN-D153N, suggesting that the association between UCN and UPF1 is not a direct (Figure 20D). This result is consistent with the fact that the UCN endonuclease was found as a RNA-dependent partner of UPF1 as presented in Chapter 1, Figure 12A. This result differs from the direct interaction observed between UPF1 and another NYN-domain endonuclease, Regnase-1 in human cells (Mino et al., 2015).

3. Subcellular localization of UCN endonuclease

UCN interacts with decapping factors DCP1 and VCS, which are known to localize in P-bodies. In order to validate that UCN is also part of P-bodies, a localization study of UCN was performed, both in transient experiments and using *Arabidopsis* lines expressing UCN fusion.

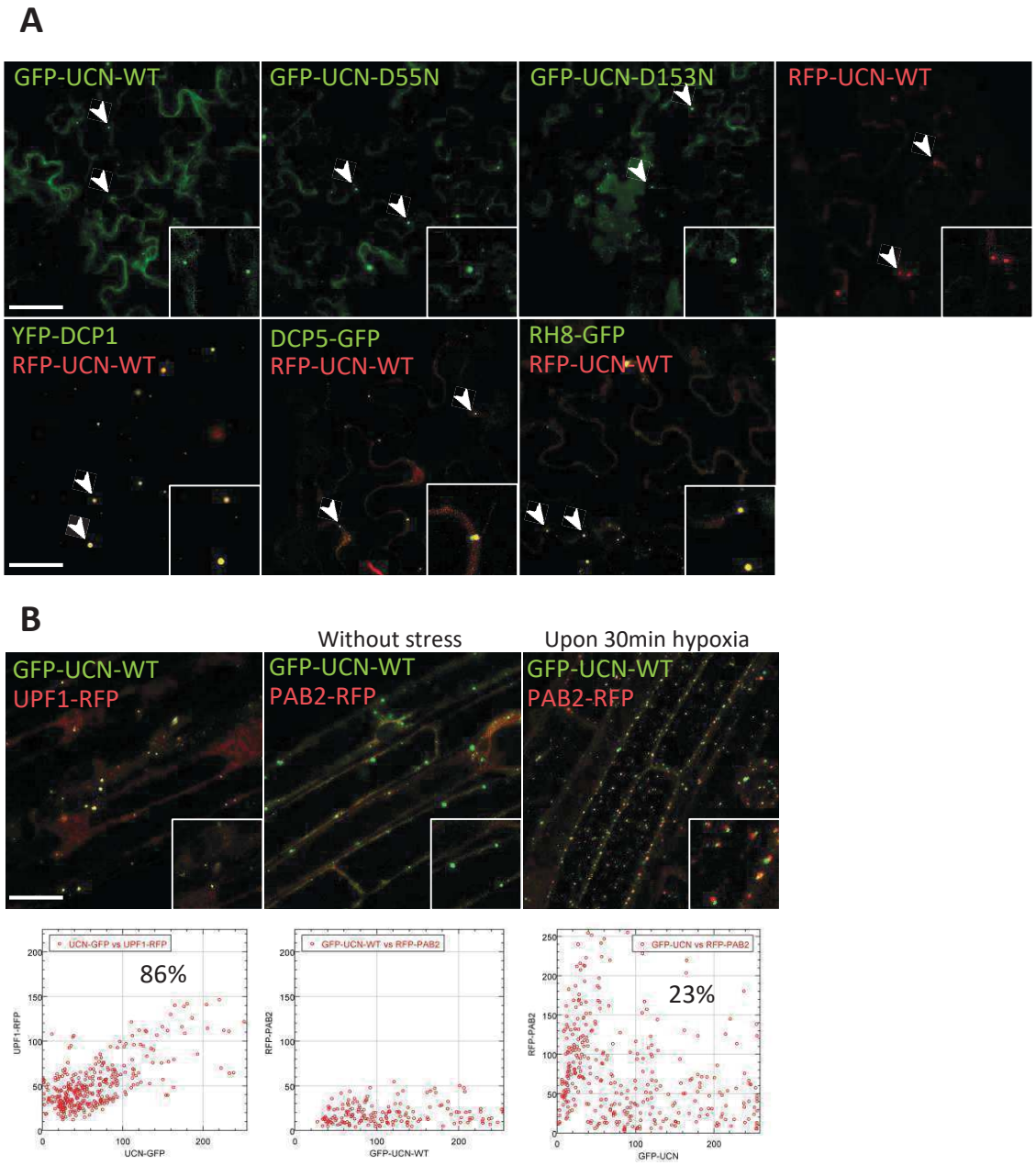


Figure 21. Subcellular localization of UCN. (A) Confocal microscopy images of GFP-UCN and RFP-UCN fusions alone or co-expressed with P-body markers YFP-DCP1, DCP5-GFP and RH8-GFP in *Nicotiana benthamniana* leaves. Localization were observed 4 days after infiltration. (B) Confocal microscopy images of epidermal root cells of stable *Arabidopsis* lines expressing both GFP-UCN-WT and the P-body component UFP1-RFP or the Stress Granules marker RFP-PAB2. Dot plot representing the quantification of the signal in both GFP and RFP channels in x or y axis respectively are represented under each confocal image. Each plot was acquired following the analysis of the foci composition of 10 randomly chosen confocal sections. The colocalization percentage between P-body markers and UPF1-RFP is indicated for each analysis. Scale bars represent 10 µm.

3.1. Subcellular localization of UCN in transient assay

In order to analyze the subcellular localization of UCN, GFP and RFP fusion were created for the WT as well as the D55N and D153N mutant versions (see part 3.3.1. for more details on mutant design). These constructs were used in agro-infiltration experiments in *N. benthamiana* and UCN subcellular localization was imaged by confocal microscopy. We observed that RFP-UCN-WT and GFP-UCN-WT are localized in the cytosol and in small cytoplasmic foci (Figure 21A). A similar localization was observed for GFP-UCN-D55N and GFP-UCN-D153N mutant versions (Figure 21A). These results indicate that the substitution of an aspartic acid into an asparagine in the catalytic site of UCN does not perturb UCN subcellular localization.

In order to definitively confirm UCN P-body localization, co-localizations were done by co-agro-infiltration of both RFP-UCN-WT and the P-body markers DCP5-GFP, YFP-DCP1 and the DDX6 homologue RH8-GFP. We observed a co-localization for each combination and validated that the UCN endonuclease is identified as a new component of P-bodies (Figure 21A-21B).

3.2. Subcellular localization of UCN in stable transgenic lines

Transient expression of *Arabidopsis* UCN protein in the *N. benthamiana* system is a rather artificial setting. For this reason, the localization study of the GFP-UCN fusion was also performed in stable transgenic plants, using GFP-UCN lines (see part 3.3.2. for more details). The subcellular localization of GFP-UCN-WT was observed by confocal microscopy in epidermal root cells using the GFP-UCN-WT lines used in the IP experiments. We observed that as reported for the transient assay analysis, GFP-UCN-WT is localized in small cytoplasmic foci. The co-localization between GFP-UCN-WT and P-bodies was assessed by the analysis of the cross between GFP-UCN-WT lines with UPF1-RFP lines. We observed that GFP-UCN-WT co-localizes with UPF1-RFP foci up to 86%, confirming in stable lines that GFP-UCN-WT is indeed localized in P-bodies (Figure 21B). In addition, time-lapse acquisitions were performed to show that GFP-UCN and UPF1-RFP move together as part of the same cytoplasmic entities (data not shown).

F1 generation resulting from the cross between a stable GFP-UCN-WT line and a plant expressing RFP-PAB2 was also observed by confocal microscopy. PAB2 is a marker for stress granules that are small cytoplasmic foci appearing following stresses including heat stress, arsenic treatment or hypoxia stress (Parker, 2007). Before stress, we observe that the GFP-UCN fusion localizes to small cytoplasmic foci, previously defined as P-bodies, while RFP-PAB2 is diffusely localized in the cytoplasm (Figure 21B). After 30 min hypoxia stress, we observe the relocalization of RFP-PAB2 into cytoplasmic foci, identifying stress granules. By contrast, GFP-UCN foci do not relocalize into stress granules upon stress, but is localized to

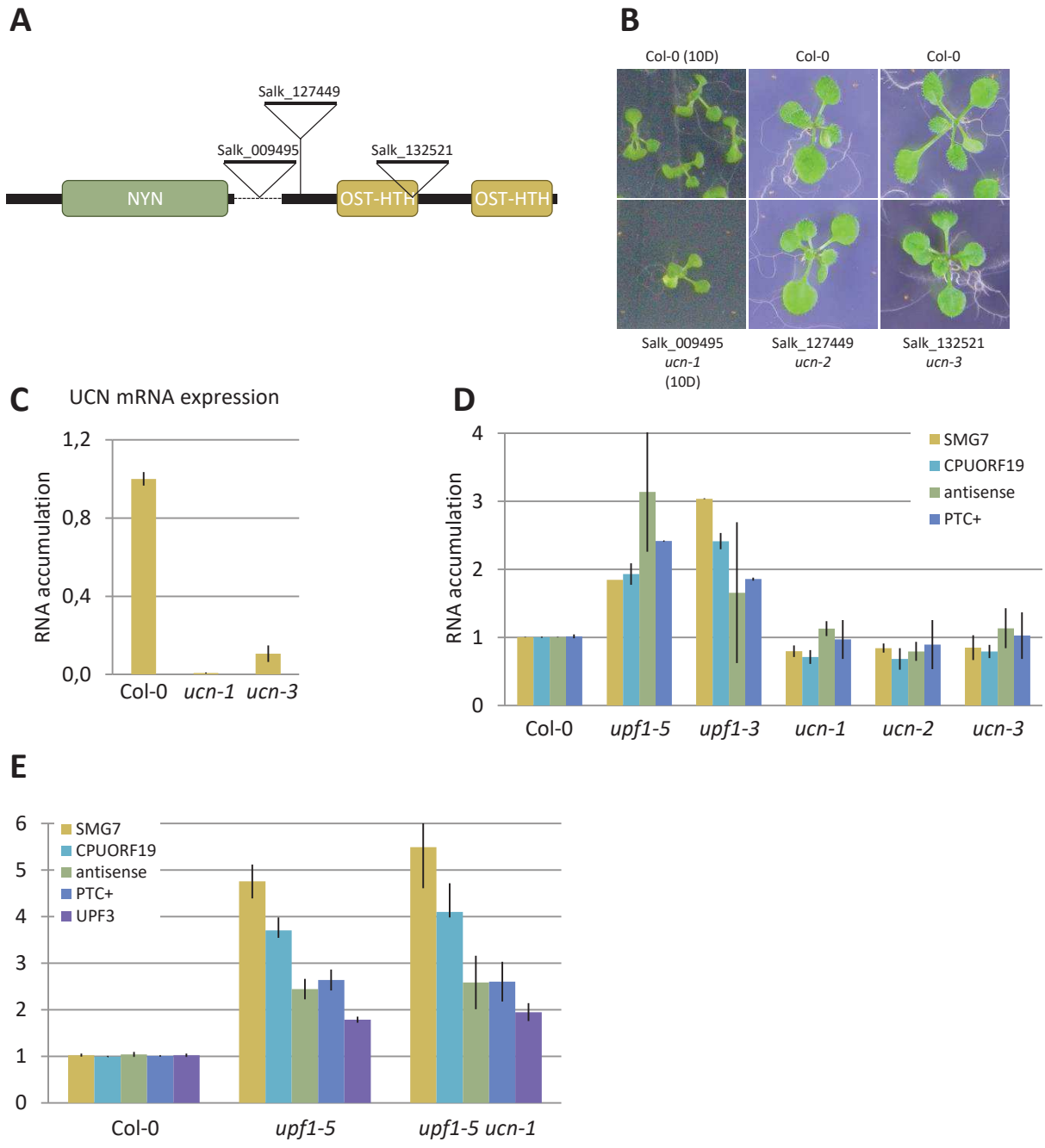


Figure 22. Characterization of mutants in UCN endonuclease. (A) Schematic representation of the localization of T-DNA insertions used in this study respectively to UCN domains. (B) Phenotype of 10 days old (10D) or 15 days old plantlets of the three selected T-DNA insertion lines compared to Col-0. (C) Real time RT-PCR analysis of UCN expression. (D) Real time RT-PCR analysis of known NMD targets transcripts in UCN mutants. (E) Real time RT-PCR analysis of known NMD targets transcripts in double mutant *upf1-5 ucn-1*. All real time RT-PCR analysis normalized to *ACTIN2* and presented as mean ± SEM. n=3

foci that do contact stress granules. The extent of colocalization between GFP-UCN-WT and RFP-PAB2 is about 23**%,** which is significantly different from GFP-UCN-WT co-localization with UPF1-RFP (86%). We can conclude that GFP-UCN-WT and RFP-PAB2 do not co-localize, but rather are present in two different populations of granules that are docked: GFP-UCN-WT localized in P-bodies and RFP-PAB2 in stress granules.

In order to clearly show that UCN endonuclease is a P-body component, we will still have to show the co-localization in stable *Arabidopsis* line between UCN and P-body markers such as DCP1 or DCP5. To this aim, stable *Arabidopsis* lines expressing the RFP-UCN fusion protein are currently produced in the laboratory.

4. Functional characterization of the UCN endonuclease

4.1. Characterization of *ucn* insertion mutants

4.1.1. Analysis of *ucn* insertion mutants

Three different T-DNA insertion mutants were retrieved in the SALK database and obtained to investigate the functional role of the UCN endonuclease (Figure 22A). No obvious developmental defect was observed at seedling stages or later stages for any of the three homozygous mutant lines compared to the Columbia control in standard growth conditions (Figure 22B). UCN expression was measured in two T-DNA mutant lines by qPCR, which validated that UCN expression is abolished in these *ucn* mutants (Figure 22C).

4.1.2. Analysis of NMD targets accumulation in *ucn* mutants

As the UCN endonuclease was identified in UPF1 immunoprecipitated fractions (Chapter I, Figure 11B), we decided to look for the accumulation of four known NMD targets in the three *ucn* alleles. No variations of the NMD targets were observed for these *ucn* mutants compared to the control (Figure 22D). If UCN has only an accessory role in NMD targets regulation, this function could be revealed only in mutant backgrounds in which NMD is already deficient. To test this hypothesis, we monitored NMD targets accumulation in the *upf1-5 ucn-1* double mutant. Similarly to the observation of the simple *ucn* mutants, no changes were observed comparing the single *upf1-5* and the double *upf1-5 ucn-1 mutants* (Figure 22E). These results suggest that the UCN endonuclease is not a limiting factor for NMD or that UCN could be involved in the regulation of a very specific subset of NMD targets, which were not tested here.

Our alternative strategy to discover the role of UCN was to express the protein under a strong a constitutive promoter and monitor the effect of this ectopic expression on plant development and gene

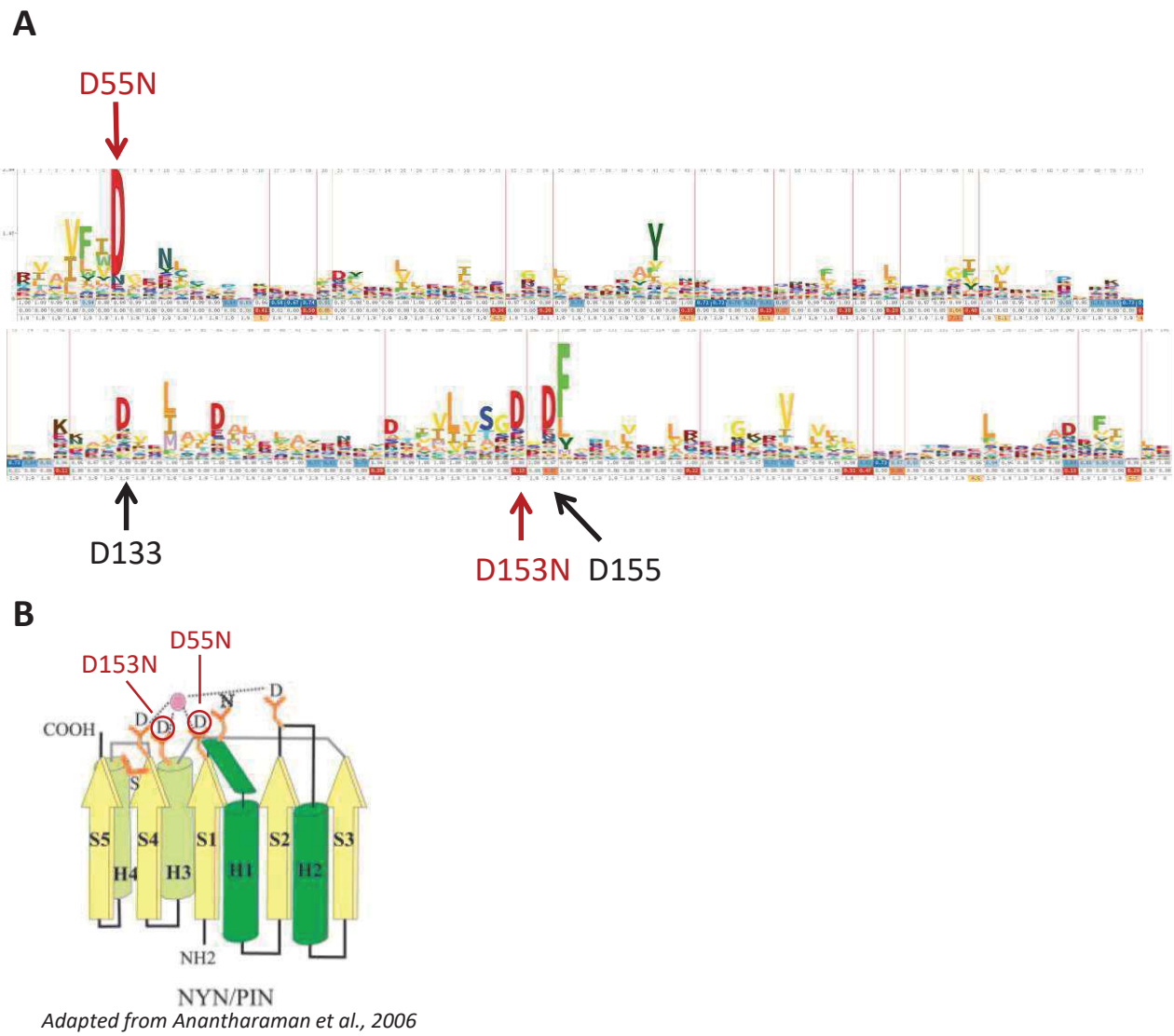


Figure 23. Design of UCN endonuclease catalytic mutants. (A) Representation of the conserved residues between 147 members of the NYN-domain family in the Pfam database (PF01936). The two aspartic acid residues (D) mutated into asparagine (N) to create catalytic mutants in this study are highlighted. (B) Residues mutated are highlighted on the topology diagram representative of the PIN/NYN domain as in Anantharaman *et al.*, 2006.

expression. For this purpose, we produced *Arabidopsis* lines expressing the GFP-UCN protein fusions, where UCN is either WT or mutated in the catalytic domain.

4.2. Production of transgenic lines expressing GFP-UCN fusion protein

4.2.1. Design of UCN endonuclease catalytic mutants

The NYN-domain endonucleases contain four conserved aspartate residues strictly required for their activity and ion-binding (Anantharaman and Aravind, 2006). In order to identify the residues important for UCN catalytic activity, we performed an alignment of the NYN-domain of 147 NYN-domain endonucleases documented on Pfam (PF01936), including UCN (Q6NQ97) (Figure S8). The four conserved aspartic acid residues at positions 55, 133, 153 and 155 were identified as the residues required for UCN NYN-domain activity (Figure 23A-23B). The work on MCPIP1, by Suzuki et al., in 2011 showed that a single mutation changing the third conserved aspartic acid into an asparagine residue of MCPIP1 NYN-domain produced a protein, which failed to cleave its pre-miRNA target. Similarly, Tsai et al., reported in 2016 that the mutation of the first conserved aspartic acid residue into an asparagine failed to complement the *rde-8* deficiency. Using this information, two mutant versions of UCN endonuclease were produced: aspartic acid residues at positions 55 and 153 were independently mutated into asparagine using the overlap extension PCR strategy (Figure 23A-23B). UCN catalytic mutants will be later named UCN-D55N and UCN-D153N respectively.

4.2.2. Characterization of stable *Arabidopsis* lines expressing GFP-UCN fusion

4.2.2.1. Phenotype analysis of GFP-UCN lines

Arabidopsis ucn-1, *ucn-2* and *ucn-3* mutants were transformed with *Agrobacterium tumefaciens* cultures carrying constructs of interest using the floral dip technique in order to produce plants expressing either the WT version or the catalytic mutant UCN-D153N fused to GFP at its N-terminal extremity.

The expression of the GFP-UCN fusions was assessed in the T2 generation by western blot using antibodies directed against GFP (Figure 24A,24D). Lines showing detectable UCN expression were selected for future analysis. Nevertheless, in the absence of antibodies directed against the endogenous UCN protein we could not assess whether the GFP-UCN fusion was expressed at physiological levels.

We noticed developmental defects appearing in several GFP-UCN lines expressing highest levels of the GFP-UCN fusion protein. These GFP-UCN lines showed growth defects and a curly leaf/cotyledon phenotype for both the GFP-UCN-WT and GFP-UCN-D153N versions compared to control plant (Figure

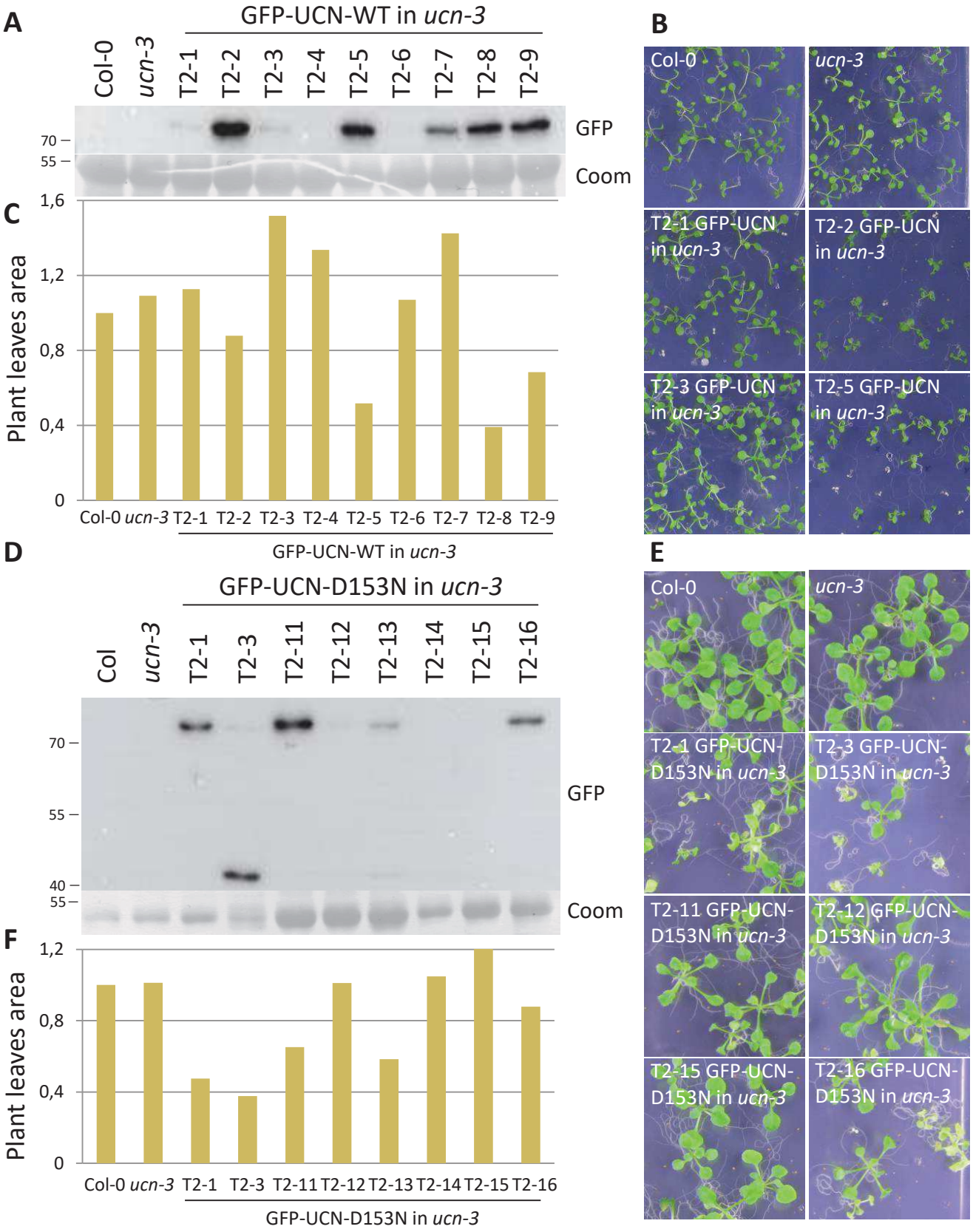


Figure 24. Characterization of *Arabidopsis* lines expressing GFP-UCN fusion. (A) Western blot analysis of the accumulation of GFP-UCN-WT in *Arabidopsis* lines expressing GFP-UCN-WT in *ucn-3* background. (B) Images of GFP-UCN-WT phenotype. 15 days old plantlets are shown on MS agar plates. (C) Quantification of plant size using plants shown in (B) using the Excessive Green plugin on ImageJ. (D) Western blot analysis of the accumulation of GFP-UCN-D153N in *Arabidopsis* lines expressing GFP-UCN-D153N in *ucn-3* background. (E) Images of UCN-GFP-D153N phenotype. 15 days old plantlets are shown on MS agar plates. (F) Quantification of plant size using plants shown in (E) using the Excessive Green plugin on ImageJ (Jérôme Mutterer).

24B,24E). The size of the aerial parts of the plants was analyzed by the measurement of plant area at 15 days after sowing on MS plates (Figure 24C,24F). The phenotypic analysis is shown for the GFP-UCN lines in the *ucn-3* background (Figure 24B,24E), the same analysis was also done for lines expressing GFP-UCN in the *ucn-1* and *ucn-2* mutant backgrounds and the same results were observed (data not shown). MS plates contained glufosinate to specifically select resistant plants, but the same phenotype was also observed without glufosinate in the growing media. In total, this developmental defect was observed in 12 independent lines expressing highest levels of GFP-UCN fusion protein in the three *ucn* mutant background but also in Col-0 background, indicating that this phenotype is independent on UCN T-DNA insertion mutants. *Arabidopsis* lines expressing the RFP-UCN fusion protein are currently being produced in the lab, the occurrence of this phenotype will be also investigated and correlated with RFP-UCN expression. The phenotypes observed here occurred for both the expression of WT and mutant forms of UCN suggesting that they are independent of UCN catalytic activity. They could be linked to the recruitment of the decapping complex to specific targets by both UCN versions.

4.2.2.2. Ethylene insensitivity in lines expressing GFP-UCN

The NMD mutants *upf1, upf2 and upf3* were shown to have an ethylene insensitivity phenotype (Merchante et al., 2015). In order to determine if the deregulation of UCN expression could affect ethylene sensitivity, we analyzed the phenotype of GFP-UCN lines regarding ethylene sensitivity. For this purpose, GFP-UCN plants were sowed on agar medium supplemented with the 1-aminocyclopropane-1-carboxylic acid (ACC) ethylene precursor. The growth of control plants on this medium results in the characteristic triple response phenotype, which consists in the shortening and thickening of hypocotyls and roots as well as an exaggerated curvature of the apical hook (Figure 25A) (Merchante et al., 2015). Typically, ethylene-insensitive plants do not show this triple response: the shortening of the hypocotyl and root is less pronounced, and the apical hook is less curved. This ethylene-insensitivity phenotype is observed in *xrn4* mutants, as well as NMD mutants including *upf1*, *upf2*, and *upf3* (Figure 25B) (Merchante et al., 2015).

The ethylene response phenotype was quantified in lines expressing GFP-UCN upon the measurement of plantlet length and the counting of exaggerated curved apical hook after ACC exposure (Figure 25B-25C). We observed that the T2-2, T2-5, T2-8 and T2-9 lines expressing GFP-UCN-WT and the T2-1 and T2-3 lines expressing GFP-UCN-D153N show a reduced triple response upon growth on ACC compared to the control Col-0 or the *ucn-3* mutant. Only lines showing hardly detectable or no GFP-UCN expression (T2-1 line expressing GFP-UCN-WT and the T2-2 line expressing GFP-UCN-D153N) showed wild type ethylene

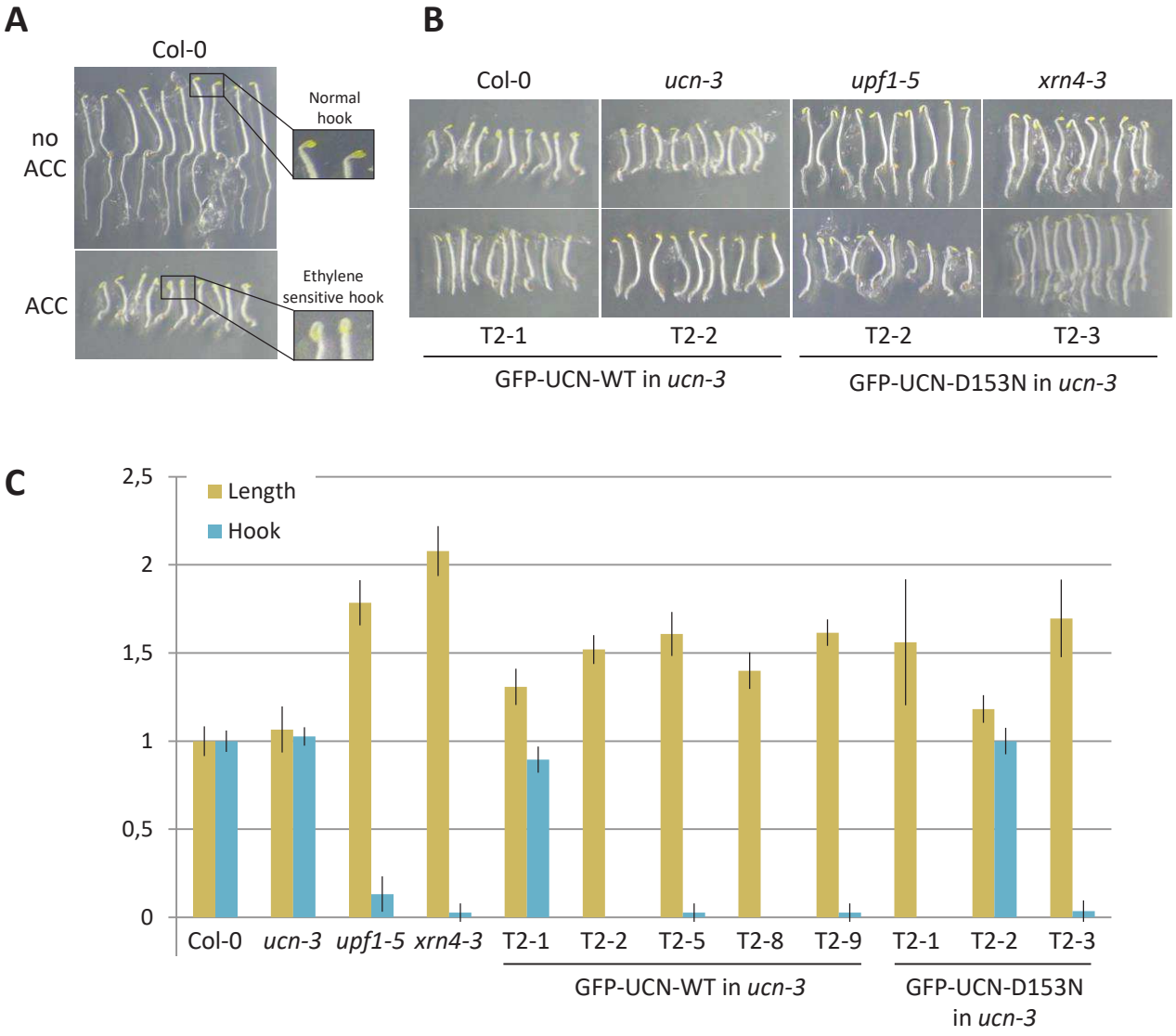


Figure 25. Ethylene insensitivity phenotype of GFP-UCN *Arabidopsis* lines. (A) Ethylene precursor (ACC) treatment induces the formation of exaggerated curved apical hook and shortening of hypocotyl and roots. (B) Example of ethylene sensitivity tests for *Arabidopsis* lines expressing GFP-UCN-WT or GFP-UCN-D153N grown on MS medium supplemented with ACC. (C) Quantification of the length of plantlets and the number of exaggerated curved apical hook in GFP-UCN plants compared to ACC-sensitive Col-0 and *ucn-3* and ethylene-insensitive *upf1-5* and *xrn4-3*. 10 plantlets were measured for each replicate. n=3

sensitivity comparable to Col-0 and *ucn-3* (Figure 25C). These results suggest a clear correlation between the expression of GFP-UCN fusion detected on western blot and the ethylene-insensitivity phenotype observed. As *Arabidopsis* lines expressing UCN mutant versions also showed an ethylene insensitivity phenotype, this suggests that this phenotype is independent on UCN catalytic activity.

4.3. Investigating the functional role of the UCN endonuclease by transient assay

UCN was initially identified as a UPF1 co-purified nuclease, but NMD targets were not deregulated in the three *ucn* mutants tested. This result suggests that UCN is not a limiting factor for NMD. In order to further test the potential influence of UCN on the accumulation of transcripts targeted by NMD, we switched to an *in vivo N. benthamiana*-based transient expression assay. This system was extensively used to study proteins important for NMD (Kerényi et al., 2008; Mérai et al., 2013; Kertész et al., 2006). In this system we monitored the effect of UCN expression on co-expressed transcripts with several NMD-inducing features including long 3'-UTR (GFP-L), presence of intron in 3'UTR (GFPcLs), and NMD-sensitive viruses including PVX and TCV.

4.3.1. Transient expression of GFP-UCN-WT in *N. benthamiana* results in necrosis

Upon transient expression of GFP-UCN for confocal microscopy analysis, we identified an additional property of the UCN endonuclease. We observed that the transient expression of GFP-UCN-WT in *N. benthamiana* leaves induced a necrosis phenotype at late stages around 6 days after infiltration (Figure 26A). Interestingly, the mutant versions GFP-UCN-D55N and GFP-UCN-D153N did not induce this necrosis phenotype while expressed to similar levels than the GFP-UCN-WT version (Figure 26B). These results indicate that the necrosis phenotype observed upon GFP-UCN-WT expression is dependent on GFP-UCN catalytic activity.

The necrosis phenotype observed upon GFP-UCN expression could be related to cell death induced during Hypersensitive Response (HR) in *N. benthamiana*. The pathogenesis-related factor PR1 is a factor whose expression is induced by a vast number of pathogens but also upon stresses, but its precise role is not yet understood. Some PR1 proteins are secreted as part of the HR to pathogens (Rivière et al., 2008). In order to test if PR1 expression is induced upon the transient expression of GFP-UCN, we analyzed PR1 protein accumulation by western blot using antibodies directed against PR1. We observed no significant changes between PR1 protein levels upon GFP-UCN-WT expression, which induce a necrosis phenotype compared to the expression of GFP-UCN-D55N and GFP-UCN-D153N catalytic mutant versions that do not show any necrosis symptom (Figure 26C). We conclude that the necrosis phenotype observed for GFP-UCN infiltration is independent of PR1 expression, but dependent on UCN catalytic activity.

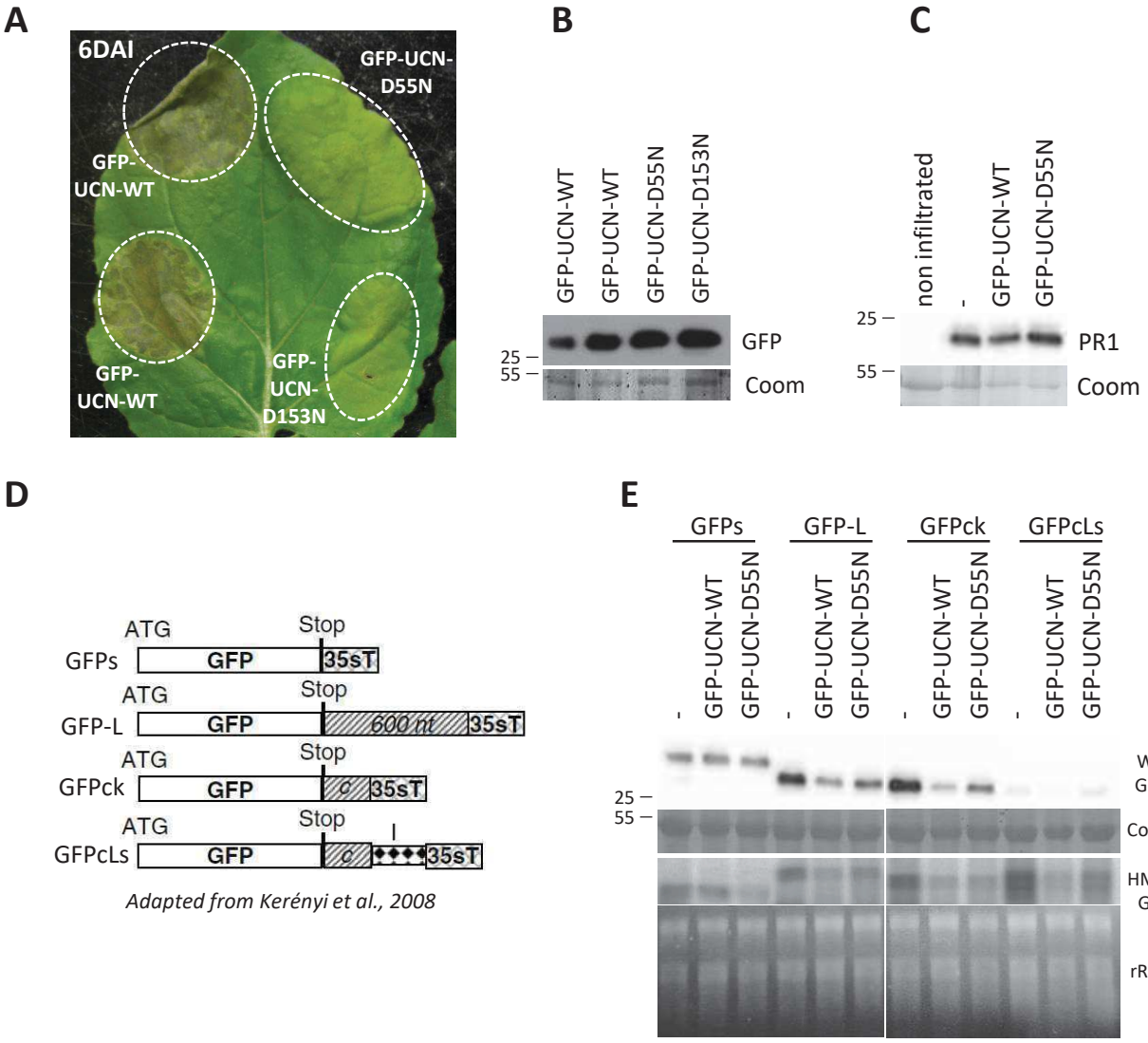


Figure 26. Influence of GFP-UCN transient expression. (A) *N. benthamiana* leaf agro-infiltrated with the indicated constructions 6DAI. (B) Western blot analysis showing the accumulation of different versions of GFP-UCN in *N. benthamiana* leaf as shown in (A). (C) Western blot analysis showing the accumulation of PR1 upon the agro-infiltration of the different GFP-UCN versions. (D) UTR organization of GFP sensors used in this study, from Kerényi et al., 2008. (E) Western blot (upper part) and Northern blot (lower part) analysis of the accumulation of GFP sensors after agro-infiltration alone or with different GFP-UCN versions.

4.3.2. Effect of UCN transient expression on NMD-sensitive RNAs

Modified GFP transcripts have been designed to test the involvement of proteins of interest in RNA decay pathways such as NMD (Paillusson et al., 2005; Mérai et al., 2013; Kertész et al., 2006). As UCN co-purifies with UPF1, we decided to test the effect of UCN expression in transient assays on the accumulation of several GFP sensors containing NMD-inducing features such as a long 3'-UTR (GFP-L) or an intron in 3'UTR (GFPcLs) (Figure 26D). All these NMD sensors accumulated to higher levels when NMD was impaired upon agro-infiltration of a dominant negative version of UPF1 (Kertész et al., 2006; Garcia et al., 2014).

Each NMD sensor used was infiltrated either alone or with the GFP-UCN-WT or the GFP-UCN-D55N catalytic mutant. In parallel, the same *N. benthamiana* leaf was infiltrated with a control GFP not targeted by NMD (Figure 26D ; GFPs used as a control for GFP-L and GFPck for GFPcLs). If UCN stimulates NMD, we expect to see a down-regulation of the NMD sensors sensitive to NMD (GFP-L and GFPcLs) and not for the controls (GFPs and GFPck). A down-regulation of GFP sensors was observed for GFP-UCN-WT agro-infiltration with GFP-L, GFPck and GFPcLs both at the protein and mRNA levels compared to the agro-infiltration of the GFP construct alone (Figure 26E). Co-infiltration with GFP-UCN-D55N mutant version also induce a GFP-L and GFPck down-regulation, this suggests that part of UCN action on GFP-L and GFPcLs may be independent on its catalytic activity. Interestingly, the down-regulation of GFP-L, GFPck and GFPcLs was more pronounced upon infiltration with GFP-UCN-WT than with GFP-UCN-D55N both at the protein and mRNA levels, indicating the possibility of a UCN effect stimulated by its catalytic activity (Figure 26E). These results suggest that UCN may target mRNAs and induce their destabilization in a NYN-domain and catalytic activity dependent but also independent manner.

UCN apparently does not target all mRNA as the GFPs version was not affected by GFP-UCN expression. An effect of UCN expression on NMD target accumulation was effectively observed, as GFP-L and GFPcLs are down-regulated, but this effect was not strictly restricted to NMD targets as GFPck, which is not targeted by NMD was also deregulated. These results suggest that there is some specificity in UCN deregulated RNAs but the precise features recognized by UCN remain to be determined. Another possible explanation here would be that the addition of the GFP tag perturbed UCN activity, and that we mainly observe the non-catalytic effects of UCN overexpression in these experiments. These effects could be rather linked with UCN interaction with the decapping activator DCP1. These experiments will be repeated using a smaller tag to compare the influence of a small compared to a bigger tag on the down-regulation of GFP targets upon UCN transient expression.

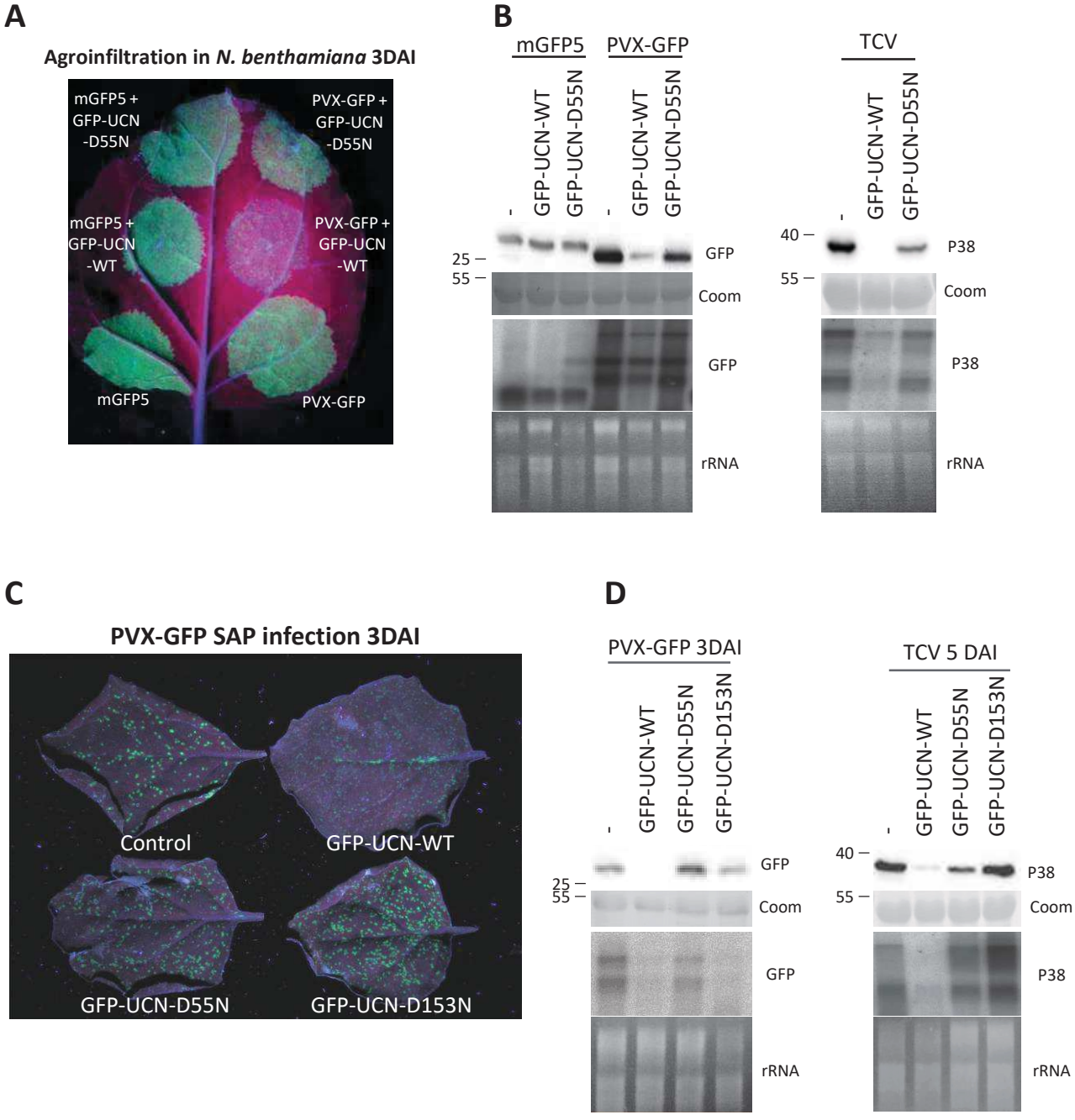


Figure 27. GFP-UCN-WT expression restricts virus accumulation. (A) *N. benthamiana* leaf agro-infiltrated with the indicated constructions 3DAI. (B) Western blot (upper panel) and Northern blot (lower panel) analysis of the accumulation of PVX-GFP and TCV (P38) accumulation after agro-infiltration alone or with different GFP-UCN versions shown in (A). (C) *N. benthamiana* leaves infiltrated with the different versions of UCN-GFP and SAP infected with PVX-GFP. 3DAI. (D) Western blot (upper panel) and northern blot (lower panel) analysis of the accumulation of PVX-GFP and TCV (P38) accumulation after SAP infection of PVX-GFP or TCV on *N. benthamiana* plants infiltrated with different GFP-UCN versions.

4.3.3. Effects of UCN expression upon viral infection

4.3.3.1. Effect of co-infiltration of RNA viruses with GFP-UCN

Some RNA viruses have characteristics, which make them sensitive to the action of NMD, including PVX and TCV (Garcia et al., 2014). Binary expression vectors containing PVX-GFP and TCV were agro-infiltrated alone or with either the GFP-UCN-WT or the GFP-UCN-D55N catalytic mutant version. mGFP5 was used as a control for PVX-GFP. A strong down-regulation was observed for PXV-GFP and TCV both at the protein and mRNA levels upon GFP-UCN-WT infiltration and a milder down-regulation was observed upon the infiltration of the GFP-UCN-D55N catalytic mutant version (Figure 27A-27B). TuMV-GFP, which is immune to NMD (Garcia et al., 2014) was also tested but no consistent variations were observed (data not shown). This result suggests an activity-dependent effect of UCN expression on the accumulation of viral RNAs expressed by PVX and TCV.

4.3.3.2. Effect of UCN expression upon SAP viral infections

In order to test more natural infection methods, independent on agro-inoculations, we used crude virions SAP prepared from *N. benthamiana* infected plants. SAP allows infections of single cells, leading to discrete infection foci, a very different setting compared to agro-infiltrations in which all agro-infected cells contain high levels of viral RNA continuously produced by the transgene. SAP infections were performed using PVX-GFP, TuMV-GFP and TCV viruses on *N. benthamiana* plants infiltrated one day earlier with GFP-UCN-WT, GFP-UCN-D153N or GFP-UCN-D55N versions.

We observed that PVX-GFP and TCV accumulation was reduced upon GFP-UCN-WT expression but not upon GFP-UCN-D55N and GFP-UCN-D153N infiltration (Figure 27C-27D). In contrast, no reproducible variations were observed for TuMV-GFP infection (data not shown). As already mentioned, PVX-GFP and TCV are both targeted by NMD, suggesting a possible effect of UCN in targeting viral RNAs depending on their sensitivity to NMD. These results obtained upon viral infections, suggest that the manipulation of UCN expression can lead to enhanced resistance to viral infection. It will be very interesting to further study this phenomenon to better understand how UCN expression affects the accumulation of RNA viruses. One of the first experiments would be to determine if UCN expression in stable lines also influence the susceptibility to viral infection in *Arabidospsis*.

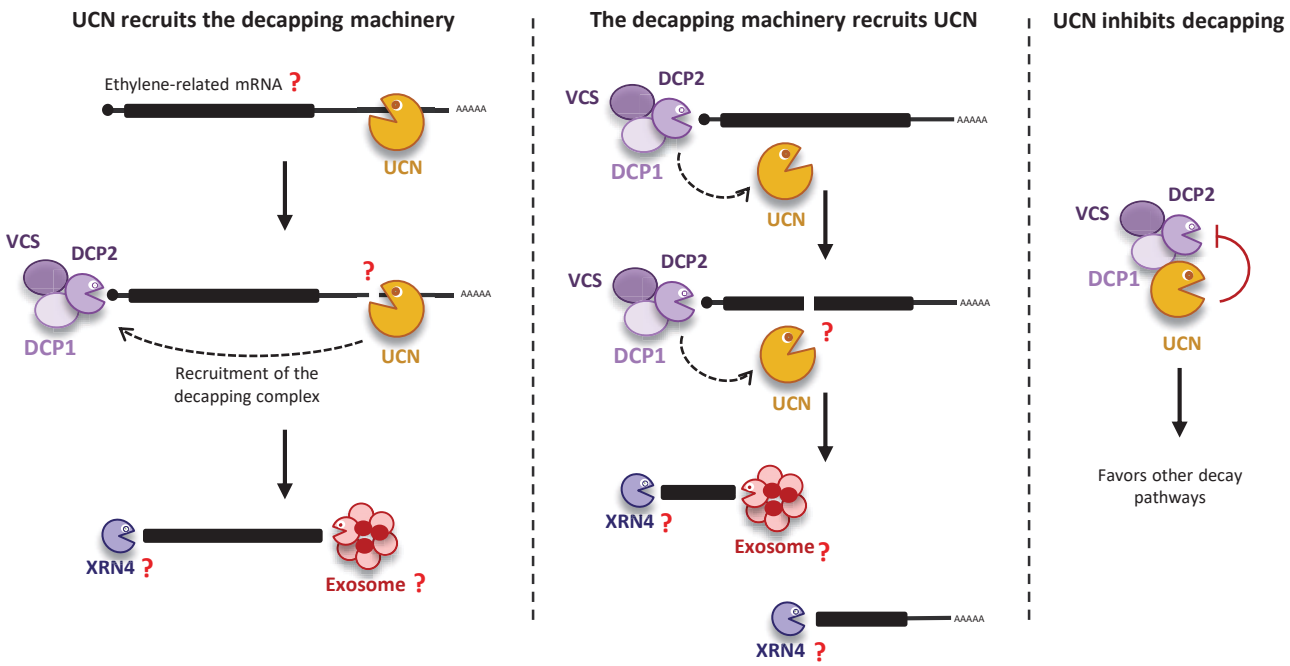


Figure 28. Proposed model for UCN mode of action. The UCN endonuclease recognizes its targets and binds their 3'UTR. UCN targets may be related to ethylene regulation. UCN recruits the decapping machinery via a direct interaction with DCP1 to mediate transcript decay, possibly via XRN4 and exosome activities (left panel). Another model could be that UCN is recruited by the decapping complex on specific targets to mediate accelerated decay (middle panel). Finally, upon UCN direct interaction with DCP1, UCN could inhibit DCP2 decapping activity, which could favor other decay pathways (right panel).

IV. Discussion

1. The UCN interactome

The study of UCN interactome revealed the association between UCN and the decapping activators DCP1 and VCS. A direct interaction between UCN and DCP1 was observed in yeast two-hybrid experiments, and UCN was observed to localize in P-bodies where decapping activators DCP1 and VCS also accumulate. These results suggest a possible role for UCN in regulating mRNA stability related to decapping. Interestingly, the closest mammal homolog of UCN, MARF1 directly interacts with the VCS homolog Ge-1 through its C-terminal region (Bloch et al., 2014). MARF1 has been shown to have a role in mRNA stability as it regulates the expression of a transcripts involved in the cellular response to Type I interferons (Bloch et al., 2014). Yeast two-hybrid experiments will be performed to identify the domains of UCN involved in its interaction with DCP1. To this aim, truncated versions of UCN, *i.e.* either the NYN or the OST-HTH domain, are currently being cloned independently to determine which of them mediates the interaction with DCP1. To further study the role of UCN in mRNA decay, it will be of interest to identify the capping status of RNAs sensitive to UCN action. One of the most sensitive RNA in transient expression assays was PVX-GFP, a capped and polyadenylated virus. The used of 5' RACE experiments on the PVX-GFP transcripts agro-infiltrated with UCN could allow the quantification of capped versus uncapped transcripts. Several scenarios are possible: either UCN is recruited on decapped mRNA by the DCP1/VCS complex to promote mRNA decay or UCN itself recruits the decapping machinery to stimulate decay after target recognition (Figure 28). Another possibility is that UCN could bind to the decapping machinery and inhibit decapping activity, which would favor other decay pathways.

Other proteins were retrieved in UCN immunoprecipitations, such as HSP70 and TUB isoforms. Of note, HSP70 isoforms were retrieved with very important number of spectra: almost half of the number of spectra of the UCN bait. This result suggests that UCN or a complex formed by UCN and its partner might need HSP70 proteins for proper folding. To test this hypothesis, IP with RNase treatment will be necessary to decipher RNA-dependent associations from protein-mediated interactions. Four TUB isoforms were identified in UCN IPs but also a protein associated with microtubules, TORTIFOLIA 1 (Buschmann et al., 2004). UCN is a P-body component and P-body movement in mammals was shown to be dependent on microtubules, while in *Arabidopsis*, P-body movement is mediated by acto-myosin (Aizer et al., 2008; Loschi et al., 2009; Hamada et al., 2017; Steffens et al., 2014). The association between the P-body factor UCN and TUB factors suggests a possible involvement of UCN in P-bodies movement, which may also be mediated by microtubules in *Arabidopsis*.

2. UCN substrate specificity

Preliminary experiments were performed to test the effect of UCN expression on NMD targets. Different GFP sensors showed reduced accumulation upon infiltration of both GFP-UCN-WT and GFP-UCN-D55N versions in transient experiments. These assays suggest that UCN expression can influence some targets independently of its catalytic activity. Some NMD targets were deregulated (GFP-L and GFPcLs) but another GFP sensor, GFPck, not targeted by NMD was also downregulated. These results suggest that the influence of UCN is not specific to NMD targets, which is consistent with the fact that *ucn* mutants did not show overaccumulation of NMD targets (Figure 22D). Two GFP sensors were not affected by GFP-UCN infiltration, GFPs and mGFP5 indicating that UCN does not equally influence all transcripts despite the fact that the same promoter drives their expression. This result strongly supports a model in which UCN does not act at the level of transcription, but instead at the post-transcriptional level. GFPs and mGFP5 are the only two GFP sensors tested that do not possess any 3'-UTR extension suggesting that UCN targeting could be influenced by the presence of an extended 3'-UTR. This is reminiscent to the mode of action of another NYN-domain endonuclease, MCPIP1, which recognizes a stem loop in the 3'-UTR of its targets (Li et al., 2012). Whether endogenous UCN is able to globally recognize all 3'-UTRs or a motif found in some 3'-UTR remains to be determined. We propose a model for which UCN recognition on mRNA 3'-UTR may recruit the decapping machinery via a direct interaction between UCN and DCP1 to promote decay (Figure 28). Another possible model would be that the decapping complex recruits UCN to accelerate mRNA decay. Finally, an alternative model could be that UCN inhibits decapping activity upon its binding to DCP1, which may favor other decay pathways.

Of notice, the effect of GFP-UCN infiltration on PVX-GFP and TCV was stronger for the WT than for the catalytic mutant version both in agro-infiltration and virus infection, suggesting an impact of the catalytic activity of UCN on these viruses. The catalytic function of UCN as an endonuclease is currently being tested in the lab. WT and catalytic mutant versions of UCN were expressed in bacteria and protein purification was performed. These purified protein versions were used in activity test using several RNA or DNA targets. This study showed the UCN specificity for RNA over DNA. Further test by Thin Layer Chromatography (TLC) demonstrated that UCN carries an endonuclease activity and no exoribonuclease activity.

The next key question will be to determine UCN target specificity. The identification of direct targets bound to UCN could be realized by RNA immunoprecipitation (RIP). For this experiment we can anticipate the value of the catalytic mutants, which could interact but stay trapped on their targets in the absence

of the residues required for the nuclease activity. In parallel, a comparative transcriptomic analysis of *ucn* and transgenic lines with stable expression of WT and mutant UCN will identify genetic pathways affected upon UCN deregulation. These approaches will likely provide more information on UCN target specificity, biological roles and on the potential agronomic interest of modulating the expression of this novel NYN domain endonuclease.

3. Effect of UCN expression on immunity

The strongest effect observed upon UCN infiltration was described for the PVX-GFP and TCV viral RNAs. PVX-GFP and TCV accumulations were reduced upon GFP-UCN-WT expression but not for GFP-UCN-D55N or GFP-UCN-D153N. PVX-GFP and TCV RNAs are both targeted by NMD as they possess an unusually long 3'-UTR due to the presence of internal termination codons (Garcia et al., 2014). Interestingly, TuMV-GFP, which does not possess any internal termination codon was not affected by UCN infiltration. These results suggest that UCN expression can influence the multiplication of some RNA viruses in plants. Of note, the NYN-domain MCPIP1 is also involved in pathogen defense as it suppresses hepatitis C replication (Lin et al., 2017). Infections are currently being tested in the lab on GFP-UCN lines with TCV and TuMV-GFP to determine if ectopic expression of GFP-UCN could impair viral accumulation in *Arabidopsis* stable transgenic plants.

PVX-GFP and TCV do not share similar features at the 3' and 5' extremities of their genomic RNAs: PVX-GFP carries a cap structure and a poly-A tail while TCV possesses a stem loop structure in 5' and a 3'OH extremity. The hypothesis of the involvement of the decapping complex to induce the decay of UCN targets may be challenged by the fact that TCV, which does not possess any cap structure is also negatively affected by transient UCN expression. Alternatively, this restriction observed for these two viruses may be indirect and due to UCN-induced cellular defects, such as deregulations leading to cell death at late stages of UCN transient expression in *N. benthamiana*.

The cell death phenotype is observed at late stages of GFP-UCN-WT transient expression but not upon GFP-UCN-D55N or D153N catalytic mutants. This phenotype could be related to hypersensitive response or *Agrobacterium*-induced necrosis (Bashir et al., 2013; Kuta and Tripathi, 2005). A first level of plant immunity corresponds to the PAMP-triggered immunity (PTI) where pathogen-associated molecular patterns (PAMPs) are recognized by extracellular receptors called PRR for Pattern Recognition Receptor. PAMPs are diverse molecules found in bacteria and fungi pathogens such as the flagellin or the chitin for example (Brendolise et al., 2017). A second layer of plant immunity corresponds to the effector-triggered

immunity (ETI) in which plants R proteins such as Leucine Rich Repeat (LRR) receptors recognize pathogens effector factors capable of suppressing PTI. Upon the recognition of the effector, a cascade of reactions is induced leading to a localized programmed cell death to restrict pathogen spreading (Brendolise et al., 2017; Bashir et al., 2013). One example of programmed cell death is the hypersensitive response, which is highly controlled and characterized by the production of reactive oxygen species (ROS) (Bashir et al., 2013). The overexpression of GFP-UCN-WT could induce an hypersensitive response, which would lead to the observed cell death phenotype at late steps and induce pathogen restriction. *Agrobacterium* strains used for agro-infiltration are non-pathogenic stains and do not normally induce hypersensitive responses. The overexpression of GFP-UCN-WT could either independently activate cellular pathway leading to hypersensitive response or could enhance the cellular response to non-pathogenic *Agrobacterium* strains.

In addition to these observations, experiments currently undergoing in the lab suggest that the lines expressing GFP-UCN-WT stimulate dark-induced senescence but not lines expressing GFP-UCN-D153N. In plants, senescence is a highly controlled process, which requires a massive transcriptional reprogramming and is essential for the adaptation to an environment with limiting condition (Liebsch and Keech, 2016). UCN may be involved in this transcriptional reprogramming by promoting the decay of specific mRNA. The dark induced senescence phenotype of GFP-UCN expressing plants is reminiscent of the phenotype of ETHYLENE-INSENSITIVE 3 overexpressing lines (EIN3-OX) (Li et al., 2013). EIN3 is known to be required for dark induced senescence, and overexpression of EIN3 increases senescence related phenotypes (Li et al., 2013). In addition, it is worth to note that UCN is overexpressed in EIN3-OX lines (Peng et al., 2014). The common phenotype between EIN3-OX and lines expressing GFP-UCN, as well as the ethylene insensitivity phenotypes observed, suggest a role for UCN in the regulation of ethylene signaling. Ethylene is known to be involved in immunity (Huang et al., 2017; van Loon et al., 2006), it will be interesting to investigate if GFP-UCN expression influences ethylene signaling, which modulates the resistance to pathogens including viruses and to determine the potential impact of GFP-UCN expression on other pathogens like fungi or bacteria.

Supplementary Figures Chapter II

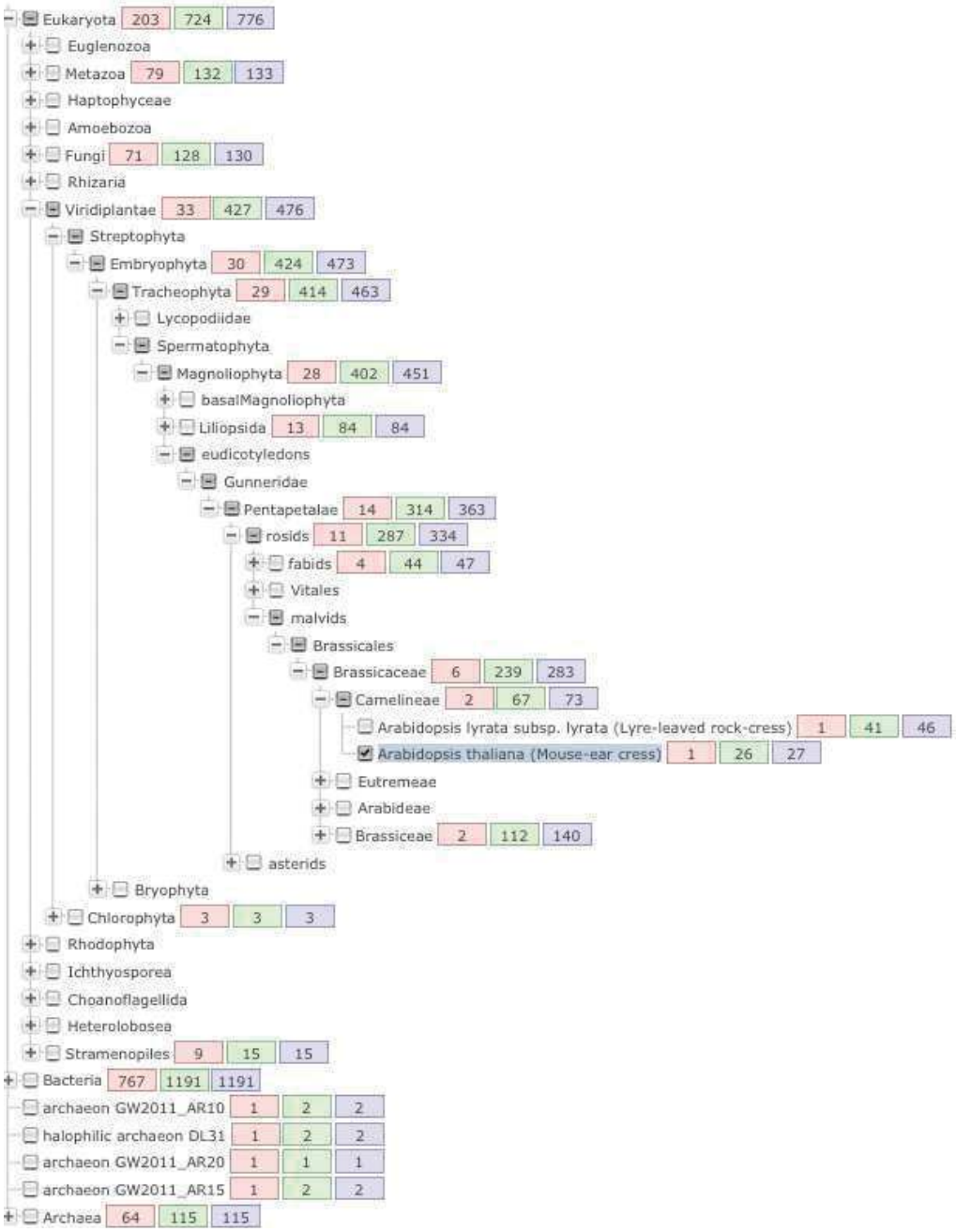


Figure S6. Phylogenetic tree of the conservation of NYN-domain endonucleases. *Arabidopsis thaliana* is highlighted in the middle. From Pfam database (PF01936). For each phylum, the number of species in indicated in red, the number of sequences in green and the number of regions in blue.

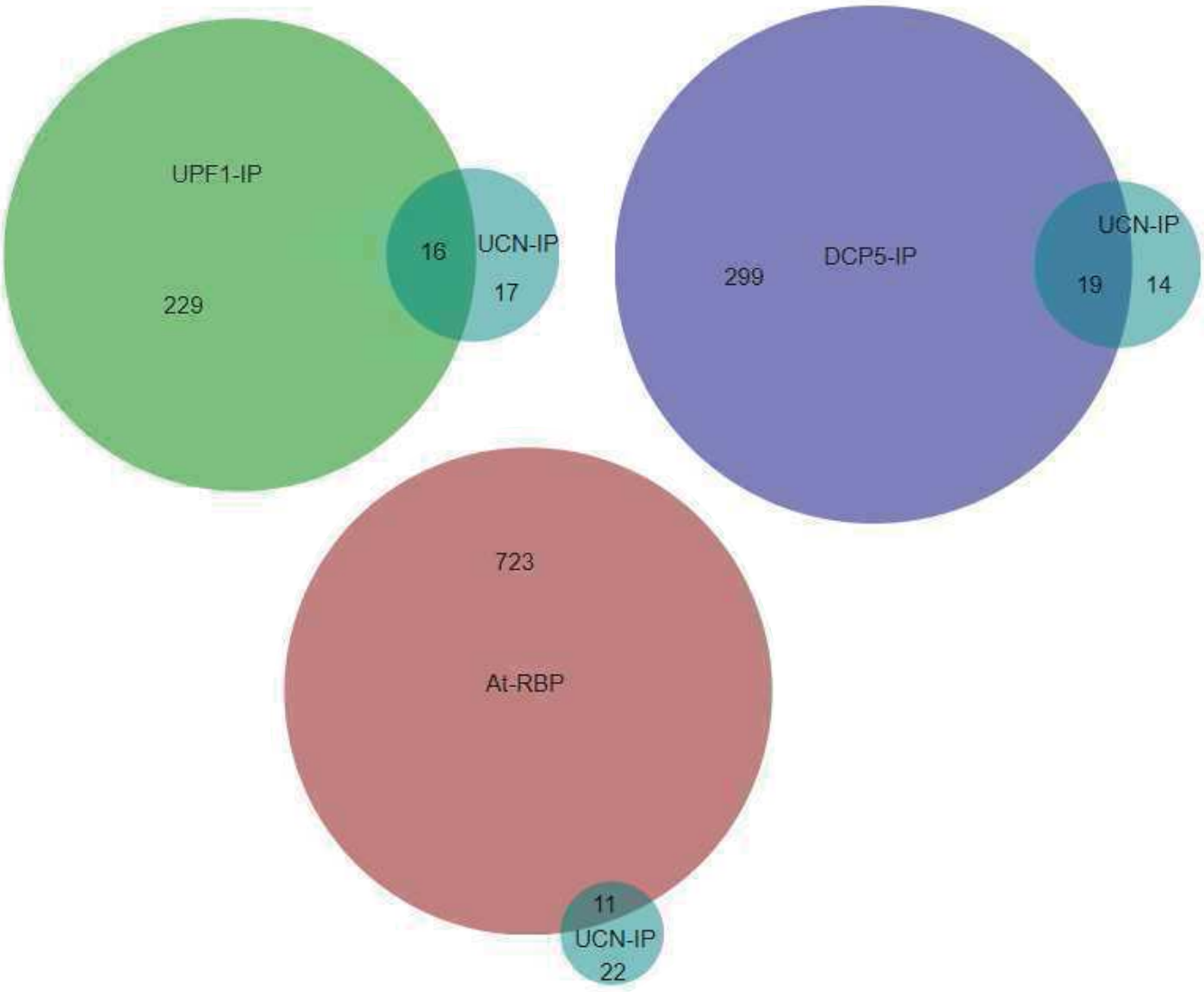
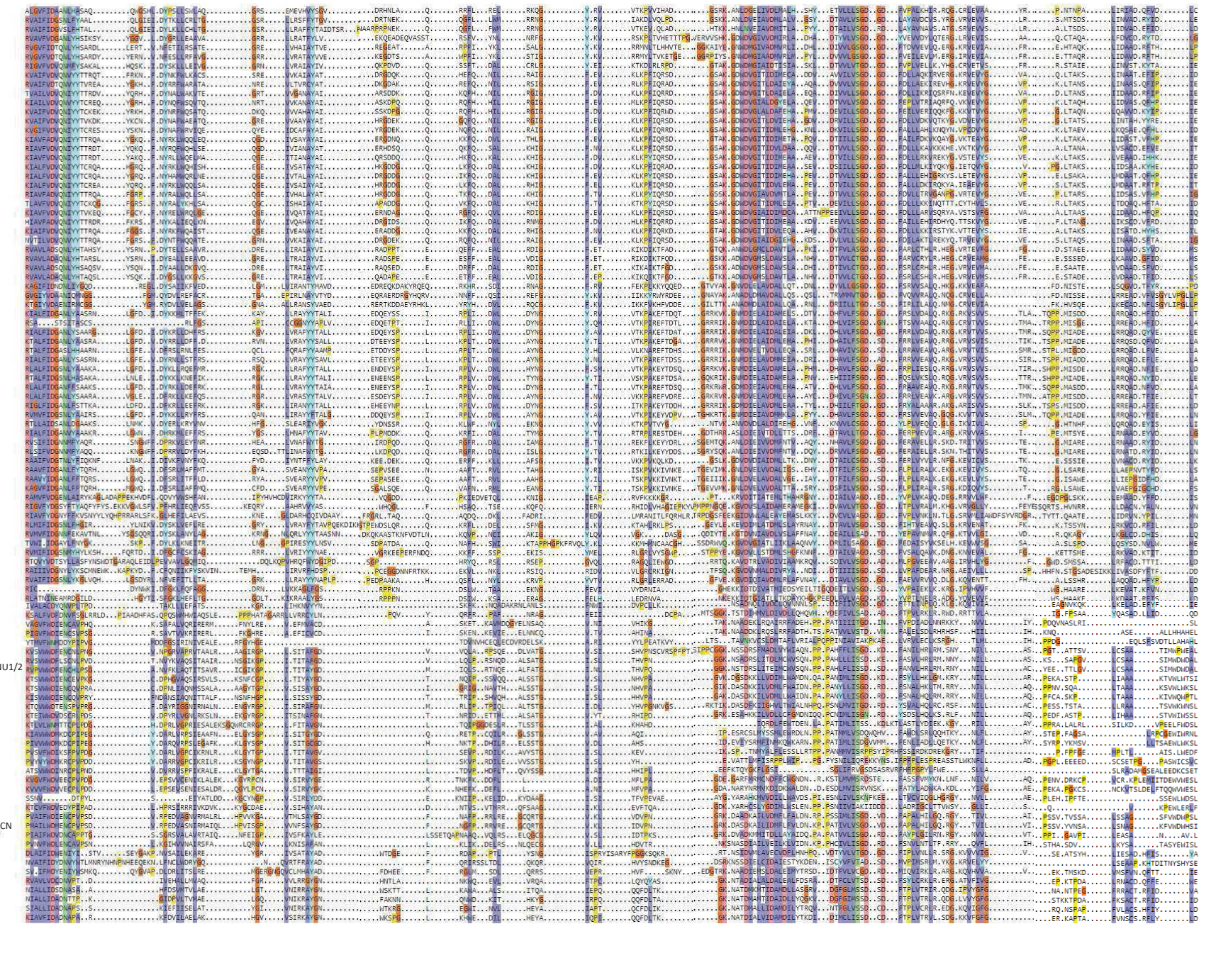


Figure S7. Comparison of UCN partners with UPF1 and DCP5 partners and At-RBP (Reichel *et al.,* 2016).

Figure S8. Alignment of NYN-domain proteins. 147 NYN-domain endonucleases documented on Pfam (PF01936) are aligned via Jalview. UCN (Q6NQ97) is highlighted. The logo of conserved residues in NYN-domain in Figure 23 was created thanks to this alignment.



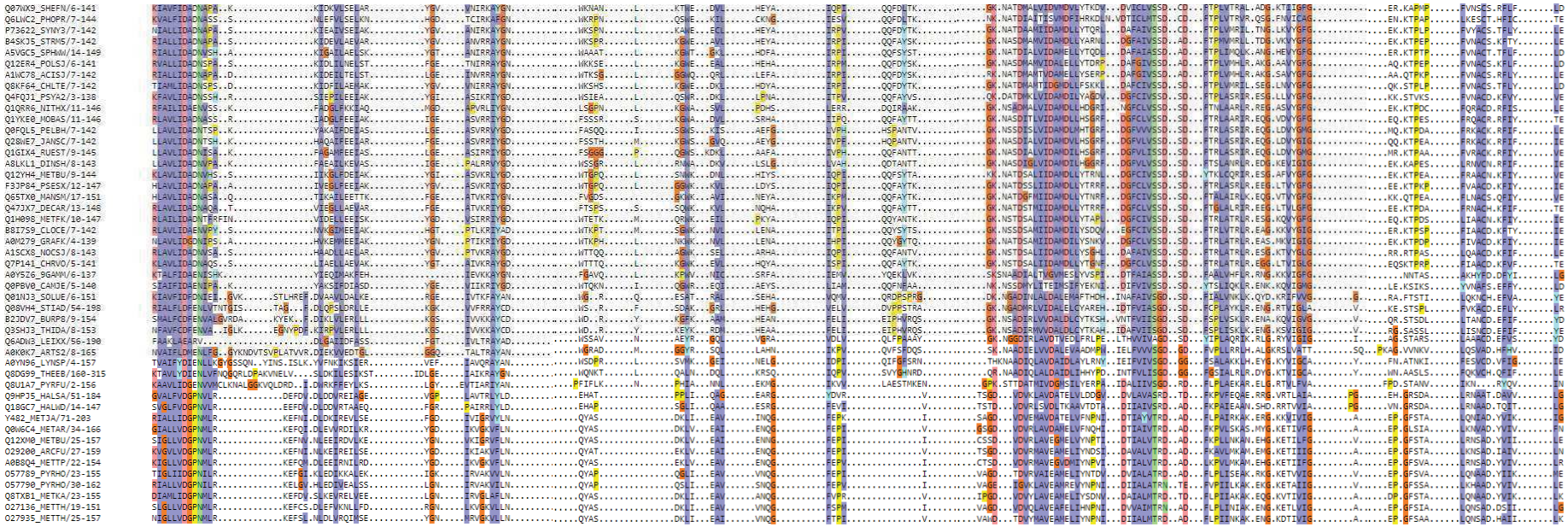

Q07WX9_SHEFN/6-141
Q6LNC2_PHOPR/7-144
P73622_SYNY3/7-142
B4SKJ5_STRM5/7-142
A5VGC5_SPHWW/14-149
Q12ER4_POLSJ/6-141
A1WC78_ACISJ/7-142
Q8KF64_CHLTE/7-142
Q4FQJ1_PSYA2/3-138
Q1QRR6_NITHX/11-146
Q1YKE0_MOBAS/11-146
Q0FQL5_PELBH/7-142
Q28WE7_JANSC/7-142
Q1GIX4_RUEST/9-145
A8LKL1_DINSH/8-143
Q12YH4_METBU/9-144
F3JP84_PSESX/12-147
Q65TX0_MANSM/17-151
Q47JX7_DECAR/13-148
Q1H098_METFK/10-147
B8I7S9_CLOCE/7-142
A0M279_GRAFK/4-139
A1SCX8_NOCSJ/8-143
Q7P141_CHRVO/5-141
A0YSZ6_9GAMM/6-137
Q0PBV0_CAMJE/5-140
Q01UJ3_SOLUE/6-151
Q08VH4_STIAD/54-198
B2JDV7_BURPS/9-154
Q3SH33_THIDA/8-153
Q6ADW3_LEIXX/56-190
A0K0K7_ARTS2/8-165
A0YN96_LYNSP/4-157
Q8DG99_THEEB/160-315
Q8U1A7_PYRFU/2-156
Q9HPJ5_HALSA/51-184
Q18GC7_HALMD/14-147
Y482_METJA/71-203
Q0W6C4_METAR/34-166
Q12XM0_METBU/25-157
O29200_ARCFU/27-159
A0B8Q4_METTP/22-154
O57789_PYRHO/23-155
O57790_PYRHO/30-162
Q8TXB1_METKA/23-155
O27136_METTH/19-151
O27935_METTH/25-157

Table S7. Common partners between UPF1 and DCP5. Comparison of proteins enriched in UPF1-IPs and DCP5-IPs

AT1G01300	"Eukaryotic aspartyl protease family protein; FUNC
AT1G02800	"cellulase 2 (CEL2); FUNCTIONS IN: cellulase activ
AT1G04680	"Pectin lyase-like superfamily protein; FUNCTIONS
AT1G07310	"Calcium-dependent lipid-binding (CaLB domain) fam
AT1G08370	"decapping 1 (DCP1); CONTAINS InterPro DOMAIN/s: D
AT1G10170	"NF-X-like 1 (NFXL1); FUNCTIONS IN: zinc ion bindi
AT1G13190	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT1G14710	"hydroxyproline-rich glycoprotein family protein;
AT1G22760	"poly(A) binding protein 3 (PAB3); CONTAINS InterP
AT1G26110	"decapping 5 (DCP5); CONTAINS InterPro DOMAIN/s: D
AT1G27090	"glycine-rich protein; BEST Arabidopsis thaliana p
AT1G29250	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT1G32790	"CTC-interacting domain 11 (CID11); FUNCTIONS IN:
AT1G33680	"KH domain-containing protein; FUNCTIONS IN: RNA b
AT1G43190	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT1G47490	"RNA-binding protein 47C (RBP47C); CONTAINS InterP
AT1G48410	"ARGONAUTE 1 (AGO1); FUNCTIONS IN: protein binding
AT1G49760	"poly(A) binding protein 8 (PAB8); FUNCTIONS IN: R
AT1G55500	"evolutionarily conserved C-terminal region 4 (ECT
AT1G64390	"glycosyl hydrolase 9C2 (GH9C2); FUNCTIONS IN: car
AT1G70710	"glycosyl hydrolase 9B1 (GH9B1); FUNCTIONS IN: cel
AT1G71770	"poly(A)-binding protein 5 (PAB5); FUNCTIONS IN: R
AT1G75660	"5'-3' exoribonuclease 3 (XRN3); CONTAINS InterPro
AT1G76010	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G06850	"xyloglucan endotransglucosylase/hydrolase 4 (XTH4
AT2G15560	"Putative endonuclease or glycosyl hydrolase; BEST
AT2G17870	"cold shock domain protein 3 (CSP3); FUNCTIONS IN:
AT2G21060	"glycine-rich protein 2B (GRP2B); FUNCTIONS IN: DN
AT2G23350	"poly(A) binding protein 4 (PAB4); FUNCTIONS IN: R
AT2G26280	"CID7; FUNCTIONS IN: damaged DNA binding, protein
AT2G27100	"SERRATE (SE); FUNCTIONS IN: DNA binding, sequence
AT2G28950	"expansin A6 (EXPA6); CONTAINS InterPro DOMAIN/s:
AT2G29140	"pumilio 3 (PUM3); FUNCTIONS IN: RNA binding, bind
AT2G29190	"pumilio 2 (PUM2); FUNCTIONS IN: mRNA binding, RNA
AT2G29200	"pumilio 1 (PUM1); FUNCTIONS IN: RNA binding, bind
AT2G31810	"ACT domain-containing small subunit of acetolacta
AT2G34160	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G34810	"FAD-binding Berberine family protein; FUNCTIONS I
AT2G36870	"xyloglucan endotransglucosylase/hydrolase 32 (XTH
AT2G38110	"glycerol-3-phosphate acyltransferase 6 (GPAT6); C
AT2G38610	"RNA-binding KH domain-containing protein; FUNCTIO
AT2G42520	"P-loop containing nucleoside triphosphate hydrola
AT2G42570	"TRICHOME BIREFRINGENCE-LIKE 39 (TBL39); INVOLVED
AT2G43970	"RNA-binding protein; FUNCTIONS IN: RNA binding, n
AT2G45810	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT2G46020	"BRAHMA (BRM); FUNCTIONS IN: helicase activity, tr
AT2G46780	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT2G47250	"RNA helicase family protein; FUNCTIONS IN: in 7 f
AT3G01540	"DEAD box RNA helicase 1 (DRH1); FUNCTIONS IN: ATP
AT3G03950	"evolutionarily conserved C-terminal region 1 (ECT
AT3G04610	"flowering locus KH domain (FLK); FUNCTIONS IN: RN
AT3G06480	"DEAD box RNA helicase family protein; FUNCTIONS I
AT3G07010	"Pectin lyase-like superfamily protein; FUNCTIONS
AT3G07030	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G13060	"evolutionarily conserved C-terminal region 5 (ECT
AT3G13290	"varicose-related (VCR); FUNCTIONS IN: nucleotide
AT3G13300	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT3G13460	"evolutionarily conserved C-terminal region 2 (ECT
AT3G14100	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G15000	"cobalt ion binding; FUNCTIONS IN: cobalt ion bind
AT3G15010	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT3G19130	"RNA-binding protein 47B (RBP47B); FUNCTIONS IN: R
AT3G22310	"putative mitochondrial RNA helicase 1 (PMH1); CON
AT3G22330	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT3G23990	"heat shock protein 60 (HSP60); FUNCTIONS IN: copp
AT3G25150	"Nuclear transport factor 2 (NTF2) family protein
AT3G26420	"ATRZ-1A; CONTAINS InterPro DOMAIN/s: RNA recognit
AT3G44110	"DNAJ homologue 3 (ATJ3); FUNCTIONS IN: unfolded p
AT3G49390	"CTC-interacting domain 10 (CID10); FUNCTIONS IN:
AT3G49490	"unknown protein; Has 722 Blast hits to 186 protei
AT3G50370	"unknown protein; FUNCTIONS IN: molecular_function
AT3G51950	"Zinc finger (CCCH-type) family protein / RNA reco
AT3G54230	"suppressor of abi3-5 (SUA); FUNCTIONS IN: nucleot
AT3G58510	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G58570	"P-loop containing nucleoside triphosphate hydrola
AT3G59770	"SUPPRESSOR OF ACTIN 9 (SAC9); FUNCTIONS IN: inosi
AT3G60240	"eukaryotic translation initiation factor 4G (EIF4
AT3G60830	"actin-related protein 7 (ARP7); CONTAINS InterPro
AT3G61240	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G61690	"nucleotidyltransferases; FUNCTIONS IN: nucleotidy
AT3G61820	"Eukaryotic aspartyl protease family protein; FUNC
AT4G00660	"RNAhelicase-like 8 (RH8); FUNCTIONS IN: helicase
AT4G02290	"glycosyl hydrolase 9B13 (GH9B13); FUNCTIONS IN: h
AT4G03110	"RNA-binding protein-defense related 1 (RBP-DR1);
AT4G08350	"global transcription factor group A2 (GTA2); FUNC
AT4G10610	"CTC-interacting domain 12 (CID12); CONTAINS Inter
AT4G21670	"C-terminal domain phosphatase-like 1 (CPL1); FUNC
AT4G21710	"NRPB2; CONTAINS InterPro DOMAIN/s: DNA-directed R
AT4G22010	"SKU5 similar 4 (sks4); FUNCTIONS IN: oxidoreduct
AT4G25550	"Cleavage/polyadenylation specificity factor, 25kD
AT4G27060	"TORTIFOLIA 1 (TOR1); CONTAINS InterPro DOMAIN/s:
AT4G29010	"ABNORMAL INFLORESCENCE MERISTEM (AIM1); FUNCTIONS
AT4G31770	"debranching enzyme 1 (DBR1); CONTAINS InterPro DO
AT4G34110	"poly(A) binding protein 2 (PAB2); FUNCTIONS IN: R
AT4G34980	"subtilisin-like serine protease 2 (SLP2); FUNCTIO
AT4G35890	"winged-helix DNA-binding transcription factor fam
AT4G36960	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT4G38680	"glycine rich protein 2 (GRP2); FUNCTIONS IN: doub
AT4G39260	"cold, circadian rhythm, and RNA binding 1 (CCR1);
AT5G02530	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G04280	"RNA-binding (RRM/RBD/RNP motifs) family protein w
AT5G14610	"DEAD box RNA helicase family protein; FUNCTIONS I
AT5G20490	"XIK; FUNCTIONS IN: motor activity; INVOLVED IN: i
AT5G22060	"DNAJ homologue 2 (J2); FUNCTIONS IN: unfolded pro
AT5G22770	"alpha-adaptin (alpha-ADR); FUNCTIONS IN: protein
AT5G26742	"embryo defective 1138 (emb1138); FUNCTIONS IN: in
AT5G37720	"ALWAYS EARLY 4 (ALY4); FUNCTIONS IN: nucleotide b
AT5G40490	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G43960	"Nuclear transport factor 2 (NTF2) family protein
AT5G44780	"unknown protein; BEST Arabidopsis thaliana protei
AT5G47010	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in
AT5G48650	"Nuclear transport factor 2 (NTF2) family protein
AT5G48900	"Pectin lyase-like superfamily protein; FUNCTIONS
AT5G51750	"subtilase 1.3 (SBT1.3); FUNCTIONS IN: identical p
AT5G53060	"RNA-binding KH domain-containing protein; FUNCTIO
AT5G54900	"RNA-binding protein 45A (RBP45A); FUNCTIONS IN: R
AT5G55670	"RNA-binding (RRM/RBD/RNP motifs) family protein;
AT5G58470	"TBP-associated factor 15B (TAF15b); FUNCTIONS IN:
AT5G61020	"evolutionarily conserved C-terminal region 3 (ECT
AT5G63120	"P-loop containing nucleoside triphosphate hydrola
AT5G64960	"cyclin dependent kinase group C2 (CDKC2); FUNCTIO

Table S8. Proteins enriched in UCN-IP. In this table are indicated the informations used to create the volcano plot in Chapter II, Figure 20B (only proteins with an *adjp<0.08* are listed here). Mean of spectral counts for control or UCN IPs, log2(Fold Change), adjusted p-value (adjp) are detailed. DEP indicates whether the considered protein passed the statistical test (TRUE) or not (FALSE). Proteins in grey are not significantly enriched in UCN-IPs.

AGI	Mean spectral count **Control**	Mean spectral count **UCN**	LogFC	adjp	DEP	Description
AT2G15560	0	563,8	11,54	5,7E-279	TRUE	"Putative endonuclease or glycosyl hydrolase; BEST
AT5G02500	18,1	330,2	2,903	1,6E-99	TRUE	"heat shock cognate protein 70-1 (HSC70-1); FUNCTI
AT3G09440	13,9	211,6	2,658	2,96E-73	TRUE	"Heat shock protein 70 (Hsp 70) family protein; FU
AT5G02490	13,4	209,6	2,678	2,2E-62	TRUE	"Heat shock protein 70 (Hsp 70) family protein; FU
AT3G12580	12,7	159,8	2,363	4,04E-48	TRUE	"heat shock protein 70 (HSP70); FUNCTIONS IN: ATP
AT1G08370	0	40,4	7,741	5,35E-39	TRUE	"decapping 1 (DCP1); CONTAINS InterPro DOMAIN/s: D
AT3G62250	1,6	29,8	2,924	3,77E-16	TRUE	"ubiquitin 5 (UBQ5); FUNCTIONS IN: structural cons
AT3G13300	0,4	21	4,15	3,12E-12	TRUE	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT4G27060	0	16,4	6,443	4,86E-12	TRUE	"TORTIFOLIA 1 (TOR1); CONTAINS InterPro DOMAIN/s:
AT5G47010	0	5,4	4,876	2,95E-05	TRUE	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in
AT3G13290	0,3	6,8	3,004	0,000149	TRUE	"varicose-related (VCR); FUNCTIONS IN: nucleotide
AT4G24780	0	4	4,464	0,000258	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT3G49490	0	4	4,463	0,000264	TRUE	"unknown protein; Has 722 Blast hits to 186 protei
AT1G43190	0	4	4,468	0,000277	TRUE	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT3G53750	0,6	7,6	2,271	0,001129	TRUE	"actin 3 (ACT3); FUNCTIONS IN: structural constitu
AT4G14960	0,4	6,6	2,433	0,001548	TRUE	"TUA6; FUNCTIONS IN: protein binding, structural c
AT1G50010	0,4	6,6	2,433	0,001548	TRUE	"tubulin alpha-2 chain (TUA2); FUNCTIONS IN: struc
AT5G62700	0,7	13	2,433	0,00241	TRUE	"tubulin beta chain 3 (TUB3); FUNCTIONS IN: struct
AT3G58510	2	13,2	1,457	0,00277	TRUE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G58570	1,6	11,6	1,571	0,00277	TRUE	"P-loop containing nucleoside triphosphate hydrola
AT2G32080	0	2,8	3,976	0,002776	TRUE	"purin-rich alpha 1 (PUR ALPHA-1); FUNCTIONS IN: n
AT3G20820	4,3	31	1,303	0,00299	TRUE	"Leucine-rich repeat (LRR) family protein; INVOLVE
AT3G07010	0,1	4	2,96	0,003835	TRUE	"Pectin lyase-like superfamily protein; FUNCTIONS
AT2G34160	0	2,6	3,876	0,004083	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G22330	0,4	6,6	2,331	0,007247	TRUE	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT1G16030	1,1	42	2,727	0,009359	TRUE	"heat shock protein 70B (Hsp70b); FUNCTIONS IN: AT
AT3G22310	0	2,2	3,65	0,0101	TRUE	"putative mitochondrial RNA helicase 1 (PMH1); CON
AT5G64260	0,6	5,8	1,918	0,01048	TRUE	"EXORDIUM like 2 (EXL2); FUNCTIONS IN: molecular_f
AT4G32020	0	2	3,529	0,01321	TRUE	"unknown protein; FUNCTIONS IN: molecular_function
AT5G48650	0	1,8	3,389	0,01989	FALSE	"Nuclear transport factor 2 (NTF2) family protein
AT2G17870	0	1,8	3,389	0,02005	FALSE	"cold shock domain protein 3 (CSP3); FUNCTIONS IN:
AT3G06980	0	1,8	3,392	0,02005	FALSE	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT1G02080	0	1,8	3,395	0,02005	FALSE	"transcription regulators; FUNCTIONS IN: transcrip
AT1G76010	0,3	4	2,31	0,02334	TRUE	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G53420	0,3	4	2,205	0,02356	TRUE	"plasma membrane intrinsic protein 2A (PIP2A); FUN
AT5G57110	0	1,6	3,237	0,02869	FALSE	"autoinhibited Ca2+ -ATPase, isoform 8 (ACA8); FUN
AT1G72610	0	1,6	3,237	0,02869	FALSE	"germin-like protein 1 (GER1); INVOLVED IN: biolog
AT5G19780	0,4	5	1,98	0,02869	TRUE	"tubulin alpha-5 (TUA5); FUNCTIONS IN: structural
AT1G24260	0	1,6	3,234	0,03676	FALSE	"SEPALLATA3 (SEP3); FUNCTIONS IN: protein binding,
AT3G51950	0	1,4	3,068	0,04346	FALSE	"Zinc finger (CCCH-type) family protein / RNA reco
AT5G54900	0	1,4	3,067	0,04346	FALSE	"RNA-binding protein 45A (RBP45A); FUNCTIONS IN: R
AT1G21780	0	1,4	3,067	0,04346	FALSE	"BTB/POZ domain-containing protein; CONTAINS Inter
AT4G22380	0	1,4	3,066	0,04346	FALSE	"Ribosomal protein L7Ae/L30e/S12e/Gadd45 family pr
AT1G29920	0,3	4,8	2,286	0,04393	TRUE	"chlorophyll A/B-binding protein 2 (CAB2); FUNCTIO
AT5G20490	0	1,4	3,066	0,04511	FALSE	"XIK; FUNCTIONS IN: motor activity; INVOLVED IN: i
AT5G63120	0,1	2,4	2,299	0,054	FALSE	"P-loop containing nucleoside triphosphate hydrola
AT3G12915	0	1,2	2,874	0,06219	FALSE	"Ribosomal protein S5/Elongation factor G/III/V fa
AT1G36240	0	1,2	2,874	0,06219	FALSE	"Ribosomal protein L7Ae/L30e/S12e/Gadd45 family pr
AT1G67750	0	1,2	2,873	0,06219	FALSE	"Pectate lyase family protein; FUNCTIONS IN: pecta
AT2G44640	0	1,2	2,873	0,06219	FALSE	"FUNCTIONS IN: molecular_function unknown; INVOLVE
AT3G56800	0	1,2	2,868	0,06352	FALSE	"calmodulin 3 (CAM3); FUNCTIONS IN: calcium ion bi
AT1G10290	2,3	13,4	1,076	0,06533	FALSE	"dynamin-like protein 6 (ADL6); FUNCTIONS IN: GTPa

Table S9. Common protein set between UCN, DCP5, UPF1 partners and At-RBP. Comparison of proteins enriched in UCN1-IPs with UPF1-IPs, DCP5-IPs and *Arabidopsis* RNA Binding Proteins (At-RBP) from Reichel *et al.*, 2016.

Common partners between UPF1 and UCN

AT1G08370	"decapping 1 (DCP1); CONTAINS InterPro DOMAIN/s: D
AT1G43190	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT1G76010	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G15560	"Putative endonuclease or glycosyl hydrolase; BEST
AT2G34160	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G07010	"Pectin lyase-like superfamily protein; FUNCTIONS
AT3G13290	"varicose-related (VCR); FUNCTIONS IN: nucleotide
AT3G13300	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT3G22310	"putative mitochondrial RNA helicase 1 (PMH1); CON
AT3G22330	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT3G49490	"unknown protein; Has 722 Blast hits to 186 protei
AT3G58510	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G58570	"P-loop containing nucleoside triphosphate hydrola
AT4G27060	"TORTIFOLIA 1 (TOR1); CONTAINS InterPro DOMAIN/s:
AT5G47010	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in
AT5G62700	"tubulin beta chain 3 (TUB3); FUNCTIONS IN: struct

Common partners between DCP5 and UCN

AT1G08370	"decapping 1 (DCP1); CONTAINS InterPro DOMAIN/s: D
AT1G43190	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT1G76010	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G15560	"Putative endonuclease or glycosyl hydrolase; BEST
AT2G32080	"purin-rich alpha 1 (PUR ALPHA-1); FUNCTIONS IN: n
AT2G34160	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G07010	"Pectin lyase-like superfamily protein; FUNCTIONS
AT3G13290	"varicose-related (VCR); FUNCTIONS IN: nucleotide
AT3G13300	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT3G22310	"putative mitochondrial RNA helicase 1 (PMH1); CON
AT3G22330	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT3G49490	"unknown protein; Has 722 Blast hits to 186 protei
AT3G53420	"plasma membrane intrinsic protein 2A (PIP2A); FUN
AT3G58510	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G58570	"P-loop containing nucleoside triphosphate hydrola
AT3G62250	"ubiquitin 5 (UBQ5); FUNCTIONS IN: structural cons
AT4G24780	"Pectin lyase-like superfamily protein; FUNCTIONS
AT4G27060	"TORTIFOLIA 1 (TOR1); CONTAINS InterPro DOMAIN/s:
AT5G47010	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in

Common protein set between UCN-IPs and At-RBPP

At1g43190	"polypyrimidine tract-binding protein 3 (PTB3); FU
AT1G76010	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT2G32080	"purin-rich alpha 1 (PUR ALPHA-1); FUNCTIONS IN: n
AT2G34160	"Alba DNA/RNA-binding protein; FUNCTIONS IN: nucle
AT3G13300	"VARICOSE (VCS); FUNCTIONS IN: protein homodimeriz
AT3G22330	"putative mitochondrial RNA helicase 2 (PMH2); FUN
AT3G53420	"plasma membrane intrinsic protein 2A (PIP2A); FUN
AT3G58510	"DEA(D/H)-box RNA helicase family protein; FUNCTIO
AT3G58570	"P-loop containing nucleoside triphosphate hydrola
AT3G62250	"ubiquitin 5 (UBQ5); FUNCTIONS IN: structural cons
AT5G02500	"heat shock cognate protein 70-1 (HSC70-1); FUNCTI
AT5G47010	"LOW-LEVEL BETA-AMYLASE 1 (LBA1); FUNCTIONS IN: in

General Discussion

Nonsense Mediated Decay (NMD) is a well conserved eukaryotic mechanism, which recognizes and degrades mRNA showing premature translation termination, but also represents a potent mechanism in the regulation of gene expression (Lykke-Andersen and Jensen, 2015). In plants, NMD has been shown to have a role in the regulation of flowering time or in response to bacterial or virus infection (Nasim et al., 2017; Gloggnitzer et al., 2014; Garcia et al., 2014). The RNA helicase UPF1 is the major component of NMD, it recognizes NMD targets and induces the cascade of reaction leading to NMD target degradation. Degradation downstream of UPF1 action occurs by both exoribonucleolytic and endoribonucleolytic pathways in animals, but no experimental data link UPF1 action with decay factors in plants (Lejeune, 2003). In addition, UPF1 is also involved in other RNA decay pathways not linked to NMD in mammals, such as Regnase-Mediated Decay, or Staufen-Mediated Decay (Mino et al., 2015; Kim et al., 2005a). The existence of such UPF1-dependent but NMD-independent pathways in plants remains to be explored.

To gain insight on NMD in plants, the objective of this work was to identify UPF1 protein partners and study their function in NMD-dependent or -independent mechanisms in plants. Using an immuno-affinity approach coupled to mass spectrometry, we identified the first large set of proteins associated with UPF1 in plants. This study identified a wide number of RNA binding proteins involved in different processes, including RNA degradation, translation repression and RNA maturation (Chapter I, Figure 11D, Table 1). Among these factors we noticed proteins related to the decapping complex, DCP1, DCP5, VCS. Association of UPF1 with decapping factors was well described in animals, this work represents the first description of such a link in plant (Isken et al., 2008). Surprisingly, whereas the inactivation of DCP2 in mammals stabilizes NMD targets, in *Arabidopsis*, mutants in DCP1 and VCS, did not change their accumulation. In our efforts to identify RNA decay pathways acting downstream of NMD, we also tested mutants for factors of either the 5'-3' or the 3'-5' degradation pathway such as *xrn3-3*, *xrn4-3* and *ski2-3* but none of them showed enhanced NMD targets accumulation (Chapter I, Figure 14A, Figure S4). One interesting possibility is that in plants, the 5'-3' and the 3'-5' can cooperate to degrade NMD targets. This cooperation would lead to compensation mechanisms in simple mutants. To better understand this unexpected result, it would be great, in the future to test multiple mutants affecting both the 5'-3' and 3'-5' pathways. Such a mutant combination is only viable in the presence of a mutation in the RNA silencing factor RDR6, the evaluation of NMD target accumulation in this *xrn4 ski2 rdr6* triple mutant could reveal the redundancy between 5'-3' and 3'-5' pathways in NMD target degradation (Zhang et al., 2015).

The monitoring of NMD targets accumulation in 22 mutants affected in UPF1 partners identified DCP5, RH14 and eIF4G as putative NMD inhibiting factors as they showed systematic reduction of NMD targets

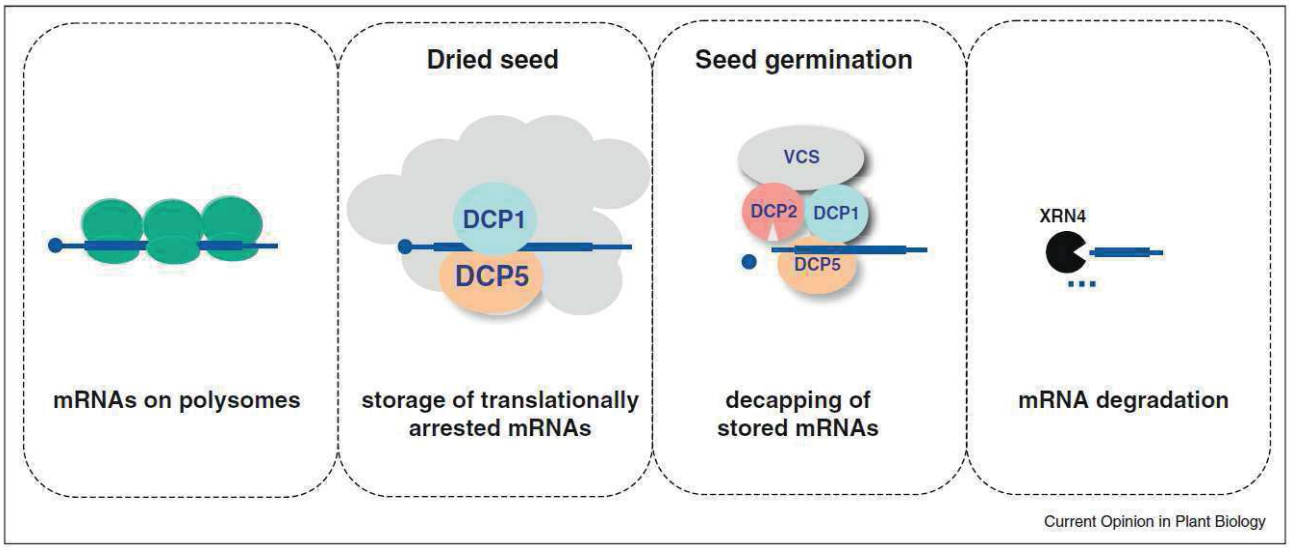


Figure 29. Model proposed by Xu et al., 2011 for DCP5 role in seed germination. See discussion for more information.

accumulation in three independent biological replicates. By studying *upf1 dcp5* double mutants we showed that the effect of *dcp5* is dependent on UPF1 action (Chapter I, Figure 14C). The comparison of our UPF1 interactome with factors enriched in DCP5 IPs revealed that UPF1 shares half of DCP5 partners, suggesting that they are part of the same mRNP. This mRNP is composed of many RNA helicases, including the three homologs of the DDX6 protein. These factors were of particular interest as they were among the most enriched UPF1 RNA-independent partners (Chapter I, Figure 12A). Additionally, we could demonstrate that both UPF1 and DCP5 directly interact with members of the DDX6 family of translation repressors (Chapter I, Figure 16).

We propose a model in which the association of UPF1 and DCP5 to DDX6 homologs leads to the translational arrest and storage of a fraction of the NMD targets in P-bodies (Chapter I, Figure 17). This process is perturbed in the *dcp5* mutant possibly due to P-bodies formation defects (Xu and Chua, 2009), leading to an increased translation. As NMD is a translation-dependent process, this leads to an accelerated decay of NMD targets. Our model integrates the function of DCP5 in translational repression as previously proposed in the context of seed germination (Figure 29) (Xu and Chua, 2009, 2011). DCP5, associated to DCP1, was suggested to recruit DCP2 and VCS to induce the decay of translationally repressed mRNAs (Xu and Chua, 2011); in our model, DCP5 would rather protect than induce the decay of targeted mRNAs. It would be interesting to test if DDX6 homologs are involved in the translation inhibition step observed during seed maturation. As the *rh6 rh8 rh12* triple mutant is lethal, because these genes probably have essential redundant functions (Cecile Bousquet-Antonelli, personal communication), RH6/8/12 RNAi lines or lines expressing a dominant negative version of one of these RNA helicases could represent interesting tools to investigate their function. In the course of this study, a predicted dominant negative version of RH12, mutated in its ATPase domain (RH12-DN), was produced. Transgenic lines expressing this RH12-DN could be useful to study the role of DDX6 homologues in NMD and seed development. Other candidates present in UPF1 IPs could also exert a protective action on some endogenous transcripts, this is the case of plants homologs to known NMD inhibiting factors, such as PUMILIO proteins or PTB3 (Ruiz-Echevarria and Peltz, 2000; Brazão et al., 2012).

In addition, the reduction of NMD targets accumulation in *dcp5-1* mutant was also observed for two other mutants, *rh14* and *eif4g* (Chapter I, Figure 14A). Translation inhibition mediated by Scd6, the DCP5 homolog in yeast, is performed upon the binding of Scd6 on eIF4G (Rajyaguru and Parker, 2012). To investigate the influence of eIF4G and RH14 on DCP5-mediated NMD targets protection, double mutants are currently being produced in the lab: *upf1-5 rh14* and *upf1-5 eif4g* as well as *dcp5-1 eif4g* and *dcp5-1*

rh14. If the down-regulation observed in *rh14* and *eif4g* is dependent on UPF1, as this is the case for dcp5-1, this would suggest a common mechanism between DCP5, RH14 and eIF4G. Moreover, the analysis of NMD target accumulation in *dcp5-1 eif4g* and *dcp5-1 rh14* will indicate if the down-regulations observed in *rh14* and *eif4g* are dependent on DCP5.

Among proteins co-purifing with both UPF1 and DCP5, we identified six novel P-body components: the UCN endonuclease, a TNTase, RH11 and the DDX6 homologues RH6, RH8 and RH12, which were validated by confocal microscopy (Chapter I, Figure 11D, 13A, Table 1). Recently, the purification and analysis of P-bodies by mass spectrometry and RNA sequencing identified a network of mRNA and regulatory proteins associated with P-bodies in human epithelial cells (Hubstenberger et al., 2017). This study identified a function of P-bodies in translation repression rather than RNA decay. Our data, and novel data that could easily be obtained from immunoprecipitations of other known core P-body factors, such as DCP1 and DDX6 homologs, could provide a great resource to identify more P-body components in plants. The study of the role of these new P-body factors may give a better understanding of P-body function. Finally, it would also be interesting to implement direct P-body purification techniques using approaches independent of IPs as it was done in mammals with the FAPS technique (Hubstenberger et al., 2017).

The detailed study of the novel P-body components TNTase and RH11 could identify new regulations or new decay pathways. TNTases are known to have a role in either the degradation or the stabilization of transcripts depending on the protein considered, the organism or cell types (Scheer et al., 2016). For example, CutA in *Aspergilus nidulans* was shown to participate in the degradation of NMD targets (Morozov et al., 2012). On the contrary, plant URT1 was proposed to restore a size of 16 nt for the binding of PABP and to favor 5'-3' degradation pathway (Zuber et al., 2016). The study of this TNTase could lead to the identification of a subset of NMD targets specifically regulated by terminal nucleotide addition or identify other more general functions. RNA helicases are involved in RNA metabolism, such as translation, RNA maturation or translation repression. RH11 is one of the three homologs to mammals DDX3X, which has been related to different functions, such as translation repression, nuclear export or translation of viral RNAs (Shih et al., 2012; Lai et al., 2013). The study of RH11 could allow the identification of the role of the homologs of DDX3X in plant, with possible implications in the previously mentioned processes.

The NYN domain UCN endonuclease is another P-body component identified in this study. The closest mammalian homolog of UCN is MARF1, known to be involved in oogenesis and neuronal differentiation (Su et al., 2012; Kanemitsu et al., 2017). Interestingly, MARF1 regulates specific transcripts, has been shown to accumulate in P-bodies and to directly interact with Ge-1, the decapping stimulator homolog to

VCS in plants (Bloch et al., 2014). Our work on the UCN interactome identified the decapping factors DCP1 and VCS as UCN partners (Chapter II, Figure 20B). Moreover, yeast two-hybrid experiments identified DCP1 as a direct interactor of UCN in plants (Chapter II, Figure 20D). These results suggest that, as it was described in mammals for MARF1, UCN might identify a novel RNA degradation pathway linked to the decapping machinery. The preliminary characterization of UCN showed that ectopic expression of this endonuclease seems to preferentially influence the accumulation of mRNAs with longer 3'-UTR (Chapter II, Figure 26D, 26E). In addition, the transgenic expression of UCN perturbed ethylene signaling and the accumulation of two distinct RNA viruses (Chapter II, Figure 27). An interesting possibility is that UCN may recruit the decapping complex to stimulate the degradation of its endogenous targets (Chapter II, Figures 20B, 20D, 28). One major challenge will be now to identify targets regulated by UCN. For this purpose, both RNA-immunoprecipitation analysis and transcriptome analysis of *ucn* mutant and UCN-expressing lines are planned. In these studies, the correlation between the deregulations observed in UCN mutants and ectopic expression lines will be used to determine the cellular functions controlled by UCN. As UCN ectopic expression led to phenotypes related to hormone signaling and defense responses, we anticipate that the study of UCN endogenous targets will likely identify processes controlling traits of agronomical interest.

To conclude, eukaryotic mRNAs are in a constant dynamic equilibrium between different subcellular compartments: being translated on polysomes, translationally repressed in P-bodies, stalled in translation in stress granules, or degraded. All compartments constantly exchange mRNAs and associated factors in order to regulate specific pathways and face fast changes in the environment. Our proposed model including DCP5 and DDX6 in NMD targets regulation is in line with this dynamic view of mRNA fate. Our work identified a new NMD inhibiting pathway in plants, we can postulate that other factors are probably limiting the impact of the NMD pathway on many endogenous mRNAs. We propose that this pathway involving the action of DCP5 and DDX6 homologs influence the fate of NMD targets, likely through their role in translation repression in P-bodies. We identified new P-body components, by the immunopurification of several P-body proteins. The detailed study of the identified new P-bodies factors, including the UCN endonuclease, will likely lead to the identification of new pathways and factors involved in mRNA decay or translation repression.

Material and methods

I. Material

1. Plant material

1.1. *Arabidopsis thaliana*

1.1.1. Ecotype

All *Arabidopsis thaliana* plants used in this study are Columbia ecotype (Col-0).

1.1.2. Standard growing conditions

Arabidopsis thaliana plants used in this study were grown during three weeks in 12 h light / 12 h dark under LED illumination before being transfered to 16 h light / 8 h dark until seed harvesting.

1.1.3. T-DNA insertion mutants

T-DNA (Transfert DNA) insertion mutants were ordered to Salk Institute Genomic Analysis Laboratory, Syngenta Arabidopsis Insertion Library and GABI-Kat (Kleinboelting et al., 2012) from the Nottingham Arabidopsis Stock Center (NASC). The mutant lines used in this study are listed in Table S10.

1.2. *Nicotiana benthamiana*

Nicotiana benthamiana plants used in this study were grown 4 weeks in 16 h light / 8 h dark before being used for transient expression studies.

2. Bacterial strains

2.1. *Escherichia coli* DH5α

Cloning was achieved in the *Escherichia coli* DH5α™ (Invitrogen) bacterial strain to allows plasmid amplification. The DH5α™ strain possess the following genotype : F– Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK–, mK+) phoA supE44 λ– thi-1 gyrA96 relA1.

2.2. *Agrobacterium tumefaciens* GV3101

The *Agrobacterium tumefaciens* strain used for transformation is the GV3101 strain whose chromosome contains a spontaneous mutation conferring resistance to Rifampicine. This bacterial strain also carries the Ti-plasmid (Tumor inducing plasmid), which encodes for genes allowing the excision of the T-DNA followed by its insertion in the genome of the transformed plant. Both stable *Arabidopsis thaliana* transformation by the floral dip technique (see part 6.2 for more details) and transient transformation of *N. benthamiana* leaves were performed with this strain (see part 6.1 for more details).

2.3. Vectors used in this study

2.3.1. pCTL235

A sequence comprising UPF1 genomic sequence and 1,6kb of upstream regulatory sequences was cloned in the pCTL235-RFP, pCTL235-GFP and pCTL235-FlagHA vectors for C-terminal fusions of UPF1 with RFP, GFP or a double FlagHA tag respectively. Those vectors carry a kanamycin resistance gene for bacteria selection and the *hph* genes allowing a resistance to hygromycin for plants selection.

2.3.2. pDONR™221

The pDONR™221 vector was used as a donor vector for the cloning using the Gateway™ technique. This plasmid carries the *ccdB* gene between the recombination sites attP1 and attP2. This gene encodes for the CcdB protein, which is toxic for the cell by inhibiting the DNA gyrase. In this system, the *ccdB* gene is used as a positive selection marker: the insertion by recombination of a sequence of interest between the attP sites allows its replacement, enabling the survival of the recombinant bacteria.

The pDONR™221 vector carries the kanamycin resistance gene (*KanR*) allowing its selection. Kanamycin is an aminoglycoside molecule that binds to the bacterial ribosome 30 S subunit causing mistranslation. The *kanR* gene present in the pDONR™221 vector encodes for the aminoglycoside 3'-phosphotransférase protein able to inactivate the kanamycin molecule by phosphorylation. The specifications of this plasmid are listed in the Table S11.

2.3.3. pB7WGF2 and pH7WGR2 for fusions with GFP and RFP

pB7WGF2 and pH7WGR2 binary vectors were used in this study in order to express the proteins of interest under a constitutive promoter – the 35S promoter from the CaMV (Cauliflower Mosaic Virus) – and to produce fusions of the proteins of interest with either GFP (Green Fluorescent Protein) or RFP (Red Fluorescent Protein) at the amino-terminal end. Those vectors carry the adenylyltransferase (*aadA*) gene conferring a resistance to spectinomycin for bacterial selection. Spectinomycin is an aminoglycoside able to bind the bacterial 16 S rRNA of the 30 S ribosome subunit, causing a translocation arrest. As for the pDONR221 these plasmids are based on the Gateway™ system and contain the *ccdB* gene between the attR1 and attR2 sites. The replacement of the *ccdB* gene with the sequence of interest produces in frame fusions with GFP or RFP.

Finally, those vectors carry either the *bar* or the *hph* genes allowing a resistance to Glufosinate or hygromycin respectively enabling transgenic plant selection. The specifications of this plasmid are listed in the Table S11.

2.3.4. pGWB618 for fusions with 4xMyc

The Gateway™ binary destination vector pGWB617 allows the expression of proteins of interest carrying a quadruple Myc tag located at their amino-terminal end. This vector carries spectinomycin resistance gene (*aad1*) for bacterial selection. This vector carries the *bar* gene for selection of stable transformants using Glufosinate. The specifications of this plasmid are listed in the Table S11.

2.3.5. pGADT7 and pGBT9 for yeast two-hybrid experiments

pGADT7 and pGBT9 were used as destination vectors using the Gateway™ system for yeast two-hybrid experiments. They contain the sequence of the GAL4 Activation and Binding Domain respectively. Both confer resistance to ampicillin in bacteria. Yeasts containing the pGADT7 vector are selected on SD -Leu medium while yeast containing the pGBT9 vector are selected on SD-Trp medium.

II. Methods

1. Techniques related to deoxyribonucleic acid

1.1. Total extraction of genomic DNA

One small *A. thaliana* leaf is grinded in an Eppendorf tube containing a metal bead and 400 µL of Edward's buffer (Edwards, 1991 ; 200 mM Tris-HCl pH7.5 ; 250 mM NaCl ; 25 mM EDTA ; 0.5% SDS) by the Tissue Lyser II (Qiagen) device during two times 1 min at 30 movements per min. The mixture is then centrifuged at 18 000 g for 1 min to get rid of the cell debris. Three hundred µL of the supernatant are then collected and added to 300 µL of isopropanol to precipitate DNA. After several tube inversions to homogenize the solution, the preparation is centrifuged for 5 min at 18 000 g. The supernatant is discarded and the tubes are dried upside down for 15 min before resuspending the DNA pellet into 100 µL of water.

One µL of this DNA extraction is then used for amplification by PCR (Polymerase Chain Reaction).

1.2. Genomic DNA amplification by PCR

The Polymerase Chain Reaction (PCR) was performed by the succession of 25 to 35 cycles of 3 steps: the denaturing step of the DNA strands (94°C); the annealing step (from 50 to 60°C depending on the primers chosen); the elongation step by a DNA-dependent DNA polymerase (72°C).

The DNA polymerase GoTaq® (Promega) was used to perform PCR for genotyping whereas we used the proofreading capacities of the DNA polymerase Phusion® (NEB) for the accurate amplification of DNA fragment for cloning experiments.

1.3. Analysis of PCR products on agarose gel electrophoresis

In order to control the amplification efficiency and the size of the amplified products after PCR, 10 µL of PCR product are deposited on a 1% (w/v) agarose gel to separate the amplified fragments by electrophoresis in a TBE buffer (45 mM Tris-Borate ; 1 mM EDTA). This electrophoresis allows to separate the amplified fragments according to their size: negatively charged DNA fragments migrate towards the positive electrodes, smaller fragments migrate faster than larger fragments. A size marker (MassRuler DNA Ladder mix®, Promega) was used to evaluate the size of the amplified fragments.

2. Cloning experiments

2.1. Amplification of gene of interest for cloning

In order to clone different sequences of interest into expression vectors, the genomic or complementary DNA (cDNA) sequence is first amplified by the DNA polymerase Phusion® (NEB) High-Fidelity PCR Master Mix.

The amplified product is then loaded on a 1% (w/v) agarose gel and products are separated by electrophoresis. The band corresponding to the right amplification size are excised from the gel and purified using the Nucleospin® Gel and PCR Clean up kit (Macherey-Nagel): DNA is retained on a silica column then purified with an elution buffer after a washing step.

For Gateway™ cloning, the attB1 and attB2 sequences were added to the amplification primers.

2.2. Mutagenesis of the UCN gene by Overlap Extension PCR

The Overlap Extension PCR technique was used to introduce point mutations in the catalytic domain of the UCN gene. The gene is amplified in two parts – named PCR1 and PCR2 – using overlapping specific primers containing the desired point mutation. PCR products are then deposited on an 1% agarose gel (w/v) and the band corresponding to the right amplification size are excised from the gel and purified using the Nucleospin® Gel and PCR Clean up kit (Macherey-Nagel). After OD measurement with the NanoDrop (Thermo Scientific), 50ng of the obtained overlapping PCR products PCR1 and PCR2 are annealed and extended in presence of 10 µL of Phusion High-Fidelity PCR Master Mix with the following PCR program:

98°C	30 sec	
98°C	10 sec	15x
55°C	30 sec	
72°C	2 min	
72°C	10 min	

The annealed and extended product is named PCR3 and it follows an amplification of the entire fragment of interest using 0.5 µM of the primers at the extremities of the fragment to amplify; 2 µL of PCR3 and 25 µL of Phusion Master Mix with the following PCR program:

98°C	30 sec	
98°C	10 sec	25x
60°C	30 sec	
72°C	4 min	
72°C	10 min	

The extreme primers designed contain the attB1 and attB2 recombination sites of the λ phage allowing the direct use of this fragment for cloning using the Gateway™ technique (for more details, see part 2.3).

Point mutations in the UCN gene created the substitution of an acid aspartic (D) residue into an asparagine (N) residue either at the position 55 or 153 of the UCN protein, those mutant versions of UCN were thus named UCN-D55N and UCN-D153N respectively. Those acid aspartic residues are part of four conserved residues required for UCN activity and ion binding (Anantharaman, 2006).

2.3. Cloning using the Gateway™ method

The Gateway™ cloning (Invitrogen™) system, based on the phage λ recombination system, was used to clone genes or parts of genes to allow their expression in plant. The advantage of the Gateway™ technology consists in a quicker cloning compared with conventional cloning methods; a lower proportion of false positives; and the possibility to use a unique Entry vector to recombine the fragment of interest in a wide number of Destination vectors.

The Gateway™ cloning allows in a first step to create an Entry vector from a Donnor vector (pDONR) containing the attP1 and attP2 recombination sites and the sequence of interest amplified by PCR flanked by the attB1 and attB2 sequences. This recombination reaction is catalyzed by the BP clonase enzyme. In a second step, the LR clonase enzyme catalyzes the recombination reaction between the attL and attR sites present in the Entry vector (pENTRY) previously produced and the desired Destination vector (pDEST) respectively allowing the creation of an Expression vector (pEXPR) containing the sequence of interest.

The BP reaction is executed as follows in a total volume of 10 µL: 150 ng of pDONR of choice; 150 ng of purified PCR product and 1 µL of BP clonase are incubated during 1 h at 25°C. One µL of proteinase K is

then added and the mix is incubated during 10 min at 37°C to stop the reaction. The LR reaction is executed as follows in a total volume of 10 µL: 150 ng of pENTRY of choice; 150 ng of pDEST of choice and 1 µL of LR clonase are incubated during 1 h at 25°C. One µL of proteinase K is then added and the mix is incubated during 10 min at 37°C to stop the reaction.

Following each recombination reaction (either BP or LR), a bacterial transformation is performed is an *Escherichia coli* DH5α strain in order to select clones that successfully integrated the recombined vector to allow its amplification in bacteria.

2.4. Heat shock transformation of competent *Escherichia coli* DH5α

An aliquot of 100 µL of competent *Escherichia coli* DH5α is transformed with the 10 µL from a BP or LR reaction. This mix is incubated 50 sec in a water bath warmed at 42°C and then put on ice during 10 min. One mL of LB (Lysogeny Broth: bacto-tryptone 1% ; yeast extracts 0.5% ; NaCl 0.5%) is added to the bacteria and incubated during 1 h at 37°C upon slow agitation to ensure bacterial growth. Bacteria are then plated on solid LB media (bacto-tryptone 1% ; yeast extract 0.5% ; NaCl 0.5% ; agar 1%) supplemented with the proper antibiotic allowing selection pressure on the vector of interest.

The next day, independent colonies are picked up, diluted in water and a PCR amplification is performed to check for the presence of the vector of interest using insert-specific primers. The PCR products are separated on a 1% agarose gel. The positive clones are used to inoculate a culture of 3 mL to grow over night at 37°C. This culture is then used to purify plasmid DNA with the NucleoSpin® Plasmid QuickPure kit (Macherey-Nagel).

2.5. Sequencing of positive clones

After the validation of positive clones by PCR amplification, vector of interest is sequenced to verify the sequence of the DNA fragment inserted during the BP reaction. This sequencing is performed by Malek Alioua at the IBMP sequencing platform according to Sanger's method (Sanger, 1977) using the ABI PRISM 3100 Genetic Analyzer capillary electrophoresis device.

3. Yeast two-hybrid

3.1. Yeast co-transformation

A pre-culture of 50 mL of YPD medium (Clontech) is inoculated with AH109 yeast strain grown on YPD agar (Clontech) plate for up to 7 days. This solution is incubated under shaking overnight at 30°C.

Three hundred mL of YPD medium (Clontech) is inoculated with the pre-culture to reach a OD_{600} of approximately 0.3 and is then incubated for up to 4 h at 30°C under shaking. If the OD_{600} has reach 0.5, the culture is centrifuged at 1 000 g for 5 min. The yeast pellet is resuspended in 1.5 mL of sterile TE/LiAc (1 mM Tris-HCl ; 0.1 mM EDTA pH8.0 ; 0.1 M LiAc). 0.1 µg of GAL4 Activation Domain-plasmid, 0.1 µg of GAL4 Binding Domain-plasmid and 0.1 mg of salmon sperm are mixed with 100 µL competent yeast cells and 600 µL sterile TE/LiAc/PEG (1 mM Tris-HCl ; 0.1 mM EDTA pH8.0 ; 0.1 M LiAc; 50% (w/v) PEG MW[3500]). After 30 min incubation at 30°C under shaking, 70 µL DMSO are added, yeast cells are then heat shocked for 15 min at 42°C and chilled on ice. Cell suspensions are centrifuged at 10 000 g for 30 sec and pellets are resuspended in 500 µL sterile TE buffer (1 mM Tris-HCl; 0.1 mM EDTA pH8.0). One hundred of this suspension is plated on each selective medium plate (Clontech) : SD -Leu -Trp, SD -Leu -Trp -His, SD -Leu -Trp -His -Ade . SD plates are incubated at 30°C for up to 7 days.

3.2. Droplet plates

Two mL of SD -Leu -Trp (Clontech) is inoculated with grown colonies and incubated at 30°C under shaking overnight. Yeast culture is then centrifuged at 2 000 g for 2 min, and pellet is resuspended in 1 mL sterile water. OD_{600} is measured and adjusted to 1.5 with sterile water. Ten µL of 1 ; 10^{-1} ; 10^{-2} and 10^{-3} dilutions are deposited on SD -Leu -Trp, SD -Leu -Trp -His, SD -Leu -Trp -His -Ade agar plates (Clontech). Plates are incubated at 30°C and photographs are taken 3 to 10 days after droplet deposition. In the case of autoactivation of an AD-plasmid or a BD-plasmid, the 3-amino-1,2,4-triazole (3-AT ; Sigma) competitive inhibitor of yeast *HIS3* protein was used. If used, 3-AT concentrations are indicated in the figures and/or in figure legends.

4. Methods related to ribonucleic acids

4.1. Total RNA extraction

Frozen plant material, either *Arabidopsis thaliana* 15 days old plantlets or a *Nicotiana benthamiana* leaf disc is supplemented with a volume of 100 µL of 4 mm diameter glass beads (Carl Roth) and is then grinded with the Silamat® S6 (Ivoclar Vivadent) during 8 sec with a mean speed of 4 500 movement per min. Five hundred µL of Tri-Reagent® (Sigma-Aldrich) is added to the powder produced and the sample is again grinded with the Silamat® during 16 sec. It follows a 4°C centrifugation at 18 000 g speed allowing the separation between the cellular debris and the supernatant. The aqueous phase is transferred into a new tube and is supplemented with 100 µL of chloroform. The sample is homogenized by vortex and then centrifuged for 10 min at 13 500 g. This step allows the sample to be divided in three parts: the upper

aqueous phase that contains the RNA, the intermediate phase that contains DNA and the lower phenolic phase, which contains proteins. The upper aqueous phase in collected in a new tube and 100% isopropanol at a 1:1 ratio is added to precipitate RNA, the sample is homogenized by multiple tube inversions and then centrifuged for 15 min at 18 000 g to pellet RNA. After a washing step with 70% (v/v) ethanol, RNAs are resuspended in ultra-pure water (Milli-Q® by Millipore filtration system) for a use in Reverse Transcription (RT) or in 50% formamide for a use in northern blotting. RNAs are then quantified by spectrophotometric measurement with the NanoDrop (Thermo Scientific), the presence of a contaminant is controlled by looking at the ratios 260/280 nm and 260/230 nm. Finally, we ensure the quality of extracted RNA by loading 600 ng of extracted RNA additionned with RNA loading buffer 2 X (Thermo Scientific – 95% formamide; 0.025% SDS ; 0.025% Bromophenol Blue ; 0.025% xylene cyanol ; 0.025% ethidium bromide) on a 1% (w/v) agarose gel after 5 min denaturation of the samples at 65°C.

4.2. Total RNAs high molecular weight northern blot

4.2.1. RNA migration on formaldehyde denaturing gel

RNAs are extracted with Tri-Reagent® (Sigma-Aldrich) and resuspended in formamide in a small volume to concentrate RNA. One to five µg of RNA in 10 µL of 50% formamide are supplemented with 30 µL of RNA loading buffer (1X HEPES ; 50% formamide ; 5% formaldehyde ; 0.05% Bromophenol Blue ; 0.025% ethidium bromide).

Samples are denatured at 65°C during 10 min, deposited on a 1% (w/v) agarose; 1X HEPES; 6% formaldehyde gel and migrated in a 1X HEPES buffer.

4.2.2. Capillary transfer and UV cross-linking

The migrated samples are transferred from the gel to a HybondNX (Amersham) membrane by capillary force overnight in a 20X SSC buffer (3 M NaCl ; 0.3 M Sodium citrate). Cross-linking of RNA onto the membrane is performed by two successive 30 sec UV exposure at 120 mJ.

4.2.3. Random-priming radioactive probe labelling

A fragment of about 400 bp of the sequence of interest is amplified by PCR and purified with the Nucleospin® Gel and PCR Clean up kit. One hundred ng of this DNA template is then used for labelling using the random-priming method with the DECAprime™ II DNA Labeling kit (Ambion®). This method allows homogenous labeling of DNA template as the Klenow polymerase incorporates gamma-phosphate radioactively labelled dCTP at multiple positions upon the binding of random decamers.

4.2.4. Hybridization

Prior to hybridization with a radioactive probe, the membrane is first incubated in the hybridization buffer (PerfectHyb™) during 30 min 65°C. The radioactive probe is denatured at 95°C for 5 min and added to the hybridization buffer; hybridization is performed overnight at 65°C.

Three washes of 20 min are performed at 65°C with a 0.1% (w/v) SDS ; 2X SSC buffer (0,3 M NaCl ; 0.03 M Sodium citrate) for high molecular Northern blot. The membrane is then dried and the signal is revealed using the PhosphorImager Typhoon FLA 7000.

4.3. cDNA synthesis

A cDNA synthesis is performed after total RNA extraction for use in quantitative PCR (qPCR).

4.3.1. DNAse treatment

A DNAse treatment is first performed to remove contaminations with DNA molecules in total RNAs. In this step, 1 µg of total RNA is incubated during 30 min at 37°C with 2 U of DNAse RQ1 (RNA Qualified 1) without RNase (Promega®) ; 1 µL 10X DNAse RQ1 buffer (400 mM Tris-HCl pH 8.0 ; 100 mM $MgSO_4$; 10 mM $CaCl_2$) in a total volume of 10 µL ; 4 U recombinant ribonuclease inhibitor RNaseOUT™ (Invitrogen™). Inactivation of the DNAse enzyme is performed by the addition of 2 µL of Stop solution (20 mM EGTA pH 8.0) and incubation at 65°C for 10 min. After DNAse treatment, 4 µL are kept to use as a negative control of reverse transcription (RT-).

4.3.2. Reverse transcription

Complementary first strand DNA (cDNA) synthesis is performed by a Reverse Transcriptase (RT - DNA polymerase RNA-dependent). The RT used is the SuperScript® IV (Invitrogen™), it catalyzes the polymerization reaction forming phosphodiester bonds between free nucleotides and the newly synthetized strand by base complementarity with the template strand. The SuperScript® IV (Invitrogen™) possesses a better processivity than its predecessors while being more stable to high temperatures thus allowing faster reactions.

The 8 µL of DNAse-treated total RNAs are supplemented with 0.5 mM of each dinucleotide triphosphate (dNTP – dATP, dCTP, dTTP, dGTP) ; 2,5 µM d'oligo-$d(T)_{17}$; 0.4 µL of hexadeoxynucleotides of random sequence (Promega) in a total reaction volume of 13 µL. This mix is incubated 5 min at 65°C, then 1 µL of 100 mM DTT ; 40 U of nuclease inhibitor RNaseOUT™ (Invitrogen™) ; 200 U of SuperScript® IV and 4 µL of 5X SuperScript® IV buffer are added for a total reaction volume of 20 µL. This solution is then incubated 5

min at 25°C and then 20 min at 50°C before inactivation step of the enzyme with an incubation of 15 min at 70°C.

One µL of cDNA or 1 µL of total RNA after DNAse treatment (RT-) are then used for amplification by PCR to ensure the correct cDNA synthesis using primers hybridizing on the *ACTIN2* gene reference (see list of oligonucleotides used in this study in Table S12). The *ACTIN2* amplification on cDNA is validated upon the migration of the PCR amplification on 1% (w/v) agarose gel.

One µL of cDNA is used for each quantitative PCR (qPCR) replicate.

4.3.3. Quantitative PCR

The quantitative PCR allows the precise measurement of the accumulation of specific cDNAs in different samples compared to the control. The amplification at each cycle is measured by the fluorescent quantification of the SYBR Green dsDNA binding dye. The SYBR Green is used in this experiment as a non-specific fluorophore DNA marker. A Cycle Threshold (Ct) is determined for each sample according to the number of amplification cycles that were necessary to reach an established fluorescent threshold.

The reaction is performed in 384-well optical reaction plates in a total volume of 10 µL composed of 1X LightCycler® 480 SYBR Green I Master (Roche) buffer, 0.25 µM of each sense and antisense primer and 1 µL of cDNA. The amplification reaction is performed by the LightCycler 480 II (Roche) device with a first denaturation step of 10 min at 95°C followed by 45 cycles alternating annealing steps of 20 sec at 60°C and elongation steps for 40 sec at 72°C. Each assay is performed in triplicate on the same plate. Results are normalized to the *ACTIN2* reference gene.

5. Methods related to protein analysis

5.1. Total protein extraction – rapid method

One *Nicotiana benthamiana* leaf disc supplemented with a volume of 100 µL of 4 mm diameter glass beads (Carl Roth) is grinded with the Silamat® S6 (Ivoclar Vivadent) during 8 sec with a mean speed of 4500 movement per min. The powder is then supplemented with 100 µL of SDS-urea buffer (8 mM Tris-HCl pH 6.8 ; 4 M urea ; 2% SDS ; 10% glycerol ; 100 mM DTT) and grinded again for 8 sec with the Silamat® S6. A centrifugation of 10 min at 13 500 g is performed to remove the cellular debris. The supernatant is transferred to a new tube and 4X Laemmli loading buffer is added (0.2 M Tris pH 6.8 ; 2% (w/v) SDS ; 40% glycerol (v/v) ; 20% (v/v) β-mercaptoethanol ; 0.02% (w/v) Bromophenol blue).

5.2. Total proteins extraction – Tri-Reagent® protein extraction

Frozen plant material, either *Arabidopsis thaliana* 15 days old plantlets or a *N. benthamiana* leaf disc, is grinded with 100 µL of 4 mm diameter glass beads (Carl Roth) with the Silamat® S6 (Ivoclar Vivadent) during 8 sec with a mean speed of 4500 movement per min. Three hundred µL of Tri-Reagent® (Sigma-Aldrich) is added to the powder produced and the sample is again grinded with the Silamat® during 16 sec. The sample is centrifuged at 4°C and 18 000 g allowing the separation between the cellular debris and the supernatant. The supernatant is transferred into a new tube with 60 µL of chloroform, homogenized by vortex and centrifuged for 10 min at 13 500 g. After centrifugation, the lower phenolic phase corresponds to the protein fraction. The supernatant and the intermediate phase are discarded, 100 µL of 100% (v/v) ethanol is added to the phenol phase and the solution is homogenized by vortex. Samples are centrifuged during 10 min and the supernatant, which contains proteins is transferred to another tube and with 3 volumes of 100% (v/v) acetone. Samples are incubated 15 min on ice to precipitate proteins and centrifuged 1 min at 5 000 g. After a washing step with 80% (v/v) acetone, the protein pellet is resuspended in 100 µL SDS-urea buffer (8 mM Tris-HCl pH 6.8 ; 4 M urea ; 2% SDS ; 10% glycerol ; 100 mM DTT) and 4X Laemmli loading buffer is added (0.2 M Tris pH 6.8 ; 2% (w/v) SDS ; 40% glycerol (v/v) ; 20% (v/v) β-mercaptoethanol ; 0.02% (w/v) Bromophenol blue).

5.3. Western blot

5.3.1. Protein separation on a denaturing gel

After total proteins extraction, proteins are separated according to their molecular weight by electrophoresis on a SDS-PAGE denaturing gel following the Laemmli method (Laemmli, 1970). Proteins first migrate through a stacking gel (0.6 M Tris pH6.8 ; 0.5 % (w/v) SDS ; 4% (w/v) 37,5/1 acrylamide/bisacrylamide), creating a thin migration front. Proteins migrate then in a resolving gel composed of acrylamide concentration ranging from 8 to 15 % (w/v) (1.1 M Tris pH8.8 ; 0.3 % (w/v) SDS) in which proteins migrate according to their molecular weight, smaller proteins migrating more rapidly than wider proteins.

Proteins are first denatured for 5 min at 95°C before being deposited on SDS-PAGE. The migration is done in a Mini-PROTEAN® Tetra Cell (Bio-Rad) device at 100 V during 1 h 30 min in a Tris-Glycine-SDS buffer composed of 25 mM Tris ; 192 mM Glycine ; 0.1% SDS buffer (w/v), pH 8.6.

5.3.2. Transfer on a PVDF membrane

After SDS-PAGE migration, proteins present on the gel are transferred on a PVDF (Immobilon-P – Merck Millipore) membrane using the Mini Trans-blot® Cell (Bio-Rad) system after activating the membrane by incubation during 5 min in 100% (v/v) ethanol. The transfer is performed at 4°C during 1 h 30 min at 80 V in a Tris-Glycine-SDS buffer composed of 25 mM Tris ; 192 mM Glycine ; 40 % (v/v) ethanol.

5.3.3. Immunodetection

After the transfer, non-specific sites of the membrane are saturated during 30 min in 5% (w/v) milk diluted in 50 mM Tris-HCl pH 7.5 ; 150 mM NaCl ; 0.1% (v/v) Tween-20. The membrane is then incubated over night with primary antibodies directed against the protein of interest or the tag of interest diluted in 5% milk (w/v); 50 mM Tris-HCl pH 7.5 ; 150 mM NaCl ; 0.1% (v/v) Tween-20. The different antibodies used in this study and their specifications are indicated in the Table S13. The next day, 3 washes of 15 min are performed with 50 mM Tris-HCl pH 7.5 ; 150 mM NaCl ; 0.1% (v/v) Tween-20 buffer to remove unbound antibodies. The membrane is then incubated during 1 h at room temperature with secondary antibodies coupled to HRP (Horseradish Peroxidase) diluted at 1/15 000 in 5% (w/v) milk ; 50 mM Tris-HCl pH 7.5 ; 150 mM NaCl ; 0.1% (v/v) Tween-20. The detection of the protein or fusion of interest is performed by chemiluminescense with the Lumi-Light Western Blotting Substrate (Roche®).

5.3.4. Membrane coloration with Coomassie blue

After immunodetection, the membrane is colored with Coomassie blue solution (40% (v/v) ethanol ; 10% (v/v) acetic acid ; 0.02% (w/v) Brillant Blue R). The membrane is then destained with a 40% (v/v) ethanol; 10% (v/v) acetic acid buffer leading to the visualization of the more abundant proteins, including Rubisco (ribulose-1,5-bisphosphate), used as a loading control.

5.4. Immunopurification and protein detection by mass spectrometry

5.4.1. Immunoprecipitation

Three hundred mg of frozen flowers are grinded in a mortar with 1.5 mL of lysis buffer (50 mM Tris-HCl pH 8.0 ; 150 mM NaCl ; 1% Triton supplemented with protease inhibitor cocktail (Roche)) until the lysate appears homogenous. This solution is then centrifuged two times 10 min at 13 500 g at 4°C to remove the cellular debris. A volume of 60 µL corresponding to the Input is collected and 4X Laemmli loading buffer is added (0.2 M Tris pH 6.8 ; 2% (w/v) SDS ; 40% glycerol (v/v) ; 20% (v/v) β-mercaptoethanol ; 0.02% (w/v) Bromophenol blue). A volume of 50µL of µMACS magnetic beads (Miletenyi Biotec), coupled to antibodies directed against the tag of interest are added to the cellular lysate and incubated on a wheel during 30

min at 4°C. For immunoprecipitations with a RNase A treatment, RNaseA/RNase T1 mix (ThermoFisher) was used at a final concentration of 200 µg/mL, supplemented to the lysate and incubated 1 h before the addition of beads; controls are also incubated without any supplemental addition. A magnetic M column (Miltenyi Biotec) is placed on the magnetic stand and prepared with 200 µL of lysis buffer, then the lysate is deposited onto the column 200 µL at a time until no lysate remains. During lysate elution, a volume of 60 µL corresponding to the flow-through is collected and 4X Laemmli loading buffer is added (0.2 M Tris pH 6.8 ; 2% (w/v) SDS ; 40% glycerol (v/v) ; 20% (v/v) β-mercaptoethanol ; 0.02% (w/v) Bromophenol blue). Four washes are performed on the column with wash buffer (20 mM Tris-HCl pH 8.0 ; 0.1% Triton). The elution step is performed with 1X Laemmli buffer pre-warmed at 95°C during 5 min.

For robust mass spectrometry detection of protein partners, three replicates are provided for each sample.

5.4.2. Mass spectrometry

The mass spectrometry analysis of proteins present in IP fractions was performed at the Plateforme Protéomique Strasbourg – Esplanade located at the Institut de Biologie Moléculaire et Cellulaire, Strasbourg.

5.4.3. Statistical analysis of partners with the R software

The identification of significantly enriched proteins was performed by a statistical analysis based on spectral counts using the msmsTest R package (Gregori et al., 2013; Robinson et al., 2010). p-value were then adjusted using the Benjamini & Hochberg method. Proteins that were over-represented in UPF1 IP were visualized as a volcano plot that displays log2(fold-change) and -log10(adjusted-p.value) on the x and y axes, respectively.

6. Methods related to plant manipulation

6.1. *Nicotiana benthamiana* agroinfiltration

Transient expression of proteins of interest from binary expression vector was performed with the GV3101 *Agrobacterium tumefaciens* strain, by infiltration of *N. benthamiana* leaves with *Agrobacterium* suspension.

For the bacterial culture, 50 µL of an *Agrobacterium* glycerol stock corresponding to bacterial culture in 40% (v/v) glycerol at a 1:1 ratio are inoculated to 10 mL of LB (Lysogeny Broth – 1% (w/v) bacto-tryptone ; 0.5% (w/v) yeast extract ; 0.5% (w/v) NaCl). The culture media contains the rifampicine antibiotic (30

µg/mL) to select the GV3101 strain and a specific antibiotic to select bacteria containing the binary vector. The culture is incubated over night at 28°C.

After 24 hours, the culture is centrifuged 10 min at 4 000 g and then resuspended in 10mM $MgCl_2$; 100 µM acetosyringone. The absorbance at 600 nm of the solution is measured and the volume of the solution is adjusted with 10mM $MgCl_2$; 100 µM acetosyringone buffer to adjust the absorbance at 600 nm to 1. The different bacterial cultures containing the different sequences of interest to transiently co-express are mixed to a 1:1 ratio. A volume of culture containing *Agrobacterium* expressing the P14 suppressor of silencing is systematically added to each infiltration mix in order to reduce the silencing of the expressed transgenes. Constructs used for agroinfiltration experiments are either produced in the laboratory or provided by Daniel Silhavy (Kerényi, 2008). This preparation is syringe-infiltrated into the abaxial side of the leaves of 4 weeks *N. benthamiana* plants lightly wounded by a needle. The abaxial face of the leaves is observed and photographed 3 to 5 days after infiltration (DAI) under UV light. For imaging of fluorescent fusion proteins, a disc of infiltrated leaf is directly observed by confocal microscopy 3 to 6 days after infiltration (DAI).

6.2. Virus sap infection

Crude virion sap extract is performed by grinding 0.2 g of *N. benthamiana* leaves transiently expressing the PVX-GFP, TCV or TuMV-GFP constructs in KH_2PO_4 0.1 M. Forty µL of crude virion sap extract was used for infection of each *N. benthamiana* leaf transiently expressing a GFP-UCN version. Sap-infected leaves were observed 3 to 5 days after infection.

6.3. Stable *Arabidopsis thaliana* transformation by floral dip

We used the floral dip technique to produce stable transgenic *Arabidopsis* plants expressing our proteins of interest. This method is based on the insertion of a T-DNA in the developing ovule to ensure its transmission to the progeny.

A volume of 250 mL of LB (Lysogeny Broth – 1% (w/v) bacto-tryptone ; 0.5% (w/v) yeast extract ; 0.5% (w/v) NaCl) are inoculated with an *Agrobacterium* culture containing the sequence of interest inserted in a binary vector. The culture media is supplemented with rifampicine (30 µg/mL) to select the *Agrobacterium* GV3101 strain and a specific antibiotic to select bacteria containing the plasmid carrying the sequence of interest. This culture is incubated over night at 28°C. The next day, the culture is centrifuged 10 min at 4 000 g and then resuspended in 100 mL of infiltration buffer composed of Murashige & Skoog 255 medium (Duchefa) ; 5% (w/v) saccharose ; 100 µM acetosyringone ; 0.02% (w/v)

Silwet. An *Arabidopsis thaliana* plant showing a high number of young open and closed flowers are dipped in the *Agrobacterium* culture resuspended in infiltration buffer during 1 min 30 sec. The floral stem is then gently dried on absorbent paper and placed overnight in a dark growing chamber. The next day, the dark chamber is removed and stakes are placed to support plant growth.

6.4. *Arabidopis in vitro* culture

6.4.1. Seed sterilization and sowing

Prior to sowing, *Arabidopsis thaliana* seeds are sterilized to avoid contaminations. Sterilization is performed by incubating 50 µL of seeds (prepared in an 1,5 mL Eppendorf tube) in 1 mL of a Bleach/Tween-20 solution (32% (v/v) Bleach ; 0.05% (v/v) Tween-20) for 15 min on wheel. Samples are centrifuged 10 sec at 4 000 g to bring all the seeds down of the tube. The supernatant is discarded and 1 mL 100% (v/v) ethanol is added. Tubes are homogenized by several tube inversions to resuspend the seeds. It follows 4 washes with 1 mL of sterile water before sowing the seeds on *in vitro* plates (Murashige & Skoog 255 medium (Duchefa) ; 1% (w/v) sucrose ; 1% (w/v) agar). After 48 h stratification in the dark at 4°C, plates are transferred to growing chambers at 21°C with a light cycle of 16 h light / 8 h dark.

6.4.2. Ethylene insensitivity experiments

For ethylene insensitivity experiments, the precursor of ethylene, ACC (1-aminocyclopropane-1-carboxylic acid) was added to the MS agar media at a final concentration of 10 µM. After 72 h stratification in the dark at 4°C, plates are transferred to a growing chamber during 8 h, before being transferred in the dark for 72 h. Plantlets are gently deposited on new MS agar plates (Murashige & Skoog 255 medium (Duchefa) ; 1% (w/v) sucrose ; 1% (w/v) agar) before taking photographs. Control growth without ACC treatment is performed in parallel.

6.4.3. *In vitro* selection of transformed plants

Following the floral dip, transformed seeds are selected on MS agar plates (Murashige & Skoog 255 medium (Duchefa) ; 1% (w/v) sucrose ; 1% (w/v) agar) supplemented with either Hygromycin (25µg/mL) or Glufosinate (30 µg/mL). Plants are observed after 2 or 3 weeks to discriminate sensitive from resistant plants and a counting is done if necessary. The counting of T2 resistant and sensitive plants on selective medium was performed to analyze the number of transgene insertion (monolocus or multilocus). The progeny of monolocus lines were harvested and homozygous lines showing 100% resistant plants on selective medium were selected.

6.5. Crossings

Arabidopsis mutant lines were crossed to produce multiple mutant combinations. For this purpose, open flowers and siliques of the plant chosen as female are removed from a floral stem. The petals, sepals and stamens are removed from at least three remaining unopened flowers exposing the gynaeceum. The next morning, stamen from the male parent are used to deposit pollen on the stigmatic papilla on the top of the gynaeceum of the female parent. The siliques produced are gently collected after complete drying.

7. Confocal microscopy imaging

Subcellular localization analysis of fluorescent fusion proteins was performed in *N. benthaminana* leaves transiently expressing protein fusions or in root epidermal cells of 7 days old stably transformed *Arabidopsis* plants using a LSM780 confocal microscope from Carl Zeiss with a 40x objective. Confocal microscopy images were acquired with a 488 laser for GFP fusion, and a 561 laser for RFP fusion. In order to avoid crosstalk between the emission of GFP and the excitation of RFP, sequential images were acquired. As the conventional frame-by-frame acquisition yield shifted detection for fast moving objects between the GFP and RFP channel, line-by-line acquisition was rather used. An ImageJ macro created by Jérome Mutterer (Institut de Biologie Moléculaire des Plantes) was used to quantify the extent of colocalization between two populations of cytoplasmic granules. Foci in images were determined with a user-supervised local maxima detection method. 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 [I_{GFP} vs I_{RFP}] scatter plot for further qualitative assessment of fluorescent spot content correlation.

Supplementary tables

Table S10. Plant material used and generated in this study.

Gene AGI	Gene name	Mutant		Origin
AT5G47010	UPF1	*upf1-5*	SALK_112922	NASC Stock Center ; described in Riehs-Kearnan et al., 2012
AT5G47010	UPF1	*upf1-3*	SALK_081178	NASC Stock Center ; described in Riehs-Kearnan et al., 2012
AT1G33980	UPF3	*upf3-1*	SALK_025175	NASC Stock Center ; described in Riehs-Kearnan et al., 2012
AT1G54490	XRN4	*xrn4-3*	Salk_014209	NASC Stock Center
AT1G08370	DCP1	*dcp1-3*	SAIL_377_B10	NASC Stock Center
AT1G26110	DCP5	*dcp5-1*	SALK_008881	NASC Stock Center
AT1G75660	XRN3	*xrn3-3*	SAIL_1172_C07	NASC Stock Center ; described in Gy et al., 2007
AT2G15560	Endonuclease	*endo-2*	SALK_127449	NASC Stock Center
AT3G61690	TNTase	*nt transf*	GABI_749E07	NASC Stock Center
AT3G60240	eIFG4	*eif4g*	SALK_112882	NASC Stock Center
AT3G58510	RH11	*rh11*	SALK_203665	NASC Stock Center
AT3G61240	RH12	*rh12*	SALK_016921	NASC Stock Center
AT3G01540	RH14	*rh14*	SALK_073018	NASC Stock Center
AT4G00660	RH8	*rh8*	GABI_447H07	NASC Stock Center
AT3G13300	VARICOSE	*vcs-8*	SAIL_1257_H12	NASC Stock Center
AT5G14610	RH46	*rh46*	SAIL_116644	NASC Stock Center
AT1G33680	AT1G33680	*AT1G33680*	GABI_536H12	NASC Stock Center
AT5G63120	RH30	*rh30*	GABI_101D04	NASC Stock Center
AT4G34110	PAB2	*pab2*		C. Bousquet-Antonelli team, Perpignan
AT2G23350	PAB4	*pab4*		C. Bousquet-Antonelli team, Perpignan
AT1G49760	PAB8	*pab8*		C. Bousquet-Antonelli team, Perpignan
AT3G46960	SKI2	*ski2-6*	SALK_122393	NASC Stock Center
AT4G21670	CPL1	*cpl1-9*	GABI_849A10	NASC Stock Center
AT3G06480	RH40	*rh40*	SAIL_573_C12	NASC Stock Center
AT2G15560	UCN	*ucn-1*	SALK_009495	NASC Stock Center
AT2G15560	UCN	*ucn-2*	SALK_127449	NASC Stock Center
AT2G15560	UCN	*ucn-3*	SALK_132521	NASC Stock Center

Gene AGI	Line name	Vector used	Background	Origin
AT1G26110	DCP5-GFP	pUBCGFP	*dcp5-1*	this study, Hélène Scheer
AT2G45810	RH6-GFP	pUBCGFP	Col-0	this study, Hélène Scheer
AT5G47010	UPF1-RFP	pCTL-RFP	*upf1-5*	this study, Shahinez Garcia
AT5G47010	UPF1-GFP	pCTL-RFP	*upf1-5*	this study, Shahinez Garcia
AT5G47010	UPF1-Flag-HA	pCTL-FHA	*upf1-3*	this study, Shahinez Garcia
AT2G15560	GFP-UCN-WT	pB7WGF2	*ucn-2*	this study
AT2G15560	GFP-UCN-D153N	pB7WGF2	*ucn-2*	this study
AT1G08370	YFP-DCP1		*dcp1-3*	C. Bousquet-Antonelli team, Perpignan
AT4G34110	RFP-PAB2		*pab2*	C. Bousquet-Antonelli team, Perpignan

Table S11. Vectors used in this study.

Name	Type	Use	Tag in	Bacterial resistance	Plant resistance	Wad used for
pDONR221	pDONR	pENTRY production	-	Kanamycin	-	all constructs
pB7WGF2	pDEST	Transient and stable expression	Nter	Spectinomycin	Glufosinate	**Transient expression :** PTB3-GFP ; GRP7-GFP ; TNTase-GFP ; RH11-GFP ; GFP-UCN-WT ; GFP-UCN-D55N ; GFP-UCN-D153N **Stable expression :** GFP-UCN-WT in *ucn-3* and GFP-UCN-D153N in *ucn-3*
pH7WGR2	pDEST	Transient expression	Nter	Spectinomycin	Hygromycin	**Transient expression :** RFP-UCN-WT, RFP-UCN-D55N and RFP-UCN-D153N
pGWB618	pDEST	Transient expression	Nter	Spectinomycin	Glufosinate	**Transient expression :** myc-UCN-WT, myc-UCN-D55N and myc-UCN-D153N
pUBCGFP	pDEST	Transient and stable expression	Cter	Spectinomycin	Glufosinate	**Transient expression :** RH6-GFP ; RH8-GFP ; RH12-GFP **Stable expression :** RH6-GFP in Col-0
pGADT7	pDEST	Yeast two-hybrid	Nter	Ampicilline	-	UPF2-Cter ; RH6 ; RH8 ; DCP1
pGBT9	pDEST	Yeast two-hybrid	Nter	Ampicilline	-	UPF1 ; DCP5 ; UCN-D153N

Table S12. DNA oligonucleotides used in this study.

Gene AGI	Gene name	Mutant		Name	Sequence	Application
-	-	-	-	SALK_LBb1.3	ATTTTGCCGATTTCGGAAC	Genotyping
-	-	-	-	SAIL_LB	GCCTTTTCAGAAATGGATAAATAGCC	Genotyping
-	-	-	-	GABI_LB_08760	GGGCTACACTGAATTGGTAGCTC	Genotyping
AT5G47010	UPF1	*upf1-5*	SALK_112922	SALK_112922-LP	ACAATCCAAATCTTCAGTCTCA	Genotyping
				SALK_112922-RP	AGGGACAACAAAATCATGTCG	Genotyping
AT5G47010	UPF1	*upf1-3*	SALK_081178	SALK_081178-LP	GATGAGTCTACTCAAGCAACAG	Genotyping
				SALK_081178-RP	GAATCTCAAGACTCTCACCTG	Genotyping
AT1G33980	UPF3	*upf3-1*	SALK_025175	SALK_025175-LP	ACTTCTATTGTTGATCTCTGG	Genotyping
				SALK_025175-RP	ATGCTGTTCCGGTTGTGGTGG	Genotyping
AT1G54490	XRN4	*xrn4-3*	Salk_014209	SALK_014209-LP	AGGTGTATGCTCTTGGCAATG	Genotyping
				SALK_014209-RP	AACTGCCATGAAAACTGATGG	Genotyping
AT1G08370	DCP1	*dcp1-3*	SAIL_377_B10	SAIL_377_B10-LP	CCTCCAGCTCTCACTTCACAG	Genotyping
				SAIL_377_B10-RP	TACAATCACCACCGCCTCTAC	Genotyping
AT1G26110	DCP5	*dcp5-1*	SALK_008881	SALK_008881-LP	CCATCAGCAGAGGATGAAGAG	Genotyping
				SALK_008881-RP	GTCCCAAAATTCAAGGCCTAG	Genotyping
AT1G75660	XRN3	*xrn3-3*	SAIL_1172_C07	SAIL_1172_C07 -LP	GCCTTCGATTTCAACAGGC	Genotyping
				SAIL_1172_C07 -RP	GAAATCGAACACAAATCCG	Genotyping
AT2G15560	Endonuclease	*endo-2*	SALK_127449	SALK_127449-LP	GCAAGGAGCTTACATTGCTTG	Genotyping
				SALK_127449-RP	GCGACTATCGTTCTCGTATCG	Genotyping
AT3G61690	TNTase	*nt transf*	GABI_749E07	GABI_749E07-LP	CAACATGCCTCTTGCTTTCTC	Genotyping
				GABI_749E07-RP	ATATGGCTCAGGATCTCCCAG	Genotyping
AT3G60240	eIFG4	*eif4g*	SALK_112882	SALK_112882-LP	GAACGCACCAGAGTGCTTATC	Genotyping
				SALK_112882-RP	AGGTTCATGTTGATCAATGCC	Genotyping
AT3G58510	RH11	*rh11*	SALK_203665	SALK_203665-LP	ATGGCGATAGAACACAACAGG	Genotyping
				SALK_203665-RP	TGGAAGGAAGTGAGCTTGAAG	Genotyping
AT3G61240	RH12	*rh12*	SALK_016921	SALK_016921-LP	TGTGTTTTGAAAGACTGTGCG	Genotyping
				SALK_016921-RP	TACCAAATTCACAGCCAATCC	Genotyping
AT3G01540	RH14	*rh14*	SALK_073018	SALK_073018-LP	ATATCGTGGTTGCAACTCCTG	Genotyping
				SALK_073018-RP	TACCTCTGCCACCATAACCAG	Genotyping
AT4G00660	RH8	*rh8*	GABI_447H07	GABI_447H07-LP	ATGGAACACTCTGTTTCGGTG	Genotyping
				GABI_447H07-RP	CAAAGCTGGCGTCGTAAATAG	Genotyping
AT3G13300	VARICOSE	*vcs-8*	SAIL_1257_H12	SAIL_1257_H12-LP	GTTTCTTCTTCAGTTGATGGCACG	Genotyping
				SAIL_1257_H12-RP	GGACCCTATGAGAGGAACATTGC	Genotyping
AT5G14610	RH46	*rh46*	SAIL_116644	SAIL_116644-LP	GTGATGGTCTTTGTCATTGGG	Genotyping
				SAIL_116644-RP	ATTTCATACCTCCCGCAGAAG	Genotyping
AT1G33680	AT1G33680	*AT1G33680*	GABI_536H12	GABI_536H12-LP	TCCGATTTATGTAACGTTGGC	Genotyping
				GABI_536H12-RP	TGAACTCGGGACATCTATTCG	Genotyping
AT5G63120	RH30	*rh30*	GABI_101D04	GABI_101D04-LP	AAATGGTGGGACAAATTAGGG	Genotyping
				GABI_101D04-RP	ACTAACGTGAACCAATGCTGG	Genotyping
AT3G46960	SKI2	*ski2-6*	SALK_122393	SALK_122393-LP	TTTCTCATTTGAACGTACCCG	Genotyping
				SALK_122393-RP	CGCCAAGCTTTTTGTAGTCTC	Genotyping
AT4G21670	CPL1	*cpl1-9*	GABI_849A10	GABI_849A10-LP	CGGTTAAAGAAAATGAAGGGG	Genotyping
				GABI_849A10-RP	CCCAAGAGCATTACTGCTGTC	Genotyping
AT3G06480	RH40	*rh40*	SAIL_573_C12	SAIL_573_C12-LP	TCAACTCCAGTTGGAAAATCG	Genotyping
				SAIL_573_C12-RP	GGATACTTGATCCCTGCCTTC	Genotyping
AT2G15560	UCN	*ucn-1*	SALK_009495	SALK_009495-LP	GGAGAATTGTCCCGTTCCTAG	Genotyping
				SALK_009495-RP	TCCTCCATTGGTAGTTACCCC	Genotyping
AT2G15560	UCN	*ucn-2*	SALK_127449	SALK_127449-LP	GCGACTATCGTTCTCGTATCG	Genotyping
				SALK_127449-RP	GCAAGGAGCTTACATTGCTTG	Genotyping
AT2G15560	UCN	*ucn-3*	SALK_132521	SALK_132521-LP	GCGACTATCGTTCTCGTATCG	Genotyping
				SALK_132521-RP	GCAAGGAGCTTACATTGCTTG	Genotyping

Gene AGI	Gene name	Mutant		Name	Sequence	Application
AT2G45810	RH6	-	-	attB1-RH6	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGAATAATAATAATAATAATAGAGGAAGATTTC	Amplification for cloning
		-	-	attB2-RH6-noSTOP	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTGACAGTAGATTGCCTTGTCG	Amplification for cloning
AT4G00660	RH8	-	-	attB1-RH8	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGAACAATCGAGGAAGGTACC	Amplification for cloning
		-	-	attB2-RH8-noSTOP	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTGGCAATAAATTGCCTGATC	Amplification for cloning
AT3G61240	RH12	-	-	attB1-RH12	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGAATACTAACAGAGGAAGATATCC	Amplification for cloning
		-	-	attB2-RH12-noSTOP	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTGACAGTAGATTGCTTGATCG	Amplification for cloning
AT1G26110	DCP5	-	-	attB1-DCP5	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGGCGGCTGATAATACGG	Amplification for cloning
		-	-	attB2-DCP5-noSTOP	GGGGACCACTTTGTACAAGAAAGCTGGGTAGGTAGTACGATTTGATACG	Amplification for cloning
AT5G47010	UPF1	-	-	UPF1-SalI-prom-F	GTCGACCGTGTGGTTGTTTTTCGTTT	Amplification for cloning
		-	-	UPF1-SmaI-noSTOP-R	CCCGGGGCCATTGTAAGGATGTTTTGG	Amplification for cloning
AT1G43190	PTB3	-	-	attB1-PTB3	GGGGACAAGTTTGTACAAAAAAGCAGGCTGGATGGCGGAATCTTCCAAAGTCG	Amplification for cloning
		-	-	attB2-PTB3	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAAATCGTCTGTAGCTGGG	Amplification for cloning
AT2G21660	GRP7	-	-	attB1-GRP7	GGGGACAAGTTTGTACAAAAAAGCAGGCTGGATGGCGTCCGGTGATGTTGAG	Amplification for cloning
		-	-	attB2-GRP7	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTACCATCCTCCACCACCAC	Amplification for cloning
AT2G15560	endonuclease / UCN	-	-	attB1-endonuclease	GGGGACAAGTTTGTACAAAAAAGCAGGCTGGATGATACAAAACGCTATGTC	Amplification for cloning
		-	-	attB2-endonuclease	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAAACCGGGCTGATGAGTTTC	Amplification for cloning
AT3G61690	TNTase	-	-	attB1-TNTase	GGGGACAAGTTTGTACAAAAAAGCAGGCTGGATGGGTGAGCATGAATCATG	Amplification for cloning
		-	-	attB2-TNTase	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTACTTTCCCCTGTTTAGTG	Amplification for cloning
AT3G58510	RH11	-	-	attB1-RH11	GGGGACAAGTTTGTACAAAAAAGCAGGCTGGATGAGTGCATCATGGGCAG	Amplification for cloning
		-	-	attB2-Rh11	GGGGACCACTTTGTACAAGAAAGCTGGGTGCTAATCCCAAGCACTGGTCAC	Amplification for cloning
AT1G33980	UPF3	-	-	UPF3q-F	GGGAGGTTGATCAAGGGAAT	qPCR
		-	-	UPF3q-R	TCTTTCCCTGCTCGAGTGTT	qPCR
AT5G19400	SMG7	-	-	SMG7q-F	TCCTAGTGGAGGCTTCAGGA	qPCR
		-	-	SMG7q-R	TGCAGGCACTTGAATACTCG	qPCR
AT1G60510	antisense	-	-	antisenseq-F	ACACGGTTGACCGTCTAAGG	qPCR
		-	-	antisenseq-R	TGAAGTCGCATGACAAGAGG	qPCR
AT1G36730	CPuORF19	-	-	CPuORF19q-F	AAGATGCCGAGGATGATGAC	qPCR
		-	-	CPuORF19q-R	AAGTGTCATGAGCCCCATTC	qPCR
AT2G45670	NMD2 PTC+	-	-	NMD2 PTC+q-F	AAGGGATGTAAATACAGAGAAAAGCTTC	qPCR
		-	-	NMD2 PTC+q-R	AGGGTAACCAGGGATGAAAGC	qPCR

Table S13. Antibodies used in this study.

Antibody	Dilution	Secondary antibody	Origin	Protein immuno-detected
@HA-HRP	1/10 000	-	Sigma	UPF1-Flag-HA
@GFP	1/30 000	GAR	Kindly provided by David Gilmer	GFP-UCN-WT ; GFP-UCN-D55N ; GFP-UCN-D153N ; UPF1-GFP ; DCP5-GFP ; RH6-GFP ; RH8-GFP ; RH12-GFP ; YFP-DCP1
@PR1	1/5 000	GAR	Kindly provided by Thierry Heitz	endogenous PR1 from *N. benthamiana*
@DCP1	1/3 000	GAGP	Kindly provided by Markus Fauth	endogenous DCP1 from *A. thaliana*
@UPF1	1/5 000	GAR	Raised against QPNQSSQNPKHPYNG and GVDDEPQPVQPKYED peptides and affinity purified following the double X protocol (Eurogentec SA)	endogenous UPF1 from *A. thaliana*
@P38	1/300 000	GAR	Azeredo *et al.* , 2010	P38 from TCV
@PAB2	1/3 000	GAR	Kindly provided by Jean-François Laliberté	endogenous PAB2 from *A. thaliana*
@AGO1	1/10 000	GAR	Kindly provided by Nicolas Baumberger	endogenous AGO1 from *A. thaliana*

Etude de l'interactome de UPF1, un acteur central du Nonsense-Mediated Decay chez *Arabidopsis thaliana*

Résumé de thèse

Clara Chicois – 2014-2016
Equipe RNA Degradation – Institut de Biologie Moléculaire des Plantes, Strasbourg
Sous la direction de Damien Garcia

Introduction

Le Nonsense-Mediated Decay (NMD) est un mécanisme conservé chez tous les eucaryotes permettant la reconnaissance et la dégradation d'ARN présentant une terminaison prématurée de la traduction. Ce mécanisme met en jeu l'action d'un complexe protéique composé des trois protéines UPF1, UPF2 et UPF3. UPF1 est un facteur central du NMD, étant responsable de la reconnaissance des ARN ciblés et participant au recrutement des facteurs permettant l'induction de la dégradation de ces ARN. Le NMD peut affecter différents types d'ARN, chacun possédant des caractéristiques précises, comme un long 3'-UTR (Untranslated Region), la présence d'un intron en 3' du codon stop ou encore la présence d'une petite ORF (Open Reading Frame) en amont de l'ORF principale.

Les différentes étapes du NMD sont bien décrites chez les mammifères, où il fait intervenir le complexe SURF (SMG1, UPF1, eRF3) permettant la reconnaissance d'ARN cible puis le complexe DECID (Decay-Inducing Complex) qui promeut la phosphorylation de UPF1 induisant une cascade d'événements permettant d'accélérer la dégradation des ARN ciblés par déadénylation, décapping et l'action de nucléases (Lejeune et al., 2003). Cependant, le NMD n'est pas aussi bien décrit chez les plantes, sachant que de nombreux facteurs connus pour leur rôle dans le NMD chez les mammifères sont absents du génome d'*Arabidopsis thaliana*. UPF1 est également impliqué dans d'autres voies de dégradation chez les mammifères comme le Staufen-Mediated Decay (Kim et al., 2005) (SMD) ou le Regnase-Mediated Decay (Mino et al. 2015) (RMD); mais chez les plantes aucune voie de dégradation secondaire impliquant UPF1 n'a jusqu'alors été décrite.

Dans ce contexte, mes objectifs de thèse sont les suivants :

- D'identifier de nouveaux partenaires protéiques d'UPF1 chez *Arabidopsis*
- De déterminer dans quelle mesure ces facteurs participent à la régulation des cibles du NMD
- D'étudier leur localisation sub-cellulaire ainsi que celle de UPF1
- D'identifier les facteurs impliqués dans une voie dépendante de UPF1 mais indépendante du NMD.

Pour répondre à ces problématiques, la stratégie utilisée s'est dans un premier temps basée sur une étude de l'interactome de UPF1 en utilisant une méthode non biaisée couplant immuno-précipitation et spectrométrie de masse. Cette approche a permis l'identification de 245 facteurs associés à UPF1 chez les plantes, parmi lesquels des facteurs de dégradation ou de maturation des ARNs ainsi que des

répresseurs de traduction. Nous avons testé l'accumulation de cibles du NMD dans 22 mutants affecté dans l'un des partenaires nouvellement identifiés. Plusieurs mutants ont montré une accumulation de cibles du NMD réduite. Parmi ces mutants, nous nous sommes focalisés sur le répresseur de la traduction DCP5. Nous avons identifié que UPF1 partage 50% de ces partenaires protéiques avec DCP5, indiquant leur présence au sein de mRNPs similaires. Nous avons identifié que UPF1 et DCP5 interagissent directement avec des ARN hélicases homologues à DDX6, qui est un répresseur de la traduction connu chez les mammifères. Nous avons démontré que toutes ces protéines co-localisent au sein des P-bodies, nous proposons donc un modèle dans lequel DCP5 et les homologues de DDX6 protègent les cibles du NMD de la dégradation via une re-localisation au sein des P-bodies.

A la suite de cette étude, nous nous sommes focalisés sur une endonucléase co-purifiant avec UPF1, que nous avons appelé UCN pour UPF1 Co-purified Nuclease. Une analyse réverse par spectrométrie de masse a permis l'identification des partenaires protéiques de UCN, comme UPF1 et des composants du complexe de decapping, DCP1 et VCS. L'utilisation de la technique de double hybride a identifié une interaction directe entre UCN et DCP1. Lors d'études de co-localisation, nous avons validé l'hypothèse que UCN est un nouveau composant des P-bodies. Enfin, l'analyse fonctionnelle de la fonction d'UCN suggère que son expression pourrait influencer les voies de signalisation hormonales, la mort cellulaire ou les mécanismes de défense. L'étude de ce nouveau partenaire de UPF1 et composant des P-bodies pourrait permettre d'identifier une nouvelle voie de dégradation des ARNs chez les plantes.

Ce travail permettrait d'identifier de nouveaux mécanismes de régulation des cibles du NMD chez les plantes en relation avec le rôle des P-bodies dans l'équilibre entre la traduction des ARNs, leur stockage et leur dégradation. Ce travail permet également l'identification de nouveaux composants de ces structures cytosoliques, qui seront d'une utilité capitale pour la compréhension des fonctions des P-bodies chez les plantes.

Résultats

Chapitre 1 : Identification de facteurs associés à UFP1

Afin d'identifier les facteurs associés à UPF1, des lignées stables d'*Arabidopsis thaliana* exprimant la protéine UPF1 couplée à différentes étiquettes peptidiques ont été produites. Nous avons pu valider que l'expression des protéines recombinantes correspondantes permettent une restauration de la régulation de cibles connues du NMD, indiquant que ces fusions protéiques sont fonctionnelles.

Ces lignées ont constitué le matériel indispensable pour effectuer des immunoprécipitations de UPF1 couplées à une détection de partenaires protéiques par spectrométrie de masse. Les résultats de spectrométrie de masse pour un total de 11 réplicats effectués avec deux antigènes différents ont été analysés afin d'identifier les protéines co-purifiant spécifiquement avec UPF1. Parmi ces facteurs, on retrouve une grande proportion de protéines interagissant avec l'ARN, incluant des membres de plusieurs familles d'ARN hélicase. Ces immunoprécipitations de UPF1 ont permis d'isoler UPF1 associé à ses partenaires protéiques mais également ses ARN ciblés. Etant donné la prépondérance de facteurs associés à l'ARN dans la fraction enrichie en UPF1, on peut se demander si les interactions mises en évidence par l'immunoprécipitation de UPF1 sont dépendantes ou indépendantes de l'ARN. Pour tester cette hypothèse, une immunoprécipitation de UPF1 avec ou sans traitement à la RNAse a été effectuée et une analyse des facteurs associés à UFP1 a permis de déterminer les facteurs associés de façon dépendante ou indépendante de l'ARN.

La localisation de UPF1 relative aux P-bodies est encore controversée chez les plantes, certaines études indiquent que UPF1 est localisé dans les P-bodies alors que d'autre montrent qu'il en est exclu (Mérai et al. 2013; Moreno et al. 2013). Afin de clarifier dans quel compartiment se localise UPF1, une analyse de la localisation cellulaire de UPF1 a ensuite été effectuée. La localisation des partenaires de UPF1 a également été étudiée, permettant l'identification de nouveaux facteurs se localisant dans les Processing-bodies (P-bodies). Cette étude a été effectuée à la fois en expression transitoire et en utilisant des lignées stables d'*Arabidopsis thaliana* en exprimant UPF1 et certains de ces partenaires couplés à un fluorophore. Pour cette étude, des paramètres d'acquisition permettant de limiter les erreurs de localisation liées au mouvement rapide des P-bodies entre l'acquisition des deux canaux ont été utilisés. Ceci a permis de démontrer clairement que UPF1 est localisé à la fois dans le cytoplasme et dans les P-bodies chez *Arabidopsis thaliana*.

Enfin, l'accumulation des cibles connues du NMD a été analysée dans des mutants de facteurs co-purifiant significativement avec UPF1 afin d'identifier quels sont les facteurs associés à UPF1 ayant un rôle dans le processus du NMD. Des facteurs associés à la dégradation des ARN, ainsi que des protéines interagissant avec l'ARN et co-purifiant avec UPF1 ont été testés. De façon surprenante, aucun des mutants testés n'entraine une accumulation plus forte des cibles du NMD, comme ce serait le cas pour un facteur stimulateur du NMD, indiquant qu'aucun des facteurs testés n'est un facteur limitant du NMD. Au contraire, le facteur DCP5 associé au complexe de decapping a été identifié par cette analyse comme un potentiel inhibiteur du NMD car une diminution de cibles connues du NMD est observée

dans le mutant *dcp5*. Afin de tester si cette diminution de l'accumulation de cibles du NMD est liée à l'action de UPF1, le facteur majeur du NMD, le double mutant *upf1 dcp5* a été produit puis l'accumulation de cibles connues du NMD a été mesurée. La diminution de l'accumulation des cibles du NMD observée dans le simple mutant *dcp5* est abolie dans le double mutant *upf1 dcp5* indiquant que cette diminution est dépendante de UPF1. Ces résultats suggèrent que DCP5 a une action protectrice sur les cibles du NMD, limitant leur dégradation par UPF1. De plus, une interaction directe entre UPF1 ou DCP5 et un homologue de DDX6 a été indépendamment mise en évidence par la technique de double hybride.

Ces résultats nous ont permis de proposer un modèle dans lequel DCP5 et les homologues de DDX6 sont impliqués dans la protection de cibles du NMD en induisant leur re-localisation dans les P-bodies (Fig1).

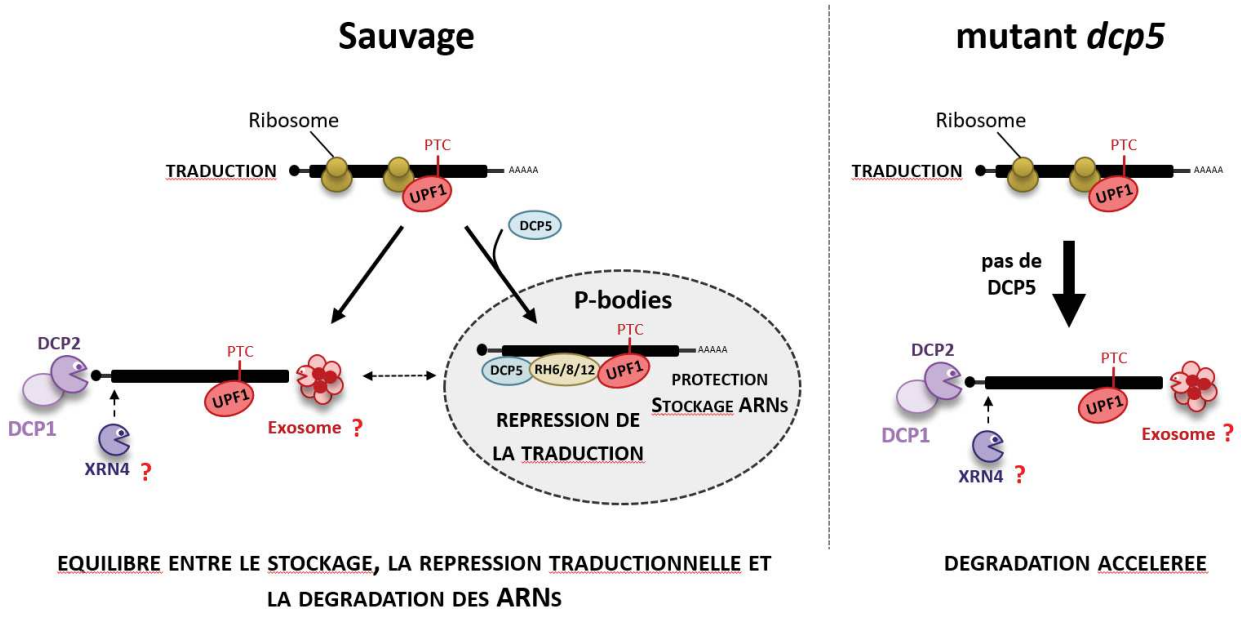


FIGURE 1. Modèle du lien entre DCP5 et UPF1 dans le NMD. Une cible typique du NMD possédant un PTC est représentée. Cette cible est reconnue par UPF1 lors de la traduction, permettant la dégradation accélérée, potentiellement par les activités d'exoribonucléase XRN et/ou de l'exosome. Lors de la liaison de DCP5 et des homologues de DDX6, RH6/8/12, une fraction des cibles du NMD sont relocalisées dans les P-bodies et échappent à la dégradation possiblement grâce une répression de la traduction (panel de gauche). Lors de la réduction de l'action de DCP5, par exemple dans un mutant *dcp5* (right panel), la

protection induite par DCP5 est levée, induisant l'augmentation de la dégradation des cibles du NMD de façon dépendante de UPF1.

Chapitre 2 : Etude d'un nouveau facteur des Processing-bodies

Dans une deuxième partie, un des nouveaux facteurs se localisant dans les P-bodies – une endonucléase putative nommée UCN pour UPF1 Co-purified Nuclease – a été étudié plus en détails. UCN est une endonucléase à domaine NYN ; plusieurs facteurs possédant ce domaine ont montré une activité endonucléolytique, comme la protéine Regnase-1 chez la souris par exemple (Suzuki et al. 2011). Des mutants catalytiques de l'endonucléase ont été produits et des lignées stables d'*Arabidopsis thaliana* exprimant soit la version sauvage soit une version mutante couplée à la GFP ont été produites.

Une immunoprécipitation de UCN a été effectuée couplée à une détection de partenaires protéiques par spectrométrie de masse. Après traitement statistique, les facteurs de decapping DCP1 et VARICOSE (VCS) ont été identifiés comme les partenaires majeurs de l'endonucléase UCN. De plus, UPF1 a été identifié dans la fraction immuno-purifiée de UCN, permettant de valider l'association entre ces deux protéines.

L'expression transitoire de versions mutantes ou de la version sauvage de UCN a permis d'observer un phénotype de nécrose cellulaire restreinte à la version sauvage, indiquant que ce phénotype de nécrose est potentiellement lié à l'activité de l'endonucléase UCN. Ce type de nécrose peut être dû à une augmentation de la défense de la plante, et est parfois observé lors d'infections virales. L'accumulation du facteur PATHOGENESIS-RELATED 1 (PR1), qui est sur-exprimé lors de la nécrose cellulaire liée à la réponse hypersensible (Rivière et al. 2017) ne varie pas en fonction de la présence de la version sauvage ou mutante de UCN. Ceci indique que la nécrose observée n'est pas due à une activation de la défense par induction de l'expression de PR1.

Afin d'étudier le rôle fonctionnel de UCN, une étude de co-expression transitoire dans *Nicotiana benthamiana* a été effectuée. Pour cela, l'endonucléase UCN a été co-exprimée avec différentes versions de la GFP, certaines étant ciblées par le NMD. Cette étude a pu mettre en évidence le fait que l'expression de la version sauvage de UCN est responsable de la diminution de l'accumulation de certaines cibles GFP au niveau ARN et protéique, alors que la version mutante n'a peu ou pas d'effet sur l'accumulation de ces cibles. Plusieurs ARNs viraux subissent également cette diminution lors de leur co-infiltration avec UCN, ceci suggérant une implication de UCN dans la défense antivirale.

Enfin, une analyse phénotypique des lignées d'*Arabidopsis thaliana* exprimant de façon ectopique l'endonucléase UCN sauvage ou mutante a été effectuée. Ces plantes transgéniques exprimant l'endonucléase UCN présentent une taille réduite ainsi qu'un phénotype d'insensibilité partielle de réponse à l'éthylène, tout comme les mutants *upf1* et *xrn4*.

Ces résultats nous ont permis de proposer un modèle préliminaire quant au mode d'action de UCN. UCN reconnait probablement ses cibles par liaison à la région 3'-UTR. Cette reconnaissance, ou potentiellement le clivage de l'ARN par UCN permet le recrutement du complexe de decapping permettant d'accélérer la dégradation des ARNs ciblés (Fig2).

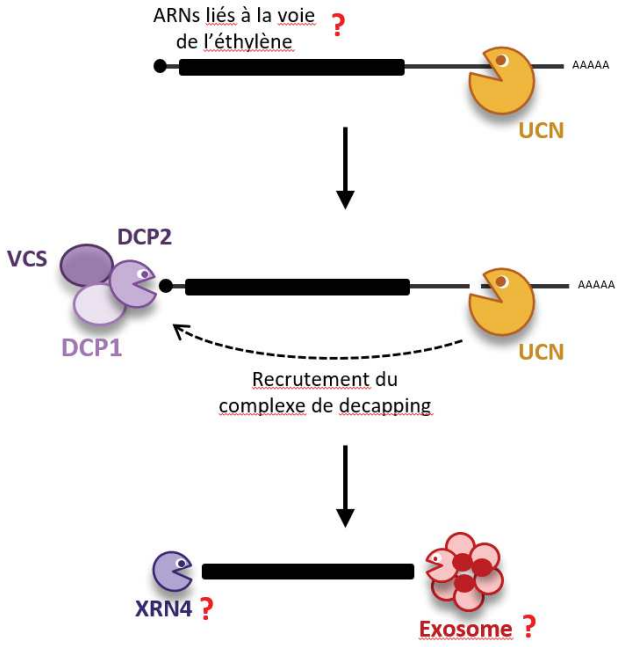


FIGURE 2. Modèle proposé pour le mode d'action de UCN. L'endonucléase UCN reconnaît ses cibles potentiellement par leur région 3'UTR, ces cibles étant potentiellement liées à la voie de signalisation de l'éthylène. UCN recrute le complexe de decapping par l'intermédiaire d'une interaction directe avec DCP1 afin d'induire la dégradation de ses ARNs ciblés, potentiellement par l'intermédiaire des activités de l'exoribonucléase XRN4 et de l'exosome.

Conclusion

Ce travail de thèse a pu mettre en évidence, pour la première fois chez les plantes l'identité des partenaires protéiques de UPF1 lui étant associés soit de façon dépendante soit de façon indépendante de l'ARN. Parmi ces facteurs, de nombreuses protéines ayant des fonctions liées à l'ARN ont été identifiées. Une étude de la localisation sub-cellulaire de ces facteurs a permis l'identification de nouveaux facteurs localisés dans les P-bodies et dont la fonction n'a encore jamais été décrite chez les plantes. Ceux-ci représentent des candidats prometteurs qui pourraient avoir un rôle encore inconnu dans la stabilité des ARN. Le facteur DCP5 associé au complexe de decapping a été identifié comme un facteur inhibiteur du NMD. Un lien direct entre UPF1, DCP5 et les homologues de DDX6 suggère l'implication de ces répresseurs de la traduction dans la régulation de la stabilité cibles du NMD. Le modèle proposé implique la protection des cibles du NMD au sein des P-bodies grâce à DCP5 et les homologues de DDX6. L'étude détaillée des deux autres inhibiteurs potentiels du NMD, le facteur d'initiation de la traduction eIF4G et l'ARN hélicase RH14, permettra de mieux comprendre le mécanisme cellulaire de cette inhibition du NMD.

Dans une deuxième partie, la fonction de la protéine de fonction inconnue UCN localisée dans les P-bodies a été étudiée. Nos résultats préliminaires suggèrent que la modulation de l'activité de UCN permet de pourrait stimuler la défense antivirale. L'identification des cibles directes de UCN et l'étude de l'implication de ses partenaires protéiques principaux DCP1 et VCS permettra d'améliorer notre compréhension du mode d'action de cette endonucléase.

Ainsi, l'étude d'un facteur connu et conservé comme UPF1 ou encore inconnu comme l'endonucléase UCN impliqués dans la dégradation des ARN pourrait permettre l'identification de nouvelles voies de régulation de la stabilité des ARNs chez *Arabidopsis thaliana*.

Liste des articles en cours de rédaction :

- Chicois Clara, Scheer Hélène, Garcia Shahïnez, Zuber Hélène, Mutterer Jérôme, Chicher Johanna, Hammann Philippe, Gagliardi Dominique and Garcia Damien - Interaction between UPF1 and translation repressors identifies novel regulatory networks in plant NMD

Liste des communications orales :

- Comprehensive analysis of UPF1 interactors in *Arabidopsis thaliana*. Chicois C., Garcia S.,Hammann P., Garcia D.The complex life of mRNA EMBO | EMBL Symposia, Heidelberg, Germany (POSTER)

Bibliographie :

Kim, Yoon Ki, Luc Furic, Luc Desgroseillers, and Lynne E Maquat. 2005. “Mammalian Staufen1 Recruits Upf1 to Specific mRNA 3’UTRs so as to Elicit mRNA Decay.” *Cell* 120: 195–208.

Lejeune, Fabrice, Xiaojie Li, and Lynne E Maquat. 2003. “Nonsense-Mediated mRNA Decay in Mammalian Cells Involves Decapping, Deadenylating and Exonucleolytic Activities.” *Molecular Cell* 12: 675–87.

Mérai, Zsuzsanna et al. 2013. “The Late Steps of Plant Nonsense-Mediated mRNA Decay.” *The Plant Journal* 73: 50–62.

Mino, Takashi et al. 2015. “Regnase-1 and Roquin Regulate a Common Element in Inflammatory mRNAs by Spatiotemporally Distinct Article Regnase-1 and Roquin Regulate a Common Element in Inflammatory mRNAs by Spatiotemporally Distinct Mechanisms.” *Cell* 161(5): 1058–73.

Moreno, Ana Beatriz et al. 2013. “Cytoplasmic and Nuclear Quality Control and Turnover of Single-Stranded RNA Modulate Post-Transcriptional Gene Silencing in Plants.” *Nucleic Acids Research* 41(8): 4699–4708.

Rivière, Marie-Pierre et al. 2017. “Silencing of Acidic Pathogenesis-Related PR-1 Genes Increases Extracellular β-(1→3)-Glucanase Activity at the Onset of Tobacco Defence Reactions.” *Journal of Experimental Botany* 59(6): 1225–39.

Suzuki, Hiroshi I et al. 2011. “MCPIP1 Ribonuclease Antagonizes Dicer and Terminates MicroRNA Biogenesis through Precursor MicroRNA Degradation.” *Molecular Cell* 44(3): 424–36. http://dx.doi.org/10.1016/j.molcel.2011.09.012.

Bibliography

Aizer, A., Brody, Y., Ler, L.W., Sonenberg, N., Singer, R.H., and Shav-tal, Y. (2008). The Dynamics of Mammalian P Body Transport , Assembly , and Disassembly In Vivo. Mol. Biol. Cell **19**: 4154–4166.

Ajamian, L., Abrahamyan, L., Milev, M., Ivanov, P. V, Kulozik, A.E., Gehring, N.H., and Mouland, A.J. (2008). Unexpected roles for UPF1 in HIV-1 RNA metabolism and translation. RNA **14**: 914–927.

Alberti, S. (2017). Phase separation in biology. Curr. Biol. **27**: R1097–R1102.

Amrani, N., Ganesan, R., Kervestin, S., and Mangus, D.A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense- mediated mRNA decay. Nature **432**: 112–118.

Anantharaman, V. and Aravind, L. (2006). The NYN Domains Novel Predicted RNAses with a PIN Domain-Like Fold. RNA Biol. **3**: 18–27.

Anantharaman, V., Zhang, D., and Aravind, L. (2010). OST-HTH : a novel predicted RNA-binding domain. Biol. Direct **5**.

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.

Banani, S.F., Rice, A.M., Peeples, W.B., Lin, Y., Jain, S., Parker, R., and Rosen, M.K. (2016). Compositional Control of Phase-Separated Cellular Bodies. Cell **166**: 651–663.

Bashir, Z., Ahmad, A., Shafique, S., Anjum, T., Shafique, S., and Akram, W. (2013). HYPERSENSITIVE RESPONSE – A BIOPHYSICAL PHENOMENON OF PRODUCERS. Eur. J. Microbiol. Immunol. **3**: 105–110.

Beckham, C., Hilliker, A., Cziko, A., Noueiry, A., Ramaswami, M., and Parker, R. (2008). The DEAD-Box RNA Helicase Ded1p Affects and Accumulates in Saccharomyces cerevisiae P-Bodies. Mol. Biol. Cell **19**: 984–993.

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.

Bloch, D.B., Li, P., Bloch, E.G., Berenson, D.F., Galdos, R.L., Arora, P., Malhotra, R., Wu, C., and Yang, W. (2014). LMKB / MARF1 Localizes to mRNA Processing Bodies , Interacts with Ge-1, and Regulates IFI44L Gene Expression. PLoS One **9**: e94784.

Boehm, V., Haberman, N., Ottens, F., Ule, J., and Gehring, N.H. (2014). 3' UTR Length and Messenger Ribonucleoprotein Composition Determine Endocleavage Efficiencies at Termination Codons. Cell Rep. **9**: 555–568.

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.

Brazão, T.F., Demmers, J., Ijcken, W. Van, Strouboulis, J., Fornerod, M., Romão, L., and Grosveld, F.G. (2012). A new function of ROD1 in nonsense-mediated mRNA decay. FEBS Lett. **586**: 1101–1110.

Brendolise, C., Montefiori, M., Dinis, R., Peeters, N., Storey, R.D., and Rikkerink, E.H. (2017). A novel hairpin library-based approach to identify NBS–LRR genes required for effector-triggered hypersensitive response in Nicotiana benthamiana. Plant Methods **13**: 32.

Brengues, M., Teixeira, D., and Parker, R. (2007). Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. Science (80-.). **310**: 486–489.

Brogna, S., Mcleod, T., and Petric, M. (2016). The Meaning of NMD : Translate or Perish. Trends Genet. **32**: 395–407.

Brogna, S., Ramanathan, P., and Wen, J. (2008). UPF1 P-body localization. Biochem. Soc. Trans. **36**: 698–700.

Brogna, S. and Wen, J. (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. Nat. Struct. Mol. Biol. **16**: 107–113.

Bruno, I.G., Karam, R., Huang, L., Bhardwaj, A., Lou, C.H., Shum, E.Y., Song, H., Corbett, M.A., Gifford, W.D., Gecz, J., Pfaff, S.L., and Wilkinson, M.F. (2011). Identification of a MicroRNA that Activates Gene Expression by Repressing Nonsense-Mediated RNA Decay. Mol. Cell **42**: 500–510.

Buschmann, H., Fabri, C.O., Hauptmann, M., Hutzler, P., Laux, T., Lloyd, C.W., and Schäffner, A.R. (2004). Helical Growth of the Arabidopsis Mutant tortifolia1 Reveals a Plant-Specific Microtubule-Associated Protein. Curr. Biol. **14**: 1515–1521.

Carroll, J.S., Munchel, S.E., and Weis, K. (2007). The DExD/H box ATPase Dhh1 functions in translation repression, mRNA decay, and processing body dynamics. J. Cell Biol. **194**: 527–537.

Caudy, A.A., Ketting, R.F., Hammond, S.M., Denlu, A.M., Bathoorn, A.M.P., Tops, B.B.J., Silva, J.M., Myers, M.M., Hannon, G.J., and Plasterk, R.H.A. (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature **425**: 411–414.

Celik, A., Baker, R., He, F., and Jacobson, A. (2017a). High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection. RNA **23**: 735–748.

Celik, A., He, F., and Jacobson, A. (2017b). NMD monitors translational fidelity 24/7. Curr. Genet. **63**: 1007–1010.

Chakrabarti, S., Bonneau, F., Schüssler, S., Eppinger, E., and Conti, E. (2014). interactions of the helicase UPF1 with the NMD factors. Nucleic Acids Res. **42**: 9447–9460.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (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.

Chang, C.-T., Bercovich, N., Loh, B., Jonas, S., and Izaurralde, E. (2014a). The activation of the decapping

enzyme DCP2 by DCP1 occurs on the EDC4 scaffold and involves a conserved loop in DCP1. Nucleic Acids Res. **42**: 5217–5233.

Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014b). TAIL-seq: Genome-wide Determination of Poly(A) Tail Length and 3' EndModification. Mol. Cell **53**: 1044–1052.

Chang, J.C. and Kan, Y.W. (1979). β0 thalassemia,thalassemia, a nonsense mutation in man. Proc. Natl. Acad. Sci. **76**: 2886–2889.

Chen, C.A., Ezzeddine, N., and Shyu, A. (2008). Messenger RNA half-life measurements in mammalian cells. Methods Enzymol. **448**: 335–357.

Chen, C.A. and Shyu, A. (2011). Mechanisms of deadenylation- dependent decay. Wiley Interdiscip. Rev. RNA: 167–183.

Chlebowski, A., Lubas, M., Jensen, T.H., and Dziembowski, A. (2013). RNA decay machines : The exosome. Biochim. Biophys. Acta **1829**: 552–560.

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.

Chu, C. and Rana, T.M. (2006). Translation Repression in Human Cells by MicroRNA-Induced Gene Silencing Requires RCK / p54. PLoS Biol. **4**: 1122–1136.

Coller, J. and Parker, R. (2005). General translation repression by activators of mRNA decapping. Cell **122**: 875–886.

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

Cooke, A., Prigge, A., and Wickens, M. (2010). Translational Repression by Deadenylases. J. Biol. Chem. **285**: 28506–28513.

Cougot, N., Babajko, S., and Séraphin, B. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. J. Cell Biol. **165**: 31–40.

Decker, C.J., Teixeira, D., and Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in. J. Cell Biol. **179**: 437–449.

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.

Dunster, K., Lai, F.P.L., and Sentry, J.W. (2005). Limkain b 1 , a novel human autoantigen localized to a subset of ABCD 3 and PXF marked peroxisomes. Clin. Exp. Immunol. **140**: 556–563.

Durand, S., Cougot, N., Mahuteau-Betzer, F., Nguyen, C.H., Grierson, D.S., Bertrand, E., Tazi, J., and Lejeune, F. (2007). Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. J. Cell Biol. **178**: 1145–1160.

Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., and Mühlemann, O. (2008). Posttranscriptional Gene Regulation by Spatial Rearrangement of the 3 9 Untranslated Region. PLoS Biol. **6**: 849–859.

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.

Fatscher, T., Boehm, V., Weiche, B., and Gehring, N.H. (2014). The interaction of cytoplasmic poly (A) - binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. RNA **20**: 1579–1592.

Fatscher, T. and Gehring, N.H. (2016). Harnessing short poly(A)-binding protein-interacting peptides for the suppression of nonsense-mediated mRNA decay. Sci. Rep. **6**: 1–12.

Fenger-Grøn, M., Fillman, C., Norrild, B., and Lykke-Andersen, J. (2005). Multiple Processing Body Factors and the ARE Binding Protein TTP Activate mRNA Decapping. Mol. Cell **20**: 905–915.

Fiorini, F., Bagchi, D., Le Hir, H., and Croquette, V. (2015). Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities. Nat. Commun. **6**: 7581.

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.

Franks, T.M. and Lykke-andersen, J. (2008). The Control of mRNA Decapping and P-Body Formation. Mol. Cell **32**: 605–615.

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.

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.

Gardner, L.B. (2008). Hypoxic Inhibition of Nonsense-Mediated RNA Decay Regulates Gene Expression and the Integrated Stress Response. Mol. Cell. Biol. **28**: 3729–3741.

Gatfield, D. and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature **429**: 575–578.

Gazzani, S., Lawrenson, T., Woodward, C., Headon, D., and Sablowski, R. (2004). A Link Between mRNA Turnover and RNA Interference in Arabidopsis. Science (80-.). **306**: 1046–1049.

Ge, Z., Quek, B.L., Beemon, K.L., and Hogg, J.R. (2016). Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. Elife: e11155.

Gehring, N.H., Kunz, J.B., Neu-yilik, G., Breit, S., Viegas, M.H., Hentze, M.W., and Kulozik, A.E. (2005). Exon-Junction Complex Components Specify Distinct Routes of Nonsense-Mediated mRNA Decay with Differential Cofactor Requirements. Mol. Cell **20**: 65–75.

Gera, J.F. and Baker, E.J. (1998). Deadenylation-Dependent and -Independent Decay Pathways for α1-

Tubulin mRNA in Chlamydomonas reinhardtii. Mol. Cell. Biol. **18**: 1498–1505.

Glavan, F., Behm-Ansmant, I., Izaurralde, E., and Conti, E. (2006). Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. EMBO J. **25**: 5117–5125.

Gloggnitzer, J., Akimcheva, S., Srinivasan, A., Kusenda, B., Riehs, N., Stampfl, H., Bautor, J., Dekrout, B., Jonak, C., Jiménez-Gomez, J., Parker, J.E., and Riha, K. (2014). Nonsense-Mediated mRNA Decay Modulates Immune Receptor Levels to Regulate Plant Antibacterial Defense. Cell Host Microbe **16**: 376–390.

Gobert, A., Gutmann, B., Taschner, A., Gößringer, M., Holzmann, J., Hartmann, R.K., Rossmanith, W., and Giegé, P. (2010). A single Arabidopsis organellar protein has RNase P activity. Nat. Publ. Gr. **17**: 740–744.

Gregersen, L.H., Schueler, M., Munschauer, M., Mastrobuoni, G., Chen, W., Kempa, S., Dieterich, C., and Landthaler, M. (2014). Contributing to UPF1 mRNA Target Degradation by Translocation along 3′ UTRs. Mol. Cell **54**: 573–585.

Gregori, J., Sanchez, A., and Villanueva, J. (2013). msmsTests: LC-MS/MS Differential Expression Tests. R package version 1.14.0.

Gutmann, B., Gobert, A., and Giegé, P. (2012). PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. Genes Dev. **26**: 1022–1027.

Gy, I., Gasciolli, V., Lauressergues, D., Morel, J.-B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C. (2007). Arabidopsis FIERY1, XRN2, and XRN3 Are Endogenous RNA Silencing Suppressors. Plant Cell **19**: 3451–3461.

Habacher, C. and Ciosk, R. (2017). ZC3H12A/MCPIP1/Regnase-1-related endonucleases: An evolutionary perspective on molecular mechanisms and biological functions. Bioessays **39**: 1700051.

Hamada, T., Tominaga, M., Fukaya, T., Nakamura, M., Nakano, A., Watanabe, Y., Hashimoto, T., and Baskin, T.I. (2017). RNA Processing Bodies , Peroxisomes , Golgi Bodies , Mitochondria , and Endoplasmic Reticulum Tubule Junctions Frequently Pause at Cortical Microtubules. Plant Cell Physiol. **53**: 699–708.

Hamid, F.M. and Makeyev, E. V (2014). Emerging functions of alternative splicing coupled with nonsense-mediated decay. Biochem. Soc. Trans. **42**: 1168–1173.

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.

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. Mol. Cell **12**: 1439–1452.

Hirayama, T., Matsuura, T., Ushiyama, S., Narusaka, M., Kurihara, Y., Yasuda, M., Ohtani, M., Seki, M., Demura, T., Nakashita, H., Narusaka, Y., and Hayashi, S. (2013). A poly(A)-specific ribonuclease directly regulates the poly(A) status of mitochondrial mRNA in Arabidopsis. Nat. Commun. **4**: ncomms3247.

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.

Hogg, J.R. (2016). Viral evasion and manipulation of host quality control pathways. J. Virol. **90**: 7010–7018.

Hogg, J.R. and Goff, S.P. (2010). Upf1 Senses 3'UTR Length to Potentiate mRNA Decay. Cell **143**: 379–389.

Holzmann, J., Frank, P., Löffler, E., Bennett, K.L., Gerner, C., and Rossmanith, W. (2008). RNase P without RNA : Identification and Functional Reconstitution of the Human Mitochondrial tRNA Processing Enzyme. Cell **135**: 462–474.

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: 615–625.

Houseley, J. and Tollervey, D. (2009). The Many Pathways of RNA Degradation. Cell **136**: 763–776.

Howard, M.J., Lim, W.H., Fierke, C.A., and Koutmos, M. (2012). Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for. PNAS **109**: 16149–16154.

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, P., Catinot, J., and Zimmerli, L. (2017). Ethylene response factors in Arabidopsis immunity. J. Exp. Bot. **67**: 1231–1241.

Hubstenberger, A. et al. (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. Mol. Cell **68**: 144–157.

Hurt, J.A., Robertson, A.D., and Burge, C.B. (2013). Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. Genome Res. **23**: 1636–1650.

Isken, O., Kim, Y.K., Hosoda, N., Mayeur, G.L., Hershey, J.W.B., 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.

Jeong, H., Kim, Y.J., Kim, S.H., Kim, Y., Lee, I., Kim, Y.K., and Shin, J.S. (2011). Nonsense-Mediated mRNA Decay Factors, UPF1 and UPF3, Contribute to Plant Defense. Plant Cell Physiol. **52**: 2147–2156.

Jeske, M., Müller, C.W., and Ephrussi, A. (2017). The LOTUS domain is a conserved DEAD-box RNA helicase regulator essential for the recruitment of Vasa to the germ plasm and nuage. Genes Dev. **31**: 939–952.

Jin, H., Suh, M.R., Han, J., Yeom, K., Lee, Y., Heo, I., Ha, M., Hyun, S., and Kim, V.N. (2009). Human UPF1 Participates in Small RNA-Induced mRNA Downregulation. Mol. Cell. Biol. **29**: 5789–5799.

Kalisiak, K., Kulinski, T.M., 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.

Kalyna, M. et al. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. Nucleic Acids Res. **40**: 2454–2469.

Kanemitsu, Y., Fujitani, M., Fujita, Y., Zhang, S., Su, Y., Kawahara, Y., and Yamashita, T. (2017). The RNA-binding protein MARF1 promotes cortical neurogenesis through its RNase activity domain. Sci. Rep. **7**: 1155.

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.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev. **20**: 355–367.

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.

Kaygun, H. and Marzluff, W.F. (2005). Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. Nat. Struct. Mol. Biol. **12**: 794–800.

Kebaara, B.W. and Atkin, A.L. (2009). Long 3'-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in Saccharomyces cerevisiae. Nucleic Acids Res. **37**: 2771–2778.

Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. J. Cell Biol. **169**: 871–884.

Kerényi, F., Wawer, I., Sikorski, P.J., Kufel, J., and Silhavy, D. (2013). Phosphorylation of the N- and C-terminal UPF1 domains plays a critical role in plant nonsense-mediated mRNA decay. Plant J. **76**: 836–848.

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.

Kervestin, S. and Jacobson, A. (2012). NMD : a multifaceted response to premature translational termination. Nat. Rev. Mol. Cell Biol. **13**: 700–712.

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.

Kim, S.H., Koroleva, O.A., Lewandowska, D., Pendle, A.F., Clark, G.P., Simpson, C.G., Shaw, P.J., and Brown, J.W.S. (2009). Aberrant mRNA Transcripts and the Nonsense-Mediated Decay Proteins UPF2 and UPF3 Are Enriched in the Arabidopsis Nucleolus. Plant Cell **21**: 2045–2057.

Kim, Y.K., Furic, L., Desgroseillers, L., and Maquat, L.E. (2005a). Mammalian Staufen1 Recruits Upf1 to Specific mRNA 3'UTRs so as to Elicit mRNA Decay. Cell **120**: 195–208.

Kim, Y.K., Furic, L., DesGroseillers, L., and Maquat, L.E. (2005b). Mammalian Staufen1 Recruits Upf1 to Specific mRNA 3'UTRs so as to Elicit mRNA Decay. Cell **120**: 195–208.

Kleinboelting, N., Huep, G., Kloetgen, A., Viehoever, P., and Weisshaar, B. (2012). GABI-Kat SimpleSearch : new features of the Arabidopsis thaliana T-DNA mutant database. Nucleic Acids Res. **40**: 1211–1215.

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**: e79219.

Kurihara, Y. et al. (2009). Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in Arabidopsis. PNAS **106**: 2453–2458.

Kuta, D.D. and Tripathi, L. (2005). Agrobacterium-induced hypersensitive necrotic reaction in plant cells: a resistance response against Agrobacterium-mediated DNA transfer. African J. Biotechnol. **4**: 752–757.

Łabno, A., Tomecki, R., and Dziembowski, A. (2016). Cytoplasmic RNA decay pathways - enzymes and mechanisms. BBA - Mol. Cell Res. **1863**: 3125–3147.

Lai, M.-C., Wang, S.-W., Cheng, L., Tarn, W.-Y., Tsai, S.-J., and Sun, H.S. (2013). Human DDX3 Interacts with the HIV-1 Tat Protein to Facilitate Viral mRNA Translation. PLoS One **8**: e68665.

Lange, H., Holec, S., Cognat, V., Pieuchot, L., Ret, M. Le, 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.

Lechner, M., Rossmanith, W., Hartmann, R.K., Thölken, C., Gutmann, B., Giegé, P., and Gobert, A. (2017). Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya. Mol. Biol. Evol. **32**: 3186–3193.

Lee, S.R., Pratt, G.A., Martinez, F.J., Yeo, G.W., and Lykke-Andersen, J. (2015). Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity. Mol. Cell **59**: 413–425.

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, M., Cao, W., Liu, H., Zhang, W., Liu, X., Cai, Z., Guo, J., Wang, X., Hui, Z., Zhang, H., Wang, J., and

Wang, L. (2012). MCPIP1 Down-Regulates IL-2 Expression through an ARE-Independent Pathway. PLoS One **7**: e49841.

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, Z., Peng, J., Wen, X., and Guo, H. (2013). ETHYLENE-INSENSITIVE3 Is a Senescence-Associated Gene That Accelerates Age-Dependent Leaf Senescence by Directly Repressing miR164 Transcription in Arabidopsis. Plant Cell **25**: 3311–3328.

Liebsch, D. and Keech, O. (2016). Dark-induced leaf senescence : new insights into a complex light-dependent regulatory pathway. New Phytol. **212**: 563–570.

Lin, R.-J., Chien, H.-L., Lin, S.-Y., Chang, B.-L., Yu, H.-P., Tang, W.-C., and Lin, Y.-L. (2017). MCPIP1 ribonuclease exhibits broad-spectrum antiviral effects through viral RNA binding and degradation. Nucleic Acids Res. **41**: 3314–3326.

Liu, J., Valencia-sanchez, M.A., Hannon, G.J., and Parker, R. (2005). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. Nat. Cell Biol. **7**: 719–723.

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

Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5 – SMG7 heterodimer directly recruits the CCR4 – NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev. **27**: 2125–2138.

van Loon, L.C., Geraats, B.P.J., and Linthorst, H.J.M. (2006). Ethylene as a modulator of disease resistance in plants. Trends Plant Sci. **11**: 184–191.

Lorenzo, L. De, Sorenson, R., Bailey-Serres, J., and Hunt, A.G. (2017). Noncanonical Alternative Polyadenylation Contributes to Gene Regulation in Response to Hypoxia. Plant Cell **29**: 1262–1277.

Loschi, M., Leishman, C.C., Berardone, N., and Boccaccio, G.L. (2009). Dynein and kinesin regulate stress-granule and P-body dynamics. J. Cell Sci. **122**: 3973–3982.

Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. EMBO J.: 1855–1868.

Lykke-Andersen, S. and Jensen, T.H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat. Rev. Mol. Cell Biol. **16**: 665–677.

Maldonado-Bonilla, L.D. (2014). Composition and function of P bodies in Arabidopsis thaliana. Front. Plant Sci. **5**: 201.

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.

Mao, R., Yang, R., Chen, X., Harhaj, E.W., Wang, X., and Fan, Y. (2017). Regnase-1, a rapid response ribonuclease regulating inflammation and stress responses. Cell. Mol. Immunol. **14**: 412–422.

Marzluff, W.F. and Koreski, K.P. (2017). Birth and Death of Histone mRNAs. Trends Genet. **33**: 745–759.

Matelska, D., Steczkiewicz, K., and Ginalski, K. (2017). Comprehensive classification of the PIN domain-like superfamily. Nucleic Acids Res. **45**: 6995–7020.

Matsushita, K., Takeuchi, O., Standley, D.M., Kumagai, Y., Kawagoe, T., Miyake, T., Satoh, T., Kato, H., Tsujimura, T., Nakamura, H., and Akira, S. (2009). Zc3h12a is an RNase essential for controlling immune responses by regulating mRNA decay. Nature **458**: 1185–1190.

McGlincy, N.J. and Smith, C.W.J. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem. Sci. **144**: 385–393.

Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum. Mol. Genet. **10**: 99–105.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet. **14**: 447–459.

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., Hu, Q., Spencer, K.R., Enriquez, P., Binder, B.M., Heber, S., Stepanova, A.N., and Alonso, J.M. (2015). Gene-Specific Translation Regulation Mediated by the Hormone-Signaling Molecule EIN2. Cell **163**: 684–697.

Metze, S., Herzog, V.A., Ruepp, M.-D., and Mühlemann, O. (2013). Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. RNA **3**: 1432–1448.

Miller, J.N. and Pearce, D.A. (2014). Nonsense-mediated decay in genetic disease : Friend or foe ? Mutat. Res. **762**: 52–64.

Min, E.E., Roy, B., Amrani, N., He, F., and Jacobson, A. (2013). Yeast Upf1 CH domain interacts with Rps26 of the 40S ribosomal subunit. RNA **19**: 1105–1115.

Mino, T. 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., Cluet, D., Robin, J.-P., Terme, J.-M., Dodon, M.D., Wittmann, J., Morris, C., Le Hir, H., Ciminale, V., and Jalinot, P. (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.

Moreno, A.B., Martinez de Alba, A.E., Bardou, F., Crespi, M.D., Vaucheret, H., Maizel, A., and Mallory,

A.C. (2013). Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-transcriptional gene silencing in plants. Nucleic Acids Res. **41**: 4699–4708.

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 Promotes Ribosome Dissociation. Mol. Cell. Biol. **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. Mol. Cell. Biol. **30**: 460–469.

Mugler, C.F., Hondele, M., Heinrich, S., Sachdev, R., Vallotton, P., Koek, A.Y., Chan, L.Y., and Weis, K. (2016). ATPase activity of the DEAD-box protein Dhh1 controls processing body formation. Elife **5**: e18746.

Mühlemann, O. and Lykke-Andersen, J. (2010). How and where are nonsense mRNAs degraded in mammalian cells? RNA Biol. **7**: 28–32.

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

Muhlrad, D. and Parker, R. (1999). Recognition of Yeast mRNAs as "Nonsense Containing" Leads to Both Inhibition of mRNA Translation and mRNA Degradation: Implications for the Control of mRNA Decapping. Mol. Biol. Cell **11**: 3971–3978.

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.

Nagy, E. and Maquat, L.E. (1998). A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Frontlines **4**: 198–199.

Nakano, K., Ando, T., Yamagishi, M., Yokoyama, K., Ishida, T., Ohsugi, T., Tanaka, Y., Brighty, D.W., and Watanabe, T. (2013). Viral interference with host mRNA surveillance , the nonsense-mediated mRNA decay (NMD) pathway , through a new function of HTLV-1 Rex: implications for retroviral replication. Microbes Infect. **15**: 491–505.

Narayanan, K. and Makino, S. (2013). Interplay between viruses and host mRNA degradation. BBA - Gene Regul. Mech. **1829**: 732–741.

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**: 30342–30347.

Nasim, Z., Fahim, M., and Ahn, J.H. (2017). Possible Role of MADS AFFECTING FLOWERING 3 and B-BOX DOMAIN PROTEIN 19 in Flowering Time Regulation of Arabidopsis Mutants with Defects in Nonsense-Mediated mRNA Decay. Front. Plant Sci. **8**: 191.

Nelson, J.O., Moore, K.A., Chapin, A., Hollien, J., and Metzstein, M.M. (2016). Degradation of Gadd45 mRNA by nonsense-mediated decay is essential for viability. Elife **5**: e12876.

Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010). Decapping Activators in Saccharomyces cerevisiae Act by Multiple Mechanisms. Mol. Cell **39**: 773–783.

Nyikó, T., Kerényi, F., Szabadkai, L., Benkovics, A.H., Major, P., Sonkoly, B., Mérai, Z., Barta, E., Niemiec, E., Kufel, J., and Silhavy, D. (2013). Plant nonsense-mediated mRNA decay is controlled by different autoregulatory circuits and can be induced by an EJC-like complex. Nucleic Acids Res. **41**: 6715–6728.

Ozgur, S., Basquin, J., Kamenska, A., Filipowicz, W., Standart, N., and Conti, E. (2015). Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex. Cell Rep. **13**: 703–711.

Paillusson, A., Hirschi, N., Vallan, C., Azzalin, C.M., and Mühlemann, O. (2005). A GFP-based reporter system to monitor nonsense-mediated mRNA decay. Nucleic Acids Res. **33**: e54.

Palusa, S.G. and Reddy, A.S.N. (2010). Extensive coupling of alternative splicing of pre-mRNAs of serine/arginine (SR) genes with nonsense-mediated decay. New Phytol. **185**: 83–89.

Park, E. and Maquat, L.E. (2013). Staufen-mediated mRNA decay. Wiley Interdiscip. Rev. RNA **4**: 423–435.

Park, O.H., Park, J., Yu, M., An, H.-T., Ko, J., and Kim, Y.K. (2016). Identification and molecular characterization of cellular factors required for glucocorticoid receptor-mediated mRNA decay. Genes Dev. **30**: 2093–2105.

Parker, R. and Sheth, U. (2007). P Bodies and the Control of mRNA Translation and Degradation. Mol. Cell **25**: 635–646.

Peixeiro, I., Inacio, A., Barbosa, C., Silva, A.L., Liebhaber, S.A., and Romão, L. (2012). Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. Nucleic Acids Res. **40**: 1160–1173.

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

Peng, J., Li, Z., Wen, X., Li, W., Shi, H., Yang, L., and Zhu, H. (2014). Salt-Induced Stabilization of EIN3/EIL1 Confers Salinity Tolerance by Deterring ROS Accumulation in Arabidopsis. PloS Genet. **10**: e1004664.

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., Alhusaini, N., Chen, Y., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., and Coller, J. (2015). Codon Optimality Is a Major Determinant of mRNA Stability. Cell **160**: 1111–1124.

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

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

Rajyaguru, P. and Parker, R. (2012). RGG motif proteins: Modulators of mRNA functional states. Cell Cycle **11**: 2594–2599.

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.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. RNA **11**: 1530–1544.

Reichel, M., Liao, Y., Rettel, M., Ragan, C., Evers, M., Alleaume, A.-M., Horos, R., Hentze, M.W., Preiss, T., and Millar, A.A. (2016). In Planta Determination of the mRNA-binding Proteome of Arabidopsis Etiolated Seedlings. Plant Cell **28**: 2435–2452.

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.

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.

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**: 616–624.

Rivière, M.-P., Marais, A., Ponchet, M., Willats, W., and Galiana, E. (2008). Silencing of acidic pathogenesis-related PR-1 genes increases extracellular β-(1→3)-glucanase activity at the onset of tobacco defence reactions. J. Exp. Bot. **59**: 1225–1239.

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.

Rühl, C., Stauffer, E., Kahles, A., Wagner, G., Drechsel, G., Rätsch, G., and Wachter, A. (2012). Polypyrimidine Tract Binding Protein Homologs from Arabidopsis Are Key Regulators of Alternative Splicing with Implications in Fundamental Developmental Processes. Plant Cell **24**: 4360–4375.

Ruiz-Echevarria, M.J. and Peltz, S.W. (2000). The RNA Binding Protein Pub1 Modulates the Stability of Transcripts Containing Upstream Open Reading Frames. Cell **101**: 741–751.

Sanchez, S.E., Petrillo, E., Kornblihtt, A.R., and Yanovsky, M.J. (2011). Alternative splicing at the right time. RNA Biol. **8**: 954–959.

Schaeffer, D., Tsanova, B., Barbas, A., Reis, F.P., Dastidar, E.G., Sanchez-Rotunno, M., Arraiano, C.M., and van Hoof, A. (2009). The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat. Struct. Mol. Biol. **16**: 56–62.

Scheer, H., Zuber, H., De Almeida, C., and Gagliardi, D. (2016). Uridylation Earmarks mRNAs for Degradation... and More. Trends Genet. **32**: 607–619.

Schell, T., Köcher, T., Wilm, M., Seraphin, B., Kulozik, A.E., and Hentze, M.W. (2003). Complexes between the nonsense-mediated mRNA decay pathway factor human upf1 (up-frameshift protein 1) and essential nonsense-mediated mRNA decay factors in HeLa cells. Biochem. J. **783**: 775–783.

Schneider, C., Kudla, G., Wlotzka, W., Tuck, A., and Tollervey, D. (2012). Transcriptome-wide Analysis of Exosome Targets. Mol. Cell **48**: 422–433.

Schneider, C., Leung, E., Brown, J., and Tollervey, D. (2009). The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res. **37**: 1127–1140.

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.

Schwartz, D.C. and Parker, R. (1999). Mutations in Translation Initiation Factors Lead to Increased Rates of Deadenylation and Decapping of mRNAs in Saccharomyces cerevisiae. Mol. Cell. Biol. **19**: 5247–5256.

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.

Serquiña, A.K.P., Das, S.R., Popova, E., Ojelabi, O.A., Roy, C.K., and Göttlinger, G. (2013). UPF1 Is Crucial for the Infectivity of Human Immunodeficiency Virus Type 1 Progeny Virions. J. Virol. **87**: 8853–8861.

Sheth, U. and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science (80-.). **300**: 805–808.

Sheth, U. and Parker, R. (2006). Targeting of aberrant mRNAs to cytoplasmic processing bodies. Cell **125**: 1095–1109.

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. Biochem. J. **129**: 119–129.

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

Singh, G., Rebbapragada, I., and Lykke-andersen, J. (2008). A Competition between Stimulators and Antagonists of Upf Complex Recruitment Governs Human Nonsense-Mediated mRNA Decay. PLoS Biol. **6**: e111.

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.

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.

Stalder, L. and Mühlemann, 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.

Steffens, A., Jaegle, B., Tresch, A., Hülskamp, M., and Jakoby, M. (2014). Processing-Body Movement in Arabidopsis Depends on an Interaction between Myosins and DECAPPING PROTEIN1. Plant Physiol. **164**: 1879–1892.

Stoll, B. and Binder, S. (2016). Two NYN domain containing putative nucleases are involved in transcript maturation in Arabidopsis mitochondria. Plant J. **85**: 278–288.

Su, Y.-Q., Sun, F., Handel, M.A., Schimenti, J.C., and Eppig, J.J. (2012). Meiosis arrest female 1 (MARF1) has nuage-like function in mammalian oocytes. Proc. Natl. Acad. Sci. **109**: 18653–18660.

Suzuki, H.I., Arase, M., Matsuyama, H., Choi, Y.L., Ueno, T., Mano, H., Sugimoto, K., and Miyazono, K. (2011). MCPIP1 Ribonuclease Antagonizes Dicer and Terminates MicroRNA Biogenesis through Precursor MicroRNA Degradation. Mol. Cell **44**: 424–436.

Sweet, T., Kovalak, C., and Coller, J. (2012). The DEAD-Box Protein Dhh1 Promotes Decapping by Slowing Ribosome Movement. PLoS Biol. **10**: e1001342.

Swisher, K.D. and Parker, R. (2011). Interactions between Upf1 and the Decapping Factors Edc3 and Pat1 in Saccharomyces cerevisiae. PLoS One **6**: e26547.

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.

Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Drazkowska, K., Pastula, A., Andersen, J.S., Stepien, P.P., Dziembowski, A., and Jensen, T.H. (2010). The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. EMBO J. **29**: 2342–2357.

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. Mol. Cell **33**: 661–668.

Tsai, H.-Y., Chen, C.-C.G., Conte, D., Moresco, J.J., Chaves, D.A., Mitani, S., Yates III, J.R., Tsai, M.-D., and Mello, C.C. (2015). A Ribonuclease Coordinates siRNA Amplification and mRNA Cleavage

during RNAi. Cell **160**: 407–419.

Tutucci, E., Vera, M., Biswas, J., Garcia, J., Parker, R., and Singer, R.H. (2017). An improved MS2 system for accurate reporting of the mRNA life cycle. Nat. Publ. Gr.: doi: 10.1038/nmeth.4502.

Unterholzner, L. and Izaurralde, E. (2004). SMG7 Acts as a Molecular Link between mRNA Surveillance and mRNA Decay. Mol. Cell **16**: 587–596.

Ustianenko, D., Hrossova, D., Potesil, D., Chalupnikova, K., Hrazdilova, K., Pachernik, J., Cetkovska, K., Uldrijan, S., Zdrahal, Z., and Vanacova, S. (2013). Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. RNA **19**: 1632–1638.

Vexler, K., Cymerman, M.A., Berezin, I., Fridman, A., Golani, L., Lasnoy, M., Saul, H., and Shaul, O. (2016). The Arabidopsis NMD Factor UPF3 Is Feedback-Regulated at Multiple Levels and Plays a Role in Plant Response to Salt Stress. Front. Plant Sci. **7**: 1376.

Wachter, A., Rühl, C., and Stauffer, E. (2012). The role of polypyrimidine tract-binding proteins and other hnRNP proteins in plant splicing regulation. Front. Plant Sci. **3**: 81.

Weber, S.C. and Brangwynne, C.P. (2012). Getting RNA and Protein in Phase. Cell **149**: 1188–1191.

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.

Weil, J.E., Hadjithomas, M., and Beemon, K.L. (2009). Structural characterization of the Rous sarcoma virus RNA stability element. J. Virol. **83**: 2119–29.

Wollerton, M.C., Gooding, C., Wagner, E.J., Garcia-blanco, M.A., and Smith, C.W.J. (2004). Autoregulation of Polypyrimidine Tract Binding Protein by Alternative Splicing Leading to Nonsense-Mediated Decay. Mol. Cell **13**: 91–100.

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.

Xu, J. and Chua, N.-H. (2011). Processing bodies and plant development. Curr. Opin. Plant Biol. **14**: 88–93.

Yamashita, A., Chang, T., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.A., and Shyu, A. (2005). Concerted action of poly (A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat. Struct. Mol. Biol. **12**: 1054–1063.

Yamashita, A. and Takeuchi, O. (2017). Translational control of mRNAs by 3’-Untranslated region binding proteins. BMB Rep. **50**: 194–200.

Yang, E., Nimwegen, E. Van, Zavolan, M., Rajewsky, N., Schroeder, M., Magnasco, M., and Darnell Jr, J.E. (2003). Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes. Genome Res. **13**: 1863–1872.

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.

Yokogawa, M., Tsushima, T., Noda, N.N., Kumeta, H., Enokizono, Y., Yamashita, K., Standley, D.M., Takeuchi, O., Akira, S., and Inagaki, F. (2016). Structural basis for the regulation of enzymatic activity of Regnase-1 by domain-domain interactions. Sci. Rep. **6**: 22324.

Yu, X., Willmann, M.R., Anderson, S.J., and Gregory, B.D. (2016). Genome-Wide Mapping of Uncapped and Cleaved Transcripts Reveals a Role for the Nuclear mRNA Cap-Binding Complex in Cotranslational RNA Decay in Arabidopsis. Plant Cell **28**: 2385–2397.

Zakrzewska-Placzek, M., Souret, F.F., Sobczyk, G.J., Green, P.J., and Kufel, J. (2010). Arabidopsis thaliana XRN2 is required for primary cleavage in the pre-ribosomal RNA. Nucleic Acids Res. **38**: 4487–4502.

Zhang, W., Murphy, C., and Sieburth, L.E. (2010). Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. PNAS **107**: 15981–15985.

Zhang, X. et al. (2015). Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. Science (80-.). **348**: 120–124.

Zhang, Z., Hu, L., and Kong, X. (2013). MicroRNA or NMD: Why Have Two RNA Silencing Systems? J. Genet. Genomics **40**: 497–513.

Zuber, H., Scheer, H., Ferrier, E., Sement, F.M., Mercier, P., Stupfler, B., and Gagliardi, D. (2016). Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in Arabidopsis. Cell Rep. **14**: 2707–2717.

Université de Strasbourg

Clara CHICOIS

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

Study of the interactome of UPF1, a key factor of Nonsense-Mediated Decay in *Arabidopsis thaliana*

Résumé

L'ARN hélicase UPF1 est un facteur clé du Nonsense-Mediated Decay (NMD), un mécanisme impliqué dans le contrôle de la qualité des ARNm et la régulation de l'expression des gènes. Malgré d'importantes fonctions chez les plantes, le NMD y est peu décrit. Cette thèse présente l'indentification et l'étude des protéines interagissant avec UPF1 chez *Arabidopsis*. Nous avons identifié un nouveau réseau d'interaction protéine-protéine entre UPF1 et des répresseurs de traduction dans les P-bodies. Nous proposons un modèle dans lequel la répression traductionnelle exerce une action protectrice sur les cibles du NMD. Notre approche a également identifié de nouveaux composants des P-bodies, comme l'endonucléase UCN. Son étude détaillée a révélé un lien direct avec la machinerie de decapping ainsi que de possibles rôles dans la signalisation hormonale ou les mécanismes de défense, suggérant que la modulation de l'expression d'UCN pourrait influencer d'importantes caractéristiques agronomiques. Ce travail décrit des facteurs associés à UPF1 jusqu'alors inconnus, leur étude permettra de découvrir de nouveaux mécanismes impliqués dans l'équilibre entre la traduction, le stockage et la dégradation des ARNm chez les plantes.

Mots-clés : NMD, UPF1, répression traductionnelle, *Arabidopsis thaliana*

Abstract

The RNA helicase UPF1 is a key factor of Nonsense-Mediated Decay (NMD), a paneukaryotic mechanism involved in mRNA quality control and fine-tuning of gene expression. Despite important biological functions in plants, NMD is poorly described compared to other eukaryotes. This thesis presents the identification and study of UPF1 interacting proteins in *Arabidopsis*. Using approaches based on immunoaffinity and mass spectrometry, we identified a novel protein-protein interaction network between UPF1 and translation repressors in P-bodies. We propose a model in which translation repression exerts a protective action on NMD targets in plants. Our approach also identified novel P-body components, including the UCN endonuclease. A detailed study revealed its direct link with the decapping machinery and possible roles in hormone signaling and defense mechanisms, suggesting that the modulation of UCN expression could influence important agronomical traits. This work describes hitherto unknown UPF1 associated factors, their study will provide novel insights into the mechanisms involved in the balance between mRNA translation, storage and decay in plants.

Key words: NMD, UPF1, translation repression, *Arabidopsis thaliana*