

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

Institut de Biologie Moléculaire des Plantes – CNRS – UPR2357

# THÈSE

Présentée par :

**Marlene SCHIAFFINI**

Soutenue le 12 Juin 2020

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

Discipline / Spécialité : Aspects Moléculaires et Cellulaires de la Biologie

**Étude du rôle d'une nouvelle  
endoribonucléase à domaine NYN dans la  
dégradation des ARN chez *Arabidopsis  
thaliana***

Devant le jury composé de :

Rapporteurs :

**Pr Peter BRODERSEN**  
**Dr Thierry LAGRANGE**

Université de Copenhague  
CNRS, université de Perpignan

Examinatrice :

**Dr Magali FRUGIER**

CNRS, université de Strasbourg

Directeur de thèse :

**Dr Damien GARCIA**

CNRS, université de Strasbourg

# REMERCIEMENTS

Tant de personnes m'ont aidé à avancer au cours de ces quatre ans de thèse qu'il me serait bien difficile de les remercier tous. Ces remerciements sont donc non exhaustifs, mais je n'oublierai aucune des rencontres faites au cours de ces années.

Tout d'abord merci aux membres du jury Pr Peter Brodersen, Dr Magali Frugier et Dr Thierry Lagrange pour avoir accepté d'évaluer et discuter mon travail. Merci au LabEX NetRNA pour avoir financé ma thèse.

Toute ma gratitude pour Damien, mon directeur de thèse, merci de m'avoir guidé avec patience et bienveillance à travers les joies et les doutes propres à une thésarde en biologie.

Merci à Gag, pour m'avoir accueilli dans l'équipe, pour ton aide et tes conseils précieux.

Merci Hélène pour avoir toujours été présente, à l'écoute et pour tes sages suggestions. Merci à Heike pour tes conseils et ton humour aux nuances uniques.

Un grand merci à Caroline et Hélène, c'était vraiment un plaisir pour moi de partager tous ces moments dans le labo et en dehors avec vous. On est un trio d'enfer ! Merci les filles, les Gag's girls, pour la bonne ambiance, la bonne musique pendant les journées à la paillasse et pour le respect des moments de solitude. Merci Caroline pour ta gentillesse, tu m'as beaucoup aidé scientifiquement et moralement pendant ces quatre ans. Merci Hélène pour avoir partagé avec moi tes connaissances scientifiques, pour les fous rires et pour m'avoir donné ce beau surnom...

Merci à Clara pour m'avoir aidé au début de la thèse. Merci à Aude pour ton aide précieux avec les manip pendant la dernière année. Je te souhaite le meilleur pour ta carrière scientifique !

Merci à Marie-Edith, Kamel, Philippe, Sébastien, Anthony et Todd pour les enrichissantes discussions scientifiques. Merci à Philippe, Lauriane et Johanna de la plateforme de protéomique de l'IBMC, à Sandrine de la plateforme d'expression génique et à Nicolas de la plateforme de production des protéines.

Merci à tout le personnel de l'IBMP des services techniques et de l'administration pour l'aide et la gentillesse avec laquelle vous m'avez traitée pendant ces années. Merci à Laurence du magasin, pour ta disponibilité et gentillesse, nos discussions vont me manquer !

Merci à mes amis Guillaume et Pierre pour tous les bons moments passés ensemble. Merci Arnaud pour m'avoir soutenu et aidé à garder un niveau de théine ;) constante pendant l'écriture.

Grazie a Simone, al mio fianco dal primo giorno, ma anche prima, di questa lunga avventura. Grazie perché senza di te non sarei qui. Grazie a Davide che sa sempre come tirarmi su di morale, per il suo aiuto costante, soprattutto negli ultimi mesi di redazione. Un ringraziamento speciale alle mie amiche di sempre Francesca, Sabrina e Giulia, sempre vicine nonostante la distanza. Ringrazio con amore i miei genitori, Mauro e Anna, per il loro sostegno immanente e la loro fiducia, e la mia sorellina, Isotta, sempre e per sempre mia complice e ispirazione. Grazie al mio caro Alexandre per la pazienza, per avermi supportato e sopportato, per non aver mai dubitato nelle mie capacità e per avermi preso per mano nelle giornate uggiose, rischiarandole con il tuo sorriso e la tua musica.

"(...) sic nos in luce timemus

interdum, nilo quae sunt metuenda magis quam  
quae pueri in tenebris pavitant finguntque futura.

Hunc igitur terrorem animi tenebrasque necessest

non radii solis neque lucida tela diei

discutiant, sed naturae species ratioque."

Lucrezio, *De rerum natura* Libro II (vv. 56-61)

## Table des matières

|  |          |
|--|----------|
| <b>Introduction .....</b>  | <b>5</b> |
| <b>1.1 Structure des ARNm.....</b>   | <b>7</b> |
| 1.1.1 La coiffe.....   | 7        |
| 1.1.2 La queue poly(A).....  | 8        |
| <b>1.2 Voie générales de dégradation des ARN messagers.....</b>  | <b>9</b> |
| 1.2.1 La déadénylation.....  | 9        |
| 1.2.2 L'uridylation.....   | 11       |
| 1.2.3 La dégradation exoribonucléolytique 3'-5'.....   | 12       |
| L'exosome.....   | 12       |
| DIS3L2/SOV.....  | 12       |
| 1.2.4 La voie de dégradation 5'-3' des ARNm .....  | 13       |
| 1.2.4.1 Le decapping .....   | 13       |
| Fonction biologique du decapping.....  | 14       |
| L'enzyme de decapping DCP2.....  | 15       |
| Les activateurs de decapping.....  | 16       |
| DCP1 (Decapping Protein 1) .....   | 16       |
| Ge-1/VCS .....   | 17       |
| Le complexe Pat1/LSm1-7 .....  | 17       |
| LSm14/Scd6/DCP5 .....  | 18       |
| DDX6/Dhh1/RH6, 8, 12.....  | 18       |
| Le cas particulier des enzymes DXO (decapping exonuclease) .....   | 19       |
| 1.2.4.2 La Dégradation exoribonucléolytique 5'-3' .....  | 20       |
| 1.2.5 Le Decapping Independent de la déadénylation .....   | 20       |
| 1.2.5.1 Le Nonsense Mediated Decay (NMD).....  | 21       |
| 1.2.5.2 Le Staufen Mediated Decay .....  | 23       |
| 1.2.5.3 La dégradation dépendante de la réplication des ARNm codant pour les histones chez les métazoaires ..... | 24       |
| 1.2.6 Dégradation dépendante d'un clivage endoribonucléolytique .....  | 24       |
| 1.2.6.1 Le RNA silencing dans la régulation de l'expression des gènes .....                                      | 25       |
| Les microARN.....  | 25       |
| Les trans-acting siRNA.....  | 26       |
| Compétition entre RNA silencing et la voie de dégradation 5'-3' des ARN.....                                     | 26       |
| 1.2.6.1 Les endoribonucléase à domaine PIN/NYN.....  | 27       |
| SMG6.....  | 28       |

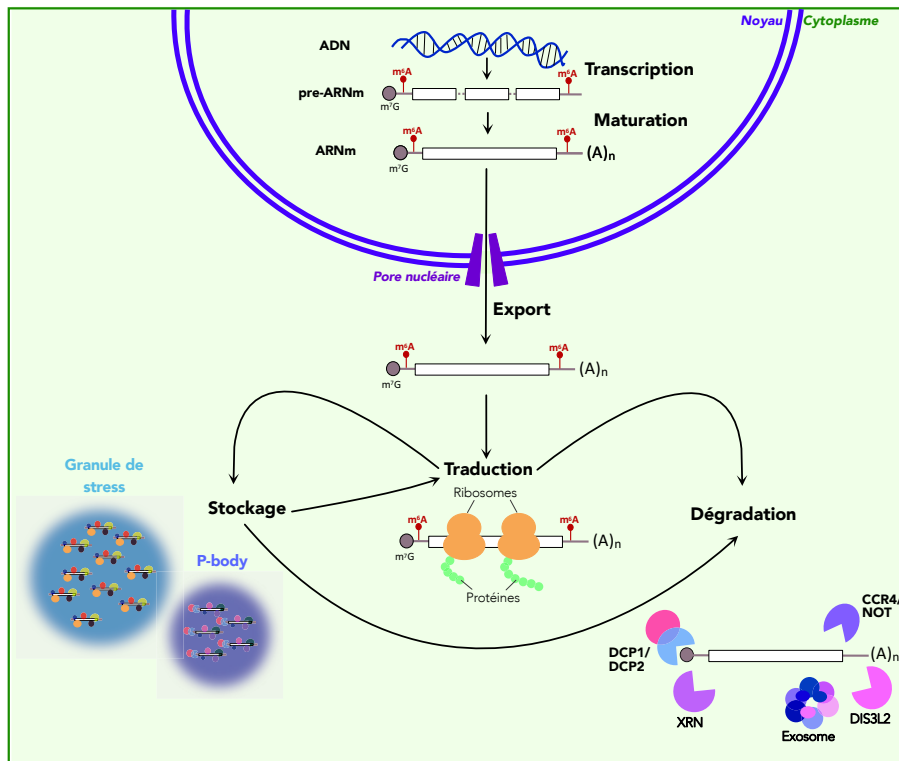
|  |           |
|--|-----------|
| PRORP.....   | 29        |
| MCPIP1/Regnase.....  | 29        |
| RDE-8.....   | 30        |
| MARF1/LimkainB.....  | 30        |
| MNU1/MNU2.....   | 31        |
| <b>1.3 Les granules d'ARN cytosoliques et leur rôle dans le métabolisme des ARNm .....</b>                                       | <b>32</b> |
| 1.3.1 Les granules de stress .....   | 33        |
| 1.3.2 Les processing bodies.....   | 34        |
| <b>Objectifs.....</b>  | <b>37</b> |
| <b>Résultats .....</b>   | <b>38</b> |
| 1.4 DAE co-purifie avec DCP1 et VCS, composants du complexe de decapping .....   | 39        |
| 1.5 DAE est conservée chez les végétaux des algues jusqu'aux plantes à fleurs .....  | 40        |
| 1.6 Le plus proche homologue de DAE chez les mammifères est l'endonucléase MARF1 .....   | 41        |
| 1.7 L'expression du gène DAE est perturbée dans les mutants dae-1, dae-2 et dae-3 .....  | 42        |
| 1.8 DAE co-localise avec DCP1 dans les p-bodies .....  | 43        |
| 1.9 La partie N-terminale de DAE contenant le domaine NYN entraîne un localisation<br>spécifique dans les p-bodies.....          | 44        |
| 1.10 DAE possède une activité ribonucléase.....  | 45        |
| 1.11 DAE est une endoribonucléase .....  | 46        |
| 1.12 DAE diminue l'accumulation d'un ARNm co-exprimé .....   | 47        |
| 1.13 DAE interagit directement avec DCP1 indépendamment de son activité catalytique .....  | 49        |
| 1.14 DCP2 co-purifie avec VCS et VCR.....  | 50        |
| 1.15 DAE est présent dans un complexe protéique basé autour de DCP1 et VCS mais ne<br>contenant pas DCP2 .....                   | 51        |
| 1.16 La surexpression de GFP-DAE et GFP-D153N entraîne des défauts de développement ...  | 52        |
| 1.17 Analyse de l'impact de la dérégulation de DAE sur le transcriptome .....  | 53        |
| 1.18 La surexpression de DAE influence l'accumulation de gènes sensibles à un dérégulation<br>de la machinerie du decapping..... | 54        |
| <b>Discussion &amp; Perspectives .....</b>   | <b>55</b> |

|                                       |  |           |
|---------------------------------------|--|-----------|
| <b>1.19</b>                           | <b>DAE une endoribonuclease dans la dégradation des ARNm .....</b> | <b>56</b> |
| 1.19.1                                | DAE est une endoribonuclease active.....                           | 56        |
| 1.19.2                                | L'interactome de DAE .....   | 58        |
| 1.19.3                                | DAE, ses liens avec la dégradation des ARNm .....                  | 59        |
| 1.19.4                                | Un modèle pour le mode d'action de DAE .....                       | 61        |
| 1.19.5                                | Quelle est la fonction biologique de DAE? .....                    | 62        |
| <b>Conclusion.....</b>                |  | <b>63</b> |
| <b>Matériels &amp; Méthodes .....</b> |  | <b>64</b> |
| <b>1.20</b>                           | <b>Souches bactériennes .....</b>                                  | <b>65</b> |
| <b>1.21</b>                           | <b>Matériel végétal.....</b>                                       | <b>65</b> |
| <b>1.22</b>                           | <b>Plasmides.....</b>  | <b>66</b> |
| <b>1.23</b>                           | <b>Méthodes relatives à l'ADN .....</b>                            | <b>68</b> |
| 1.23.1                                | Extraction d'ADN génomique .....                                   | 68        |
| 1.23.2                                | Amplification par PCR .....  | 68        |
| 1.23.3                                | Analyse des produits PCR par électrophorèse sur gel d'agarose..... | 68        |
| 1.23.4                                | Purification de l'ADN.....   | 69        |
| 1.23.5                                | Marquage de l'ADN avec des nucléotides radioactifs.....            | 69        |
| <b>1.24</b>                           | <b>Techniques de clonage.....</b>                                  | <b>70</b> |
| <b>1.25</b>                           | <b>Méthodes relatives à l'ARN.....</b>                             | <b>72</b> |
| 1.25.1                                | Transcription <i>in vitro</i> .....                                | 72        |
| 1.25.2                                | Extraction d'ARN.....  | 72        |
| 1.25.3                                | QPCR.....  | 73        |
| 1.25.4                                | Chromatographie en couche mince (CCM) .....                        | 73        |
| 1.25.5                                | Northern blot.....   | 74        |
| 1.25.6                                | Analyses RNAseq .....  | 75        |
| <b>1.26</b>                           | <b>Méthodes relatives aux protéines .....</b>                      | <b>77</b> |
| 1.26.1                                | Extraction des protéines .....                                     | 77        |
| 1.26.2                                | Western blot.....  | 77        |
| 1.26.3                                | Immunoprécipitation et spectrométrie de masse .....                | 78        |
| 1.26.4                                | Production et purification des protéines recombinantes .....       | 79        |
| 1.26.5                                | Test d'activité enzymatique .....                                  | 81        |
| 1.26.6                                | Double-hybride de levure.....                                      | 81        |
| <b>1.27</b>                           | <b>Méthodes relatives à la manipulation des plantes.....</b>       | <b>82</b> |

|                             |  |           |
|-----------------------------|--|-----------|
| 1.27.1                      | Agroinfiltration des feuilles de <i>N.Benthamiana</i> .....          | 82        |
| 1.27.2                      | Transformation stable d' <i>A.thaliana</i> par trempage floral ..... | 82        |
| 1.27.3                      | Culture des plantules <i>in vitro</i> .....                          | 83        |
| <b>1.28</b>                 | <b>Microscopie confocale .....</b>                                   | <b>83</b> |
| <b>1.29</b>                 | <b>Arbre phylogénétique .....</b>                                    | <b>83</b> |
| <b>1.30</b>                 | <b>Liste Amorces.....</b>  | <b>84</b> |
| <b><i>Bibliographie</i></b> | <b>.....</b>   | <b>87</b> |

# INTRODUCTION



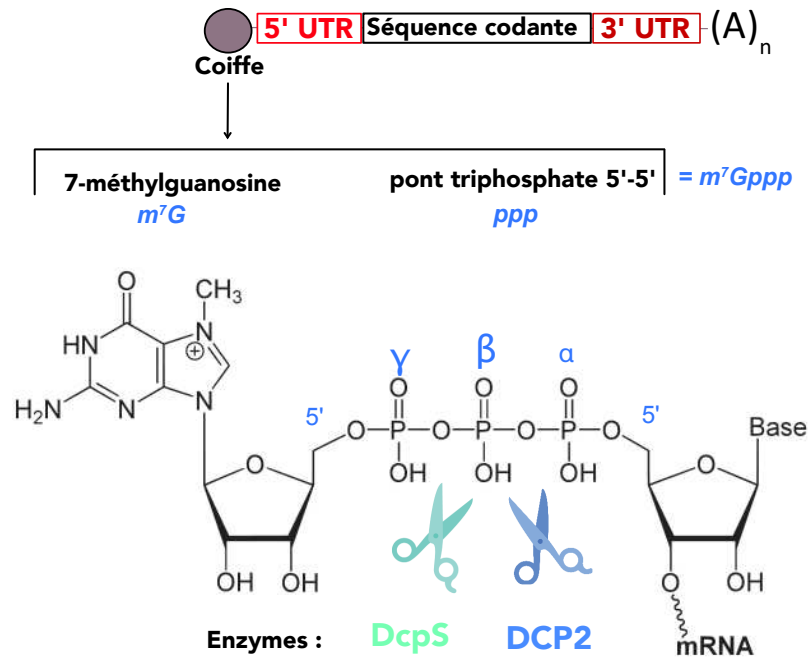


**Figure 1** Le destin d'une molécule d'ARN messager dans une cellule eucaryote (Inspiré de Das et al., 2017)

La première étape de la vie d'un ARNm est la transcription dans le noyau à partir de la matrice à ADN. Le pré-ARNm est modifié co-transcriptionnellement avec l'ajout de la coiffe en 5' et, éventuellement, avec la méthylation en position 6 des adénines ensuite la queue de poly(A) est ajoutée en 3', la structure des ARN est décrite dans le **chapitre 1.1**. Après l'épissage l'ARNm mature est exporté vers le cytoplasme où il peut s'associer aux ribosomes pour être traduit en protéine. La traduction de l'ARNm peut être inhibée dans les granules à ARN, la composition et la fonction des granules de stress et des p-bodies sont décrites dans le **chapitre 1.3**. Le transcrit peut sortir de ces compartiments pour être ensuite dégradé ou bien pour s'associer à nouveau avec les ribosomes, les voies générales de dégradation des ARNm sont décrites dans le **chapitre 1.2**.

Les acides ribonucléiques messagers ou ARNm sont des molécules qui ont une fonction cruciale au sein de la cellule. En effet ceux-ci sont nécessaires à la production des protéines indispensables au fonctionnement cellulaire. Les ARNm ont pour rôle de transmettre l'information génétique contenue dans le génome à la machinerie cellulaire responsable de la traduction, qui utilise les ribosomes pour décoder cette séquence et produire toutes les protéines cellulaires. Les ARNm sont synthétisés dans le noyau au cours de la transcription, maturés et ensuite exportés dans le cytosol, où a lieu la traduction (**Figure 1**). Afin de pouvoir répondre efficacement à des modifications de l'environnement ou faire évoluer le transcriptome au cours du développement, les ARNm doivent être éliminés par des mécanismes spécifiques. Ainsi, la régulation de la dégradation des ARNm est un mécanisme fondamental dans le maintien de l'homéostasie cellulaire. Les ARNm peuvent être aussi réprimés au niveau traductionnel et stockés sous forme de complexes ribonucléoprotéiques (RNP) dans des granules cytosoliques (**Figure 1**). Ceux-ci, formés par transition de phase liquide-liquide (Standart and Weil, 2018), incluent notamment les P-bodies et les granules de stress. Ces compartiments cytosoliques contiennent un grand nombre de protéines de liaison à l'ARN ainsi que des protéines impliquées dans leur dégradation ou l'inhibition de la traduction (Chantarachot and Bailey-Serres, 2018).

Dans les paragraphes qui suivent je vais aborder l'état des connaissances concernant la dégradation des ARNm et ses rôles biologiques, avec une emphase particulière sur la biologie des plantes, car elle représente mon modèle d'étude. Je m'attarderai également plus en détails sur les mécanismes liés au decapping et à l'action des endoribonucléases qui sont directement liés aux problématiques soulevées durant ma thèse. Dans ce manuscrit j'utiliserai le terme anglo-saxon decapping pour décrire les mécanismes liés au retrait de la coiffe des ARNm.



**Figure 2** Structure des ARNm et de la coiffe canonique  
(Adapté de Wojtczak et al., 2018)

Un ARNm mature contient une séquence codante flanquée des UTR (Untranslated Transcribed Region) 5' et 3'. Les extrémités 5' et 3' sont respectivement protégées par la coiffe et la queue poly(A). La coiffe est composée de 7- méthyle guanosine triphosphate liée au premier nucléotide transcrit avec une liaison 5'-5'.

## 1.1 STRUCTURE DES ARNm

Chez les eucaryotes la structure d'une molécule d'ARNm mature est la suivante : la coiffe en 5'; une région non traduite, la 5' UTR (Untranslated Transcribed Region); la séquence codante; une autre région non traduite, la 3' UTR; enfin une séquence homopolymérique d'adénine ou queue poly(A) (**Figure 2**).

### 1.1.1 LA COIFFE

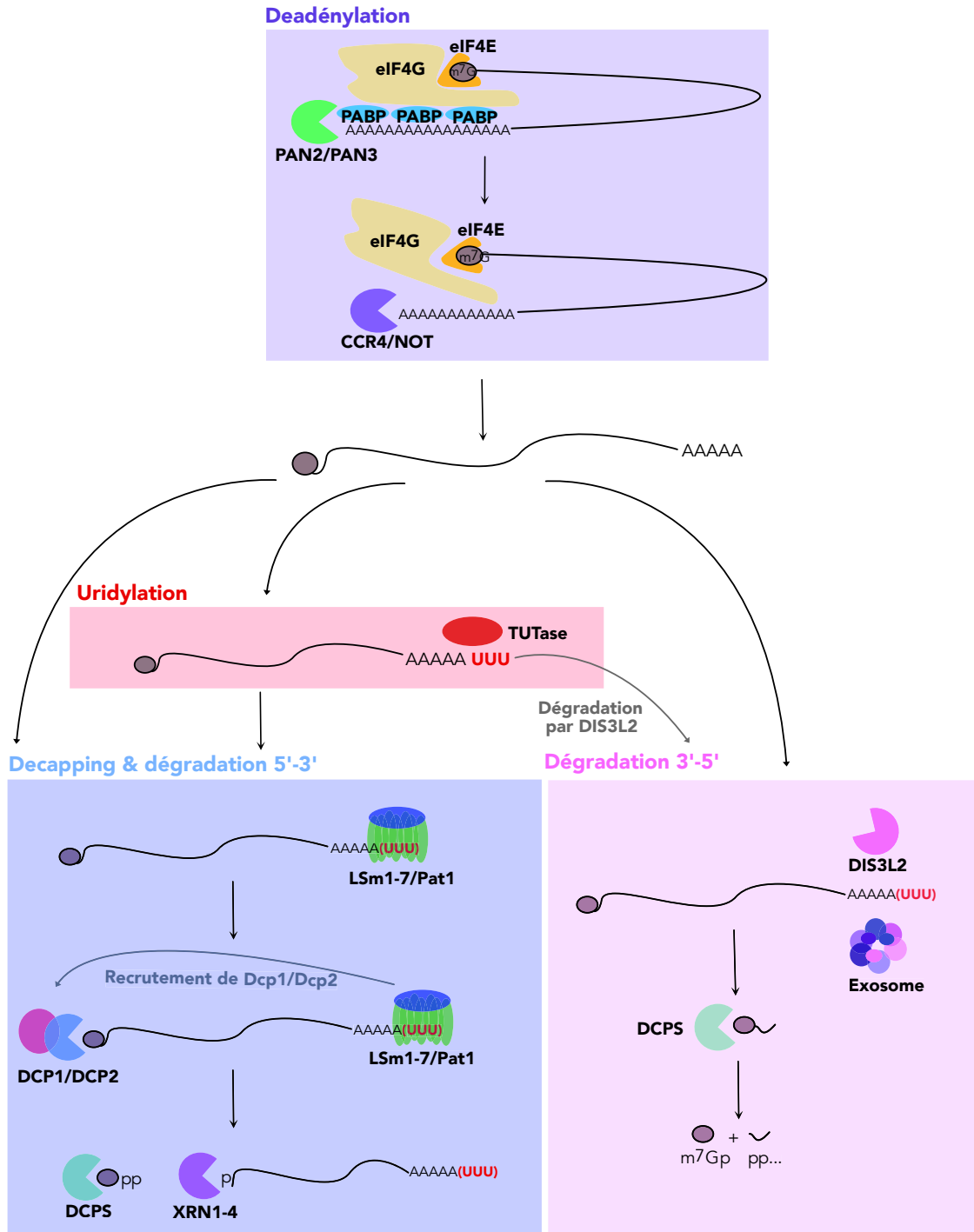
Chez les eucaryotes la coiffe est constituée par un nucléotide modifié 7-méthyl guanosine ( $m^7G$ ) à l'extrémité 5' des ARNm (**Figure 2**). Celui-ci est ajouté de façon co-transcriptionnelle lorsque le transcrit a une taille de ~20 nucléotides (Ramanathan et al., 2016). La  $m^7G$  est liée au 1<sup>er</sup> nucléotide de l'ARN par une liaison 5'-5'. Cette modification est introduite par l'action de réactions enzymatiques séquentielles. Les enzymes responsables sont recrutées par le domaine C-terminal de la ARN polymérase ADN-dépendante II (PolII). Une ARN triphosphatase catalyse le retrait du phosphate- $\gamma$  de l'extrémité 5' de l'ARNm, ensuite la formation de la liaison 5'-5' est catalysée par une guanylyltransférase à partir du GTP, le méthyle est ensuite ajouté en position 7 par la méthyltransférase en utilisant la S-Adénosylméthionine (Ramanathan et al., 2016). La fonction de la coiffe ne se limite pas à la protection du transcrit de la dégradation par les exoribonucléases 5'-3'. Cette structure participe à la régulation de toutes les étapes de la vie d'un ARNm comme l'épissage, la polyadénylation, l'export du noyau et la traduction. Cette régulation est possible via l'interaction de la coiffe avec des protéines ou des complexes protéiques. Dans le noyau la coiffe est associée au complexe nCBC (Nuclear cap binding complex) composé des sous-unités Cbp20 (Cap Binding Protein 20) et Cbp80. Cbp20 interagit directement avec la  $m^7G$ , Cbp80 stabilise cette interaction (Fortes et al., 1999; Kierzkowski et al., 2009). Le complexe nCBC permet le recrutement de facteurs nécessaires pour la maturation du transcrit, ainsi que pour son export du noyau au cytosol (Gonatopoulos-Pournatzis and Cowling, 2014). Par exemple chez les mammifères le complexe nCBC favorise l'assemblage du spliceosome, la machinerie nucléaire responsable de l'épissage du pré-ARNm, par interaction avec les petites ribonucleoprotéines nucléaires (U snRNP) qui en font partie (Lewis et al., 1996; Pabis et al., 2013). Dans le cytosol, lors de l'initiation de la traduction, le recrutement du ribosome sur l'ARNm nécessite l'interaction entre la coiffe et le facteur d'initiation de la traduction eIF4E, qui fait partie du complexe hétérotrimérique eIF4F avec l'hélicase eIF4A et la protéine "scaffold" eIF4G (Shirokikh and Preiss, 2018). Cependant chez les mammifères l'initiation de la traduction dépendante de la coiffe peut avoir lieu indépendamment de eIF4E par interaction entre l'isoforme d'eIF4G, eIF4G2 (DAP5), et le facteur d'initiation de la traduction eIF3D (de la Parra et al., 2018). On peut noter ici que certaines caractéristiques de l'ARNm, comme par exemple les

IRES (Internal Ribosomal Entry Site) ou des modifications internes de l'ARNm comme les N6-méthyl-adenosine (m<sup>6</sup>A), peuvent favoriser la traduction indépendamment de la coiffe (Coots et al., 2017).

Il est intéressant de mentionner qu'il existe dans la plupart des organismes des ARNm dont l'extrémité 5' est constituée de nicotinamide adénine dinucléotide (NAD<sup>+</sup>). En effet la nicotinamide adénine diphosphate dans sa forme oxydée peut être utilisée par PolII comme nucléotide d'initiation non canonique *in vivo* et *in vitro* (Bird et al., 2016; Julius and Yuzenkova, 2017). Cependant cette modification n'est pas très fréquente chez les eucaryotes. Chez *Arabidopsis thaliana* les ARNm NAD<sup>+</sup> représentent moins de 1% des transcrits (Zhang et al., 2019).

### 1.1.2 LA QUEUE POLY(A)

La région 3' des ARNm est constituée par un homopolymère d'adénines. Dans la phase d'élongation, la transcription du signal de polyadénylation par PolII active le recrutement des complexes protéiques CPSF (cleavage/polyadenylation specificity factor) chez l'humain, CPF chez la levure, qui induisent la coupure du pré-ARNm et l'ajout de la queue poly(A). La polyadénylation est catalysée par les poly(A) polymérase canoniques (cPAP) et elle est impliquée dans la régulation de la stabilité, de la traduction, de l'épissage et du transport nucléo-cytosolique (Shatkin and Manley, 2000). L'augmentation de l'efficacité de la traduction est due à la liaison entre les Poly(A) Binding Protein (PABP) et le facteur d'initiation de la traduction eIF4G qui induit la circularisation de l'ARNm (Nicholson and Pasquinelli, 2019; Bicknell and Ricci, 2017). Une queue de poly(A) courte a été pendant longtemps associée à une instabilité de l'ARNm qui la possède. Récemment, grâce à l'amélioration des protocoles d'analyse des queues poly(A) par différentes techniques de séquençage haut débit, il a été possible d'avoir une idée plus précise de la taille de la queue de poly(A) qui est en moyenne bien plus courte de ce qui avait été proposé, avec une taille médiane entre 50 et 100 nucléotides chez les mammifères et *C. elegans* et d'environ 30 nt chez la levure (Nicholson and Pasquinelli, 2019; Castellano and Bazzini, 2017; Lima et al., 2017). De plus il a été montré que chez les mammifères la taille de la queue de poly(A) de la  $\beta$ -globine, un transcrit très stable, est de 30 nt indiquant que une queue de poly(A) courte n'est pas forcément indicatif de l'instabilité de l'ARN qui la possède. Une étude récente montre que cette tendance est générale et que les ARNm fortement traduits ont une queue de poly(A) plutôt courte (<70nt ; (Lima et al., 2017)). Il y a cependant des exceptions, dans les embryons de *X. laevis* et *D. rerio*, les ARNm avec une queue poly(A) longue sont mieux traduits (Subtelny et al., 2014). La corrélation entre la taille de la queue poly(A) et le niveau de traduction de l'ARNm existe donc mais peut varier en fonction du stade de développement.



**Figure 3 Les voies de dégradation générales des ARNm chez les eucaryotes**

La première étape de la dégradation d'un ARNm est généralement la déadénylation catalysée par les complexes PAN2/PAN3 et CCR4/NOT. L'ARN peut être ensuite dégradé de 3' en 5' par les activités exoribonucléases de DIS3L2 et de l'exosome ou decapé par DCP1/DCP2 et dégradé de 5' en 3' par l'exoribonucléase XRN1-4. L'ARN peut être uridylé en 3' et dégradé par DIS3L2 de 3' en 5'. De plus l'uridylation favorise la liaison du complexe LSM1-7 qui recrute DCP1 et DCP2 via Pat1 ; l'ARN est, dans ce cas, decapé et dégradé en direction 5'-3' par XRN1/4. La coiffe est ensuite dégradée par DCPS.

## 1.2 VOIE GÉNÉRALES DE DÉGRADATION DES ARN MESSAGERS

La dégradation des ARNm est un mécanisme essentiel chez tous les organismes. Ces mécanismes sont très complexes, mais les étapes principales sont conservées chez les eucaryotes (**Figure 3**).

### 1.2.1 LA DEADENYLATION

La dégradation d'un ARNm est généralement, précédée par une étape de déadénylation, c'est-à-dire un raccourcissement enzymatique de la queue poly(A) à l'extrémité 3'. La déadénylation est une étape limitante et détermine la vitesse de dégradation de l'ARNm (Łabno et al., 2016). Chez les eucaryotes les principales déadénylases sont PAN2/3 (PAB-dependent polyA ribonuclease 2/3) et le complexe CCR4/NOT (Carbon Catabolite Repressor 4/Negative On TATA).

Le complexe PAN2/PAN3 actif est un hétérotrimère constitué de PAN2 et de deux sous-unités PAN3, exoribonucléases de la famille DEDD (Schäfer et al., 2019; Wolf and Passmore, 2014). Les protéines PAN sont recrutées sur l'ARNm *via* l'interaction avec les PABP, ceci favorise leur activité ribonucléase (Mangus et al., 2004). Récemment il a été montré que PAN2/3 reconnaît l'interface formée suite à l'oligomérisation des PABP sur la queue de poly(A). Le nombre des interfaces ainsi formées influence l'efficacité de la déadénylation (Schäfer et al., 2019). La fonction de PAN2/PAN3 ne semble pas être essentielle et elle est partiellement redondante avec celle du complexe CCR4/NOT (Parker and Song, 2004; Wolf and Passmore, 2014). Il est intéressant de mentionner que PAN2 est conservé chez les Chlorophytes et les Bryophytes mais il n'y a pas d'homologues chez les plantes supérieures (Pavlopoulou et al., 2013). Chez la majorité des eucaryotes le rôle principal de PAN2/PAN3 est d'initier le raccourcissement de la queue de poly(A), le complexe CCR4/NOT complète la déadénylation du transcrit (**Figure 3**) (Mangus et al., 2004; Schäfer et al., 2019; Łabno et al., 2016). Dans les cellules humaines l'inactivation de PAN2/3 n'induit pas de changement notable du transcriptome à l'état d'équilibre mais une accumulation de queues poly(A) de plus de 150 nt (Yi et al., 2018). Alors que la déplétion d'une des sous-unités du complexe CCR4/NOT impacte de manière importante le transcriptome à l'état d'équilibre, avec des changements évidents de la taille des queues poly(A) (Parker and Song, 2004; Wolf and Passmore, 2014).

Le complexe CCR4/NOT possède deux sous-unités catalytiques, la protéine CCR4 de la famille EEP et CAF1 de la famille DEDD. L'étude de l'activité *in vitro* du complexe CCR4/NOT de *S. pombe* a permis d'élucider le mécanisme de raccourcissement de la queue poly(A) (Webster et al., 2018). Cette étude montre que la déadénylation par CCR4 procède de manière progressive et que le taux de déadénylation diminue régulièrement tous les 8 nucléotides. Cette périodicité est due à la liaison des domaines RRM (RNA recognition motif) des PABP sur la queue poly(A). Les PABP contiennent 4 domaines de liaison à l'ARN de type RRM. Chacun de ces domaines lie environ 8 nucléotides. Chez la

levure et les mammifères, il a été montré que CCR4 est recruté par les PABP sur la queue poly(A) (Webster et al., 2018). Cependant la présence des PABP induit CCR4 à ralentir jusqu'au dégageant de chaque domaine RRM lié à la queue de poly(A). Ce qui explique la périodicité de 8 nt observée. *A contrario* la présence des PABP liées à la queue de poly(A) inhibe CAF1. La quantité des PABP est liée à l'optimalité des codons de la séquence codante (Webster et al., 2018). Les ARNm contenant des codons non-optimaux ont un taux de traduction bas et moins de PABP sont liées à la queue de poly(A) (Hanson and Coller, 2018). CAF1 a donc une préférence pour ces transcrits alors que CCR4 dégrade de préférence les ARNm qui ont des codons optimaux et un taux de traduction élevé (Webster et al., 2018, 2017).

Chez les plantes, une augmentation de l'expression de CAF1 et CCR4 est observée suite un stress ou lors des changements développementaux. La dérégulation de l'expression de CAF1 et CCR4 impacte l'accumulation d'ARNm codant pour des protéines impliquées dans des processus fondamentaux : le développement, le métabolisme de l'amidon et la réponse aux stress biotiques et abiotiques (Chen et al., 2016; Liang et al., 2009; Sarowar et al., 2007; Suzuki et al., 2015; Walley et al., 2010).

Deux autres déadénylases sont connues chez les eucaryotes : PARN (poly(A)-specific ribonucléase) et NOC (Nocturnin). NOC est l'enzyme responsable de la déadénylation des ARNm codant pour les protéines essentielles pour la régulation des rythmes circadiens chez les mammifères (Beta and Balatsos, 2018). L'enzyme PARN a la particularité d'interagir avec la coiffe et son activité est inhibée par la présence des protéines de liaison à la coiffe CBP (Cap Binding Proteins) (Gao et al., 2000; Martínez et al., 2001). PARN est impliqué dans la déadénylation des ARNm impliqués dans la réponse aux dommages à l'ADN et des ARNm cibles des miARN (Cevher et al., 2010; Zhang et al., 2015c). Des études récentes montrent que PARN est surexprimé dans les cellules cancéreuses et est activé par des suppresseurs des tumeurs (Maragozidis et al., 2015; Shukla et al., 2019). PARN est aussi impliqué dans la maturation et la régulation de la stabilité de plusieurs ARN non codants dans les cellules des mammifères : les ARN des Cajal bodies, les petits ARN nucléolaires, les ARN de la télomérase et les ARN Y (Moon et al., 2015; Berndt et al., 2012; Shukla and Parker, 2017; Shukla et al., 2016; Son et al., 2018; Tseng et al., 2018).

Chez *Arabidopsis* PARN est nécessaire pour le développement post-embryonnaire (Chiba et al., 2004). De plus son expression augmente lors d'une exposition à l'acide abscissique et à l'acide salicylique ce qui suggère un rôle dans la réponse aux stress abiotiques (Nishimura et al., 2009). Chez les humains, *X. laevis* et les plantes PARN se déplace entre le nucléoplasme et le cytosol (Godwin et al., 2013). Chez *A. thaliana* PARN possède une séquence d'adressage à la mitochondrie où, avec la poly(A) polymérase AGS1 régule l'état de polyadénylation des transcrits mitochondriaux (Hirayama et al., 2013).



### 1.2.2 L'URIDYLATION

L'uridylation des ARN est une modification post-transcriptionnelle conservée chez tous les eucaryotes étudiés à l'exception de *Saccharomyces cerevisiae*. Elle consiste en l'ajout d'uridine en 3' des ARN, cette réaction est catalysée par des enzymes appelées terminal uridylyltransférases (TUTases). Le développement de protocoles d'analyse à haut-débit a permis de montrer l'importance de cette modification qui cible les ARN codants et de nombreux ARN non codants influençant leur dégradation, leur activité et leur maturation (Zigáčková and Vaňáčková, 2018; De Almeida et al., 2018a, 2018b). De plus chez *C. elegans* la TUTase Cde1 est impliquée dans la restriction du virus Orsay (OrV) et uridyle les extrémités 3' des ARN viraux. Une activité similaire a été observée pour TUT4 et TUT7 sur des ARN du virus de la grippe en cellules humaines (Le Pen et al., 2018).

Le rôle principal de l'uridylation des ARNm est d'induire leur dégradation. L'uridylation facilite l'interaction du complexe LSM1-7 donc le recrutement du complexe de decapping et la dégradation 5'-3' des ARNm par Xrn1 (Song and Kiledjian, 2007; Lim et al., 2014; Rissland and Norbury, 2009). De plus, l'exoribonucléase 3'-5' Dis3L2 reconnaît les uridines terminales et induit la dégradation 3'-5' d'ARN non codants chez les mammifères et des quelques ARNm chez *S. pombe* (**Figure 3**; (Ustianenko et al., 2013, 2016; Malecki et al., 2013). Chez l'humain les TUTases TUT4 et TUT7 sont responsables de l'uridylation des ARNm et favorisent leur dégradation puisque la diminution de l'expression de ces enzymes induit une augmentation de la demi-vie des ARNm (Lim et al., 2014). De plus chez la souris l'uridylation a un rôle crucial pour l'établissement du transcriptome maternel dans les oocytes. L'action des TUTases TUT4 et TUT7 est requise pour la dégradation de transcrits spécifiques, nécessaire pour compléter la méiose I et permettre la maturation des gamètes femelles (Morgan et al., 2017).

Chez les plantes, la perturbation de la voie de dégradation 5'-3' induit l'accumulation d'ARNm uridylés, ceci indique que le rôle de l'uridylation dans la dégradation des ARNm est conservé (Zuber et al., 2016; de Almeida et al., 2018). Chez *A. thaliana* deux TUTases ont été caractérisées : URT1 et HESO1. URT1 est la principale TUTase responsable de 80% de l'uridylation des ARNm. Le rôle de l'uridylation par URT1 dans la dégradation des ARNm n'est qu'une facette de sa fonction, elle possède ainsi un rôle additionnel dans la prévention de la déadénylation excessive des ARNm. En effet, le mutant *urt1* accumule des ARNm excessivement déadénylés. De plus l'analyse par TAIL-seq des extrémités 3' révèle qu'URT1 est capable de reconstituer une extension A/U de 16 nt. Une taille suffisante pour la liaison d'une PABP (Zuber et al., 2016). La présence de transcrits uridylés dans le mutant nul d'URT1 suggère la présence d'autres enzymes capables d'exercer la même fonction. HESO1 est un bon candidat, en effet il est connu que HESO1 et URT1 peuvent agir sur les mêmes

substrats ARN, comme les fragments d'ARNm clivés par RISC et les miARN. L'uridylation par URT1 et HESO1 des miARN qui ont des extrémités non méthylées influence leur activité et peut en induire la dégradation (Ren et al., 2012; Tu et al., 2015).

### 1.2.3 LA DEGRADATION EXORIBONUCLEOLYTIQUE 3'-5'

#### L'EXOSOME

Suite à la déadénylation et à l'éventuelle uridylation du transcrit oligoadénylé, celui-ci peut être dégradé à partir de l'extrémité 3' par l'exosome cytosolique ou par l'exoribonucléase DIS3L2/SOV (Figure 3). L'exosome est un complexe constitué par un cœur de neuf sous-unités (Exo9) auquel s'associe différents cofacteurs/adaptateurs comme des exoribonucléases et des hélicases. Ces cofacteurs sont essentiels à la reconnaissance des substrats ARN de l'exosome et ils varient en fonction de la localisation nucléolaire, nucléocytosolique et cytosolique de l'exosome (Łabno et al., 2016). Chez les eucaryotes l'activité exoribonucléase est portée par un de ses cofacteurs : RRP6 dans le noyau et DIS3/RRP44 dans le noyau et le cytosol. Chez les humains, un paralogue, DIS3L, est associé à l'exosome dans le cytosol. L'exosome des plantes possède une activité catalytique supplémentaire : le cœur de l'exosome, Exo9, présente une activité phosphorolytique conférée par la sous-unité RRP41 (Sikorska et al., 2017). L'exosome est non seulement impliqué dans la dégradation des ARNm, mais aussi dans la maturation et la dégradation des ARN ribosomiques et des snoRNA, et dans la dégradation de nombre d'ARN non-codants nucléaires (Łabno et al., 2016). Les fonctions cytosoliques de l'exosome requièrent son association avec les co-facteurs SKI7 et le complexe SKI (Zinder and Lima, 2017; Marshall et al., 2018). Chez les plantes, deux nouveaux types de co-facteurs de l'exosome, RST1 et RIPR, ont récemment été identifiés et connectent le core de l'exosome au complexe SKI (Lange et al., 2019). Si RIPR semble spécifique des plantes, RST1 possède un homologue chez les humains appelé FOCAD qui est aussi impliqué dans le recrutement du complexe SKI (Tuck et al., 2020).

#### DIS3L2/SOV

La seconde 3'-5' exoribonucléase impliquée dans la dégradation des ARNm dans le cytosol est DIS3L2/SOV, une exoribonucléase de type RNaseIII RRP44-like conservée chez *S. pombe*, les métazoaires et les plantes mais absent chez *S. cerevisiae*. Contrairement à RRP44, DIS3L2 ne contient pas le domaine PIN nécessaire pour interagir avec l'exosome et est donc impliquée dans la dégradation des ARNm indépendamment de l'exosome (Lubas et al., 2013; Lv et al., 2015). Chez *A. thaliana* l'homologue de DIS3L2 appelé SUPPRESSOR OF VARICOSE (SOV) a été identifié en étudiant les variations naturelles qui influencent le phénotype d'un mutant dans l'activateur de decapping VCS (Zhang et al., 2010). Une étude récente de la vitesse de dégradation des ARNm sur le

transcriptome entier et dans un panel de mutants et multiples mutants *vcs* et *sov*, a montré que SOV est impliqué dans la dégradation cytosolique des ARNm et qu'il cible des transcrits qui sont aussi des substrats de la voie 5'-3'. Cette analyse met en lumière un phénomène souvent observé entre différentes voies de dégradation des ARNm, la redondance fonctionnelles entre deux voies de dégradation de polarité différentes 5'-3' et 3'-5' (Sorenson et al., 2018). Ces phénomènes de compensation complexifient de manière générale l'étude de la contribution de différentes voies dans le turnover des ARNm. D'après cette étude, on peut anticiper l'avantage d'utiliser un écotype comme Colo, qui possède une version mutée de SOV incapable de compléter le phénotype induit par exemple par la mutation *vcs* (Zhang et al., 2010). En effet cet écotype de référence pourrait faciliter l'identification de nouvelles voies de dégradation, en évitant la redondance avec SOV.

#### **1.2.4 LA VOIE DE DEGRADATION 5'-3' DES ARNM**

Une étude récente en cellules humaines montre que, vraisemblablement, la majorité des ARN est dégradée par la voie 5'-3' (Tuck et al., 2020). Un ARNm possède à l'extrémité 5' la coiffe, qui le protège de la dégradation par les exonucléases 5'-3', il est donc logique dans ces conditions que la première étape de la dégradation 5'-3' soit le decapping. Les ARNm ainsi décoiffés sont alors dégradés par les exonucléases de la famille XRN1 (XRN4 chez les plantes) (**Figure 3**).

##### **1.2.4.1 LE DECAPPING**

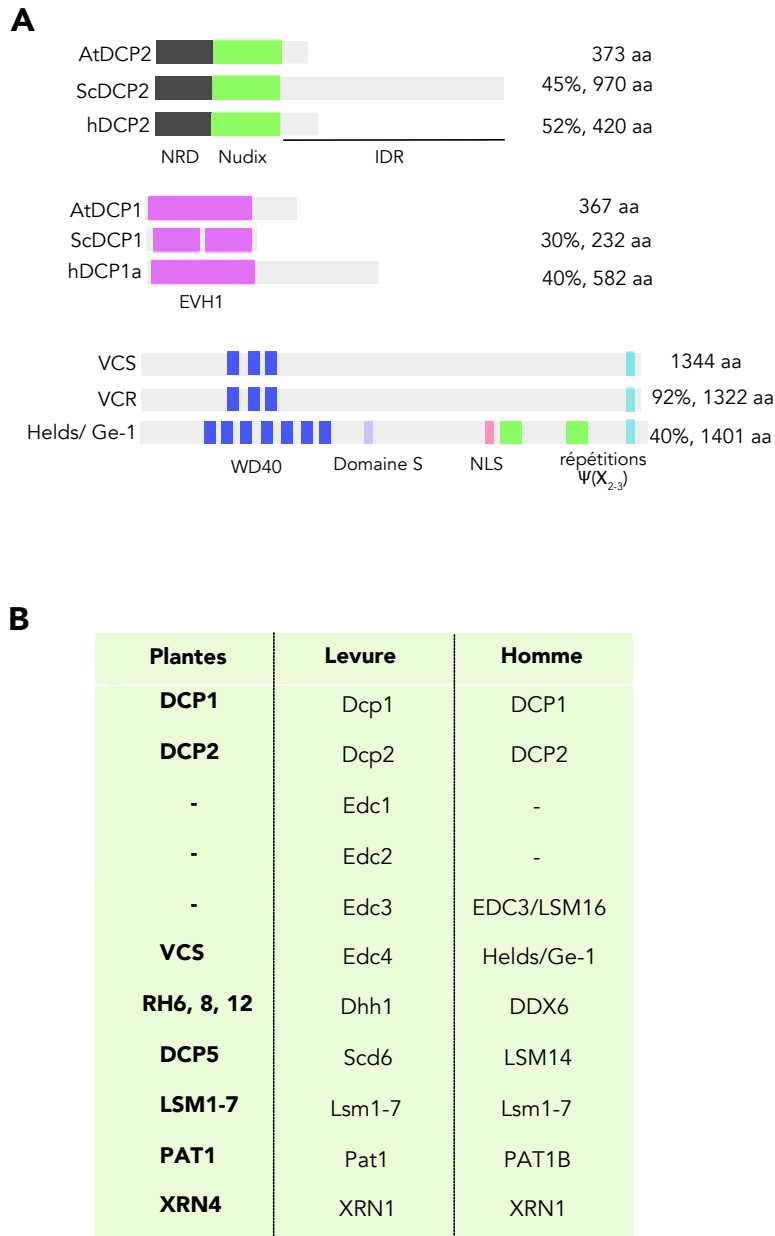
Le decapping est un processus fondamental et une étape limitante pour le turnover des ARNm cytosoliques chez les eucaryotes (Decker and Parker, 1993). Ce mécanisme est également mis en jeu dans des voies de dégradation spécifiques comme le Nonsense Mediated mRNA Decay (NMD) (He and Jacobson, 2001), le AU-rich element-mediated mRNA decay (Yamashita et al., 2005; Fenger-Grøn et al., 2005) et le RNA silencing (Behm-Ansmant et al., 2006). Bien qu'il y ait quelques études suggérant la possibilité d'un recapping cytosolique chez l'humain (Trotman and Schoenberg, 2019; Fejes-Toth et al., 2009; Otsuka et al., 2009), l'hydrolyse de la coiffe en 5' est de façon générale considérée comme une étape irréversible. La régulation fine de ce processus est possible grâce à la flexibilité conformationnelle de l'enzyme de decapping DCP2 et par la régulation de son réseau d'interactions.

La présence de la coiffe est généralement nécessaire à la traduction et de ce fait les mécanismes du decapping et la traduction peuvent être considérée comme deux mécanismes en compétition. L'interaction des facteurs de decapping avec la coiffe empêche l'interaction entre la coiffe et eIF4F, ce qui inhibe l'initiation de la traduction (Eulalio et al., 2007b). La dégradation des ARNm peut avoir lieu sur les polysomes de façon co-traductionnelle (Pelechano et al., 2015; Yu et al., 2016). De plus la

dégradation 5'-3' est influencée par la vitesse de l'élongation de la traduction (Mugridge et al., 2018a). Le cœur du complexe de decapping est conservé chez les eucaryotes. Il est composé par DCP2, porteur de l'activité enzymatique qui est activé par DCP1 et, seulement chez les eucaryotes supérieurs par VCS/ Heds/Ge-1.

#### FONCTION BIOLOGIQUE DU DECAPPING

Le decapping est crucial pour le maintien de l'homéostasie cellulaire. Chez l'homme, la mutation des gènes codant pour l'enzyme scavenger DcpS et pour l'activateur de decapping EDC3 est associée avec des handicaps mentaux héréditaires, ce qui suggère un rôle crucial du decapping dans le développement neuronal (Ahmed et al., 2014). Il est intéressant de mentionner que le decapping intervient aussi dans la réponse antivirale, par exemple chez la Drosophile il a été montré que Dcp2 est un facteur de restriction virale capable de limiter la réplication des Bunyavirus, des nouveaux virus d'intérêt médical, transmis aux mammifères par les insectes (Hopkins et al., 2013). Chez *A. thaliana*, le decapping est un mécanisme crucial pour le développement (Goeres et al., 2007). Les mutants nul *dcp1*, *dcp2*, et *vcs* sont létaux au stade post embryonnaire (Xu et al., 2006). De plus le decapping est nécessaire pour la régulation de la germination de la graine (Basbous-Serhal et al., 2017) et pour la réponse à l'hormone ABA (Wawer et al., 2018). D'après une étude récente menée dans le laboratoire de Leslie Sieburth, le decapping contribue à la dégradation à hauteur de 68% des ARNm chez *Arabidopsis* (Sorenson et al., 2018).



**Figure 4 Les protéines impliquées dans le decapping chez *Arabidopsis thaliana***

**(A)** (Adapté de Xu et al., 2006) Conservation de l'organisation en domaines et homologie de séquences des facteurs de decapping chez *Arabidopsis thaliana* (At) *Saccharomyces cerevisiae* (Sc) et *Homo sapiens* (h). Les pourcentages d'identité, le nombre d'acides aminés (aa) et les noms des domaines sont indiqués. NRD : N-ter Regulatory Domain ; Nudix : domaine catalytique ; IDR : Intrinsically Disordered Region ; EVH1 : Domaine de type Ena/Vasp Homology 1 ; WD40 : répétition de Tryptophane (W) et Acide Aspartique (D). Domaine S : région riche en Serine (S). NLS : Nuclear Localization Signal. Répétitions  $\Psi(X_{2.3})$  : répétitions de deux acides aminés hydrophobes ( $\Psi$ ) séparés par 2/3 acides aminés indéterminés. **(B)** Tableau avec les correspondances des noms des protéines impliquées dans le decapping et la voie de dégradation 5'-3' chez les plantes, la levure et l'homme.

## L'ENZYME DE DECAPPING DCP2

Chez les eucaryotes l'enzyme principale responsable du decapping est la protéine DCP2 qui possède un domaine Nudix. Elle est présente à la fois chez la levure (Dunckley and Parker, 1999) et dans les eucaryotes supérieurs, incluant le règne animal et végétal (Cohen et al., 2005; Iwasaki et al., 2007; Lin et al., 2008; Lykke-Andersen, 2002; Steiger et al., 2003; Van Dijk et al., 2002; Wang et al., 2002; Xu et al., 2006). Dcp2 interagit avec d'autres protéines impliquées dans la dégradation des ARNm. Elle catalyse l'hydrolyse de la coiffe entre les phosphates  $\alpha$  et  $\beta$  de l'ARNm coiffé (**Figure 2**). Elle a la particularité d'être une protéine de grande taille par rapport aux hydrolases Nudix (*S. cerevisiae* 109 kDa, *H. sapiens* 49 kDa, *A. thaliana* 42 kDa). Dcp2 possède une région régulatrice N-terminale (NRD), le domaine catalytique Nudix (catalytic domain, CD) et une queue C-terminale intrinsèquement désorganisée (Intrinsically disordered region, IDR), pour laquelle aucune structure secondaire n'a pu être prédite sur la base de l'analyse des séquences peptidiques (**Figure 4**). Les domaines CD et NRD sont liés par une petite séquence linker qui permet leur mouvement/réorientation. Le domaine IDR, a une taille variable selon l'organisme : sa longueur varie d'environ 100 acides aminés (aa) chez l'humain et les plantes à plus de 500 aa chez la levure (Van Dijk et al., 2002; McLennan, 2006; Wurm and Sprangers, 2019). Il contient des motifs riches en leucine (HLM, helical leucine-rich motifs) nécessaires pour l'interaction protéine-protéine (Fromm et al., 2012). Ces domaines sont absents chez les métazoaires et les plantes. L'activité catalytique de Dcp2 dépend du motif Nudix, Gx5Ex5[UA]xREx2EEExGU (PROSITE PS00893), où le U indique un résidu hydrophobe, ce motif a une structure loop-helix-loop. Les résidus de glutamate (E) coordonnent un cation bivalent essentiel pour l'activation d'une molécule d'eau, cette réaction étant nécessaire pour l'hydrolyse de la coiffe. La structure active de Dcp2 a été difficile à étudier à cause de sa flexibilité. Cependant, il a été montré, par RMN, que chez la levure, la reconnaissance de la m<sup>7</sup>G pouvait se produire seulement par reconstitution d'un site actif fractionné entre les domaines CD et NRD de Dcp2 (Floor et al., 2010). L'activité du domaine CD seul ainsi que de la protéine entière est faible *in vitro*. L'interactant le plus notable et le plus étudié de Dcp2 est Dcp1. Chez les eucaryotes le complexe de decapping contenant l'hétérodimère Dcp1/Dcp2 est responsable du decapping cytosolique des ARNm. Il a été montré que la présence du domaine NRD augmente l'activité catalytique de Dcp2 et favorise l'interaction avec l'activateur de decapping Dcp1 *in vitro* et *in vivo* (She et al., 2006; Deshmukh et al., 2008; Wurm et al., 2017). La résolution de la structure du complexe RNP Dcp1/Dcp2/Edc3/ m<sup>7</sup>Gpp (Charenton et al., 2016; Mugridge et al., 2018b) a permis de capturer Dcp2 dans une conformation active où l'orientation de ses domaines CD et NRD positionne la m<sup>7</sup>Gpp dans le site catalytique. La résolution de ce complexe a permis aussi l'identification des domaines nécessaires pour l'interaction entre Dcp1 et Dcp2 ainsi qu'entre Dcp2 et Edc3. Il est

intéressant de noter qu'une étude par RMN de la conformation de Dcp2 en solution avec Dcp1 a montré que l'enzyme est dans sa conformation inactive (Wurm et al., 2017). De plus chez la levure, des données biochimiques et génétiques ont permis de révéler la présence dans la région IDR de Dcp2 de motifs auto-inhibiteurs. Cette auto-inhibition est atténuée par l'interaction de Edc3, nécessaire pour la mise en place d'un complexe de decapping actif via l'interaction avec les autres activateurs (He and Jacobson, 2015; Paquette et al., 2018). Ceci montre l'importance des cofacteurs pour l'activation du complexe du decapping. En effet Dcp1/Dcp2 interagit avec beaucoup de protéines *in vivo* : Pat1, LSM1-7, Dhh1/DDX6 et Scd6/LSM14 ainsi qu'avec les protéines stimulatrices de l'enlèvement de la coiffe appelées globalement EDCs (Enhancer of decapping). Le domaine IDR de Dcp2 interagit directement avec Pat1, Upf1 et Edc3 via des motifs linéaires courts (SLiM, Short Linear Motif). Chez les métazoaires et les plantes l'activité de decapping requiert la présence additionnelle de EDC4 (HEDLS/Ge-1), VCS chez les plante (Xu et al., 2006). De plus le complexe de decapping interagit avec les protéines Upf (Upf1, Upf2, Upf3) impliquées dans le NMD qui cible des ARN présentant des anomalies lors de la terminaison de la traduction. Les protéines interagissant avec le complexe de decapping, par exemple Dhh/DDX6 et Scd6/LSM14, sont impliquées non seulement dans la dégradation des ARNm mais aussi dans l'inhibition de la traduction. En effet les deux mécanismes sont strictement liés et leur régulation concertée est possible via le complexe réseau d'interactions qui se met en place autour de l'holoenzyme Dcp1/Dcp2.

## LES ACTIVATEURS DE DECAPPING

### DCP1 (DECAPPING PROTEIN 1)

L'activateur de decapping qui a été décrit en premier chez la levure est Dcp1, sa découverte précède celle de Dcp2 et cette protéine avait été considérée comme l'enzyme responsable du decapping (LaGrandeur and Parker, 1998). Dcp1 possède un domaine EVH1 (enabled/VASP homology 1) conservé chez les eucaryotes, et une région intrinsèquement désorganisée qui contient des motifs HLM, présente seulement chez les métazoaires (**Figure 4**). Les domaines HLM chez la Drosophile permettent l'interaction entre DCP1 et l'homologue de Scd6 (Fromm et al., 2012). La résolution de la structure cristallographique de DCP1 chez la Drosophile et l'homme a permis de montrer que le domaine C-terminal forme des trimères et que ceci est nécessaire pour l'assemblage correct du complexe de decapping ainsi pour la réaction de decapping *in vivo*. De plus les résidus critique pour la formation des trimères sont conservés chez les eucaryotes multicellulaires analysés y compris *A. thaliana* (Tritschler et al., 2009). Chez *A. thaliana* DCP1 est impliqué dans la réponse aux stress biotique et abiotique. DCP1 est régulé post-traductionnellement par phosphorylation (Ser<sup>237</sup>) lors d'une infection bactérienne. L'interaction des kinases MAPK6 et MAPK3 avec DCP1 induit sa phosphorylation ce qui favorise l'interaction avec XRN4 et ainsi la dégradation 5'-3' de transcrits

cibles impliqués dans la réponse à l'infection par la bactérie *Pseudomonas syringae* (Yu et al., 2019). De plus lors d'un stress hydrique la phosphorylation de DCP1 est cruciale pour son interaction avec DCP5 ce qui permet le remodelage du transcriptome en réponse à la déshydratation (Xu and Chua, 2012).

### GE-1/VCS

La protéine Ge-1/VCS contient des répétitions (7 chez l'humain, 4 chez *A. thaliana*) du motif WD40 en N-terminal, ceux-ci étant impliqués dans des interactions protéine-protéine (**Figure 4**). La protéine Ge-1/Helds (human enhancer of decapping large subunit), VARICOSE chez les plantes, fait partie du complexe de decapping. Chez *A. thaliana* la drosophile et chez l'homme VCS interagit avec DCP1, DCP2 et peut aussi former des multimères grâce à sa région C-terminale (Xu et al., 2006; Jinek et al., 2008; Bloch et al., 2011). Chez la Drosophile Ge-1 est impliqué mais pas essentiel pour la formation et la localisation de l'ARNm Oskar dans l'oocyte et le développement correct de l'embryon. De plus Ge-1 est nécessaire pour la formation des P-bodies dans la lignée germinale (Fan et al., 2011). Le rôle de Ge-1/VCS dans la formation des P-bodies est conservée chez les eucaryotes ; chez la Drosophile et les mammifères la partie C-ter de Ge-1 est nécessaire pour sa localisation (Fenger-Grøn et al., 2005; Xu et al., 2006; Jiang et al., 2005; Eulalio et al., 2007a). Chez *A. thaliana* la protéine VCS est un composant essentiel du complexe de decapping et elle est nécessaire aussi pour la répression de la traduction induite par les miARN (Brodersen et al., 2008). Le mutant nul présente des défauts de développement des feuilles et du méristème apical. Le phénotype au niveau des feuilles s'accroît lorsque la voie de signalisation de l'auxine est perturbée, cependant le mutant ne montre pas de défauts *per se* dans la voie de signalisation de cette hormone. De ce fait il a été proposé que le phénotype observée est dû à la perturbation, par les inhibiteurs du transport polaire d'auxine, d'une voie indépendante de VCS qui est nécessaire pour le développement correct de la feuille (Deyholos et al., 2003). Il est intéressant de remarquer la présence d'un homologue de VCS chez *A. thaliana*, appelé VARICOSE RELATED (VCR) ; VCR partage une homologie de séquence de 92% avec VCS, néanmoins l'extrémité C-terminale est dépourvue de 22 aa (**Figure 4**). Cette observation et l'absence de phénotype chez le mutant *vcr* est à l'origine de l'hypothèse selon laquelle cette protéine n'est pas nécessaire pour le déroulement du decapping (Xu et al., 2006).

### LE COMPLEXE PAT1/LSM1-7

Chez les eucaryotes les protéines Sm-like (LSm) sont à la fois des facteurs impliqués dans la voie cytosolique de dégradation 5'-3' des ARNm et dans l'épissage nucléaire. Les complexes LSm nucléaire et cytosolique sont composés respectivement des sous-unités LSM2-8 et LSM1-7. Dans le noyau le complexe LSM2-8 protège les extrémités 3' des snRNA U6 nécessaires pour l'assemblage



du spliceosome (Achsel et al., 1999; Bouveret et al., 2000; Tharun et al., 2000; Tharun and Parker, 2001). Dans le cytosol LSM1-7 se lie de préférence aux ARN ayant l'extrémité 3' oligoadénylée et uridyliée (Chowdhury et al., 2007; Song and Kiledjian, 2007). La protéine Pat1 forme un complexe avec les protéines LSM et est impliquée dans plusieurs voies du métabolisme des ARNm, le splicing avec LSM1-8, la répression de la traduction et la dégradation (Vindry et al., 2017; Lobel et al., 2019). Chez les mutants LSM1-7 et Pat1 il y a accumulation d'ARNm déadénylés et coiffés (Bouveret et al., 2000; Tharun et al., 2000). Concernant son rôle dans la dégradation des ARNm, Pat1 stabilise la liaison du complexe LSM1-7 à l'extrémité 3' de l'ARN grâce à deux petits motifs linéaires d'interaction. Successivement Pat1 recrute Dcp1/Dcp2 ce qui produit l'hydrolyse de la coiffe. Pat1 diminue l'effet d'autoinhibition par liaison à la partie C-terminale de Dcp2 et renforce la liaison de Dcp2 à l'extrémité 5' de l'ARN (Lobel et al., 2019). Pat1 interagit aussi avec l'exoribonucléase XRN1 ainsi qu'avec Dhh1 impliquée dans l'inhibition de la traduction (Braun et al., 2010; Haas et al., 2010; Ozgur et al., 2010). Le complexe constitué de Pat1 et de LSM1-7 facilite donc l'interaction des extrémités 3' et 5' et lie la déadénylation au decapping.

#### LSM14/SCD6/DCP5

Chez les plantes, la protéine DCP5 fait partie de la famille de protéines RAP55. Elles possèdent un domaine LSM14, une région riche en serine/thréonine et des petits motifs linéaires responsables des interactions protéine-protéine. DCP5 est impliquée dans inhibition de la traduction et formation de granules d'ARN (Marnef et al., 2009). Chez *Arabidopsis* DCP5 fait partie des activateurs de decapping même si, *in vitro*, elle n'a pas d'effet sur l'activité catalytique de DCP2, ce qui suggère que son activité *in vivo* sur le decapping est indirecte. Chez le mutant knockdown *dcp5-1* le decapping des ARNm est compromis, le phénotype phénocopie les défauts du développement observés dans les mutants de decapping et la taille des P-bodies diminue (Xu and Chua, 2009). DCP5 interagit avec DCP1 lors de la réponse à l'infection bactérienne (Yu et al., 2019) et elle est essentielle pour la répression de la traduction de transcrits spécifique dans les P-bodies lors du développement post embryonnaire (Jang et al., 2019; Xu and Chua, 2009).

#### DDX6/DHH1/RH6, 8, 12

DDX6/Dhh1 est une hélicase à ARN ATP dépendante de type DEAD, elle possède deux domaines globulaires RecA connectés par une région linker flexible et interagit avec le complexe Dcp1/Dcp2. Une étude récente chez la levure suggère que Scd6 recrute Dhh1 ce qui induit l'inhibition de la traduction et active le decapping et la dégradation 5'-3' des cibles (Zeidan et al., 2018). L'interaction de DDX6/Dhh1/RH6, RH8, RH12 avec LSM14/Scd6/DCP5 est critique pour la formation des P-bodies et elle est bien conservée chez les eucaryotes (Ayache et al., 2015; Chicois et al., 2018). Chez la

drosophile l'homologue de DDX6, Me31B, est impliquée dans la répression traductionnelle des ARNm Oskar lors de leur transport des cellules nourricières vers l'oocyte durant l'oogenèse (Nakamura et al., 2001). Chez la levure une fonction importante de Dhh1 est d'induire spécifiquement la dégradation des ARNm contenant des codons non optimaux par interaction avec les ribosomes (Radhakrishnan et al., 2016).

#### LE CAS PARTICULIER DES ENZYMES DXO (DECAPPING EXONUCLEASE)

Les ARN coiffés avec du NAD<sup>+</sup>, ont été découverts chez les bactéries (Chen et al., 2009; Julius and Yuzenkova, 2019), où cette coiffe insolite stabilise les ARN, la coiffe NAD<sup>+</sup> est hydrolysée par l'enzyme NudC (Kiledjian, 2018). Ces dernières années la présence de ARNm NAD<sup>+</sup>, a été confirmée dans des études sur les cellules humaines et chez *S. cerevisiae* (Jiao et al., 2017; Walters et al., 2017) bien que les ARNm NAD<sup>+</sup> représentent une fraction très réduite du transcriptome (Jiao et al., 2017; Walters et al., 2017). Les résultats montrent que la présence d'une coiffe NAD<sup>+</sup> corrèle avec une stabilité réduite des transcrits. L'hydrolyse de la coiffe NAD<sup>+</sup> requiert des enzymes spécifiques. Chez la levure, les mammifères et les plantes les enzymes DXO sont responsables de cette hydrolyse (Kiledjian, 2018; Pan et al., 2019; Kramer and McLennan, 2019a).

Les protéines DXO chez les mammifères, avec leur double identité catalytique, pyrophosphatase et exonucléase, ont une excellente versatilité au niveau de substrats. Chez la souris DXO est impliqué dans l'hydrolyse des extrémités 5' incomplètes des ARNm. De plus elle peut dégrader les ARN avec un 5' monophosphate ou 5' -OH grâce à son activité exonucléase (Kramer and McLennan, 2019b). Chez *A. thaliana* DXO1 est aussi active sur les deux substrats avec une préférence pour le 5' monophosphate (Pan et al., 2019). Il est intéressant de remarquer que chez *A. thaliana* le mutant *dxo* a des défauts phénotypiques notables avec une déficience en pigmentation et des changements importants du transcriptome (Kwasnik et al., 2019). Cependant le phénotype observé chez le mutant n'est pas lié à l'absence des activités catalytiques de DXO. Ces défauts semblent être liés à son extension N-terminale, spécifique des plantes, qui intervient dans la régulation de la communication rétrograde du noyau au chloroplaste (Kwasnik et al., 2019).

#### 1.2.4.2 LA DEGRADATION EXORIBONUCLEOLYTIQUE 5'-3'

Une étape cruciale dans la dégradation des ARNm est la dégradation 5'-3' par l'exoribonucléase Xrn1 (Pacman chez la Drosophile). Cette enzyme est conservée chez les métazoaires et elle est responsable aussi de la dégradation de certains ARN non codants ainsi que des fragments produits par RISC (Nagarajan et al., 2013). Xrn2, le paralogue de Xrn1, localise dans le noyau et dégrade les précurseurs des ARN ribosomiaux ainsi que d'autres ARN nucléaires ayant une extrémité 5' accessible (Nagarajan et al., 2013). Chez *A. thaliana* il n'y a pas d'homologues de XRN1 mais il y a trois orthologues de XRN2 appelés XRN2, XRN3 et XRN4/EIN5 (Kastenmayer and Green, 2000). XRN2 et XRN3 ont une localisation nucléaire, elles sont responsables de la dégradation des précurseurs des ARN ribosomiaux et d'ARN non codants respectivement (Kurihara, 2017). XRN4 est la principale exoribonucléase 5'-3' cytosolique responsable de la dégradation des ARNm. Le mutant *xrn4* présente des phénotypes développementaux peu marqués, cependant il est insensible à l'hormone l'éthylène. XRN4 cible l'ARNm qui code pour un composant essentiel de la voie de signalisation de l'éthylène, EBF1, appelé aussi Fbl6. L'absence de XRN4 induit l'accumulation du transcrite EBF1 qui dérégule le turnover du facteur de transcription EIN3 nécessaire pour la réponse à l'éthylène (Olmedo et al., 2006; Potuschak et al., 2006). L'étude du transcriptome du mutant *xrn4* montre une accumulation d'ARNm decappés ainsi que de fragments 3' produits par clivage dépendant des miARN (Souret et al., 2004; Gazzani et al., 2004; Gregory et al., 2008). Le rôle de XRN4 dans la relation de compétition entre les voies générales de dégradation des ARNm et la voie du silencing sera approfondi dans un paragraphe dédié. XRN4 semble être aussi impliqué dans la voie du NMD puisque le mutant accumule des fragments 3' dérivant des cibles du NMD. Ce qui suggère aussi l'existence chez les plantes d'un homologue fonctionnel de l'endoribonucléase SMG6 (Nagarajan et al., 2019). Les ARNm decappés qui s'accumulent dans le mutant *xrn4* sont enrichis en transcrits codant pour des protéines impliquées dans des voies fondamentales pour la vie des plantes comme la photosynthèse ou le métabolisme de l'azote ou encore dans les voies hormonales et de réponse aux stress (Nagarajan et al., 2019). La redondance des voies de dégradation peut expliquer le phénotype peu marqué de *xrn4*. XRN4, associé à la protéine LARP1, est responsable de la dégradation des ARNm induite lors d'un stress thermique. Il est intéressant de mentionner que dans ces conditions XRN4 localise dans les polysomes ce qui suggère l'existence d'une dégradation co-traductionnelle des ARN ciblés par XRN4-LARP1 (Merret et al., 2013).

#### 1.2.5 LE DECAPPING INDEPENDENT DE LA DEADENYLATION

Comme décrit précédemment, chez les eucaryotes, les ARNm sont généralement dégradés par des mécanismes mettant en jeu tout d'abord la déadénylation, et ensuite le decapping. Cependant il existe également d'autres voies de dégradation qui agissent par des mécanismes indépendants de la

déadénylation. C'est le cas par exemple des ARNm cible du NMD, des cibles de la protéine Staufen (STAU) et de la dégradation dépendante de la réplication des ARNm codant pour les histones chez les mammifères, qui ne sont pas polyadénylés. Ces voies de dégradation sont toutes dépendante de la protéine UPF1 et, généralement, des facteurs de decapping (Kim and Maquat, 2019). De plus, le decapping indépendant de la déadénylation est aussi responsable de la dégradation de transcrits spécifiques avec des mécanismes uniques. C'est le cas de l'ARNm codant pour la protéine EDC1, activateur de decapping chez *S. cerevisiae*, qui contient dans sa région 3'UTR les déterminants qui induisent l'inhibition de la déadénylation et le recrutement des facteurs de decapping. Ceci favorise une boucle de rétroaction qui régule la traduction de cet activateur quand le decapping est compromis (Muhlrad and Parker, 2005).

#### **1.2.5.1 LE NONSENSE MEDIATED DECAY (NMD)**

Le NMD est un mécanisme de régulation post-transcriptionnelle de l'expression génique qui cible des transcrits spécifiques et induit leur dégradation. Cette voie, conservée dans l'évolution, participe au contrôle co-translationnel de la qualité des transcrits et à la régulation de l'expression des gènes dans des conditions physiologiques normales et en réponse aux changements environnementaux. UPF1 est la protéine clé de cette voie; il a été proposé que son rôle soit d'intégrer les informations contenues dans la molécule d'ARN (longueur de la 3' UTR, position des Exon Junction Complex, EJC), ainsi que son état de traduction et de réguler sa dégradation. Ces signaux vont influencer l'état de phosphorylation d'UPF1 et ainsi son activation et son interaction avec les facteurs de dégradation des ARNm (Kurosaki and Maquat, 2013; Kurosaki et al., 2019). Certaines caractéristiques de l'ARNm sont connues pour induire le NMD, c'est le cas d'un codon stop prématuré, de la présence d'une petite séquence codante en amont de l'ATG (uORF), de la présence d'un intron plus de 50 nt en aval d'un codon stop, d'un 3' UTR particulièrement long ou de la présence un codon de sélénocystéine reconnu comme un codon stop. L'absence de dégradation de ces ARNm peut, dans certains cas, induire la production d'une protéine tronquée avec des conséquences délétères pour la cellule et à la base d'un grand nombre de maladies (Ottens and Gehring, 2016; Kurosaki et al., 2019). Des défauts dans la voie du NMD sont liés, chez les mammifères, au développement de différents types de cancer (Fernandes et al., 2019). Chez les plantes le NMD est impliqué dans la réponse aux stress biotique et abiotique ainsi qu'au contrôle de processus fondamentaux comme la floraison et la régulation de l'horloge circadienne (Jeong et al., 2011; Rayson et al., 2012; Garcia et al., 2014; Kwon et al., 2014; Vexler et al., 2016; Sureshkumar et al., 2016; Nasim et al., 2017; Ohtani and Wachter, 2019). Les protéines clés du NMD sont UPF1, UPF2 et UPF3 qui activent UPF1 en stimulant ses activités enzymatiques d'hélicase et ATPase, et les protéines SMG (Suppressor with Morphological

effect on Genitalia) qui régulent l'état de phosphorylation de UPF1 et induisent la dégradation des cibles, directement par clivage endoribonucléolytique (SMG6) ou par recrutement du complexe de decapping (SMG5-7) (Yamashita, 2013; Okada-Katsuhata et al., 2012; Huntzinger et al., 2008; Dehecq et al., 2018; Conti and Izaurralde, 2005). Plusieurs modèles ont été proposés pour expliquer la reconnaissance des cibles du NMD; on peut noter parmi d'autres le modèle EJC dépendant et le modèle EJC indépendant. Ceux-ci afin d'expliquer le ciblage de transcrits avec ou sans introns en 3'UTR. Les ribosomes, pendant la traduction déplacent les complexes protéiques associés à l'ARNm, notamment les complexes EJC déposés sur les jonctions entre exons lors de l'épissage du pré-ARNm. La présence d'un codon stop prématuré induit le recyclage du ribosome avant qu'il puisse déplacer le/les EJC présents en 3' du codon stop prématuré. Ceci réduit l'efficacité de la terminaison de la traduction. Il a été observé que la présence du complexe EJC en aval d'un codon stop est responsable d'une forte induction du NMD sur les ARNm correspondants. Dans ce modèle EJC dépendant, UPF1 dans le complexe SURF, est phosphorylé par SMG1 et s'associe avec UPF2 et UPF3 présents sur l'ARNm, ces derniers associés à l'EJC, et ils forment le complexe DECID (Decay Inducing Complex). DECID induit l'inhibition de la traduction et la dégradation de la cible par clivage endoribonucléolytique par SMG6 ou par recrutement des facteurs de dégradation 5'-3' et 3'-5' par SMG5-SMG7 (Conti and Izaurralde, 2005; Fernandes et al., 2019).

Le modèle EJC indépendant s'applique aux transcrits qui ont une terminaison de la traduction inefficace à cause de la présence d'un long 3' UTR, qui empêche l'interaction entre les facteurs eRF1 et 2 et les PABP cytosoliques ainsi que le recyclage du ribosome. UPF1 en complexe avec SMG6, 7 et 5 est phosphorylé et induit l'inhibition de la traduction et la dégradation de la cible par SMG6 ou par recrutement des facteurs de decapping et de déadénylation par SMG5-SMG7 (Kurosaki and Maquat, 2013; Kurosaki et al., 2019). Ces deux modèles impliquent toujours l'étape de phosphorylation d'UPF1 qui permet le recrutement des facteurs de dégradation de l'ARN. L'endoribonucléase SMG6, enzyme à domaine PIN spécifique du NMD, induit la coupure proche du codon stop de l'ARN cible. Ce qui produit un fragment 5', dégradé par l'exosome et DIS3L2 et un fragment 3' dégradé par XRN1 (Boehm et al., 2014; Eberle et al., 2009; Schmidt et al., 2015; Huntzinger et al., 2008). Cependant chez les cellules humaines UPF1 peut aussi interagir avec SMG6 sans être phosphorylé (Nicholson et al., 2014). UPF1, phosphorylé, recrute directement le complexe de decapping, induisant le decapping et la dégradation 5'-3' de l'ARN cible. Chez la levure l'interaction entre Upf1 et Dcp2 se fait *via* Edc3 ou/et Pat1, ces deux protéines interagissent à la fois avec Upf1 et avec Dcp2 (Swisher and Parker, 2011). Chez les mammifères PNRC2 (Proline-Rich Nuclear Receptor Coactivator 2) permet le recrutement du complexe de decapping sur les cibles de UPF1 ensuite dégradées par XRN1 (Lai et al., 2012; Nicholson et al., 2018). Chez les plantes des orthologues de UPF1, UPF2,

UPF3 et SMG7 ont été caractérisés, mais on peut noter l'absence d'homologues des protéines SMG6 et SMG1 chez la plante modèle *A. thaliana* (Méraï et al., 2013; Shaul, 2015; Causier et al., 2017; Kerényi et al., 2008). Malgré l'absence de SMG1, il a été montré que UPF1 existe dans une forme phosphorylée ce qui suggère l'implication d'une autre kinase (Méraï et al., 2013; Lloyd and Davies, 2013). L'induction de la dégradation des cibles du NMD est actuellement mal caractérisée chez les plantes et elle semble s'éloigner du modèle chez les mammifères. Dans les cellules humaines la fonction de SMG6 dans des mutants *smg6* est complétée par la présence de SMG7 et *vice versa* ce qui suggère que les cibles du NMD peuvent être dégradées indistinctement par les deux voies (Colombo et al., 2017; Ohtani and Wachter, 2019). Chez *A. thaliana* des orthologues de SMG5, SMG6, SMG8, SMG9 et PNRC2 sont absents (Kerényi et al., 2008). Cependant il existe un homologue de SMG7 qui est impliqué dans la dégradation des cibles du NMD (Méraï et al., 2013; Kerényi et al., 2008). L'implication du complexe de decapping dans la dégradation des cibles du NMD avec un codon stop prématuré a été suggérée par l'étude de la protéine TAV (translational transactivator/viroplasmine) du virus CaMV (Cauliflower Mosaic Virus), qui interagit avec VCS ce qui empêche son interaction avec UPF1 et ainsi inhibe la dégradation des ARN viraux contenant un PTC (Lukhovitskaya and Ryabova, 2019).

#### **1.2.5.2 LE STAUFEN MEDIATED DECAY**

Chez les mammifères d'autres voies de dégradation des ARN sont dépendantes d'UPF1, incluant le Staufen Mediated Decay (SMD), et la voie de dégradation dépendante de la réplication des ARNm codant pour les histones chez les métazoaires. Ces deux voies de dégradation sont co-translationnelles. La présence d'UPF1 dans la 3' UTR des ARNm est un senseur de l'efficacité de la terminaison de la traduction. UPF1 entre en compétition avec les PABPC pour lier les facteurs eRF1 et 2 et s'il réussit à interagir avec ces facteurs induit l'inhibition de la traduction en cours et la dégradation de l'ARNm cible. En absence des complexes EJC UPF1 peut lier la 3' UTR d'un ARNm grâce à son interaction avec des protéines de liaison à l'ARN qui interagissent spécifiquement avec des structures secondaires présentes dans cette région.

Les protéines STAU1 et 2 sont des protéines de liaison à l'ARN double brin ; elles interagissent avec une structure secondaire en tige boucle dans la 3'UTR de certains ARNm (Roegiers and Jan, 2000). La protéine STAU1 a été découverte comme facteur de transport et localisation des ARNm d'origine maternelle dans les oocytes et les œufs chez la Drosophile (St Johnston et al., 1991). Chez les mammifères STAU1 et 2 sont impliquées dans la localisation des ARNm dans les oocytes et les

neurones mais aussi dans la dégradation de cibles spécifiques par un processus, le Staufen Mediated Decay (SMD), où UPF1 est impliquée.

STAU, en forme de dimère, se lie au SBS (Staufen Binding site) situé dans la 3' UTR de ces cibles où a lieu l'interaction avec UPF1 et UPF2, qui active l'hélicase (Gowravaram et al., 2019). UPF1 est ensuite phosphorylé et recrute la protéine PNRC2 ce qui permet l'interaction avec DCP1, l'induction du decapping et la dégradation 5'-3'. La structure SBS peut se former par appariement intramoléculaire des bases dans la 3' UTR ou par appariement intermoléculaire entre une séquence Alu dans la 3' UTR de l'ARNm cible et une autre séquence Alu présente dans un lncRNA (Gong and Maquat, 2011). Chez les mammifères la dégradation par la voie du SMD est cruciale dans la localisation d'ARNm spécifiques dans les axones des cellules neuronales (Gowravaram et al., 2019) ainsi que dans l'adipogénèse (Cho et al., 2012).

### **1.2.5.3 LA DEGRADATION DEPENDANTE DE LA REPLICATION DES ARNm CODANT POUR LES HISTONES CHEZ LES METAZOAIRES**

Les ARNm codant pour les histones chez les animaux ont une structure unique, ils ne possèdent pas la queue de poly(A) à l'extrémité 3' mais une structure secondaire conservée en tige boucle qui est fondamentale pour la dégradation rapide de ces ARNm à la fin de la phase S (après la duplication de l'ADN) ou dans des conditions de stress quand la réplication de l'ADN est inhibée (Marzluff and Koreski, 2017). La dégradation de ces ARNm requiert l'action concertée de la protéine de liaison à l'ARN double brin SLBP et d'UPF1. Ce processus a lieu sur des ARNm en cours de traduction. L'action de UPF1 va induire l'uridylation de l'extrémité 3' par l'uridylyltransferase TUT7, une étape importante pour cette dégradation puisqu'elle permet la restauration d'un site de liaison pour le complexe LSM1-7, qui recrute le complexe de decapping et permet la dégradation 5'-3' mais aussi le recrutement de l'exosome et la dégradation 3'-5' (Marzluff and Koreski, 2017; Łabno et al., 2016; Kim and Maquat, 2019).

### **1.2.6 DEGRADATION DEPENDANTE D'UN CLIVAGE ENDORIBONUCLEOLYTIQUE**

L'induction de la dégradation et la maturation des ARN par clivage endoribonucléolytique est un mécanisme observé à un degré divers chez les organismes, des procaryotes aux eucaryotes. Chez les eucaryotes les voies générales de dégradation des ARNm nécessitent l'activité de différentes exoribonucléases. En parallèle, d'autres voies de dégradation essentielles comme le NMD et le RNA silencing mettent cette fois-ci en jeu l'activité d'endoribonucléases, respectivement SMG6 et les protéines de la famille Argonaute (Huntzinger et al., 2008; Zhang et al., 2015a). Le clivage endoribonucléolytique des cibles de ces voies de dégradation produit deux fragments d'ARN qui

sont eux dégradés par les exoribonucléases 5'-3' et 3'-5' cytosoliques. Les endoribonucléases contiennent un domaine catalytique souvent associé à un/des domaine(s) de liaison à l'ARN. Elles peuvent aussi former des complexes avec une/des sous-unité(s) qui possèdent la capacité de liaison à l'ARN. Les endoribonucléases ont des cibles pouvant avoir une spécificité de séquence ou de structure. Leur expression est cruciale notamment lors des réponses aux stress et plus en général pour une réponse rapide aux stimuli internes et externes à l'organisme (Tomecki and Dziembowski, 2010; Gu et al., 2018).

#### **1.2.6.1 LE RNA SILENCING DANS LA REGULATION DE L'EXPRESSION DES GENES**

Le RNA silencing est un mécanisme d'extinction de gènes qui utilise des petits ARN interférents (20-30 nt, siARN ou miARN) afin de diriger l'action des protéines de la famille ARGONAUTE (AGO) sur les ARN cibles, dans le but de supprimer l'expression de gènes au niveau transcriptionnel (TGS) ou post-transcriptionnel (PTGS) de manière séquence spécifique. Ces petits ARN sont produits à partir de structures double brin, par l'action d'enzymes de la famille DICER, ou DICER-LIKE (DCL) chez les plantes ayant une activité de type RNaseIII. Les petits ARN ainsi produits sont chargés dans les endoribonucléases AGO faisant partie du RISC pour RNA Induced Silencing Complex, pour permettre la reconnaissance de la cible et leur répression. Ce mécanisme est conservé chez la plupart des eucaryotes et est essentiel dans le développement, la réponse aux stress biotique et abiotique ainsi que pour le maintien de l'intégrité du génome (Bologna and Voinnet, 2014). Dans les paragraphes suivants je vais décrire brièvement les deux classes de petits ARN impliqués dans la dégradation des ARNm par la voie du RNA silencing chez *A. thaliana* : les micro ARN et les trans-acting siARN ainsi que de la relation de compétition qui existe entre le RNA silencing et les voies générales de dégradation de l'ARN.

#### **LES MICROARN**

Les gènes de microARN (miARN) sont généralement transcrits par PolIII. Les précurseurs des miARN, les pri-miARN, sont coiffés en 5' et polyadénylés en 3'. Ces précurseurs adoptent une structure secondaire en tige-boucle avec une structure simple brin aux extrémités 5' et 3'. La molécule est ensuite maturée par l'action de DCL1 et des cofacteurs HYL1 (HYPONASTIC LEAVES 1) et SE (SERRATE) (Vazquez et al., 2004; Laubinger et al., 2008), et le miARN est exporté dans le cytosol sous forme de molécule d'ARN double brin, de 21 nt généralement, avec les extrémités 3' méthylées par HEN1 (HUA ENHANCER 1) (Yu et al., 2005). Ensuite le miARN en forme de duplex est chargé dans RISC contenant une protéine AGO, généralement AGO1. Un des deux brins est utilisé



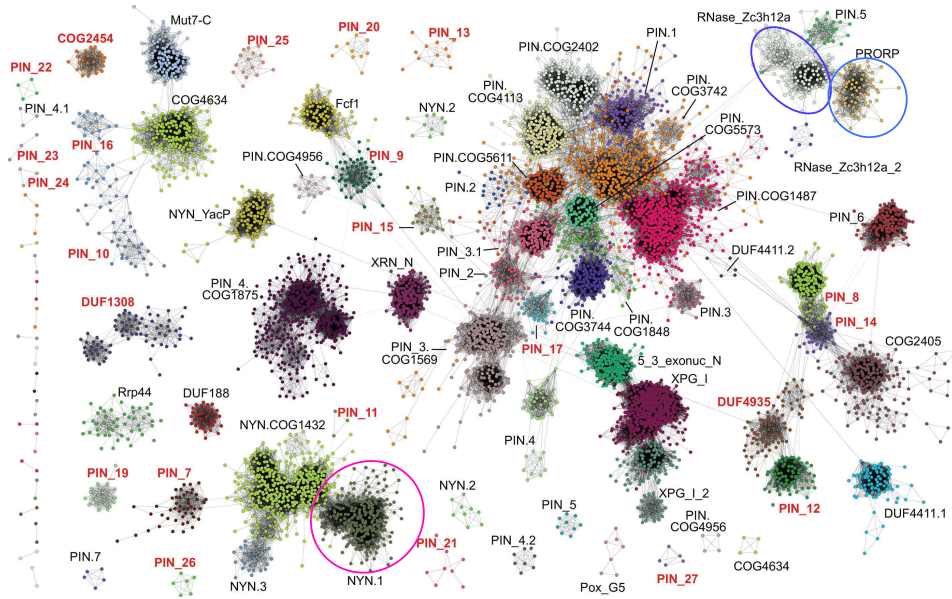
comme guide et permet au complexe RISC de reconnaître les ARN avec complémentarité de séquence alors que l'autre brin est dégradé (Xie et al., 2005). La répression de l'expression des gènes par les miARN peut avoir lieu à la fois par clivage endoribonucléolytique et par répression traductionnelle (Brodersen et al., 2008; Carbonell et al., 2012). Les produits de clivage produits par RISC sont ensuite dégradés par l'exosome en 3' et par XRN4 en 5' (Branscheid et al., 2015; Szádeczky-Kardoss et al., 2018; Souret et al., 2004). De plus récemment il a été montré l'existence de la ribonucléase RICE1 qui, en forme d'hexamère, initie la dégradation des fragments 5' uridylés produits par RISC pour permettre le recyclage des protéines AGO et ainsi le bon fonctionnement de la machinerie de silencing (Zhang et al., 2017).

### LES TRANS-ACTING siRNA

L'accumulation des tasiRNA dépend du clivage de leur ARN précurseur simple brin par des miARN spécifiques. Ce précurseur simple brin est généralement un ARN non codant, bien qu'il existe des tasiRNA issus d'ARNm (Howell et al., 2007; Zhai et al., 2011; Shivaprasad et al., 2012; Boccara et al., 2014). Chez *A. thaliana* ont été identifiés huit transcrits dérivant des gènes TAS (TAS1a-c, TAS2, TAS3a-c et TAS4). Ces transcrits sont clivés par un complexe AGO/miARN spécifique ; ils sont ensuite convertis en ARNdb par la polymérase ARN dépendante RDR6, l'ARNdb est reconnu par DCL4 pour produire des siRNA de 21 nt, en phase à partir du site de clivage reconnu par le complexe AGO/miARN. Les siRNA sont ensuite chargés dans AGO1 et vont induire le clivage simultané de cibles ARNm ayant une séquence complémentaire au tasiRNA. Parmi les cibles de cette classe de siRNA il y a les ARNm dérivant de gènes qui font partie de grandes famille multigéniques comme ceux codant pour les protéines PPR ou pour les facteurs de transcription MYB et ARF (Allen and Howell, 2010; Bologna and Voinnet, 2014).

### COMPETITION ENTRE RNA SILENCING ET LA VOIE DE DEGRADATION 5'-3' DES ARN

Chez *A. thaliana* il a été observé que lorsque les voies de dégradation 5'-3' ou 3'-5' sont compromises certaines molécules d'ARNm sont prises en charge par la machinerie du RNA silencing (Zhang et al., 2015b; De Alba et al., 2015; Branscheid et al., 2015). Dans ce contexte de dérégulation l'enzyme RDR6 produit des ARN double brin à partir d'une partie des ARNm endogènes ce qui conduit à la production de petits RNA qui peuvent avoir un impact négatif sur le transcriptome et le fitness des plantes (Tsuzuki et al., 2017). Un tel exemple concerne, entre autre, les mutants du decapping, qui ont un développement très retardé, mais pour lesquels ce phénotype est partiellement restauré dans des doubles mutants *vcs/rdr6* et *dcp2/rdr6* (De Alba et al., 2015). Cette observation suggère que le phénotype observé dépend, en partie, de la production de petits ARN secondaires par RDR6. Cette hypothèse est validée par l'étude par séquençage haut débit des petits ARN dans des simple



**Figure 5** Les protéines de la superfamille PIN-like selon la classification de Matelska et al., 2017

Entouré en rose le cluster NYN.1 très répandu chez les plantes, avec 1151 membres sur 1806 chez les Viridiplantae auquel appartient aussi la protéine mammalienne MARF1. Entouré en violet le cluster RNase\_Sc3h12a auquel appartient MCPIP1/Regnase-1. Entouré en bleu le cluster PRORP auquel appartiennent les seule protéines avec activité RNase P chez les plantes. En rouge les nouveaux cluster identifiés dans l'étude.

mutant *vcs* et *dcp2* et des doubles mutants *vcs/rdr6* et *dcp2/rdr6*, qui montre que l'accumulation de petits ARN ainsi produits, appelés dans ce cas *rqc*-siRNA (RNA quality control-siRNA) est dépendante de RDR6 (De Alba et al., 2015). Dans leur ensemble ces résultats démontrent clairement la compétition qu'il existe entre la voie de dégradation 5'-3' des ARN et le phénomène de RNA silencing amplifié.

Chez *A.thaliana* le mutant de l'exoribonucléase 5'-3' cytosolique *XRN4* accumule des transcrits décapés ainsi que des siRNA de 21 nt qui ciblent des ARNm fonctionnels (Souret et al., 2004; Gazzani et al., 2004; Gregory et al., 2008). L'accumulation de transcrits décapés, aberrants, induit la production de siRNA par une voie dépendant de RDR6 (Gazzani et al., 2004). De plus le mutant *xrn4* est moins susceptible à l'infection par le virus à ARNs(+) CMV (Cucumber Mosaic Virus ; Gy et al., 2007). En absence de *XRN4*, les ARN viraux aberrants avec une extrémité 5' accessible qui accumulent dans le cytosol sont convertis en ARNdb et pris en charge par la machinerie du silencing. Ce qui induit une diminution conséquente des ARN viraux (Gy et al., 2007). Le même mécanisme est responsable de la réduction d'ARN dérivant d'un transgène sous promoteur constitutif observé chez le mutant *xrn4* (Vogel et al., 2011; Hayashi et al., 2012).

#### **1.2.6.1 LES ENDORIBONUCLEASE A DOMAINE PIN/NYN**

Les protéines possédant un domaine PIN-like représentent une très vaste famille de nucléases, ayant des fonctions biologiques et une distribution taxonomique très variées. Selon la classification présentée dans l'article de Matelska et al. en 2017, la superfamille de protéine à domaine PIN-like est composée de 5 sous-groupes : FEN, VapC, NYN, PRORP et Mut7-C (**Figure 5**). Les enzymes à domaine NYN (Nedd4-BP1, YacP-like Nuclease) avaient été initialement identifiées *in silico* par homologie de séquence et par prédiction de structure secondaire par Anantharaman et Aravind en 2006. La dénomination NYN de ces auteurs regroupait alors les sous-groupes NYN (YacP-like) et PRORP (Nedd4-BP1) de Matelska et al. 2017. Le site actif des enzymes du sous-groupe NYN possède quatre résidus d'acides aspartiques conservés, supposés nécessaire pour coordonner un seul cation métallique divalent essentiel pour l'activité catalytique (Gobert et al., 2019; Anantharaman and Aravind, 2006). Le domaine NYN est souvent associé à des domaines de liaison à l'ARN comme le domaine PPR (Pentatricopeptide Repeat), le domaine à doigts de zinc CCCH ou le domaine OST-HTH (Anantharaman et al. 2010). Les protéines à domaine NYN sont présentes chez tous les organismes des procaryotes aux eucaryotes mais elles sont exceptionnellement abondantes chez les plantes avec 1151 gènes codant pour les protéines du cluster NYN.1 chez les Viridiplantae (Matelska et al., 2017). Les seules protéines à domaine NYN qui ont été étudiées chez *A. thaliana* sont les protéines MNU1 et MNU2 impliquée dans la maturation des ARNm mitochondriaux (Stoll

and Binder, 2016). Cependant, les données expérimentales concernant l'implication d'enzymes endoribonucléases de la superfamille PIN-like dans la dégradation des ARN, les messagers notamment, sont assez limitées. Dans les paragraphes suivants je vais décrire les connaissances les plus récentes relatives aux protéines PIN/NYN ayant une activité endoribonucléase et impliquées dans le métabolisme des ARN chez les organismes eucaryotes.

## SMG6

Comme présenté précédemment la dégradation des cibles du NMD peut être induite soit par recrutement de la machinerie du decapping soit par coupure endoribonucléolytique. La dégradation par clivage endoribonucléolytique à proximité d'un codon stop prématuré des cibles du NMD a été décrite d'abord chez la Drosophile (Gatfield and Izaurralde, 2004). L'enzyme responsable de ce clivage est appelé SMG6 (Gatfield and Izaurralde, 2004; Huntzinger et al., 2008). Elle contient un domaine 14.3.3-like en N-terminal et un domaine catalytique du sous-groupe VapC. L'importance de cette activité endoribonucléase dans la voie du NMD a été démontrée en utilisant une protéine SMG6 avec le domaine PIN d'une protéine non liée au NMD, qui est capable de rétablir un phénotype sauvage lorsqu'elle est exprimée dans un mutant *smg6*. L'activité catalytique de SMG6 est nécessaire pour la dégradation de transcrits contenant un codon stop prématuré dans les cellules humaines et sa fonction dans la voie du NMD est conservée au cours de l'évolution (Huntzinger et al., 2008). Dans les cellules humaines l'analyse des extrémités 5' des intermédiaires de dégradation par PARE (Parallel Analysis of RNA Ends) dans des mutants *upf1* et *smg6* a permis de montrer que SMG6 a une préférence pour un motif pentamérique spécifique présent sur ses cibles. De plus en absence de SMG6 il y a une accumulation de cibles du NMD dépourvues de la coiffe ce qui suggère une redondance entre les voies de dégradation dépendantes et indépendantes de SMG6 (Schmidt et al., 2015). La reconnaissance des cibles du NMD par SMG6 est possible grâce à l'interaction de l'endoribonucléase avec les EJC à travers le motifs EBM en N-terminal (Kashima et al., 2010). En plus de son rôle dans le NMD, SMG6 est impliquée dans la régulation de la maintenance des télomères dans les cellules humaines cancéreuses même si les mécanismes exact ne sont pas encore clairs (Li et al., 2015). De plus SMG6 est importante pour la différenciation des cellules souches dans l'embryon ainsi que dans la réplication du virus du VIH (Li et al., 2015; Rao et al., 2019). Aucun orthologue de SMG6 n'a pu être identifié chez les plantes sur la base des analyses de similarité de séquences.

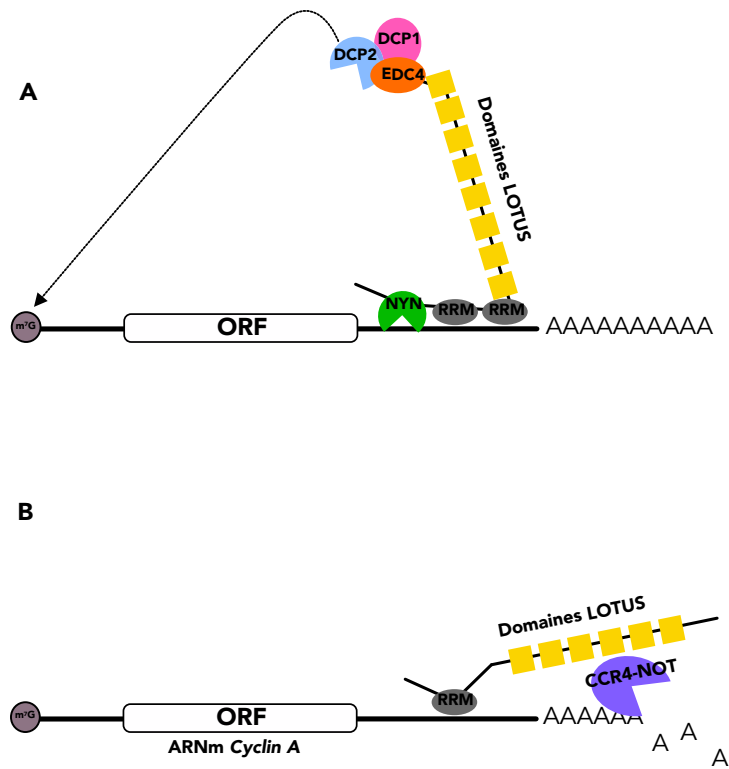
## PRORP

Les protéines PRORP (Proteinaceous RNase P) sont présentes exclusivement chez les eucaryotes. Elles sont composées d'un site catalytique du sous-groupe PRORP (**Figure 5**) précédé d'un domaine de liaison à l'ARN de type PPR responsable de la reconnaissance des substrats. Leur fonction est essentielle pour la maturation des précurseurs des ARN de transfert par clivage de la séquence leader en 5'. Ces enzymes reconnaissent aussi les TLS (tRNA-like structures) et participent à la maturation des ARNm contenant de telles structures (Gobert et al., 2010). Les protéines PRORP sont actives dans les mitochondries des métazoaires, cependant chez les plantes supérieures ce sont les seules enzymes responsables de l'activité RNase P dans le noyau, les mitochondries et les plastides (Gobert et al., 2019). Il est intéressant de noter que PRORP<sub>1</sub> interagit avec l'enzyme MNU2 aussi une enzyme PIN-like mais du sous-groupe NYN chez *A. thaliana* (Bouchoucha et al., 2019). Le rôle exact de cette interaction demeure encore inconnu.

## MCPIP<sub>1</sub>/REGNASE

MCPIP<sub>1</sub>/Regnase a été caractérisé chez l'humain et la souris. Cette enzyme fait partie d'un groupe plus étendu allant de MCPIP<sub>1</sub> à 4 chez l'homme. Ces enzymes contiennent un domaine catalytique du sous-groupe PRORP et un domaine de liaison à l'ARN de type doigt de zinc. MCPIP<sub>1</sub> a été impliqué directement dans la dégradation des ARN de plusieurs virus lorsqu'il est surexprimé et supprime la réplication de HCV tout en contribuant à réguler les défenses de l'hôte par son effet sur l'expression de gènes importants pour la réponse inflammatoire (Lin et al. 2013). En plus de cette activité MCPIP<sub>1</sub> est un suppresseur de l'activité et de la biogenèse des miARN qui agit comme antagoniste de Dicer par clivage des pré-miARN (Suzuki et al., 2011).

Son homologue chez la souris, Regnase-1, a été beaucoup étudié pour son rôle dans la régulation de transcrits impliqués dans la réponse inflammatoire et dans la réponse aux stress (Mao et al., 2017; Habacher and Ciosk, 2017a; Mino et al., 2015). La résolution de la structure de Regnase1 a permis de confirmer la présence de 4 résidus d'acide aspartique nécessaires pour la liaison de l'ARN dans le site actif ainsi que de découvrir que l'oligomérisation des domaines PIN est nécessaire pour l'activité enzymatique et que le domaine N-terminal favorise l'activité endoribonucléase (Xu et al., 2012; Yokogawa et al., 2016). Regnase-1 reconnaît une structure en tige boucle localisée à plus de 20 nt d'un codon stop dans la 3' UTR de ces cibles qui peuvent changer selon les conditions et le type cellulaire mais sont globalement associées à la régulation de la réponse inflammatoire. Regnase-1 interagit avec UPF1 sur des ARNm en cours de traduction, cette interaction est critique pour la dégradation de ces cibles dans une voie de dégradation qui a été appelée Regnase Mediated Decay (RMD) (Yokogawa et al., 2016; Mino et al., 2015).



**Figure 6** Modèles du mécanisme de dégradation des cibles de MARF1 chez l'humain et chez la Drosophile (Adapté de Nishimura et al., 2018; Zhu et al., 2018)

(A) hMARF1 induit la dégradation des cibles par coupure endoribonucléolytique et ensuite par recrutement des facteurs de decapping DCP1 et DCP2 sur l'ARN à travers son extension en C-ter (aa 1563-1742).

(B) DmMARF1 recrute le complexe CCR4-NOT à travers les domaines LOTUS et induit la déadénylation de l'ARNm de la *Cyclin A*.

## RDE-8

L'enzyme RDE-8 (RNAi Deficiency 8) a été identifiée chez *C. elegans* lors d'un crible génétique visant la caractérisation d'individus déficients dans la voie de l'interférence à ARN (Tsai et al., 2015). RDE-8 dispose d'un site catalytique classifié dans le sous-groupe NYN mais pas d'autre domaine connu pour interagir avec l'ARN. Le phénotype du mutant Rde8 a été complété par l'expression de la version catalytiquement active de RDE-8 mais pas du mutant catalytique ce qui indique l'importance de l'activité catalytique pour la fonction biologique de RDE-8. De plus il a été montré que RDE-8 est recruté par la protéine Argonaute RDE-1 afin d'induire une coupure endoribonucléolytique sur les ARN cible ; le fragment 3' ainsi produit est ensuite uridylé et pris en charge par la protéine RdRP RRF-1 responsable de l'amplification du signal de silencing (Tsai et al., 2015; Habacher and Ciosk, 2017b).

## MARF1/LIMKAINB

La protéine MARF1 (Meiosis regulator and mRNA stability factor 1), appelée aussi LimkainB est caractérisée, chez l'humain et la souris, par la présence d'un domaine NYN en N-terminal suivi par deux motifs RRM et huit domaines LOTUS en région C-terminale. Cependant, chez la Drosophile, MARF1 ne possède pas de domaine NYN mais contient un motif RRM ainsi que six domaines LOTUS en tandem. Dans ces organismes, MARF1 interagit directement avec des protéines impliquées dans les voies générales de dégradation des ARNm. Chez la Drosophile il a été montré par immunoprécipitation que MARF1 interagit spécifiquement avec le complexe CCR4/NOT responsable de la déadénylation cytosolique des ARNm et que cette interaction ne dépend pas de la présence de l'ARN (**Figure 6**). La protéine MARF1 chez la Drosophile est fondamentale pour la maturation des oocytes (Zhu et al., 2018). Ceci est possible grâce à l'induction de la déadénylation de l'ARNm de la CyclineA une étape cruciale pour la maturation des gamètes femelles. Des expériences de ancrage moléculaire (*tethering*) avec un ARN rapporteur GFP ont permis de montrer qu'un seul domaine LOTUS suffit pour induire la déadénylation d'un transcrite ciblé (Zhu et al., 2018). L'étude du réseau d'interaction de hMARF1 par BioID et immunoprécipitation couplé à la spectrométrie de masse a permis de montrer qu'elle est spécifiquement présente dans un complexe protéique qui interagit avec Ge-1, DCP1 et DCP2 (Nishimura et al., 2018a). Les auteurs proposent un modèle où hMARF1, après coupure endoribonucléolytique de ses cibles, recrute le complexe de decapping pour la dégradation du fragment 5' (**Figure 6**) Chez les mammifères l'activité catalytique de MARF1 et sa capacité à interagir avec l'ARN sont nécessaires pour la progression dans la méiose dans les oocytes ainsi que pour la surveillance des rétrotransposons (Su et al., 2012a; Yao et al., 2018). De plus MARF1 a été caractérisé comme étant nécessaire pour la régulation de transcrits spécifiques, comme par

exemple IFI44L (Interferon Induced Protein 44 Like) et sa fonction est cruciale pour la différenciation des neurones (Bloch et al., 2014; Kanemitsu et al., 2017). Une analyse initiale tendait à montrer que hMARF1 colocalisait partiellement avec les marqueurs des peroxyosomes ABCD3 et PXF (Dunster et al., 2005). Par des études ultérieures, il est maintenant clair que hMARF1 colocalise dans les P-bodies avec ses partenaires protéiques DCP1 et Ge-1 (Bloch et al., 2014; Yao et al., 2018).

### **MNU1/MNU2**

Les protéines MNU1 et MNU2 (Mitochondrial Nucleases 1 et 2) ont été caractérisées dans les mitochondries de la plante modèle *A. thaliana*. Elles sont composées d'un domaine catalytique du sous-groupe NYN suivi d'un domaine OST-HTH de liaison à l'ARN, d'un domaine WW d'interaction protéique pour MNU2 (nécessaire pour l'interaction avec la protéine PRORP1 citée auparavant), d'un domaine OHA (OST-HTH Associated domain) de fonction inconnue et finalement d'un second domaine OST-HTH. Ces deux enzymes interviennent dans la maturation par clivage endoribonucléolytique de l'extrémité 5' des transcrits mitochondriaux (Stoll and Binder, 2016).



### 1.3 LES GRANULES D'ARN CYTOSOLIQUES ET LEUR ROLE DANS LE METABOLISME DES ARNm

Les ARNm sont dans la cellule, toujours associés avec de nombreuses protéines pour former des complexes RNP. La régulation du métabolisme des ARNm dépend de la nature de ces RNP, qui vont permettre la maturation, l'export, la traduction, le stockage et la dégradation des ARNm. Dans le cytosol, certains RNP peuvent induire la formation de compartiments cellulaires dynamiques sans membrane. Ces compartiments sont appelés granules à RNP et incluent notamment les P-bodies les granules de stress ainsi que les granules à ARN spécifiques des cellules neuronales. Ces compartiments ont des propriétés et une composition en protéines et ARN différents. Les granules à RNP, conservées chez les eucaryotes, se forment par un phénomène physique appelé transition de phase liquide-liquide. Les mécanismes physiques à la base de ce processus sont l'objet de plus en plus d'études : une solution homogène, le cytosol, se sépare en deux phases liquides, une plus dense qui contient une haute concentration de protéines de liaison à l'ARN et des ARN sous forme de complexes RNP et l'autre qui possède une concentration plus faible (Alberti and Dormann, 2019). La composition et la fonction des P-bodies et des granules de stress est différente mais les protéines impliquées dans leur formation ont des caractéristiques communes. Elles possèdent notamment des régions IDR (Darling et al., 2018). La présence d'ARN et de protéines possédant des IDR est nécessaire pour la formation de ces compartiments (Teixeira et al., 2005; Youn et al., 2019). Les modifications internes de l'ARNm peuvent influencer la formation de ces compartiments, ainsi la présence de plusieurs adénosines méthylées en position 6 (m6A) dans un ARNm favorise la transition de phase liquide-liquide *in vitro* et *in vivo* (Ries et al., 2019). On peut noter, qu'il a été proposé que les granules de stress et les P-bodies peuvent fusionner et échanger des composants (Kedersha et al., 2005; Stoecklin and Kedersha, 2013).

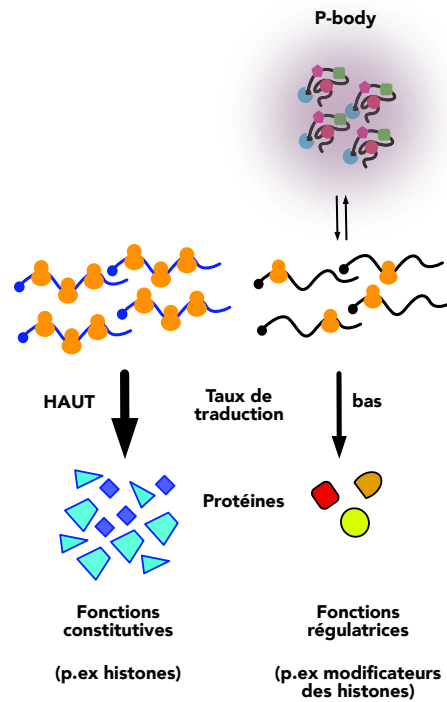
Comme mentionné précédemment, les granules RNP sont des structures dynamiques. Le transport des granules de RNP est notamment crucial pour le fonctionnement des neurones et la perturbation de ce mouvement est à la base de plusieurs maladies neurodégénératives, par exemple la maladie d'Alzheimer ou la sclérose latérale amyotrophique (Oeffinger and Zenklusen, 2019). Le mouvement à longue distance des granules à ARN est lié à l'interaction avec les microtubules alors que le mouvement à courte distance dépend de l'actine (Aizer and Shav-Tal, 2008). Chez les mammifères, le transport des P-bodies est lié à l'actine et aux microtubules (Aizer and Shav-Tal, 2008). Les microtubules interagissent avec certaines protéines présentes dans les P-bodies. Les deux intermédiaires connus pour cette interaction, dans les cellules humaines, sont la Myosine-VA qui interagit avec eIF4E (Lindsay and McCaffrey, 2011) et la Nesprin-1 qui interagit avec Matr3 (Rajgor et al., 2016). Chez les plantes, il a été démontré que le mouvement intracellulaire des P-bodies est

dépendent de l'actine et est rendu possible par l'interaction directe de DCP1 avec des membres de la famille des myosine XI (Steffens et al., 2014).

### 1.3.1 LES GRANULES DE STRESS

Lors d'un stress, la cellule répond en réprimant la traduction d'un grand nombre de d'ARNm pour permettre la production d'autres protéines spécifiques, dont enzymes et chaperonnes, afin de répondre au stress de manière efficace. Comme résultat de cette répression massive de la traduction, des RNP, formés principalement par des protéines impliquées dans l'initiation de la traduction et d'ARNm polyadénylés, se condensent et forment les granules de stress (Buchan and Parker, 2009). Ceci permet une réponse rapide de la cellule au cours du stress et ce processus, non-destructif, permet aussi la réactivation rapide de la traduction de ce pool d'ARNm lorsque le stress disparaît (Anderson and Kedersha, 2002a). Dans les cellules des mammifères la formation des granules de stress est liée à la dissociation des polysomes. La phosphorylation du facteur d'initiation de la traduction eIF2 $\alpha$  mène au remodelage du complexe d'initiation de la traduction sur des transcrits qui codent pour des protéines non nécessaires à la réponse au stress. Ainsi la présence d'un complexe d'initiation inefficace sur ces ARNm induit la relocalisation, *via* les protéines de liaison à l'ARN TIA, de ces RNP dans les granules de stress (Anderson and Kedersha, 2002b; Wolozin, 2012). Les facteurs protéiques contenus dans les granules de stress sont assez bien conservés chez les animaux, la levure et les plantes (Hubstenberger et al., 2017; Jain et al., 2016; Kosmacz et al., 2019). Le protéome des granules de stress comprend les facteurs d'initiation de la traduction eIF4F, EIF3, les poly(A) binding protein (PABP) ainsi que la petite sous-unité ribosomale (40S). De plus les granules de stress contiennent des tRNA synthétases ainsi que des facteurs de biogenèse des ribosomes. Chez les mammifères ils contiennent des protéines impliquées dans les maladies neurodégénératives (Jain et al., 2016; Wolozin, 2012). Le protéome des granules de stress est enrichi en protéines de liaison à l'ARN ayant de régions de faible complexité qui permettent les liaisons protéine-protéine et protéine-ARN nécessaires pour la formation de ces compartiments (Radó-Trilla and Albà, 2012; Harrison and Shorter, 2017). Il est intéressant de mentionner que chez les plantes certaines enzymes impliquées dans des voies spécifiques associées à la réponse au stress (par exemple métabolisme de l'éthylène et des glucosinolates) sont aussi des composants des granules de stress (Kosmacz et al., 2019).

Une étude transcriptome a permis de montrer que 80% des ARN contenus dans les granules de stress sont des ARNm. Les autres ARN présents sont des lincRNA (long intragenic non coding RNA) et des ncRNA (ncRNA) (Khong et al., 2017). Il n'y a pas d'enrichissement en ARNm avec des fonctions spécifiques dans les granules de stress, tout ARN transcrit peut y être localisé. Par contre, il a été



**Figure 7** La composition en ARN des p-bodies selon Hubstenberger et al., 2017  
(Adapté de Hubstenberger et al., 2017)

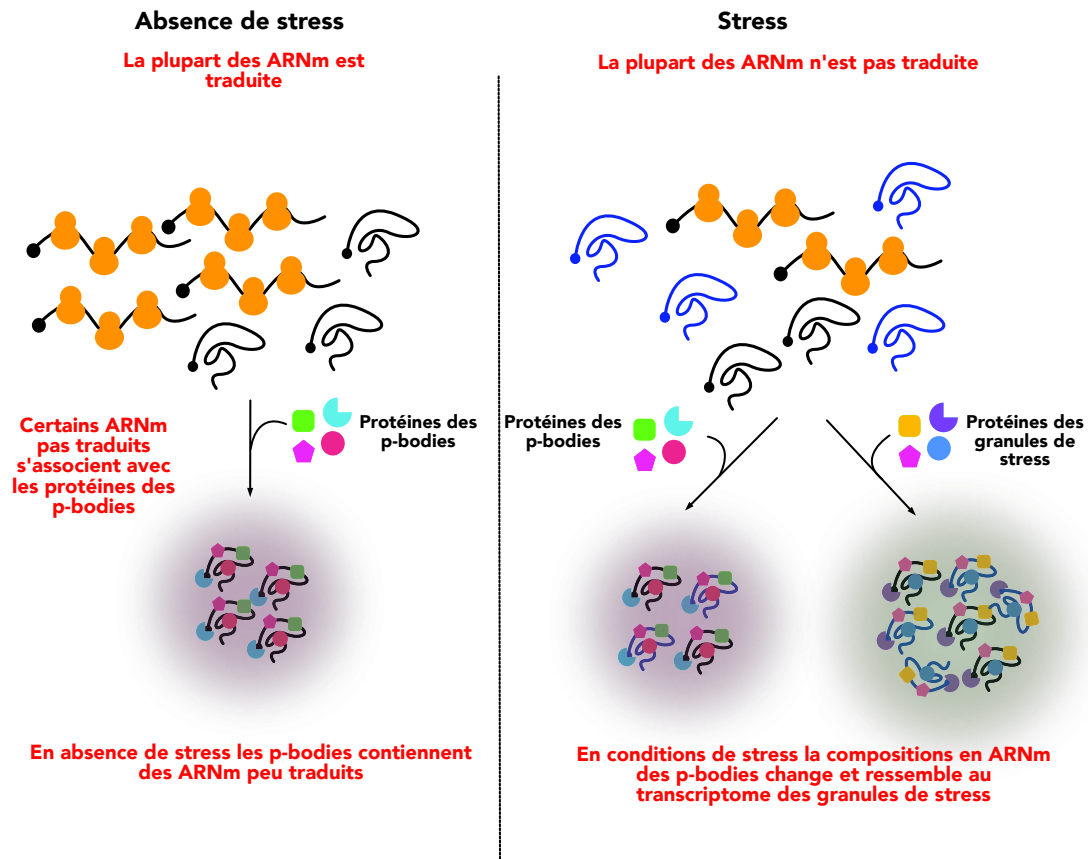
Les ARN avec un niveau de traduction bas et qui codent pour des protéines régulatrices sont stockés dans les p-bodies de manière réversible. Les ARN qui ont un engagement polysomal élevé et qui codent pour des protéines avec des fonctions constitutives sont exclus des p-bodies.

montré que le niveau de traduction et la taille de l'ARN influence sa localisation dans les granules de stress. En effet, les granules de stress sont enrichis en ARNm ayant un niveau de traduction bas et une longue séquence codante et/ou longues UTR (Khong et al., 2017).

### 1.3.2 LES PROCESSING BODIES

Les P-bodies sont des granules RNP de 100-300 nm de diamètre. Au contraire des granules de stress, ces compartiments sont déjà présents dans la cellule en absence de stress, bien que suite à un stress, leur nombre et leur dimension augmentent (Teixeira et al., 2005). Ceci semble indiquer que ces structures sont importantes pour l'homéostasie cellulaire indépendamment de l'état de la cellule, mais participent aussi à la réponse au stress. Les P-bodies ont été identifiés lors des études de localisation des protéines de decapping DCP1 et DCP2 et de la protéine GW182, impliqué dans la voie du RNA silencing. Ces études ont montré que ces protéines co-localisent dans des foci cytosoliques qui contiennent des ARNm traductionnellement inhibés, distincts des granules de stress (Parker and Sheth, 2007; Oeffinger and Zenklusen, 2019). Les P-bodies sont composés principalement par des protéines de liaison à l'ARN impliquées dans les voies de dégradation des transcrits. Notamment ils contiennent les facteurs du complexe de decapping DCP1, DCP2, EDC3/DCP5, Ge-1/VCS, l'exoribonucléase XRN1/XRN4, les facteurs du NMD, UPF1/2 et SMG7, les protéines AGO1 et AGO2 chez l'homme (pas de données à ce sujet chez les plantes actuellement), LSM14A LSM14B, les protéines impliquées dans la répression de la traduction DDX6 et 4E-T, des facteurs impliqués dans des voies spécialisées de dégradation (STAU) ainsi que d'autres protéines de liaison à l'ARN (Sheth and Parker, 2003; Kedersha et al., 2005; Hubstenberger et al., 2017). La comparaison du protéome des granules de stress et des P-bodies dans les cellules humaines indique que 75% des protéines des P-bodies et 91% des protéines des granules de stress sont spécifiques de leur compartiment respectif (Jain et al., 2016; Hubstenberger et al., 2017). A cause de la présence des facteurs impliqués dans la dégradation des ARN, notamment dans le decapping, la fonction des P-bodies a été initialement associée à la dégradation 5'-3' des ARNm (Sheth and Parker, 2003). Cependant il a été montré que la perturbation des P-bodies n'empêche pas la dégradation des ARN (Eulalio et al., 2007b). De plus, des travaux récents montrent que les ARN dans les P-bodies sont réprimés au niveau traductionnel, mais pas dégradés (Horvathova et al., 2017; Hubstenberger et al., 2017). Ces données suggèrent que les P-bodies sont un lieu de stockage et non de dégradation.

La purification des P-bodies à partir de cellules humaines par FACS (Hubstenberger et al., 2017) montre que la plupart des protéines de liaison à l'ARN sont contenues dans les P-bodies avec leurs ARN cibles et qu'il s'agit, pour la plupart, de RNP avec une fonction de régulation. De plus les P-bodies sont dépourvus de RNP contenant des ARNm codant pour des protéines ayant une fonction



**Figure 8** La composition en ARN des p-bodies en absence ou en présence de stress selon Matheny et al., 2019

(Adapté de Matheny et al., 2019)

La composition en ARN des p-bodies dépend de leur niveau de traduction. En absence de stress les ARN peu traduits sont séquestrés dans les p-bodies. Lors d'un stress le niveau de traduction des ARNm décroît de manière critique. Le grand nombre d'ARNm peu traduits est donc séquestré dans les p-bodies ou les stress granules, sans spécificité.

constitutive. Cette observation suggère que les P-bodies représentent une réserve de RNP spécifiques, disponibles dans le cytosol et qui peuvent être mobilisés rapidement selon les besoins de la cellule (**Figure 7**) (Hubstenberger et al., 2017).

La formation des P-bodies chez les mammifères dépend de la présence de l'hélicase DDX6 (RH6, RH8 et R12 chez Arabidopsis) (Ayache et al., 2015; Di Stefano et al., 2019). Chez la drosophile, l'interaction entre DDX6 et des protéines 4E-T impliquée dans l'inhibition de la traduction est nécessaire pour la formation des P-bodies contrairement aux interactions entre DDX6 et les protéines impliquées dans la dégradation des ARN (Kamenska et al., 2016). Chez Arabidopsis la protéine DCP5, impliquée dans l'inhibition de la traduction, est requise pour la formation des P-bodies (Xu and Chua, 2009).

L'analyse du transcriptome des P-bodies par Hubstengerger et al. révèle que tous les transcrits contenus dans les P-bodies ont comme caractéristique commune leur faible taux de traduction. Ces ARNm peu traduits, codent pour des protéines régulatrices, ne sont pas dégradés mais disponibles dans le cytosol pour être mobilisés au besoin. La comparaison des transcriptomes des P-bodies dans les cellules humaines en présence ou non de stress montre un changement dans le contenu en ARN des P-bodies en réponse au stress (Matheny et al., 2019). De plus la comparaison avec le transcriptome des granule de stress (Khong et al., 2017) montre que le contenu en ARN des P-bodies en cas de stress est plus similaire au transcriptome des granules de stress ( $R = 0,85$ ) que à celui des P-bodies en absence de stress ( $R = 0,55$ ) (Matheny et al., 2019). Sur la base de leurs résultats Matheny et collègues proposent que les ARNm contenus dans les P-bodies en absence de stress sont des transcrits peu traduits et que c'est leur état de traduction ainsi que leur association avec facteurs spécifiques des P-bodies, et non leur fonction, qui influence leur localisation dans les P-bodies. Ainsi ils désignent le niveau de traduction comme seul déterminant pour la localisation des ARNm dans les P-bodies en présence ou en absence de stress (**Figure 8**) (Matheny et al., 2019).

Les deux articles discutés précédemment proposent deux visions un peu différentes des P-bodies et de leurs propriétés. Cependant les expériences réalisées dans les deux études n'ont pas été réalisées avec les mêmes lignées cellulaires. Avec d'un côté des cellules HEK293, cellules épithéliales dérivées d'embryons humains et de l'autre des cellules épithéliales U-2 OS, obtenues à partir de cellules osseuses cancéreuses. Dans ces conditions, il sera très intéressant de suivre l'évolution de ce domaine encore nouveau et de comparer ces résultats initiaux avec d'autres obtenus dans d'autres cellules provenant de tissus différents. Ceci permettra de déterminer si l'on peut extrapoler des caractéristiques générales aux ARNm s'accumulant dans les P-bodies ou si l'on constate des différences très marquées en fonction des tissus et/ou de la nature des cellules utilisées. Ce type

d'étude pourrait aussi être mené dans d'autres organismes (incluant les plantes), ce qui permettrait de déterminer si ces structures ont gardées les mêmes propriétés au cours de l'évolution.

# OBJECTIFS

Chez les plantes aucune protéine à domaine NYN n'a été jusqu'alors impliquée dans la dégradation des ARNm. L'objectif de ma thèse était d'étudier une protéine à domaine NYN de fonction jusqu'alors inconnue et qui avait été identifiée comme étant associée à plusieurs protéines importantes pour la dégradation des ARNm (Thèse C. Chicois). Celle-ci a été nommée DAE pour **D**CP1 **A**ssociated **E**ndoribonuclease, à l'issue des premières expériences réalisées au cours de ma thèse.

Les objectifs de ma thèse étaient :

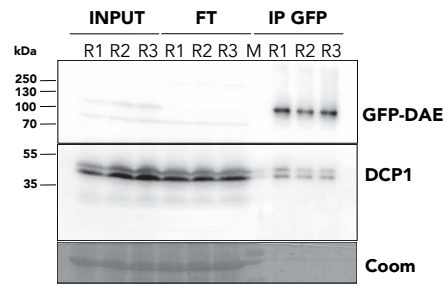
- D'analyser les **complexes protéiques** associés à DAE par une approche couplant co-immunoprécipitation (co-IP) et spectrométrie de masse
- D'étudier l'**activité enzymatique de DAE** à travers des approches *in vitro* et *in vivo*
- De déterminer la **localisation subcellulaire** de DAE et l'implication de ses domaines dans cette localisation
- D'évaluer l'impact de la dérégulation de DAE, par une **étude du transcriptome** de lignées mutantes et de surexprimeurs, par RNA-seq.

Les travaux réalisés lors de mon doctorat ont montré que DAE possède une activité enzymatique de type endoribonucléase et s'associe avec un complexe composé d'activateurs de decapping. Des liens très étroits ont pu être mis en évidence entre l'effet d'une dérégulation de *DAE* et du gène *VCS*, un acteur crucial pour le decapping. Ces résultats ont permis d'élaborer un modèle expliquant le mode d'action de DAE et son positionnement par rapport à la machinerie du decapping.

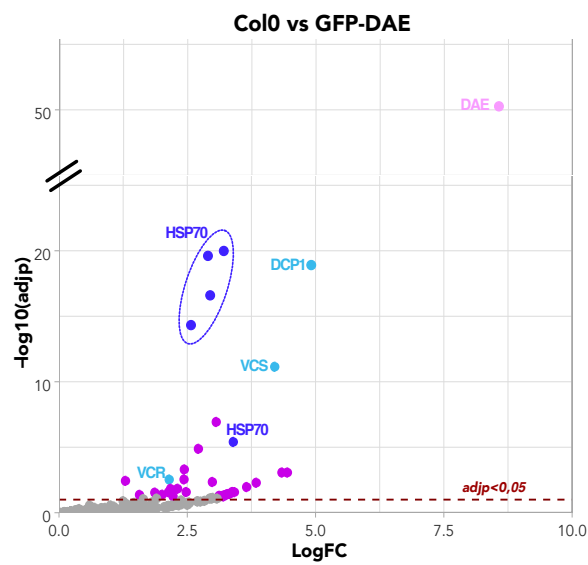


# RESULTATS

**A**



**B**

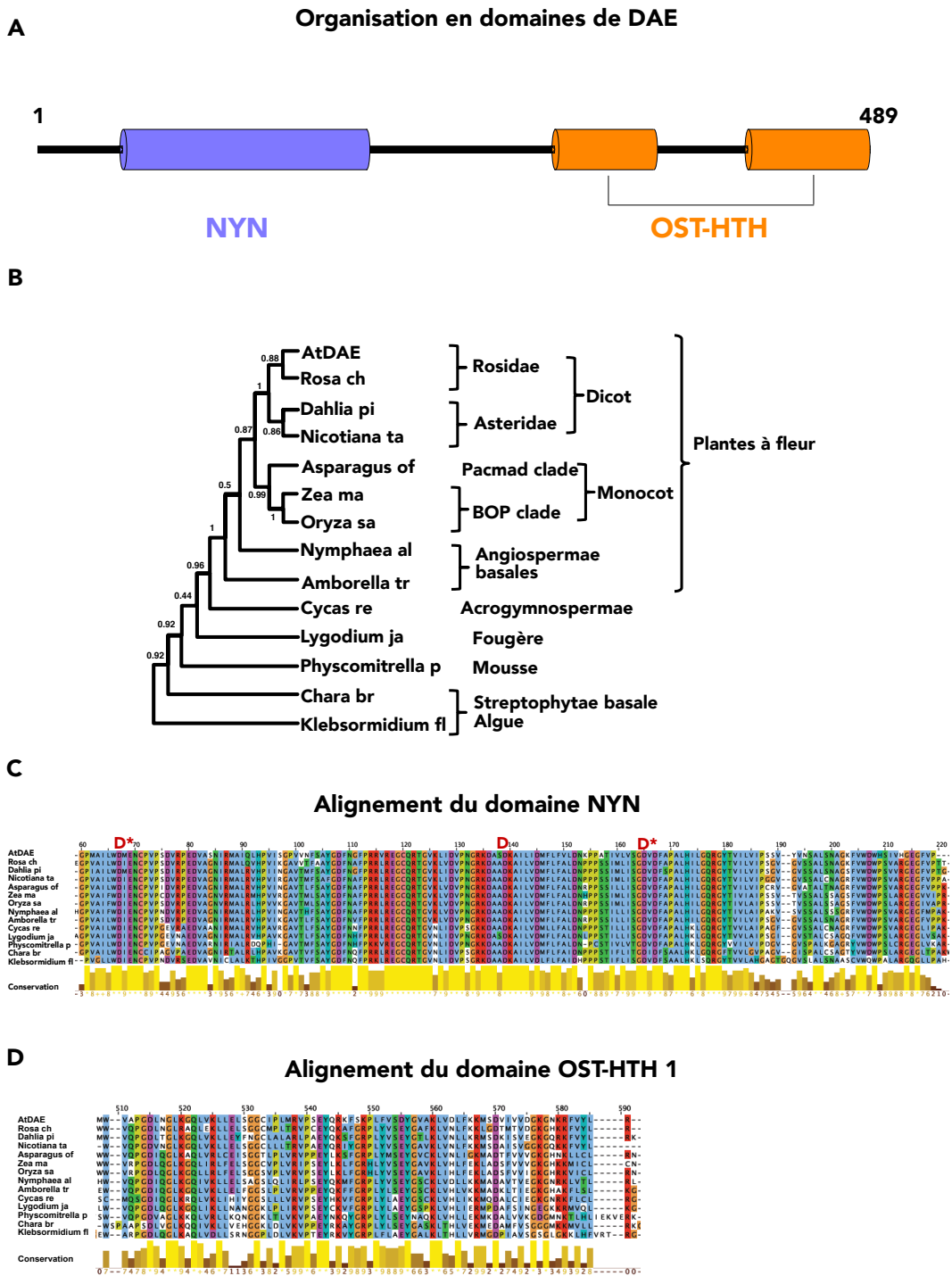


**Figure 9 DAE co-purifie avec les activateurs de decapping DCP1 et VCS.**

(A) Western blot des immunoprécipitations de GFP-DAE. GFP-DAE et DCP1 ont été détectés par hybridation de la membrane avec un antiGFP (haut de la figure) et un antiDCP1 (bas de la figure) dans les inputs, flow-through (FT) et immunoprécipitations (IP-GFP) de trois réplicats techniques (R1, R2, R3). Marqueur de taille (M) (B) Semi-volcano plot des protéines enrichies avec Log du fold change (FC) supérieur à 1 et p-value ajustée inférieure à 0,05 dans les IP DAE (3 réplicats biologiques, chacun en 3 réplicats techniques)

## 1.4 DAE CO-PURIFIE AVEC DCP1 ET VCS, COMPOSANTS DU COMPLEXE DE DECAPPING

Les travaux précédents menés au laboratoire ont permis d'identifier une protéine à domaine NYN de fonction inconnue qui co-purifie avec deux facteurs des P-bodies, UPF1, un facteur clé du NMD et DCP5 une protéine impliquée dans le decapping (Chicois et al., 2018). De plus une étude de sa localisation a montré que cette protéine se localise également dans des foci cytoplasmiques, probablement dans les P-bodies puisqu'elle co-localise avec UPF1. Ces résultats initiaux plaçaient cette endonucléase putative comme un très bon candidat pour représenter un nouveau facteur impliqué dans la dégradation des ARNm. Dans le but d'étudier le réseau d'interaction de cette protéine, nous avons utilisé une approche d'analyse protéomique non biaisé, basé sur la co-IP couplée avec une analyse par spectrométrie de masse (LC-MS/MS). Ces expériences ont été réalisées à partir d'une lignée établie précédemment au laboratoire, et correspondant à la lignée mutante *dae-3* exprimant la séquence codante *DAE* fusionnée en N-terminal à la protéine fluorescente GFP (Thèse de Clara Chicois). Trois réplicats de co-IP ont été réalisés, les éluats ont été chargés sur SDS-PAGE et la présence de GFP-DAE a été révélée avec un anticorps anti-GFP par western blot. Le résultat du western montre un enrichissement important de la protéine de fusion GFP-DAE dans la co-IP comparé aux lysats (Input) dans lesquels GFP-DAE est à la limite du seuil de détection (**Figure 9A**). Ces fractions enrichies ont été analysées par spectrométrie de masse, les résultats sont représentés sous forme de semi-volcano plot (**Figure 9B**). Après analyse statistique, 36 protéines ayant une p-value ajustée  $\leq 0,05$  ont été retenues. Nous pouvons observer que les protéines les plus significativement enrichies dans les co-IP GFP-DAE sont représentées par les composants du complexe de decapping DCP1 et VCS. L'homologue de VCS, VCR, est aussi enrichi dans les co-IP DAE. Parmi les protéines les plus enrichies nous trouvons également plusieurs membres de la famille des HSP-70 (Heat Shock Protein) (**Figure 9B**) ainsi que des protéines du cytosquelette comme TUB2 et TOR1 (**Annexe 1**). L'incubation de la même membrane avec l'anticorps anti-DCP1 valide la présence de DCP1 dans la co-IP DAE (**Figure 9B**). En conclusion le résultat de l'analyse des protéines qui co-purifient avec DAE dans des expériences de co-IP montre qu'elle est présente dans des complexes protéiques qui contiennent DCP1 et VCS.

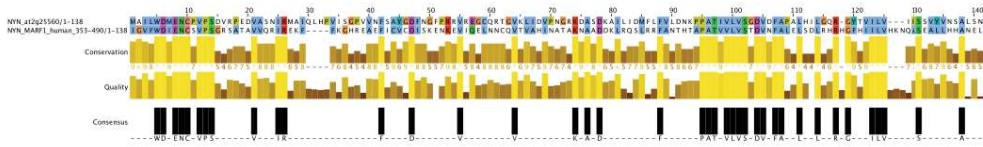
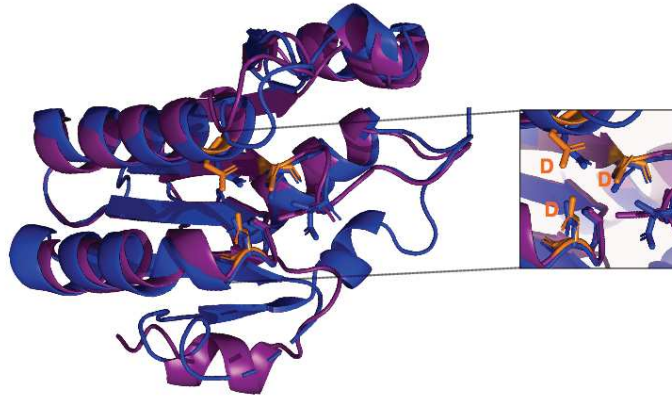
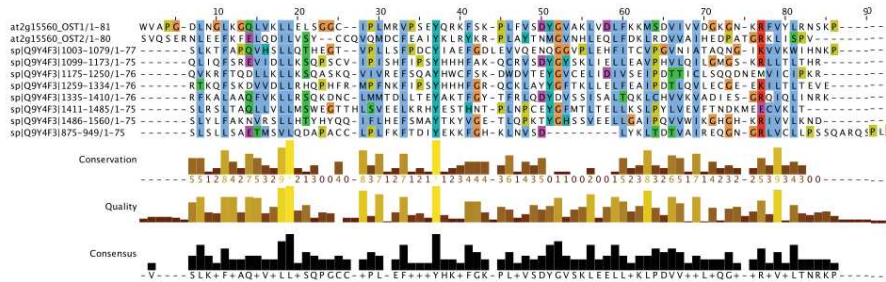
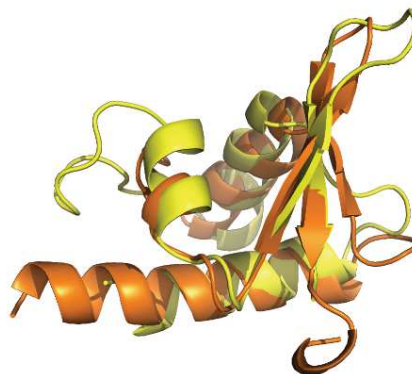


**Figure 10** DAE est conservée des algues basales aux plantes à fleur.

(A) Structure protéique de DAE. Le domaine NYN en violet et les deux domaines OST-HTH en orange (B) Pour la construction de l'arbre, les séquences ont été alignées avec MUSCLE (v3.8.31). L'arbre phylogénétique a été construit avec la méthode du maximum de vraisemblance implémentée dans le logiciel PhyML (v3.1/3.0 aLRT). La représentation graphique de l'arbre a été modifiée avec MEGA. Les valeurs de bootstrap sont indiquées pour chaque nœud. (C) et (D) Alignement des séquences protéiques des domaines NYN et OST-HTH dans les espèces considérées pour la construction de l'arbre phylogénétique.

## 1.5 DAE EST CONSERVEE CHEZ LES VEGETAUX DES ALGUES JUSQU'AUX PLANTES A FLEURS

La protéine DAE d'*A. thaliana* possède deux types de domaines prédits, un domaine endonucléase de type NYN dans la partie N-terminale suivi par deux domaines de liaison à l'ARN de type OST-HTH dans la région C-terminale (**Figure 10A**). Afin d'analyser la conservation de DAE chez les plantes nous avons fait une recherche de séquences homologues de DAE en utilisant l'outil tBLASTn. Cet outil interroge des banques de données de séquences d'ADN traduites avec une séquence protéique d'intérêt. Cette analyse réalisée dans 13 espèces végétales, des algues primitives jusqu'aux angiospermes, a permis d'identifier des homologues de DAE. Ces séquences ont été utilisées pour construire un arbre phylogénétique (**Figure 10B**), une analyse de la robustesse a été obtenue avec la méthode de bootstrap, les valeurs obtenues pour chaque branche sont présentées dans la **Figure 10B**. L'arbre phylogénétique ainsi obtenu montre que DAE est une protéine ancienne, qui a été conservée au cours de l'évolution des algues basales jusqu'aux angiospermes. Nous avons également aligné les séquences protéiques des domaines NYN et OST dans les différentes espèces végétales analysées (**Figure 10C et D**). L'alignement du domaine NYN montre que celui-ci est très bien conservé (pourcentage d'identité entre 62,26% et 82,91%), ce qui est attendu pour un domaine catalytique (**Figure 10C**). Sur cet alignement, on remarque la conservation des trois résidus d'acide aspartique nécessaires pour l'activité enzymatique des domaines NYN (en rouge dans la figure). L'alignement des domaines OST-HTH montre également une bonne conservation dans les différentes espèces végétales avec des scores un peu moins élevés allant de 46,15% à 73,68%, ce qui est attendu pour un domaine d'interaction par rapport à un domaine catalytique. Cette étude montre que DAE est une protéine très ancienne, sa séquence protéique a été conservée au cours de 460 millions d'années d'évolution et ceci suggère que DAE pourrait avoir une fonction importante chez les végétaux.

**A****B****NYN : HsMARF vs DAE****C****D****OST-HTH : MmMARF vs DAE**

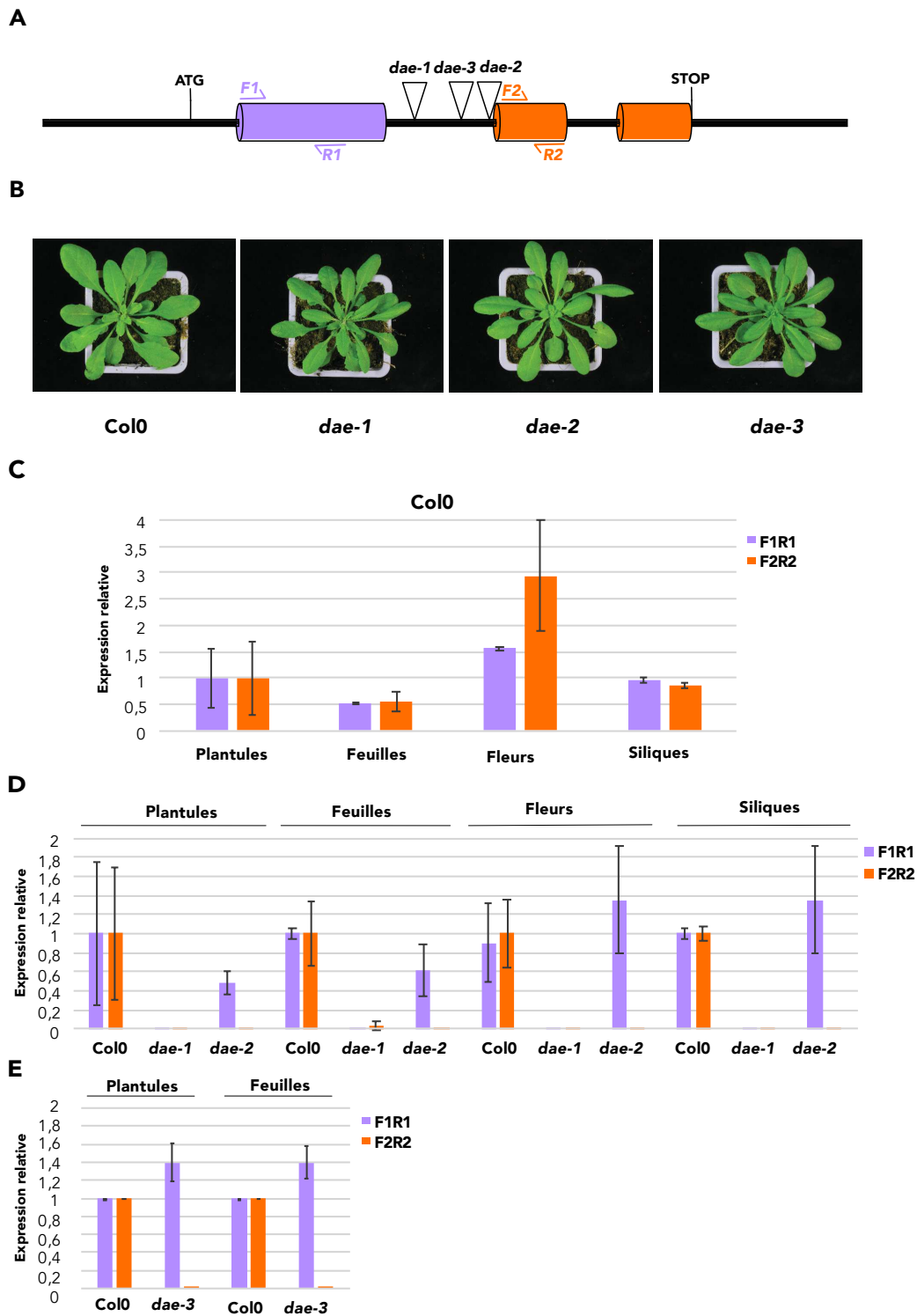
**Figure 11** La structure 3D des domaines NYN et OST-HTH1 de DAE ressemble à celle des domaines NYN et OST-HTH de MARF1 chez l'homme et la souris respectivement

(A) Alignement du domaine NYN de DAE et de MARF1 chez l'humain (HsMARF1). (B) Modélisation du domaine NYN de DAE en violet et MARF1 en bleu, magnification du site actif, les résidus d'acides aspartiques sont indiqués en orange. (C) Alignement des deux domaines OST-HTH de DAE avec les 8 domaines OST-HTH de MARF1 chez la souris (MmMARF1). (D) Modélisation du domaine OST-HTH1 de DAE en jaune avec le domaine OST-HTH1 de MmMARF1 en orange.

## 1.6 LE PLUS PROCHE HOMOLOGUE DE DAE CHEZ LES MAMMIFERES EST L'ENDONUCLEASE MARF1

Une recherche par BLASTP du plus proche homologue de DAE chez les mammifères révèle que DAE présente une homologie de séquence avec la protéine à domaine NYN MARF<sub>1</sub>, de 27,33% chez l'homme et de 26,74% chez la souris. MARF<sub>1</sub> est une endoribonucléase qui interagit avec DCP<sub>1</sub>, DCP<sub>2</sub> et avec Ge-1, l'homologue de VCS chez les mammifères (Bloch et al., 2014; Nishimura et al., 2018a). Les structures 3D des domaines NYN et OST-HTH de MARF<sub>1</sub> ont été résolues chez l'homme (HsMARF<sub>1</sub>) et la souris (MmMARF<sub>1</sub>) respectivement et ont été récemment publiées dans deux études indépendantes (Nishimura et al., 2018b; Yao et al., 2018). Avec l'aide de notre collaborateur Anthony Gobert de l'IBMP, nous avons comparé les structures 3D de MARF<sub>1</sub> et DAE en utilisant le logiciel en ligne Phyre2 (Kelley et al., 2015). Cette analyse a permis de faire une prédiction de la structure 3D du domaine NYN et du domaine OST-HTH<sub>1</sub> de DAE. Les deux structures présentées en **Figure 11A** et **B** représentent le résultat de cette modélisation. Nous pouvons constater que la structure 3D du domaine NYN de DAE (**Figure 11A**), en violet, s'ajuste parfaitement à la structure 3D du domaine NYN de HsMARF en bleu. De plus nous pouvons remarquer une conservation parfaite de l'emplacement des résidus d'acide aspartique dans le site actif, en orange. MARF<sub>1</sub> ayant été validée comme étant une authentique endoribonucléase, ce résultat suggère que DAE, comme MARF<sub>1</sub>, pourrait porter la même activité catalytique.

D'autre part, la modélisation du domaine OST-HTH de DAE (**Figure 11B**), en jaune et de MmMARF<sub>1</sub> en orange, révèle aussi une conservation de la structure 3D. Ce qui suggère que la capacité de lier l'ARN est aussi conservée.



**Figure 12** Caractérisation des mutants d'insertion T-DNA *dae-1*, *dae-2* et *dae-3*.

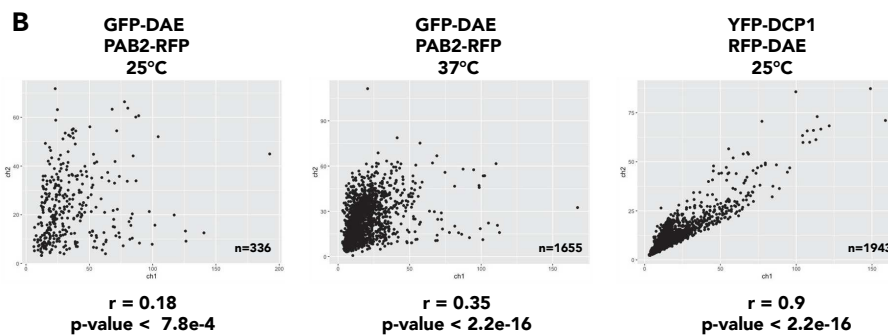
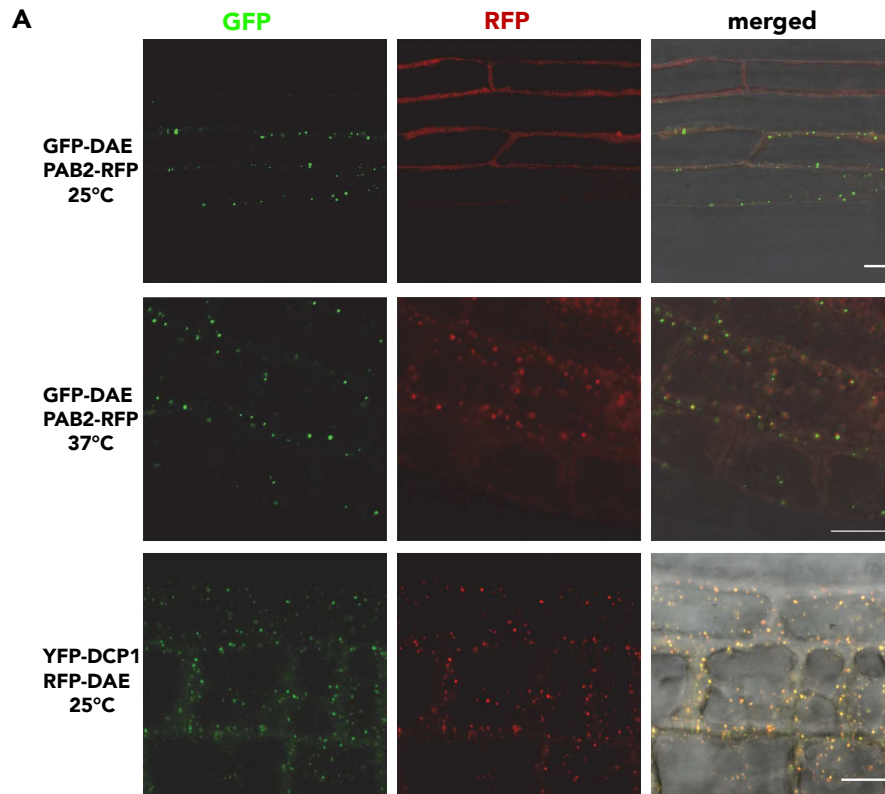
(A) Représentation graphique du gène *DAE*. La position des insertions dans les lignées *dae-1*, *dae-2* et *dae-3* est indiquée ainsi que les sites d'hybridation des deux couples d'amorces utilisées (F1R1 et F2R2). (B) Phénotype des mutants d'insertion au stade rosette, 20 jours après germination. (C) Analyse de l'expression par QPCR du gène *DAE* dans les plantes sauvages (Col0). (D) Analyse de l'expression par QPCR du gène *DAE* chez *dae-1* et *dae-2* et (E) chez *dae-3*.



## 1.7 L'EXPRESSION DU GENE DAE EST PERTURBEE DANS LES MUTANTS DAE-1, DAE-2 ET DAE-3

Afin de caractériser la fonction de la protéine DAE chez *Arabidopsis*, nous avons obtenu des lignées de mutants d'insertion T-DNA disponibles dans la collection SALK, *dae-1* (Salk\_009495), *dae-2* (Salk\_127449) et *dae-3* (Salk\_132521). Les trois lignées possèdent toutes une insertion au niveau de la séquence comprise entre le domaine NYN et le premier domaine OST-HTH (**Figure 12A**). L'observation d'individus homozygotes de ces lignées mutantes au cours de leur développement, montre qu'il n'y a pas défauts majeurs évidents pour les trois lignées, comme l'illustre les photos des plantes au stade de rosette (**Figure 12B**).

Afin d'analyser l'expression de *DAE* par QPCR nous avons utilisé deux couple d'amorces F1R1 et F2R2 qui s'hybrident respectivement au niveau du domaine NYN, avant l'insertion du T-DNA, et du premier domaine OST-HTH, après l'insertion du T-DNA comme indiqué dans la **Figure 12A**. Tout d'abord nous avons analysé l'expression du gène *DAE* dans des plantes sauvages d'écotype Columbia (Colo) dans différents tissus : plantules, feuille de rosette, fleurs et siliques vertes (**Figure 12C**). L'expérience de QPCR a été faite avec l'ARN purifié à partir de matériel végétal composé par un pool de 10 plantes de même génotype. Le résultat de cette analyse chez Colo montre que l'expression de *DAE* est assez constante dans les tissus analysés avec une augmentation légère de l'expression dans les fleurs (**Figure 12C**). Ensuite afin de vérifier l'impact des insertions de T-DNA nous avons procédé à l'analyse de l'expression du gène *DAE* chez les mutants *dae-1* et *dae-2* dans les mêmes conditions expérimentales indiquées précédemment **Figure 12D**. L'analyse QPCR chez le mutant *dae-3* présentée en **Figure 12E** est le résultat de trois réplicats biologiques. Pour l'analyse de l'expression dans les plantules chaque réplicat est fait à partir d'ARN purifié de 100 individus de même génotype regroupé en pool. Les **Figures 12D** et **E** montrent que l'expression du gène *DAE* est perturbée dans les trois lignées mutantes. Nous pouvons remarquer la présence d'expression résiduelle avant l'insertion du T-DNA dans les tissus analysés chez les mutants *dae-2* et *dae-3*. En revanche l'expression du gène *DAE* est complètement abolie, dans les tissus analysés, chez le mutant *dae-1*.

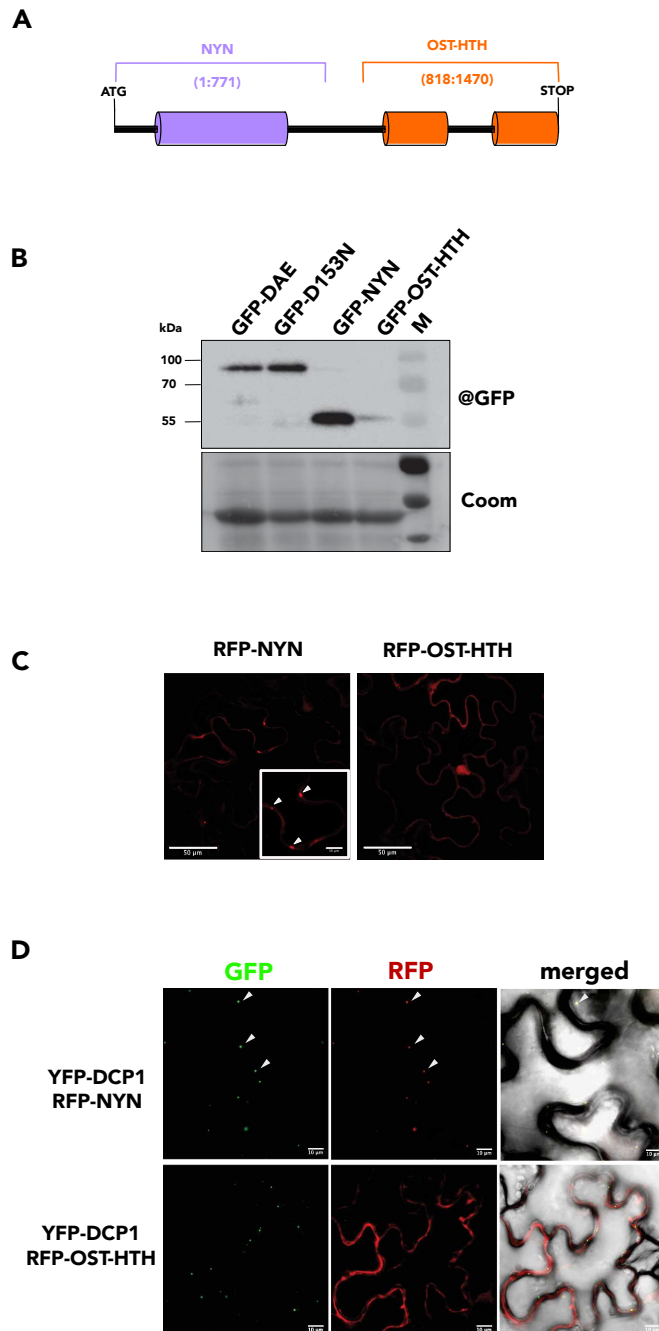


**Figure 13 DAE localise spécifiquement dans les P-bodies.**

(A) Étude de co-localisation par microscopie confocale sur des lignées stables exprimant GFP-DAE et le marqueur des granules de stress PAB2-RFP, en absence (25°C) ou présence (37°C) de stress thermique (panels supérieurs) ou RFP-DAE et le marqueur des p-bodies YFP-DCP1 à 25°C (panel inférieur) (B) Dot-plot qui représente la quantification du signal des foci obtenus par colocalisation des signaux GFP ou YFP (ch1, ordonnées) et RFP (ch2, abscisses) ; n indique le nombre de cellules analysées. Le coefficient de corrélation (r) ainsi que la p-values sont indiquées en bas de chaque graphe.

## 1.8 DAE CO-LOCALISE AVEC DCP1 DANS LES P-BODIES

L'étude par expression transitoire dans les feuilles de *N. benthamiana* co-infiltrées avec les constructions GFP-DAE et RFP-UPF1 publiées par le laboratoire en 2018 (Chicois et al., 2018), montre une co-localisation de DAE et UPF1 dans des foci cytoplasmiques suggérant fortement que DAE représente un nouveau composant des P-bodies chez les plantes. Afin de confirmer cette hypothèse et de discriminer de façon catégorique entre P-bodies et granules de stress, nous avons utilisé des lignées obtenues par croisement exprimant GFP-DAE et soit le marqueur des granules de stress RFP-PAB2, soit le marqueur des P-bodies YFP-DCP1. Les deux lignées marqueur ont été gracieusement données par JM. Deragon et C. Bousquet-Antonelli. Dans ces lignées, nous avons ensuite analysé la localisation des deux protéines de fusion dans la racine de plantules âgées de 7 jours par microscopie confocale, en présence ou absence de stress thermique (37°C) pour les granules de stress et en conditions non stressées pour les P-bodies. Les photos dans la **Figure 13A** confirment que GFP-DAE est localisé en foci cytoplasmique en présence ou absence de stress. Dans aucune condition, GFP-DAE ne co-localise avec RFP-PAB2. En revanche en absence de stress thermique nous pouvons observer une co-localisation quasi parfaite entre RFP-DAE et YFP-DCP1. La significativité de la co-localisation RFP-DAE / YFP-DCP1 et l'absence de co-localisation GFP-DAE / RFP-PAB2 a été confirmée par un test de corrélation Pearson sur une série d'images dont le résultat est présenté en forme de dot plot en **Figure 13B**. Le test donne une forte corrélation de 0,9 uniquement avec DCP1 mais pas avec PAB2 en présence ou absence de stress (0,35 et 0,18 respectivement). Ces résultats valident que DAE localise dans les P-bodies à l'instar de DCP1 son partenaire principal dans les expériences de co-IP.



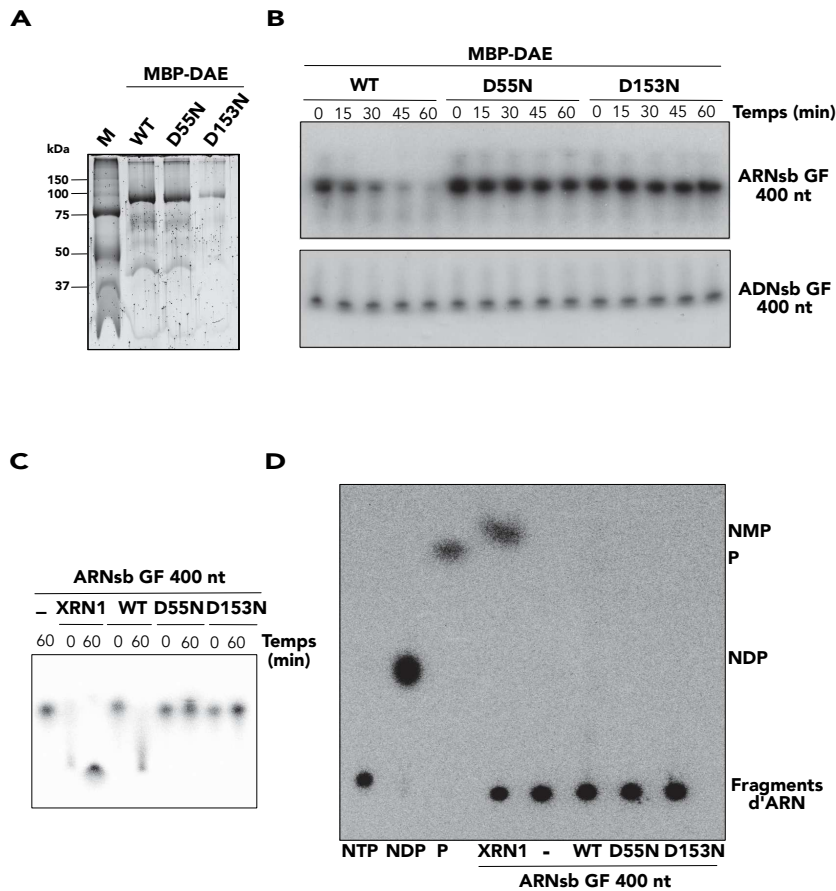
**Figure 14** La partie N-terminale de DAE contenant le domaine NYN localise dans les p-bodies.

(A) Structure protéique de DAE, les régions de la protéine utilisées pour les expériences d'expression transitoire dans les feuilles de *N. Benthamiana*. Les portions du gène DAE utilisées pour le clonage des domaines NYN et OST-HTH sont indiqués.

(B) Western blot qui montre l'accumulation de la protéine entière (GFP-DAE et GFP-D153N) et des domaines GFP-NYN et GFP-OST, détectées avec un anticorps antiGFP. (C) Images de microscopie confocale qui montrent la localisation de RFP-NYN et RFP-OST-HTH seules et (D) co-exprimés avec le marqueur des p-bodies YFP-DCP1.

## 1.9 LA PARTIE N-TERMINALE DE DAE CONTENANT LE DOMAINE NYN ENTRAINE UN LOCALISATION SPECIFIQUE DANS LES P-BODIES

Afin d'étudier la contribution des domaines NYN et OST-HTH dans la localisation cellulaire de DAE nous avons produit deux constructions : la partie N-terminale de DAE contenant le domaine NYN et la partie C-terminale de DAE contenant les deux domaines OST-HTH comme représenté dans la **Figure 14A**. Ensuite nous avons cloné les deux constructions dans un vecteur permettant une fusion avec la RFP ou la GFP en partie N-terminale. Une expression transitoire des protéines de fusion ainsi obtenues a été réalisée dans des feuilles de *N. Benthamiana* avec les constructions RFP-NYN, GFP-NYN, RFP-OST-HTH ou GFP-OST-HTH. L'expression des protéines de fusion GFP a été testée par western blot avec un anticorps dirigé contre la GFP (**Figure 14B**). Le résultat confirme l'expression des protéines de fusion dans les conditions testées. Cependant le niveau d'expression de GFP-OST est plus faible que celui de GFP-NYN. Nous n'avons pas pu tester l'expression des fusions RFP en raison de l'apparition de très forts signaux aspécifiques avec l'anticorps anti-RFP testé. L'observation par microscopie confocale des feuilles infiltrées avec RFP-NYN ou RFP-OST-HTH montre que seulement le domaine NYN a une localisation en foci cytoplasmiques alors que RFP-OST-HTH a une localisation diffuse dans le noyau et le cytoplasme (**Figure 14C**). De plus la co-infiltration de RFP-NYN avec YFP-DCP1 montre une co-localisation des deux protéines ce qui suggère que le domaine NYN est suffisant pour la localisation de DAE dans les P-bodies (**Figure 14D**). La possible co-localisation de YFP-DCP1 et RFP-OST-HTH semble due au fait que le domaine OST s'accumule dans tous les compartiments mais ne semble pas spécifique. Des études de localisation avec des lignées stables qui co-expriment RFP-NYN ou RFP-OST-HTH et YFP-DCP1 seront nécessaires afin de confirmer ces résultats.



**Figure 15 Étude de l'activité enzymatique de DAE**

(A) Gel coloré avec SYPRORuby des trois protéines purifiées, une version sauvage MBP-DAE et deux mutants catalytiques MBP-D55N et MBP-D153N (B) Analyse sur gel PAGE en conditions dénaturantes de l'activité de DAE, sauvage et mutants catalytiques, sur l'ARN simple brin (ARNsb, panel supérieur) et l'ADNsb (panel inférieur). Les protéines recombinantes (100 nM) ont été incubées avec 2 nM d'ARNsb de 400 nt transcrit in vitro et marqué internement avec de l' $\alpha^{32}\text{P}$  UTP ou 2 nM d'ADNsb marqué avec l'enzyme PNK à l'extrémité 5' avec du  $\gamma^{32}\text{P}$  ATP. (C) Test d'activité de l'exoribonucléase XRN1 et de DAE utilisé pour la chromatographie en couche mince (CCM). Indiqué avec - l'ARN incubé dans le buffer de réaction sans protéines. (D) CCM sur plaque de polyéthylèneimine dans buffer : 0,5 M LiCl 1M Acide formique, chargée avec les produits de l'activité enzymatique de XRN1 et DAE, sauvage et mutants. Nucléotides triphosphate (NTP), nucléotides diphosphate (NDP), nucléotides monophosphates (NMP) et phosphate (P) ont été utilisés comme marqueurs.

## 1.10 DAE POSSEDE UNE ACTIVITE RIBONUCLEASE

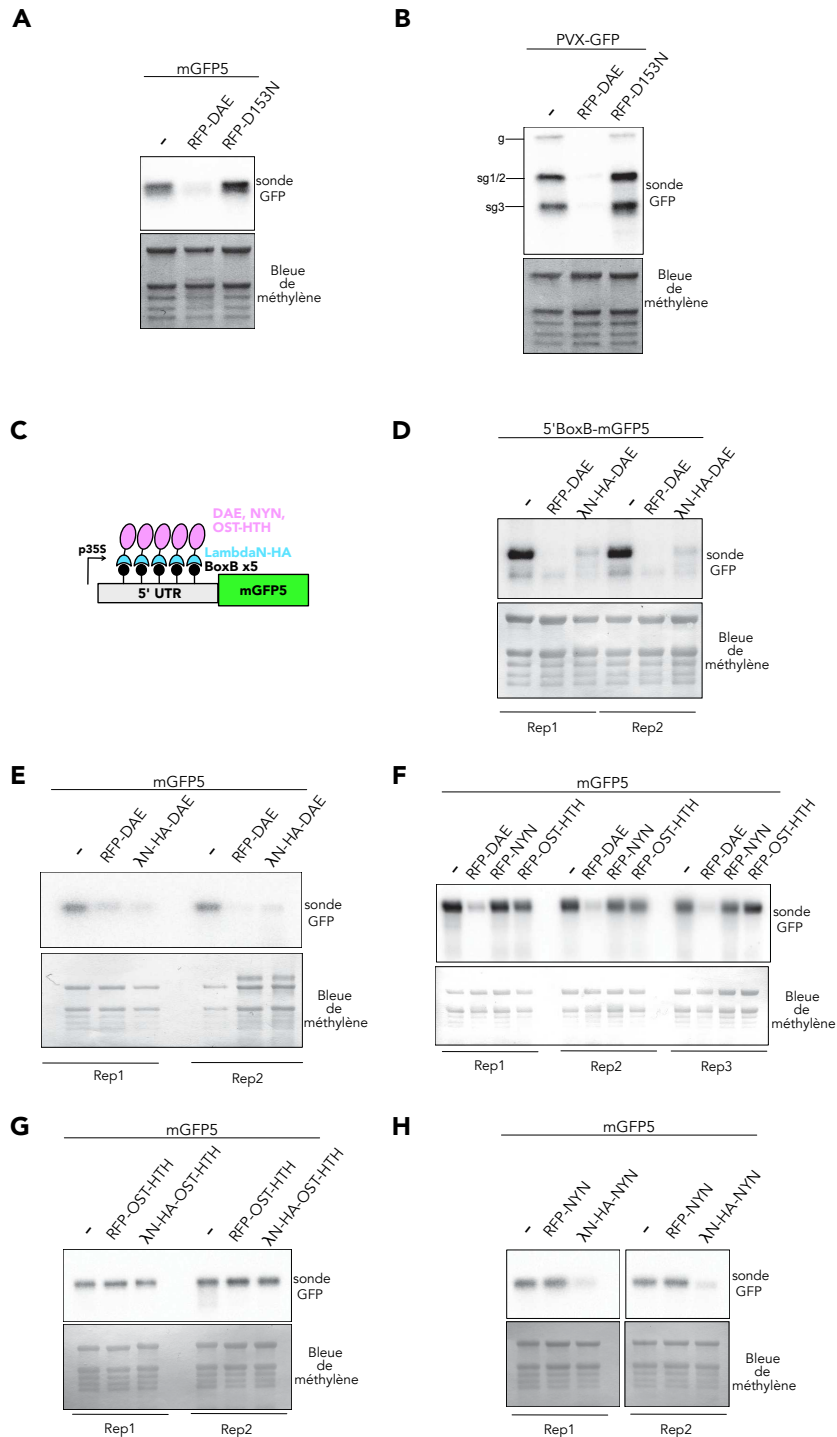
Comme mentionné précédemment (voir chap. Résultats § 2), le domaine NYN de DAE possède une structure prédite suggérant qu'elle pourrait être une authentique endoribonucléase. Afin d'étudier l'activité enzymatique de DAE nous avons choisi une stratégie de production de la protéine en système bactérien, suivi de test d'activité. Pour ce faire, j'ai produit des mutants catalytiques DAE où les résidus d'acide aspartique du site actif ont été mutés en asparagines (indiqués en rouge et marqués par un astérisque en **Figure 18C**, le changement de charge est à la base de la perte d'activité puisque il rend impossible la liaison du  $Mg^{2+}$  nécessaire pour l'activité catalytique comme cela a été montré pour d'autres protéines à domaine NYN (Tsai et al., 2015; Suzuki et al., 2011; Nishimura et al., 2018a). Différents tag en N-ter ont été testés : GST (Glutathion S-transférase), 6xHis (6 Histidines) et MBP (Maltose Binding Protein). Les constructions contenant le tag GST et 6xHis ne s'expriment pas dans *E. coli*. Finalement, trois versions de DAE, une version sauvage, MBP-DAE, et deux mutants catalytiques, MBP-D55N et MBP-D153N ont été exprimées dans *E. coli* et purifiées. La procédure de purification a consisté en une chromatographie d'affinité sur une résine de maltose en batch ou par HPLC (High Performance Liquid Chromatography) Äkta, couplée avec une purification en fonction de la taille par filtration sur gel. Avec cette double purification nous avons obtenu des extraits protéiques purifiés comme nous pouvons voir après migration sur gel de polyacrylamide en **Figure 15A**. Cependant l'analyse des chromatogrammes obtenus après filtration sur gel montre que les protéines sortent dans le volume mort ce qui nous suggère qu'elles sont présentes en solution sous forme d'agrégats solubles (**Annexe 2**). Pour exclure la présence de contaminants majeurs dans nos préparations protéiques nous avons analysé nos échantillons par spectrométrie de masse. Le résultat de l'analyse confirme l'absence de contaminants majeurs (**Annexe 3**). Pour résoudre ce problème d'agrégation plusieurs conditions de purification ont été testées. J'ai tout d'abord réduit le temps d'induction afin de diminuer la quantité de protéine produite. J'ai ensuite changé la composition des tampons utilisés pour la purification en ajoutant du  $Mg^{2+}$  ou en changeant le pH (pH = 5, 6, 7, 8, 9). Après j'ai co-exprimé les trois versions de MBP-DAE avec les chaperones dnaK, danJ,, grpE, groES et groEL. Enfin j'ai exprimé et purifié seulement la partie N-terminale de DAE contenant le domaine NYN (**Figure 14A**) avec un tag MBP en N-ter. Aucune des conditions testées n'a résolu le problème d'agrégation. Nous avons donc choisi de tester l'activité enzymatique des protéines entières par incubation des trois versions de DAE avec un ARN transcrit *in vitro* avec la polymérase T7 à partir d'un ADNc contenant la séquence du gène rapporteur GFP (GF). La **Figure 15B** (panel supérieur) montre que après 30 minutes d'incubation on observe une diminution de la quantité de transcrit avec la version sauvage MBP-DAE, mais pas avec les mutants catalytiques MBP-D55N et MBP-D153N. Une dégradation complète de l'ARN est finalement

observée avec MBP-DAE au bout d'une heure d'incubation. Ce résultat indique que DAE possède une activité qui est nécessaire pour la dégradation de l'ARN et qui est dépendante des résidus conservés dans son domaine NYN. En revanche, DAE apparaît incapable de dégrader de l'ADN simple brin (**Figure 15B**, panel inférieur). Les expériences que nous avons réalisées montrent que DAE est une ribonucléase active *in vitro*.

### 1.11 DAE EST UNE ENDORIBONUCLEASE

Les tests d'activité réalisés précédemment ne permettent pas de mettre en évidence la nature de l'activité de DAE et de distinguer entre une activité exoribonucléase et une activité endoribonucléase. Afin d'obtenir des données expérimentales permettant de discriminer entre les deux, nous avons réalisé une expérience de chromatographie en couche mince (CCM) avec une matrice de polyéthylènimine chargée avec les produits de digestion de DAE (**Figure 15D**). Cette méthode permet de séparer de façon différentielle des fragments d'ARN produits par une endoribonucléase ou des nucléosides monophosphates (NMP) produits par une exoribonucléase. La migration de ces deux produits sur une plaque de polyéthylènimine est différente du fait de leurs propriétés physico-chimiques, notamment la charge. Les NMP, les moins chargés négativement, migrent vers le haut de la plaque. Les fragments d'ARN ou les nucléotides triphosphate sont retenus sur la matrice. La plaque a été chargée avec les produits de la réaction enzymatique de DAE ainsi qu'avec les produits d'une exoribonucléase commerciale, XRN<sub>1</sub>, utilisée comme contrôle (**Figure 15C**). La **Figure 15D** montre que l'incubation de MBP-DAE produit des fragments d'ARN qui sont retenus dans la partie inférieure de la plaque, alors que les produits de XRN<sub>1</sub>, les NMP, migrent à travers la matrice. Ce résultat démontre que DAE est une endoribonucléase, comme le suggérait la présence d'un domaine NYN.





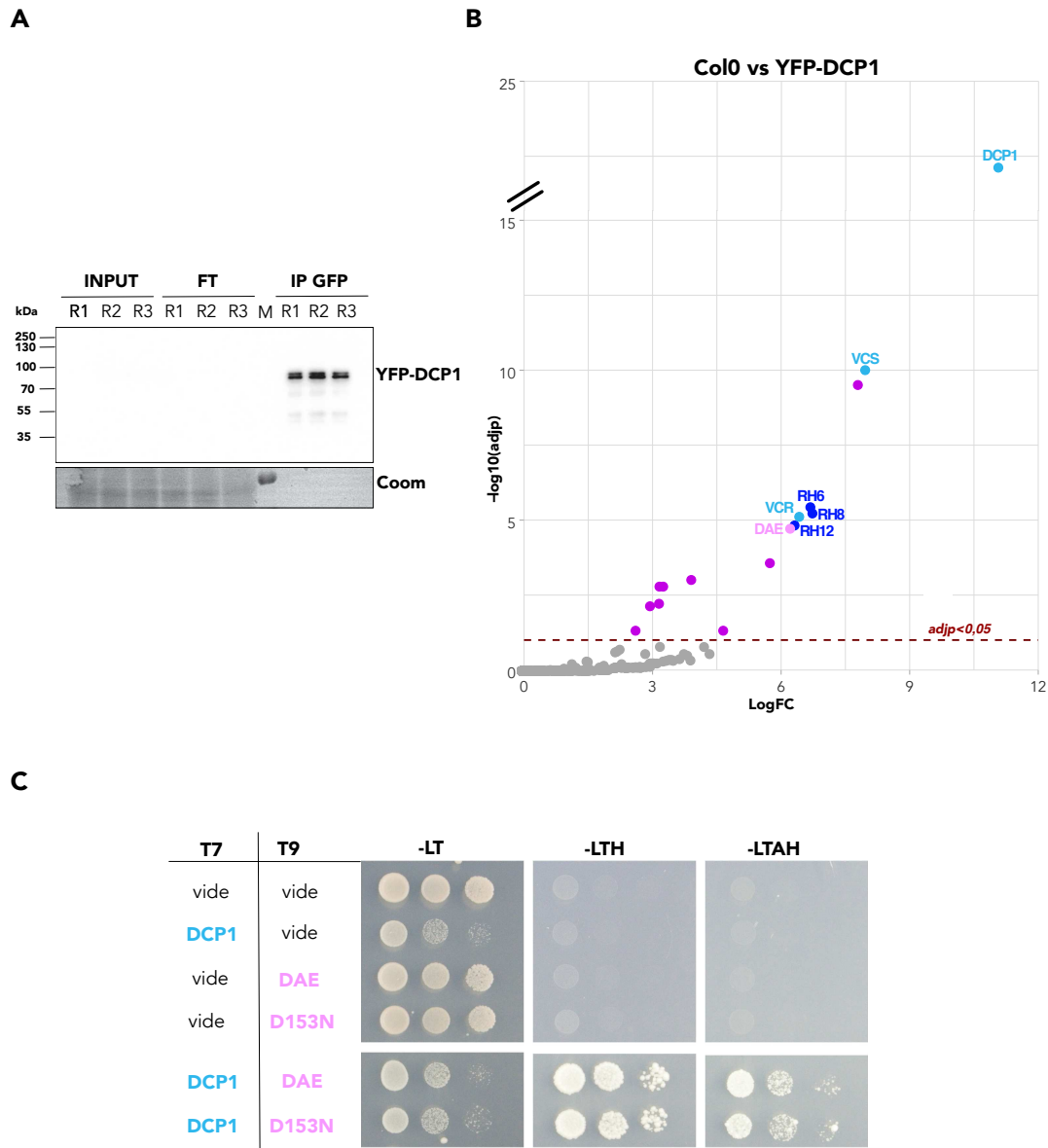
**Figure 16** Etude de l'activité de DAE *in vivo* par expression transitoire dans les feuilles de *N. benthamiana*. (A) et (B) Analyse par northern blot de l'accumulation de l'ARN rapporteur mGFP5 ou du virus PVX-GFP co-exprimés avec RFP-DAE ou RFP-D153N et P14, avec une sonde GFP. (C) Schéma de l'expérience d'ancrage moléculaire (« tethering »). (D) Analyse par northern blot de l'accumulation de l'ARN 5'BoxB-mGFP5 co-exprimé avec le plasmide vide (-), RFP-DAE ou λN-HA-DAE et P19. (E) Analyse par northern blot de l'accumulation de l'ARN mGFP5 co-exprimé avec le plasmide vide (-), RFP-DAE ou λN-HA-DAE et P19. (F) Analyse par northern blot de l'accumulation de l'ARN mGFP5 co-exprimé avec le plasmide vide (-), RFP-DAE, RFP-NYN ou RFP-OST-HTH et P19. (G) Analyse par northern blot de l'accumulation de l'ARN mGFP5 co-exprimé avec le plasmide vide (-), RFP-OST-HTH ou λN-HA-OST-HTH et P19. (H) Analyse par northern blot de l'accumulation de l'ARN mGFP5 co-exprimé avec le plasmide vide (-), RFP-NYN ou λN-HA-NYN et P19.

## 1.12 DAE DIMINUE L'ACCUMULATION D'UN ARNm CO-EXPRIME

Afin de tester la possibilité d'une activité catalytique de DAE *in vivo*, nous avons co-exprimé DAE avec l'ARNm de la GFP dans un système d'expression transitoire dans *N. benthamiana*. Dans ce test nous avons utilisé des souches d'*A. tumefaciens* possédant un vecteur binaire permettant l'expression des constructions RFP-DAE et RFP-D153N. Les feuilles de *N. benthamiana* ont été ensuite infiltrées à l'aide d'une seringue dépourvue d'aiguille avec un mélange d'agrobactéries exprimant RFP-DAE ou RFP-D153N et des souches d'agrobactérie permettant la co-expression de la mGFP5 ou du virus PVX-GFP (**Figure 16A et B**) dans les deux cas en présence d'un suppresseur viral du RNA silencing, P14 ou P19, afin d'assurer une expression optimale des construits. Deux jours et demi après l'infiltration, j'ai extrait l'ARN total à partir des feuilles infiltrées et j'ai analysé l'expression des construits co-infiltrés grâce à une sonde ciblant la GFP. La **Figure 16A** montre le résultat de cette expérience ; nous pouvons observer que la co-infiltration de RFP-DAE avec la mGFP5 conduit à réduction très forte de l'accumulation de l'ARN GFP ; en revanche la co-infiltration de RFP-D153N avec la mGFP5 n'a aucun effet. Ce résultat indique que l'activité catalytique de DAE est responsable de la baisse de l'expression d'un ARNm co-exprimé. Des résultats préliminaires obtenus au laboratoire (Thèse Clara Chicois) suggèrent que DAE pourrait avoir un effet négatif sur l'accumulation du virus PVX (Potato Virus X). Nous avons donc décidé de co-exprimer également RFP-DAE et RFP-D153N avec un vecteur permettant l'expression d'un clone infectieux du virus PVX exprimant la GFP à partir d'un de ses ARN subgénomique (**Figure 16B**). Ce virus peut également être détecté grâce à une sonde GFP qui révèle l'ARN génomique (g) et les trois ARN subgénomiques (sg1, 2 et 3). Le northern blot en **Figure 16B** montre que, à nouveau, seulement la co-expression de la version catalytiquement active de DAE, et pas du mutant catalytique D153N, réprime l'accumulation des ARN génomique et subgénomiques du clone PVX-GFP.

Afin de tester si l'effet observé est dû à un effet direct de l'activité endoribonucléase de DAE sur les ARNm co-exprimés une stratégie d'ancrage moléculaire (tethering) a été mise en place. Pour cette expérience nous avons choisi le système dérivé du phage  $\lambda$ , dans lequel la séquence qui code pour le peptide de l'antitermineur du phage  $\lambda$  appelé  $\lambda$ N interagit avec le motif BoxB (Schoenberg et al., 2004). Ce système force l'interaction d'une protéine d'intérêt avec un ARN cible, et permet donc de déterminer les effets de l'interaction directe entre la protéine et l'ARN. Nous avons donc cloné la séquence de DAE et du mutant catalytique D153N dans un plasmide contenant  $\lambda$ N avec un tag HA fusionné en partie N-terminale (**Figure 16C**) en parallèle nous avons cloné la séquence codant pour la mGFP5 avec 5 motifs BoxB en 5' ou en 3', les résultats sont équivalents pour les deux constructions et seulement ceux obtenus avec 5'BoxB-mGFP5 sont montrés. Quand nous avons testé ce système nous avons constaté que la co-expression de  $\lambda$ N-HA-DAE ou de RFP-DAE (utilisé comme contrôle

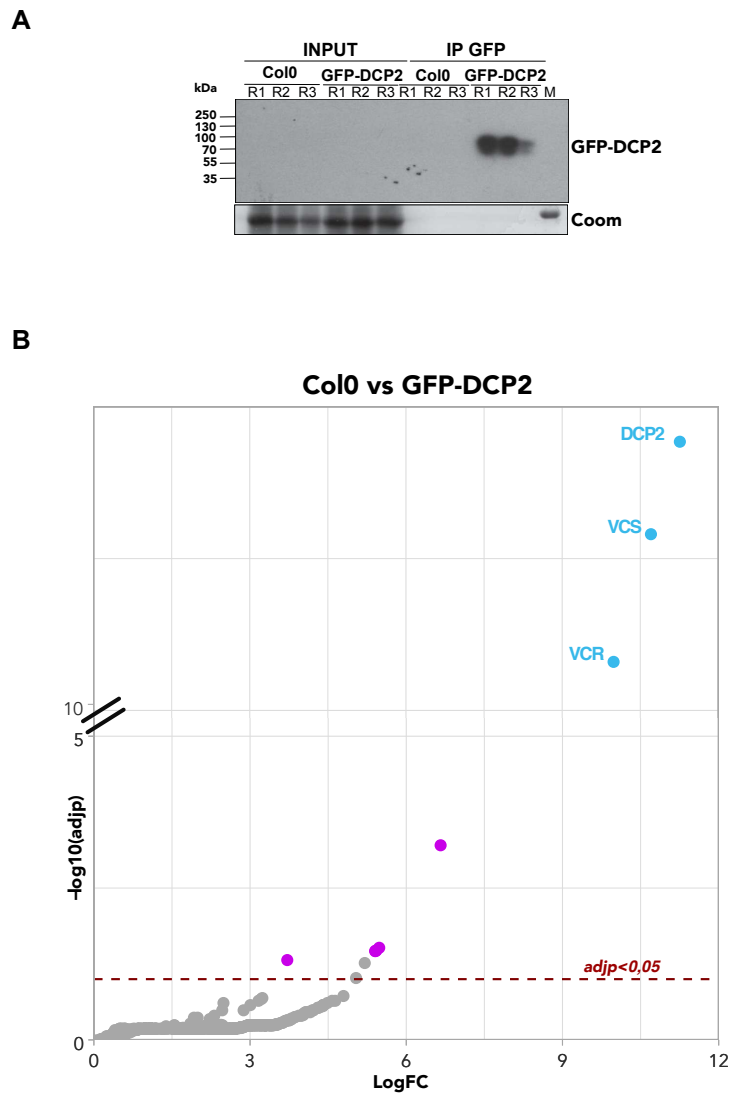
négatif) avec l'ARN cible produit une telle diminution des ARN co-exprimé qu'il est impossible de conclure quant à l'effet des motifs BoxB (**Figure 16D et E**). Pour tenter de s'affranchir de ce problème nous avons ensuite décidé de cloner uniquement les domaines de DAE : la région N-terminale contenant le domaine NYN et la région C-terminale contenant les domaines OST-HTH (**Figure 14A**) en fusion avec la RFP et de tester si nous pouvions utiliser ces constructions dans une stratégie de tethering. Dans ces expériences nous souhaitons utiliser RFP-NYN et RFP-OST comme contrôles, nous les avons donc co-exprimés avec la mGFP<sub>5</sub>. La **Figure 16F** montre que seule la présence de RFP-DAE entraîne une diminution de l'ARNm co-exprimé, indiquant que les domaines seuls pourraient permettre de tester le tethering. Rassurés par ce résultat, nous avons ensuite cloné NYN et OST dans un plasmide contenant  $\lambda$ N-HA en N-ter et nous avons testé l'activité des protéines ainsi produites sur l'ARN rapporteur mGFP<sub>5</sub> sans motifs BoxB. Les **figures 16G et H** montrent que  $\lambda$ N-HA-NYN, mais pas  $\lambda$ N-HA-OST, entraîne encore une diminution de la mGFP<sub>5</sub>, même dépourvue de motifs BoxB. Ceci indique que cette construction  $\lambda$ N-HA-NYN ne peut pas être utilisée efficacement dans le système de tethering. A cette étape seule la construction RFP-NYN semble ne pas induire une diminution de l'ARN rapporteur (**Figure 16F et H**). Pour cette raison nous prévoyions une expérience en utilisant  $\lambda$ N-HA-RFP-NYN et  $\lambda$ N-HA-RFP-NYN D153N, ce qui pourrait nous permettre de mettre en évidence *in vivo* l'activité du domaine NYN de DAE par une stratégie de tethering.



**Figure 17** DAE est une des protéines plus enrichies dans l'IP DCP1 et interagit directement avec DCP1 indépendamment de son activité catalytique (A) Western blot des immunoprécipitation de YFP-DCP1. YFP-DCP1 a été détecté par hybridation de la membrane avec un antiGFP dans les inputs, flow-through (FT) et immunoprécipitations (IP-GFP) de trois réplicats techniques (R1, R2, R3). Marqueur de taille (M) (B) Semi-volcano plot des protéines enrichies avec Log du fold change (FC) supérieur à 1 et p-value ajustée inférieure à 0,05 dans les IP DCP1 (2 réplicats biologiques, chacun en 3 réplicats techniques). En turquoise les protéines du complexe de decapping, en bleu marine les hélicases homologues de DDX6, en rose DAE, en magenta les autres protéines significativement enrichies. (C) étude de l'interaction entre DAE, sauvage et mutant catalytique (D153N), et DCP1 par double hybride. Les séquences de DCP1, DAE et D153N ont été clonées dans un plasmide contenant le domaine d'activation (T7 ; pGADT7) ou le domaine de liaison à l'ADN (T9 ; pGBT9) de l'activateur transcriptionnel GAL4. Milieu de culture riche, sans leucine et tryptophane (-LT) ; medium sans leucine, tryptophane et histidine (-LTH) et pauvre sans leucine, tryptophane, adénine et histidine (-LTAH). Toujours en présence de 5 mM de 3-AT (3-amino-1,2,4-triazole) pour augmenter la stringence.

### 1.13 DAE INTERAGIT DIRECTEMENT AVEC DCP1 INDEPENDAMMENT DE SON ACTIVITE CATALYTIQUE

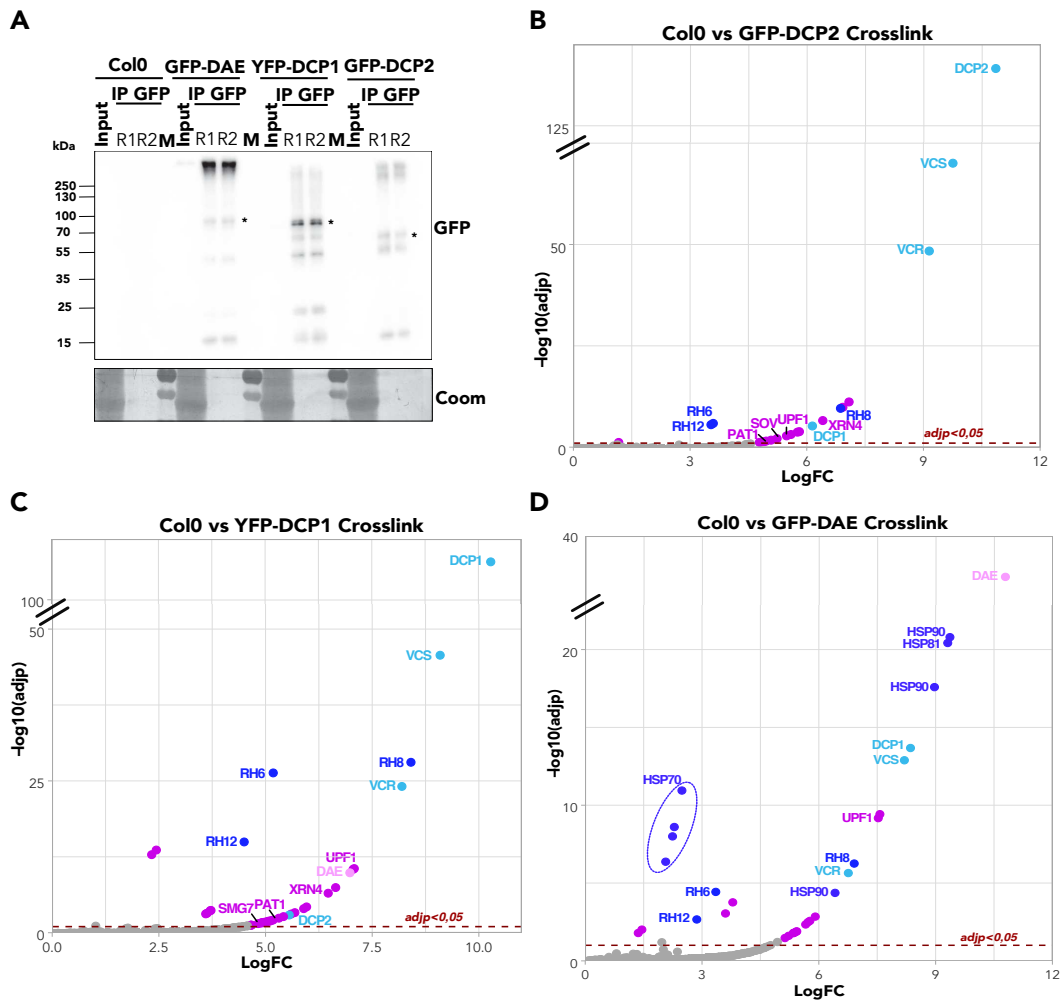
DCP1 apparait dans les co-IP DAE comme son partenaire principal. Dans le but de d'étudier spécifiquement l'interaction entre DAE et DCP1 nous avons conduit une expérience de co-immunoprécipitation dans des lignées stables qui expriment une protéine de fusion YFP-DCP1 sous promoteur endogène. La présence de DCP1 dans les fractions immunoprécipitées est révélée avec un anticorps ciblant la GFP, qui reconnaît aussi son variant YFP. Le western blot présenté en **Figure 17A** montre que nous avons obtenu des fractions où YFP-DCP1 est très enrichi par rapport à l'input. Ces extraits enrichis ont ensuite été analysés par spectrométrie de masse. Le semi-volcano plot présenté en **Figure 17B** montre que le composant du complexe de decapping VCS ainsi que les hélicases homologues de DDX6 (RH6, RH8 et RH12) sont fortement enrichis dans l'IP DCP1, ceci valide notre approche. De plus nous pouvons observer que la protéine VCR est aussi très enrichie dans les IP DCP1 ce qui suggère un rôle potentiel de cette protéine dans le complexe de decapping. DAE est l'une des protéines les plus fortement enrichies avec DCP1 ce qui place DAE en première ligne parmi ces protéines dans l'optique d'identifier un interactant direct de DCP1. Afin de tester cette hypothèse nous avons exprimé DCP1, DAE et D153N dans un système hétérologue dans une expérience de double hybride (**Figure 17C**). Le résultat montre que de façon équivalente, DAE et le mutant catalytique D153N, rétablissent la croissance des levures sur un milieu pauvre. Ceci qui suggère que DAE interagit directement avec DCP1 et ce de façon indépendante de son site catalytique.



**Figure 18 DCP2 co-purifie avec VCR et VCS mais pas avec DCP1** (A) Western blot des immunoprécipitations de GFP-DCP2. GFP-DCP2 a été détecté par hybridation de la membrane avec un antiGFP dans les inputs, flow-through (FT) et immunoprécipitations (IP-GFP) de trois réplicats techniques (R1, R2, R3). Marqueur de taille (M) (B) Semi-volcano plot des protéines enrichies avec Log du fold change (FC) supérieur à 1 et p-value ajustée inférieure à 0,05 dans les IP DCP2 (1 réplicats biologique en 3 réplicats techniques). En turquoise les protéines du complexe de decapping, en magenta les autres protéines significativement enrichies.

## 1.14 DCP2 CO-PURIFIE AVEC VCS ET VCR

Un des résultats surprenant de notre analyse par spectrométrie de masse des IP DCP1 est l'absence de DCP2, un partenaire connu de DCP1 et sous-unité catalytique responsable du decapping. Deux hypothèses peuvent expliquer ce résultat, la première, technique, est que DCP2 est difficile à détecter par spectrométrie de masse, la seconde est que l'interaction entre DCP1 et DCP2 est faible ou transitoire. Afin de tester ces hypothèses, nous avons tout d'abord fait une expérience de co-IP en utilisant des lignées stables qui expriment sous promoteur constitutif 35S une protéine de fusion fonctionnelle GFP-DCP2 (Goeres et al., 2007). Une co-immunoprécipitation en utilisant un antiGFP, a permis d'obtenir des fractions enrichies en GFP-DCP2 (**Figure 18A**), nous avons ensuite analysé ces fractions par spectrométrie de masse. Le résultat de cette analyse est présenté sous forme de semi-volcano plot en **Figure 18B**, et montre qu'il y a 6 protéines enrichies dans l'IP DCP2 incluant VCS et son homologue VCR, comme observé précédemment avec DCP1. Dans ces conditions, ni DCP1 ni DAE n'est détectée dans les IP DCP2. Cette expérience nous indique que la protéine DCP2 est détectable dans nos conditions expérimentales, cependant nous pouvons toujours remarquer l'absence de DCP1 ce qui nous a conduit à vérifier la seconde hypothèse présentée dans le paragraphe précédent : l'interaction entre DCP1 et DCP2 pourrait être faible ou transitoire.



**Figure 19** Immunoprécipitation avec crosslink de GFP-DAE, YFP-DCP1 et GFP-DCP2 (A) Western blot des immunoprécipitations de GFP-DAE, YFP-DCP1 et GFP-DCP2. Les trois appâts ont été détectés par hybridation de la membrane avec un antiGFP dans les inputs, flow-through (FT) et immunoprécipitations (IP-GFP) de trois réplicats techniques (R1, R2, R3). Marqueur de taille (M) (B), (C), (D) Semi-volcano plot des protéines enrichies avec Log du fold change (FC) supérieur à 1 et p-value ajustée inférieure à 0,05 dans les IP DCP2, DCP1 et DAE respectivement (2 réplicats biologiques chacun en 2 réplicats d'affinité). En turquoise les protéines du complexe de decapping, en bleu marine les hélicases homologues de DDX6, en violet les chaperonnes HSP, en rose DAE, en magenta les autres protéines significativement enrichies.

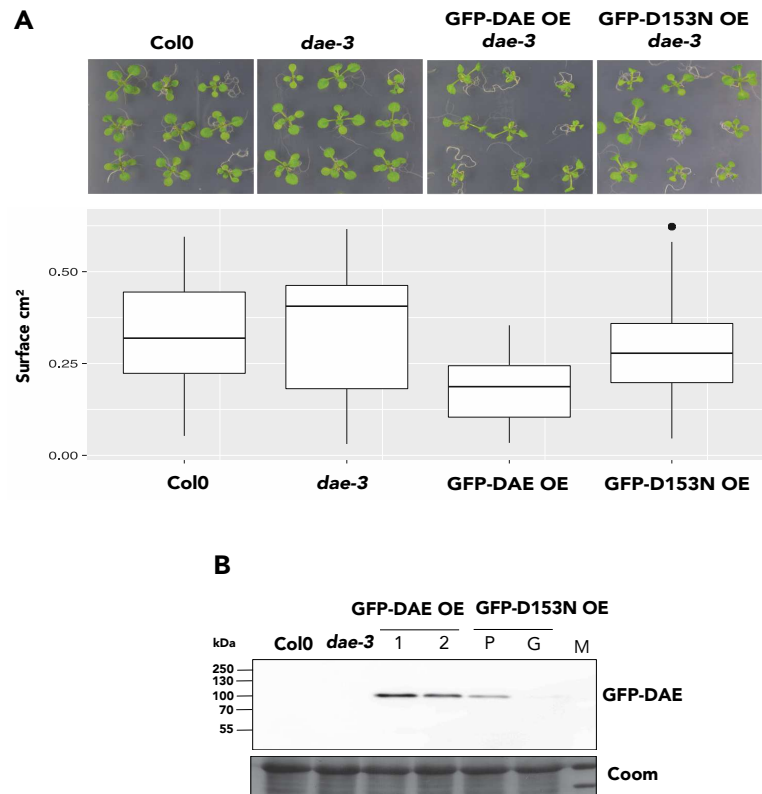


## 1.15 DAE EST PRESENT DANS UN COMPLEXE PROTEIQUE BASE AUTOUR DE DCP1 ET VCS MAIS NE CONTENANT PAS DCP2

L'absence réciproque de DCP1 dans les IP DCP2 et de DCP2 dans les IP DCP1 suggère que leur interaction est faible ou transitoire. De même l'absence de DAE des IP DCP2 pourrait également être due à une interaction du même type. Afin de résoudre ce problème les co-IP ont été répétées en ajoutant du formaldéhyde pendant le broyage du matériel végétal, un processus appelé pontage chimique ou crosslink. Cette méthode a été utilisée sur les lignées exprimant GFP-DAE, YFP-DCP1 et GFP-DCP2 déjà présentées dans les paragraphes précédents. Par cette méthode nous avons pu obtenir des extraits enrichis en GFP-DAE, YFP-DCP1 et GFP-DCP2 (**Figure 19A**). L'effet du crosslink est visible par l'apparition de signaux de poids moléculaire beaucoup plus important. Cependant nous pouvons aussi observer des bandes qui pourraient correspondre à GFP-DAE, YFP-DCP1 et GFP-DCP2 sous forme libre dans les fractions immunoprécipitées, marqués avec un astérisque. L'analyse par spectrométrie de masse des immunoprécipitation en crosslink de GFP-DCP2 (**Figure 19B**) montre à nouveau la présence des protéines VCS et VCR, validant l'efficacité de cette approche. De plus nous pouvons maintenant détecter DCP1 parmi les protéines les plus enrichies avec DCP2, ceci soutiens l'hypothèse que l'interaction entre DCP1 et DCP2 est faible ou transitoire. De façon intéressante, d'autres protéines connues pour être impliquées dans la dégradation 5'-3' et 3'-5' des ARNm, les exoribonucléases XRN4 et SOV et la protéine PAT1 notamment sont détectées, ce qui n'était pas le cas sans pontage chimique. Le semi-volcano plot produit par l'analyse des protéines les plus enrichies dans l'IP YFP-DCP1 (**Figure 19C**) montre la présence de VCS, VCR ainsi que des hélicases RH6, RH8 et RH12 ce qui valide notre approche. Nous pouvons remarquer encore une fois la présence de DAE dans les co-IP DCP1. Dans ces nouvelles conditions expérimentales nous avons pu détecter aussi la protéine clé du NMD UPF1 et finalement DCP2 l'enzyme responsable du decapping, qui n'était pas présente dans les tests sans pontage chimique.

Le semi-volcano plot issu de l'analyse par MS des IP DAE (**Figure 19D**) montre la présence des composant du complexe de decapping DCP1, VCS et VCR ainsi que des hélicases RH6, RH8 et RH12 et de UPF1. De plus nous avons pu détecter un grand nombre de chaperons moléculaires de la famille HSP de type HSP70, 90 ou 81. Malgré le crosslink, la protéine DCP2 reste absente de cette co-IP.

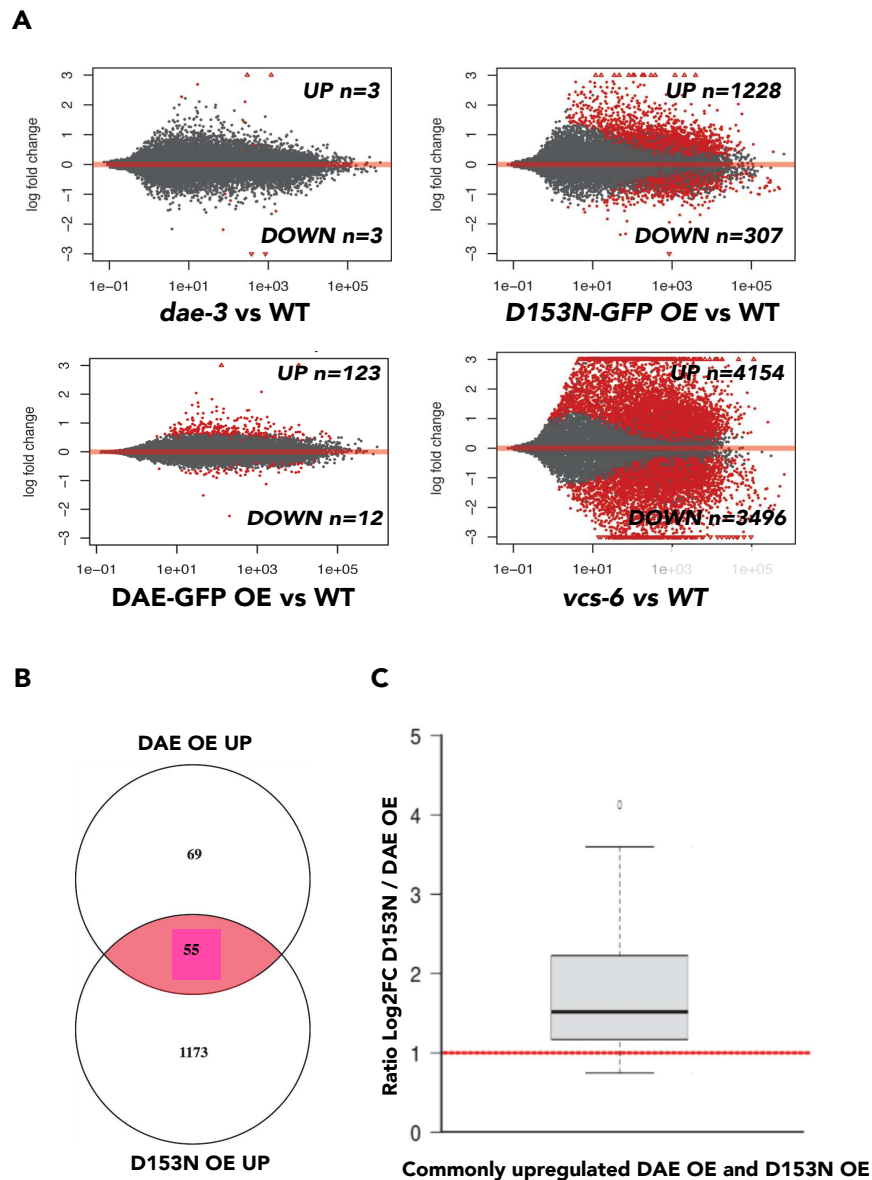
En conclusion ces expériences suggèrent l'existence d'un nouveau complexe formé par les activateurs de decapping DCP1 et VCS, liés avec l'endonucléase à domaine NYN DAE. Nos résultats suggèrent fortement que ce complexe est différent du complexe de decapping canonique composé des activateurs de decapping DCP1 et VCS et de l'enzyme de decapping DCP2.



**Figure 20** Caractérisation des lignées surexpressesurs GFP-DAE OE et GFP-D153N OE, les deux dans un fond génétique *dae-3* (A) Phénotype de plantules de 14 jours cultivées *in vitro* (panel supérieur). Les données de surface foliaire (cm<sup>2</sup>) de 100 individus par génotype sont représentés par des boîtes à moustaches (panel inférieur). (B) Western blot de l'accumulation de GFP-DAE dans deux lignées indépendantes (1 et 2) et de GFP-D153N dans des individus petits (P) et grands (G) de la même lignée. La membrane a été hybridée avec un anticorps antiGFP.

## 1.16 LA SUREXPRESSION DE GFP-DAE ET GFP-D153N ENTRAINE DES DEFATS DE DEVELOPPEMENT

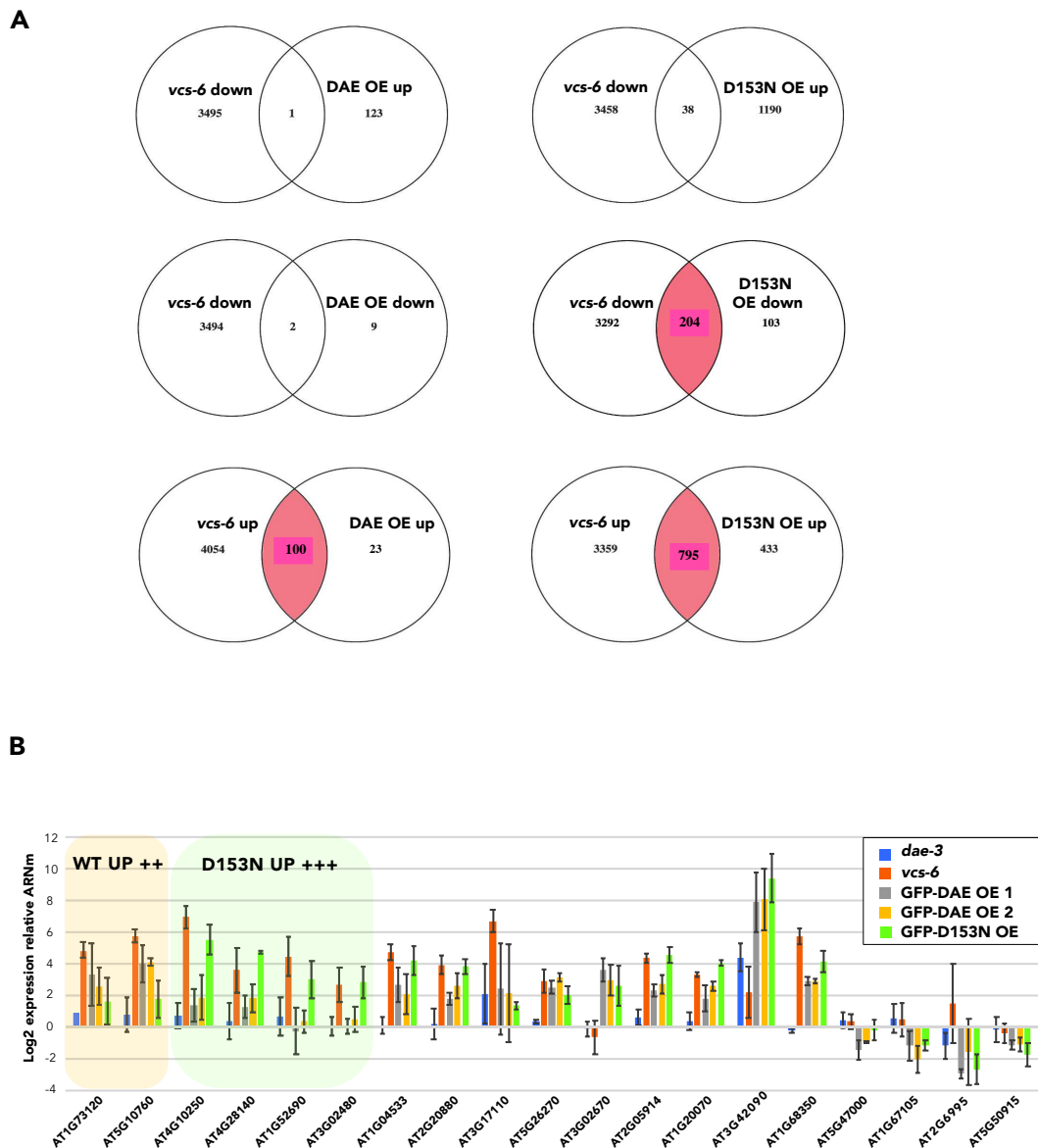
Nos résultats précédents montrent que DAE est une endoribonucléase associée avec des facteurs de dégradation des ARNm, particulièrement ceux impliqués dans le decapping et la voie de dégradation 5'-3'. Ceci indique que DAE a très probablement un rôle dans la dégradation des ARNm. Comme présenté précédemment (voir chap. Résultats § 4) nous disposons au laboratoire de lignées mutantes de *DAE*, ainsi que de lignées exprimant de façon stable des versions sauvage (GFP-DAE OE) et mutante de la protéine (GFP-D153N OE). Aucun des mutants *dae* ne présente un développement anormal dans nos conditions de culture comme cela est illustré en **Figure 12B**. Au contraire, certaines des lignées qui expriment GFP-DAE et GFP-D153N montrent un retard de croissance dès le stade plantule (**Figure 20A**). L'effet détecté est plus prononcé pour les plus forts expresseurs de GFP-D153N, qui sont létaux à l'état post-embryonnaire. Pour cette raison seule une lignée qui exprime D153N à un niveau moyen a pu être propagée (**Figure 20A**). Dans cette lignée nous pouvons observer une hétérogénéité de taille des plantules (**Figure 20A**). La quantification de la surface foliaire en **Figure 20A**, obtenue avec le plugin d'ImageJ « Excessive Green » (développé par Jérôme Mutterer à l'IBMP), montre que les plantules des lignées GFP-DAE OE sont plus petites que les plantules GFP-D153N OE et que les plantules sauvages. Nous avons ensuite analysé l'expression de GFP-D153N dans les individus qui présentent une surface foliaire petite ou grande ; le western blot en **Figure 20B** montre que la petite taille des plantules corrèle avec un niveau d'expression plus élevé de GFP-D153N qui reste quand même à un niveau inférieure par rapport aux plantules exprimant GFP-DAE dans deux lignées à forte expression indépendantes (GFP-DAE OE 1 et 2).



**Figure 21** Analyse transcriptomique de la lignée mutante *dae-3*, des lignées surexpresseurs GFP-DAE et GFP-D153N et du mutant de decapping *vcs-6* (A) MA-plots qui représentent les gènes sur-exprimés (UP) ou sous-exprimés (DOWN) dans les lignées analysées comparées avec le sauvage (Col0). En abscisse le logarithme du fold change et en ordonnée la moyenne des comptages normalisés. Les points en rouge représentent les gènes dérégulés avec une p-value inférieure à 0,05. Le nombre des gènes (n), considéré pour les comparaison, avec LogFC > 0,75 (UP) ou LogFC < -0,75 (DOWN) et p-value < 0,05 est indiqué dans chaque graphe (B) Diagramme Venn qui représente l'overlap entre les gènes UP (LogFC > 0,75, p-value < 0,05) chez GFP-DAE OE (DAE OE UP) et GFP-D153N OE (D153N OE UP). La significativité de l'overlap a été validé par un test hypergéométrique. Overlap significatif en rouge (C) Boxplot qui montre que les gènes surexprimés communs entre les lignées DAE OE et D153N OE ont un fold change plus élevé chez D153N OE. En ordonnée le ratio du log<sub>2</sub> du fold change entre D153N OE et DAE OE.

## 1.17 ANALYSE DE L'IMPACT DE LA DEREGULATION DE DAE SUR LE TRANSCRIPTOME

Le matériel décrit dans le paragraphe précédent a été analysé par RNAseq (l'ARN poly(A) + extrait à partir de plantules de 14 jours cultivées *in vitro*), afin d'étudier l'impact de la dérégulation de DAE sur le transcriptome d'*A. thaliana*. En parallèle, nous avons aussi analysé dans les mêmes conditions le transcriptome de *vcs-6* considéré comme un mutant nul pour le decapping. Les résultats de l'analyse transcriptomique sont montrés en **Figure 21A** sous forme de MA plot qui nous permet d'avoir une vision d'ensemble des gènes dérégulés dans les différents génotypes analysés par rapport à des plantes sauvages (Colo). Le nombre n indiqué sur le MA plot représentent le nombre de gènes que nous avons retenu après l'analyse statistique qui ont une p-value ajustée  $< 0,05$  et un  $\log_2$  du fold change  $> 0,75$  (UP) ou inférieur à  $-0,75$  (DOWN). Nous pouvons observer que chez le mutant *dae-3* seulement 6 gènes sont dérégulés comparé au sauvage, 3 sont plus exprimés et 3 sont moins exprimés incluant le gène *DAE*. Nous pouvons ensuite observer que la surexpression de GFP-DAE entraine une dérégulation de l'expression des gènes un peu plus importante, avec 123 gènes plus exprimés et 12 moins exprimés. Ce résultat est totalement inattendu puisque on aurait pu imaginer que la surexpression d'une endoribonucléase entraine plutôt une baisse de l'expression de ces cibles. Ensuite nous pouvons observer que la surexpression de la version mutante de DAE dans son site catalytique, GFP-D153N OE, a un effet beaucoup plus important sur le transcriptome avec 1228 gènes dont l'expression augmente, soit 10 fois plus qu'observé précédemment, et 307 gènes dont l'expression diminue. Ce résultat indique, comme suggéré lors de la production de lignées transgéniques, que l'expression de D153N a un impact plus délétère que l'expression de la version sauvage, ceci malgré le fait que dans les lignées analysées la protéine mutante s'accumule à des niveaux plus faibles. Enfin nous pouvons observer, comme attendu, une dérégulation massive de l'expression des gènes dans le mutant *vcs-6* avec 4154 gènes avec une hausse de l'expression et 3496 gènes avec une baisse de l'expression. Afin de mieux interpréter ces résultats, nous avons ensuite comparé les gènes dont nous observons une hausse de l'expression chez GFP-D153N OE et chez GFP-DAE OE, cette comparaison (**Figure 21B**) montre qu'il y a 55 gènes communs ce qui représente une fraction significative (44,71%) des gènes surexprimés chez GFP-DAE OE (test hypergéométrique  $p\text{-value} = 3,8 \cdot 10^{-48}$ ). Nous avons ensuite comparé pour chaque gène surexprimé en commun le ratio entre le niveau de dérégulation dans DAE OE et D153N OE ce résultat indique qu'une fois de plus D153N a un impact plus important que DAE, mais cette fois ci sur le niveau d'accumulation des gènes dérégulés (**Figure 21C**).



**Figure 22** La surexpression de DAE influence l'accumulation de gènes sensibles à un dérégulation de la machinerie du decapping

(A) Diagrammes Venn qui montrent les comparaisons entre les gènes sur-exprimés (UP) ou sous-exprimés (DOWN) chez le mutant *vcs-6* et les lignées surexpressesurs DAE OE et D153N OE. La significativité de l'overlap a été validée par un test hypergéométrique. Overlap significatif en rouge.

(B) Validation des expériences transcriptomiques par analyse QPCR de l'expression des gènes dérégulés dans les mutants *dae-3* et *vcs-6* et les surexpressesurs DAE OE lignées 1 et 2 et D153N OE. En jaune les gènes qui sont plus surexprimés chez DAE OE que chez D153N OE. En vert les gènes qui sont plus surexprimés chez D153N OE que dans DAE-OE.

## 1.18 LA SUREXPRESSION DE DAE INFLUENCE L'ACCUMULATION DE GENES SENSIBLES A UN DEREGULATION DE LA MACHINERIE DU DECAPPING

Nos résultats précédents montrent que DAE est associé aux protéines impliquées dans le complexe d'activation du decapping. Nous avons donc ensuite comparé les gènes dérégulés chez GFP-DAE OE et GFP-D153N OE avec ceux dérégulés chez *vcs-6* (**Figure 22A**). Concernant la comparaison entre GFP-DAE OE et *vcs-6*, les gènes surexprimés chez *vcs-6* sont enrichis dans la liste des gènes surexprimés dans DAE OE (p-value =  $5,4 \cdot 10^{-73}$ ). Pour la comparaison entre D153N OE et *vcs-6*, les gènes surexprimés chez *vcs-6* sont enrichis parmi les gènes surexprimés dans D153N OE (n=795, p-value = 0), et les gènes sous-exprimés chez *vcs-6* sont enrichis parmi les gènes sous-exprimés dans D153N OE (n = 204, p-value =  $2,5 \cdot 10^{-133}$ ). Ces résultats indiquent des liens très forts entre les listes de gènes dérégulés dans les lignées DAE OE et le mutant *vcs-6*, ainsi la surexpression de DAE et spécifiquement de sa version mutante D153N, affecte les mêmes gènes qui sont perturbés lorsque l'on affecte l'expression de VCS.

Afin de valider ces résultats nous avons choisi 18 gènes dont l'expression est dérégulée dans les différents génotypes analysés (**Annexe 10**) et nous avons testé leur expression par QPCR sur trois réplicats biologiques indépendants. Dans la **Figure 22B** les gènes analysés sont regroupés en cluster, le premier est représenté par 2 gènes ayant une hausse de l'expression plus importante chez GFP-DAE OE par rapport à GFP-D153N OE (**Figure 22B** en jaune), le deuxième regroupe les 4 gènes dont la hausse de l'expression est plus importante chez D153N OE par rapport à DAE OE (**Figure 22B** en vert). Enfin, le troisième groupe est composé des gènes ayant une hausse d'expression de niveau comparable chez GFP-DAE OE et GFP-D153N OE. Nous avons validé aussi la diminution de l'expression de 4 gènes chez GFP-DAE OE et GFP-D153N OE. De plus nous avons pu valider la hausse d'expression observée chez *vcs-6* pour 14 gènes sur 15. En conclusion les résultats obtenus nous ont permis de valider et renforcer les résultats de l'analyse transcriptomique globale et de confirmer la forte corrélation entre gènes dérégulés lors de la surexpression des différentes formes de DAE et une mutation dans le facteur de decapping VCS.

# DISCUSSION & PERSPECTIVES



## 1.19 DAE UNE ENDORIBONUCLEASE DANS LA DEGRADATION DES ARNm

Alors que les exemples d'endonucléases importantes pour la régulation des ARNm chez les animaux sont nombreux (cf. chap. Introduction § 1.2.6), les connaissances actuelles de leur contribution chez les plantes est minime. On connaît mieux les voies générales cytoplasmiques de dégradation des transcrits, qui dépendent de l'action d'exoribonucléases, incluant l'exosome et SOV pour la dégradation 3'-5' et XRN4 pour la dégradation 5'-3' (Li et al., 2018; Zhang and Guo, 2017). Le seul exemple d'endonucléase vraiment bien caractérisée pour son rôle dans la régulation des ARNm chez les plantes concerne les protéines ARGONAUTE et spécifiquement AGO1, un acteur central pour le mécanisme du RNA silencing et la régulation des cibles de miRNA et *ta*-siRNA (Poulsen et al., 2013; Garcia, 2008). Dans d'autres organismes, les endonucléases à domaine NYN, dont la découverte est plus récente (Anantharaman and Aravind, 2006), ont été étudiées pour leurs rôles dans des processus biologiques cruciaux allant de régulation de la réponse inflammatoire (Regnase), l'amplification du RNA silencing (RDE8), la méiose et la régulation des transposons (MARF1). Chez les plantes seulement deux classes de protéines à domaines NYN avaient été caractérisées avant mon travail de thèse, les PRORP, impliquées dans la maturation des tRNA et les protéines MNU1 et 2 impliquées dans la maturation des transcrits mitochondriaux (Gobert et al., 2010; Stoll and Binder, 2016; Gobert et al., 2019). Notre étude de DAE identifie une nouvelle endonucléase à domaine NYN chez les plantes, et met à jour une nouvelle voie de dégradation qui dépend des activateurs de decapping DCP1 et VCS et de l'activité enzymatique de DAE. Ceci ouvre la perspective à des nouvelles études qui permettront de mieux comprendre le rôle de cette endoribonucléase dans le maintien de l'homéostasie cellulaire des ARNm, un processus important pour la régulation de l'expression des gènes et la biologie de la plante.

### 1.19.1 DAE EST UNE ENDORIBONUCLEASE ACTIVE

J'ai étudié l'activité enzymatique de DAE et montré que celle-ci a une activité enzymatique qui dégrade spécifiquement l'ARN mais pas l'ADN simple brin *in vitro*. Cette activité est dépendante de son domaine NYN, puisque les mutations induites par conversion des acides aspartiques en asparagines (D55N et D153N) abolissent cette activité enzymatique. L'analyse CCM sur matrice de Polyéthylèneimine montre que DAE produit des fragments d'ARN, à la différence de l'exoribonucléase XRN1 qui produit des NMP. DAE est donc une endoribonucléase active *in vitro*. L'absence de fragment d'une taille spécifique dans les produits de l'activité endoribonucléase de DAE pourrait indiquer qu'elle dégrade l'ARN avec un clivage aspécifique dans les conditions de notre test. Ce n'est pas surprenant car le substrat fourni à l'enzyme n'est pas une cible validée de DAE. De plus, on peut noter que le même type de profil de dégradation est observé pour les protéines à

domaine NYN Rde-8, chez *C. elegans* ainsi que MmMARF1 chez la souris (Tsai et al., 2015; Yao et al., 2018). Malgré le test de plusieurs conditions de purification je n'ai jamais réussi à produire MBP-DAE sous forme de monomère et ceci a été un facteur limitant pour une analyse *in vitro* plus poussée de son activité catalytique. Dans toutes les conditions testées DAE sort dans le volume mort de la colonne de filtration sur gel, ce qui pourrait indiquer la présence d'agrégats solubles de la protéine. Une autre explication possible est que DAE soit purifiée en complexe avec de l'ARN, un traitement des extraits protéiques avec des ribonucléases au cours du processus de purification pourrait permettre de tester cette possibilité. Une expérience complémentaire envisageable pourrait être d'exprimer et purifier des versions de DAE dépourvues des séquences N-terminale et/ou C-terminale, une stratégie utilisée pour tester l'activité *in vitro* de MmMARF1 (Yao et al., 2018). L'obtention de la protéine sous forme monomérique et/ou sans ARN associés pourrait permettre d'envisager des tests de son activité plus approfondis, par exemple en présence de ses partenaires protéiques ainsi que la mise en œuvre d'approches structurales.

### 1.19.2 L'INTERACTOME DE DAE

Notre analyse des partenaires protéiques par IP couplée à la spectrométrie de masse de DAE indique qu'elle interagit spécifiquement avec les activateurs de decapping DCP1 et VCS mais pas avec l'enzyme de decapping DCP2 (**Figure 17 C**). L'interactome de DAE montre donc des similarités très fortes avec l'interactome de MARF1 chez les mammifères. Les deux protéines ont la même organisation en domaines (si on excepte les domaines RRM) et interagissent avec les mêmes complexes protéiques. Le modèle le plus récent du mécanisme d'action de MARF1 chez les mammifères, propose que le complexe de decapping est recruté par MARF1 pour induire le decapping du fragment 5' après clivage endoribonucléolytique par MARF1. Ce modèle est supporté par les données qui montrent qu'*in vivo*, hMARF1 dégrade un ARN rapporteur en absence de la région C-terminale nécessaire pour l'interaction avec le complexe de decapping. De plus les données issues d'une analyse par Bio-ID de l'interactome de hMARF1 montrent qu'elle interagit avec les activateurs de decapping DCP1 et Ge-1 (homologue de VCS) mais aussi avec l'enzyme DCP2. L'absence d'interaction de DAE avec DCP2 pourrait indiquer une différence de mode d'action entre MARF1 et DAE. Cette différence pourrait être aussi expliquée par les différences techniques entre les approches utilisées, co-IP ou marquage dépendant de la proximité (BioID). Il est intéressant de remarquer que l'extension C-terminale de HsMARF1 qui permet l'interaction avec DCP1 et DCP2 (Nishimura et al., 2018b) est absente chez DAE. Une expérience de double hybride en utilisant des versions tronquées de DAE pourrait permettre de trouver la région nécessaire pour l'interaction avec DCP1. La présence de DCP1 à la fois dans les IP DAE et dans les IP DCP2 suggère que DCP1 fonctionne dans deux complexes distincts, favorisant l'activité enzymatique de DCP2 (Xu et al., 2006) et partenaire de l'endoribonucléase DAE. À l'instar de son rôle d'activateur de decapping, il est possible que DCP1 favorise l'activité enzymatique de DAE sur ses cibles. Pour confirmer cette hypothèse il est envisageable de tester l'activité endoribonucléase de DAE sur un ARN rapporteur en présence ou en absence de DCP1 et ou VCS, comme suggéré précédemment.

De façon intéressante DAE avait tout d'abord été identifié dans des IP de UPF1 (Chicois et al 2018). La présence d'UPF1 dans les IP crosslink de DAE (**Figure 19D**) pourrait indiquer un rôle de DAE dans la voie du NMD chez les plantes. En effet DAE grâce à son activité endoribonucléase pourrait induire la dégradation de transcrits cibles du NMD par interaction avec UPF1 comme c'est le cas pour SMG6 chez les mammifères. Des résultats obtenus au laboratoire par analyse Q-PCR de l'accumulation de 4 cibles connues du NMD chez le simple mutant *dae-3* ne montrent pas une différence d'accumulation de ces transcrits (Thèse C.Chicois). Cependant l'effet de l'absence de DAE pourrait être masquée par compensation de la voie de dégradation dépendante de SMG7, et potentiellement par l'action de DCP2. Afin de tester cette hypothèse il est envisageable de produire des doubles

mutants *dae/dcp2* et de tester un potentiel effet synergétique sur les cibles du NMD dans ce mutant. Il est intéressant de noter qu'UPF1 est présent à la fois dans les IP crosslink de DCP1 et de DCP2 (**Figure 19B** et **C**); ce résultat suggère que les cibles du NMD ou d'autres voies dépendantes de UPF1 sont substrats du complexe de decapping ainsi que de DAE.

### 1.19.3 DAE, SES LIENS AVEC LA DEGRADATION DES ARNM

L'analyse transcriptomique du mutant *dae-3* (**Figure 21**) montre que dans des conditions naïves, il y a des changements mineurs du transcriptome; 6 gènes sont dérégulés dont 3 surexprimés et 3 sous-exprimés. Une possibilité pouvant expliquer ce résultat serait que, chez *dae-3* le complexe de decapping et l'activité de DCP2 pourraient compenser l'absence de DAE et dégrader ses cibles. La redondance des voies de dégradation est un phénomène connu (Sorenson et al., 2018) qui pourrait expliquer l'absence de changements majeurs du transcriptome chez *dae-3* à l'état d'équilibre.

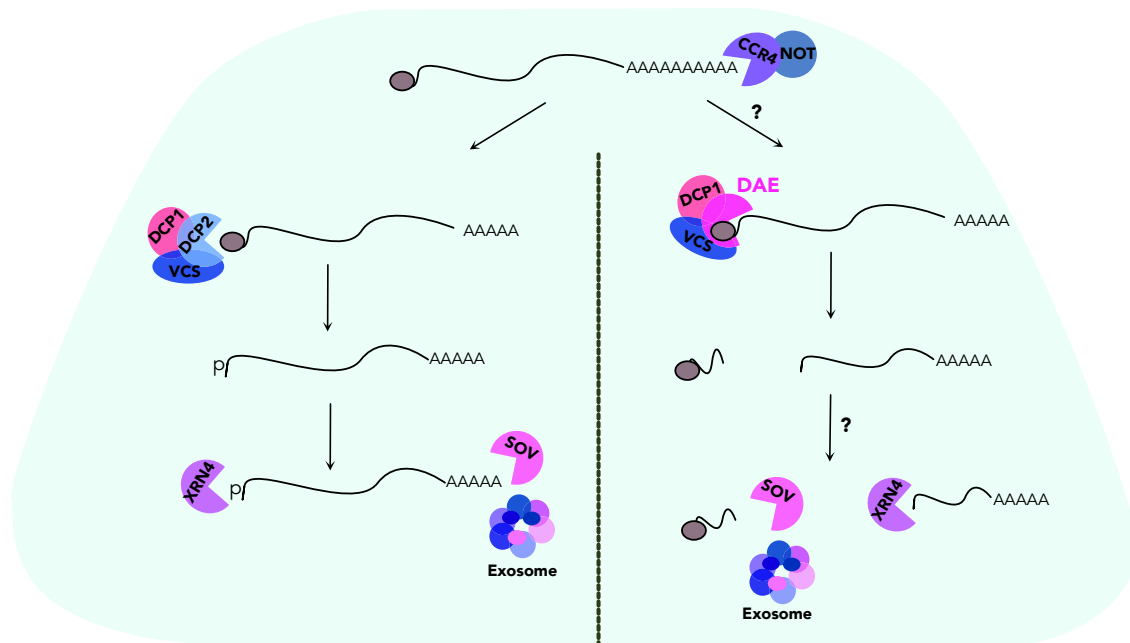
La surexpression de DAE induit une dérégulation du transcriptome, cette dérégulation est encore plus forte lorsque l'on surexprime le mutant catalytique D153N. Une grosse partie des dérégulations du transcriptome observée avec le mutant D153N est donc induite par la mutation de l'activité catalytique de DAE. De plus il est intéressant de remarquer que la plupart des gènes dérégulés dans le surexprimeur, sont surexprimés ce qui est assez contre-intuitif si on considère que DAE est une endoribonucléase. On aurait pu s'attendre à un effet négatif sur ses ARN cibles chez le surexprimeur DAE OE.

La comparaison avec les gènes dérégulés dans le mutant *vcs-6* (**Figure 22A**) montre que environ 80% des gènes surexprimés chez DAE OE sont aussi surexprimés en absence de VCS. Une interprétation possible de ce résultat est que la surexpression de DAE aboutit à une perturbation du mécanisme du decapping lui-même. DAE pourrait séquestrer les activateurs de decapping, ainsi indisponibles pour la création d'un complexe de decapping fonctionnel *via* l'activation de DCP2. Lorsqu'on surexprime la version inactive de DAE (D153N) une dérégulation plus importante du transcriptome est observée ce qui pourrait indiquer que l'activité catalytique de DAE dans les lignées DAE OE entraîne la dégradation de certaines cibles du complexe de decapping, alors que la version mutante les stabilisent. Il est pour l'instant impossible de conclure sur l'identité des cibles de DAE par rapport à celles de DCP2, cependant ces résultats suggèrent que ces deux protéines pourraient avoir des cibles en commun.

Afin d'apporter des indices génétiques des liens entre DAE et DCP2, nous avons croisé GFP-DAE OE avec le mutant hypomorphe de DCP2 appelé *its-1* (*Induced Transgene Silencing 1*) (Thran et al., 2012). Dans ces plantes une amélioration du phénotype d'*its-1* sera observée si DAE compense l'absence de DCP2. En même temps nous avons croisé *its-1* avec *dae-3*. S'il existe une redondance

entre la voie de dégradation dépendante de DAE et celle dépendante de DCP2 les doubles mutants auront un phénotype plus sévère que le simple mutant *its-1*.

Parmi les gènes surexprimés dans la lignée D153N, il y a probablement des cibles directes de DAE. Afin de mettre à jour ces cibles, il serait intéressant d'utiliser une approche de d'IP d'ARN après pontage chimique en utilisant le mutant catalytique D153N comme appât. L'identification des cibles de DAE par cette méthode permettrait d'imaginer de nouvelles expériences afin de mieux définir son rôle biologique qui reste, pour le moment, à déterminer. Il est intéressant de noter la difficulté de trouver les cibles directes de tels facteurs, ainsi les cibles directes de MARF1 chez les mammifères n'ont pas été formellement identifiées et ce malgré les huit articles portant sur la caractérisation de cette endonucléase (Yao et al., 2018; Su et al., 2012b; Bloch et al., 2014; Nishimura et al., 2018a; Kanemitsu et al., 2017; Cao et al., 2018; Su et al., 2012a; Dunster et al., 2005). L'étude de DAE pourrait aussi permettre d'imaginer des modèles alternatifs pour mieux comprendre le mécanisme d'action de MARF1 son homologue le plus proche chez les mammifères, qui régule la progression de la méiose dans les oocytes ainsi que la neurogènes et la réponse immunitaire.



**Figure 23** Model de la voie de dégradation dépendante de DAE.

Les ARNm qui doivent être dégradés sont généralement déadénylés, ensuite décapés par le complexe de decapping composé des activateurs de decapping DCP1 et VCS et de l'enzyme DCP2. Ces ARNm décapés sont ensuite dégradés par l'activité des exoribonucléases 5'-3' et 3'-5': XRN4, SOV et l'exosome. Certains ARNm peuvent, dans des conditions spécifiques, être dégradés par une voie parallèle à celle qui implique le complexe de decapping canonique.

Cette voie dépend de l'activité enzymatique de l'endoribonucléase DAE qui agit sur ses cibles en complexe avec les activateurs de decapping DCP1 et VCS. DAE induit le clivage de l'ARNm. Les fragments 5' et 3' peuvent être ensuite dégradés respectivement par SOV et l'exosome et par XRN4.

#### 1.19.4 UN MODELE POUR LE MODE D'ACTION DE DAE

Les expériences de co-IP avec pontage chimique montrent que DAE interagit spécifiquement avec les activateurs de decapping mais pas avec DCP2, sous-unité catalytique du complexe de decapping canonique. Ceci suggère l'existence de deux complexes (**Figure 23**). Les deux complexes ont en commun les activateurs de decapping mais l'enzyme responsable de l'activité catalytique est soit DAE soit DCP2. Les deux complexes pourraient agir en parallèle pour la dégradation de certains ARNm, ou spécifiquement en fonction des conditions physiologiques. L'absence de phénotype moléculaire des mutants *dae* pourrait être expliquée par la prise en charge des cibles de *dae* par le complexe de decapping canonique lorsque DAE est inactivée. L'existence des deux complexes explique aussi pourquoi lorsque l'on surexprime DAE nous observons une dérégulation des gènes sensibles à une déplétion de VCS. En effet dans ces lignées les activateurs de decapping sont séquestrés par DAE et ne sont plus disponibles pour activer DCP2. Dans un mutant nul *dcp2*, l'activité catalytique de DAE ne semble pas réussir à compenser l'absence de l'activité du complexe de decapping canonique. Ceci indique que probablement DAE et DCP2 partagent seulement une partie des cibles et que certains ARNm sont dégradés spécifiquement par le complexe DCP2 et inaccessible à DAE. Pris ensemble les résultats présentés dans ma thèse révèlent l'existence de ce complexe alternatif qui agit en parallèle du complexe de decapping canonique. On peut spéculer que les cibles de DAE sont probablement prises en charge, à la suite de leur clivage, par les exonucléases 3'-5' et 5'-3', ceci ne pourra être testé qu'après la découverte des cibles directes de DAE. Il serait par exemple intéressant d'analyser la présence de fragments stabilisés dans un double mutant *dae/xrn4* ou *dae/ski2*. Ceci pourrait aussi nous indiquer si DAE agit avec un clivage séquence spécifique, ou s'il est recruté par des structures spécifiques présentes dans l'ARNm.

### 1.19.5 QUELLE EST LA FONCTION BIOLOGIQUE DE DAE?

La conservation de DAE des algues aux plantes à fleurs indique que cette endoribonucléase était sous pression de sélection au cours de l'évolution, ce qui suggère une fonction importante chez les végétaux. Nous avons montré que son domaine catalytique est nécessaire pour la dégradation de l'ARN *in vitro* et *in vivo* (Figure 15 et 16) ce qui indique que son activité catalytique est conservée et donc importante pour sa fonction. Les résultats obtenus précédemment montraient que l'expression transitoire de DAE chez *N. benthamiana* induisait une diminution de la multiplication des phytovirus PVX et TCV (Thèse C.Chicois). Pour cette raison nous avons exploré la possibilité que DAE soit impliquée dans la réponse antivirale. Pour ce faire nous avons infecté le mutant *dae-3* et les surexprimeurs DAE OE et D153N OE avec les virus à ARNss(+) Turnip Mosaic Virus (TuMV) et Turnip Crinkle Virus (TCV) dont *A. thaliana* est un hôte naturel, sans obtenir de différences significatives (résultats non montrés). Une étudiante en Master, Marie Bonnin, a également analysé l'impact de la surexpression de DAE dans la réponse à l'infection par *Pseudomonas syringae* (Pst DC3000), pathogène modèle couramment utilisé pour tester la susceptibilité des plantes à l'infection bactérienne (Katagiri et al., 2002; Xin and He, 2013). Aucune de ces expériences n'a permis de mettre en évidence une différence entre lignée sauvage, mutante et lignée de surexpression. Il serait donc maintenant intéressant de tester la susceptibilité du mutant *dae-3* et des surexprimeurs DAE OE et D153N OE aux stress abiotiques. Les mutants de decapping étant plus sensible au stress osmotique, thermique et à la déshydratation que le sauvage (Xu and Chua, 2012; Perea-Resa et al., 2015). DAE pourrait participer à la réponse à l'un de ces stress, en complexe avec les activateurs de decapping VCS et DCP1 et ainsi contribuer au remodelage du transcriptome dans ces conditions par son activité endoribonucléase.



# CONCLUSION

Mon travail de thèse a permis de caractériser la protéine à domaine NYN DAE. Notre étude phylogénétique montre que DAE est une protéine ancienne qui a été conservée au cours de l'évolution des algues primitives jusqu'aux angiospermes. DAE est une endoribonucléase qui interagit directement avec l'activateur de decapping DCP1 et est présente dans un complexe contenant VCS mais excluant l'enzyme responsable du decapping DCP2. L'analyse de l'interactome de DAE, DCP1 et DCP2 révèle l'existence de deux complexes, un composé par DAE, DCP1 et VCS en parallèle du complexe de decapping canonique, composé par DCP2, DCP1 et VCS. Notre analyse transcriptomique montre que la surexpression de DAE induit une dérégulation des gènes sensibles à une mutation dans l'activateur du decapping VCS ; et que la surexpression de la version inactive de DAE (D153N) produit une dérégulation encore plus importante. Ceci suggère que DAE, en complexe avec DCP1 et VCS, lors de la surexpression perturbe la dégradation d'une partie des ARNm ciblés par le complexe de decapping et que son activité catalytique est probablement nécessaire pour la dégradation d'une partie de ces transcrits. Mon travail permet de proposer l'existence d'une nouvelle voie de dégradation des ARNm, ancienne et très conservée au cours de l'évolution, agissant avec les activateurs du decapping et en parallèle de l'activité du complexe de decapping canonique. L'étude de la fonction biologique de cette nouvelle voie impliquée dans la dégradation des ARNm permettra de mieux comprendre comment les mécanismes de dégradation des ARNm sont orchestrés chez les plantes. Ces processus ont un rôle majeur afin d'assurer la régulation de l'expression des gènes en condition de stress biotiques et abiotiques et l'élimination des transcrits en fin de vie pour le maintien de l'homéostasie cellulaire.

# MATERIELS & METHODES

## 1.20 SOUCHES BACTERIENNES

### **Escherichia coli**

La souche d'*E.coli* **TOP10F'** (Invitrogen™) chimio-compétente a été utilisé pour l'amplification des plasmides. Génotype : F' [*lacI*<sup>q</sup>, *Tn10*(Tet<sup>R</sup>)] *mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80*lacZ*Δ*M15* Δ*lacX74* *recA1* *araD139* Δ(*ara-leu*)7697 *galU* *galK* *rpsL*(Str<sup>R</sup>) *endA1* *nupG*.

La souche d'*E.coli* **BL21 DE3** a été utilisée pour la production des protéines (MBP-DAE, MBP-D55N, MBP-D153N). Elle contient le gène codant pour la T7 RNA polymérase du phage λ sous promoteur *lacUV5* inductible à l'IPTG (isopropyl β-D-1-thiogalactopyranoside). Génotype : λ(DE3[*lacI* *lacUV5*-T7 gene 1 *ind1* *sam7* *nin5*]).

### **Agrobacterium tumefaciens**

La souche d'*A.tumefaciens* **GV3101** (pMP90RK) a été utilisée pour la transformation des plantes (*A.thaliana* et *N. Benthamiana*). Cette souche possède un gène de résistance à la gentamycine intégré dans son génome ainsi qu'un gène de résistance à la rifampicine dans le plasmide Ti (Tumor inducing). Le plasmide Ti pMP90RK est un plasmide helper qui a été rendu avirulent. Il possède les gènes *vir* qui assurent le transfert du T-DNA (Transferred DNA) dans la cellule végétale.

## 1.21 MATERIEL VEGETAL

Toutes les plantes d'*Arabidopsis thaliana* utilisées dans cette étude sont d'écotype Columbia (Colo). Les mutants *dae-1* (Salk\_009495), *dae-2* (Salk\_127449) et *dae-3* (Salk\_132521) proviennent de la collection SIGnAL (Salk Institute Genomic Analysis Laboratory). Le mutant *vcs-6* (SAIL\_831\_Do8) provient de la collection SAIL (Syngenta Arabidopsis Insertion Library). Les plantes ont été cultivées sous lumière LED, sur terre à 21°C la journée et 18°C la nuit, trois semaines avec un cycle de 12h de lumière et 12h d'obscurité. Ensuite transférées dans des conditions de jours long, 16h de lumière et 8h d'obscurité, jusqu'à maturation des siliques.

Pour les expériences d'analyse transcriptomique, les graines ont été stérilisées et déposées sur milieu de culture solide. Après stratification pendant 48h à 4°C les boîtes ont été mises dans une chambre de culture à 21°C avec un cycle diurne de 16h et récoltées à 14 jours de culture.

Pour les expériences d'expression transitoire dans les feuilles de *Nicotiana Benthamiana*, les plantes ont été cultivées 4 semaines avec un cycle diurne de 16h. Les disques foliaires ont été prélevés entre 2,5 et 3, jours après infiltration.

## 1.22 PLASMIDES

### **pGEM-T easy**

Le vecteur pGEM-T easy (Promega) a été utilisé pour le clonage direct de produits PCR à séquencer. Ce vecteur commercial permet le clonage par la technique TA. Il est vendu sous forme linéaire avec une thymidine sortante à chaque extrémité 3' ce qui permet l'insertion de produit PCR qui ont des adénosines (A) sortantes aux extrémités 5'. Le vecteur pGEM-T easy est à haut nombre de copies. Il contient les promoteurs de la T7 et SP6 polymérase flanquant la cassette de clonage qui est insérée au milieu du gène de la  $\beta$ -galactosidase. Ceci permet un sélection blanc/bleu des colonies transformées. Ce plasmide possède un gène de résistance à l'ampicilline.

### **pDONR 207**

Ce vecteur est un plasmide donneur utilisé pour le clonage Gateway®. Il contient les sites de recombinaison attP1 et attP2 ainsi que un gène de résistance à la gentamycine et le gène ccdB entre les sites de recombinaison. La protéine associée au gène ccdB interfère avec l'ADN gyrase bactérienne et permet d'éliminer toute bactérie transformée avec un gène n'ayant pas recombinié.

### **pDONR 221**

Plasmide donneur utilisé pour le clonage Gateway®. Il contient les sites attP1 et attP2, le gène ccdB et un gène de résistance à la kanamycine.

### **pB7WGF2**

Ce plasmide permet l'expression du gène d'intérêt fusionné à la GFP en N-ter sous promoteur constitutif 35S qui provient du virus CaMV (Cauliflower Mosaic Virus) et terminateur 35S. Il a été utilisé pour l'expression stable et transitoire des protéines de fusions GFP-DAE, GFP-D153N, GFP-NYN et GFP-OST-HTH. pB7WGF2 est un plasmide de destination Gateway®. Il contient le gène ccdB entre les sites de recombinaison attR1 et attR2 et porte un gène de résistance à la spectinomycine pour la sélection des bactéries ainsi que un gène de résistance au Glufosinate (BASTA) pour la sélection des plantes.

### **pH7WGR2**

Ce plasmide permet l'expression du gène d'intérêt fusionné à la RFP en N-ter sous promoteur constitutif 35S et terminateur 35S. Il a été utilisé pour l'expression stable et transitoire des protéines de fusions RFP-DAE, RFP-D153N, RFP-NYN et RFP-OST-HTH. pH7WGR2 est un plasmide de destination Gateway®. Il contient le gène ccdB entre les sites de recombinaison attR1 et attR2 et

porte un gène de résistance à la spectinomycine pour la sélection des bactéries ainsi que un gène de résistance à l'hygromycine pour la sélection des plantes.

### **pH2GW7**

Ce vecteur permet l'expression du gène d'intérêt sans étiquette sous promoteurs constitutif 35S. Il possède aussi le terminateur 35S. pH2GW7 est un vecteur de destination Gateway®, il contient le gène ccdB entre les sites attR1 et attR2 ainsi que un gène de résistance à la spectinomycine pour la sélection des bactéries, et un gène de résistance à l'hygromycine pour la sélection des plantes. Ce vecteur a été utilisé pour l'expression transitoire de  $\lambda$ N-HA-DAE,  $\lambda$ N-HA-NYN,  $\lambda$ N-HA-OST-HTH, mGFP5-BoxB3' et 5'BoxB-mGFP5.

### **pMGWA**

Ce vecteur de destination Gateway® contient le gène ccdB entre les sites attR1 et attR2 et il est utilisé pour l'expression des protéines dans un système bactérien. Ce vecteur possède un gène de résistance à l'ampicilline ainsi que le gène lacI qui code pour le répresseur de l'opéron lactose. Ce vecteur permet l'expression de la séquence d'intérêt fusionnée à la MBP (Maltose Binding Protein) en N-ter sous promoteur T7. Ce plasmide a été utilisé pour la production de MBP-DAE, MBP-D55N et MBP-D153N.

### **pGADT7 et pGBT9**

Ces deux vecteurs de destination Gateway® ont été utilisés pour les expériences de double hybrid. pGADT7 et pGBT9 contiennent respectivement la séquence du domaine d'activation et du domaine de liaison à l'ADN de GAL4. Les deux vecteurs possèdent un gène de résistance à l'ampicilline pour la sélection des bactéries. Les levures transformées avec pGADT7 ont été sélectionnées sur milieu SD -Leu. Les levures transformées avec pGBT9 ont été sélectionnées sur milieu SD -Trp.

## **1.23 METHODES RELATIVES A L'ADN**

### **1.23.1 EXTRACTION D'ADN GENOMIQUE**

Une feuille d'*A. thaliana* est broyée dans un tube Eppendorf contenant une bille en métal et 400 µL de tampon d'extraction Edward (Edwards et al., 1991 ; Tris-HCl pH 7,5 200 mM; NaCl 250 mM; EDTA 25 mM; SDS 0,5%) avec le broyeur Tissue Lyser II (Quiagen). Deux broyages successifs à 30 mouvements par minute pendant 2 minutes (min) ont été faits. Ensuite le mélange est centrifugé à 18000g pendant 1 min. 300 µL de surnageant, ainsi dépourvu des débris cellulaires, est prélevé est mélangé avec 300 µL d'isopropanol pour permettre la précipitation de l'ADN. Le mélange est homogénéisé par plusieurs inversions du tube et ensuite centrifugé 5 min à 18000g. Le surnageant est ensuite jeté et le culot contenant l'ADN est laissé sécher à l'air pendant 15-30 min. Le pellet est ensuite remis en suspension dans 100 µL d'eau.

### **1.23.2 AMPLIFICATION PAR PCR**

La réaction de polymérisation en chaîne (PCR) consiste en l'amplification d'une séquence d'ADN spécifique de manière exponentielle. Une réaction PCR est composée de plusieurs étapes répétées. La dénaturation initiale de 2 min à 95/98°C permet de dénaturer la matrice et de casser la paroi bactérienne dans le cas de PCR sur colonies. Ensuite il y a entre 25 et 35 cycles composés de trois étapes : 1 min de dénaturation de l'ADN à 95/98°C ; 30 secondes (s) d'hybridation des amorces entre 50°C et 62°C dépendant du couple d'amorces utilisés ; enfin entre 15 s et 2 min d'élongation à 72°C dépendant de la taille du fragment à amplifier et de l'enzyme utilisée. Enfin une phase d'élongation finale (5 min à 72°C) permet de finir l'élongation de tous les brins.

Pour le génotypage des plantes et les PCR sur colonie l'enzyme GoTaq® (PROMEGA) a été utilisée. Pour l'amplification des fragments d'ADN à cloner, l'enzyme Phusion® High-Fidelity PCR Master Mix (Thermo Fischer Scientific) à haut degré de fidélité a été utilisée.

### **1.23.3 ANALYSE DES PRODUITS PCR PAR ELECTROPHORESE SUR GEL D'AGAROSE**

Afin de vérifier la réussite de l'amplification PCR et la taille des fragment amplifiés, 5 µL des produits PCR sont chargés sur gel d'agarose 1% ou 2% (p/v) selon la taille du fragment et soumis à électrophorèse dans le buffer TBE (Tris-Borate 45 mM; EDTA 1 mM). Ceci permet la séparation des fragments amplifiés selon leur taille. L'ADN chargé négativement migre vers les électrodes positifs, les petits fragments migrent plus vite que les grands fragments. L'estimation de la taille des produits PCR est possible grâce à l'utilisation d'un marqueur de taille (MassRuler DNA Ladder mix®, PROMEGA).

#### **1.23.4 PURIFICATION DE L'ADN**

Pour les clonages il est nécessaire de purifier le produit PCR. Si l'amplification par PCR produit un seul amplicon ce fragment est purifié directement avec le kit Nucleospin™ PCR and Gel cleanup (Macherey-Nagel) sur une colonne de silice et élué avec de l'eau (pour le séquençage) ou avec le tampon d'éluion (Tris-HCl 5mM pH8 ; EDTA 0,1 mM) après lavages avec une solution contenant de l'éthanol.

Si l'amplification PCR produit des fragments de taille différent, la bande de taille attendue est visualisée lumière UV et excisée du gel avec un scalpel et le fragment d'ADN purifié avec le kit Nucleospin™ PCR and Gel cleanup (Macherey-Nagel). L'agarose contenant le fragment d'intérêt est fondu dans un tampon qui contient du perchlorate de sodium. L'ADN, ainsi libéré en solution, est fixé sur colonne de silice et élué avec de l'eau ou avec le tampon d'éluion après lavages avec une solution éthanolique.

#### **1.23.5 MARQUAGE DE L'ADN AVEC DES NUCLEOTIDES RADIOACTIFS**

##### **Marquage de l'extrémité 5'**

Cette méthode de marquage de l'extrémité 5' de l'ADN a été utilisée pour produire le substrat ADN utilisé pour tester l'activité enzymatique de DAE. Dans un volume final de 25 µL, 200 ng d'ADN ont été marqués avec 25 µCi de  $\gamma$ -<sup>32</sup>P dATP en présence de 10U de T<sub>4</sub> PNK et 1/10 (v/v) de buffer T<sub>4</sub> PNK. Le mélange a été incubé 30 min à 37°C. L'ADN radiomarqué a été ensuite dénaturé 5 min à 98°C et gardé dans la glace.

##### **Marquage interne**

Cette méthode de marquage radioactif de l'ADN a été utilisée pour produire les sondes utilisées pour les expériences de Northern Blot. L'ADN matrice d'une taille d'environ 500 bp est amplifié par PCR avec l'enzyme GoTaq® et purifié sur gel d'agarose avec le kit Nucleospin™ PCR and Gel cleanup (Macherey-Nagel). Ensuite 100 ng d'ADN purifié sont utilisés pour la production de sondes radioactives avec le kit Decalabel® DNA Labeling kit (Thermo Scientific). La sonde est marquée avec du  $\alpha$ -<sup>32</sup>P dCTP suivant les instructions du kit. Cette méthode consiste en l'incorporation du  $\alpha$ -<sup>32</sup>P dCTP dans le brin complémentaire à la matrice ADN synthétisé par le fragment de Klenow en utilisant des amorces random qui s'hybrident à la matrice ADN. La sonde radioactive ainsi produite est purifiée par chromatographie d'exclusion sur une résine Sephadex G-50.

## 1.24 TECHNIQUES DE CLONAGE

### Overlap extension PCR

La technique de la PCR d'extension-chevauchement (Overlap PCR) a été utilisée pour la création des fusions  $\lambda$ N-HA-DAE,  $\lambda$ N-HA-NYN,  $\lambda$ N-HA-OST-HTH, mGFP<sub>5</sub>-BoxB<sub>3'</sub> et 5'BoxB-mGFP<sub>5</sub>. Les séquences  $\lambda$ N-HA et BoxB ont été mobilisées par PCR des plasmides pCI-Neo- $\lambda$ N-HA et pCI-Neo-BoxB (gracieusement fournis par Dominique Weil) en utilisant des amorces qui ajoutent une extension correspondante à la séquence de DAE, NYN ou OST, pour  $\lambda$ N-HA ou de la mGFP<sub>5</sub> pour BoxB (PCR<sub>1</sub>). *Viceversa*, les séquences de DAE, NYN, OST-HTH et mGFP<sub>5</sub> ont été mobilisées par PCR en ajoutant des extensions correspondantes à la séquence de  $\lambda$ N-HA ou BoxB (PCR<sub>2</sub>). Les produits de la PCR<sub>1</sub> et 2 ont été chargés sur gel d'agarose 1% (p/v) et purifiés par excision de la bande à la bonne taille (cf. chap. Matériels et méthodes § 1.4.4). Après quantification des fragments d'ADN par NanoDrop® (Thermo Scientific), 50 ng de PCR<sub>1</sub> et 2 ont été mélangés. Les deux fragments ont été ensuite annelés et étendus par PCR en présence de 10  $\mu$ L de Phusion® High-Fidelity PCR Master Mix (Thermo Fischer Scientific). Le produit, PCR<sub>3</sub>, a été utilisé directement pour une ultérieure amplification PCR avec des amorces qui s'hybrident aux extrémités du gène de fusion complet correspondant. Ces amorces contiennent les sites attB pour permettre le clonage Gateway® dans le vecteur donneur pDONR207/pDONR221.

### A-tailing et ligation avec pGEM-T easy

L'enzyme Phusion® utilisée pour l'amplification des fragments à cloner génère des extrémités franches. Afin d'insérer le fragment PCR dans le plasmide pGEM-T easy il est nécessaire d'ajouter des adénosines aux extrémités 3' -OH (A-tailing) avec une Taq ADN polymérase classique. Dans cette étude l'enzyme GoTaq® a été utilisée. Le milieu réactionnel est composé de : 100 ng de produit PCR purifié, 1/5 (v/v) de tampon PCR non coloré 5x, 5U de GoTaq® et 0,2 mM de dATP dans un volume final de 10  $\mu$ L. Le mélange est ensuite incubé 30 min à 72°C. 4  $\mu$ L du produit de la réaction de A-tailing sont utilisés pour la ligation dans le pGEM-T®. Le milieu réactionnel est composé de : 50 ng de plasmide linéaire pGEM-T, 1/2 (v/v) de tampon de ligation rapide (PROMEGA) et 3U de T<sub>4</sub> ADN ligase dans un volume final de 10  $\mu$ L. La réaction se fait pendant la nuit jusqu'au lendemain à 4°C.

### Clonage Gateway®

Le clonage Gateway® (Invitrogen) se base sur le système de recombinaison du phage  $\lambda$ . Cette méthode a été utilisée pour le clonage de DAE ou des fragments 5' et 3' contenant respectivement le domaine NYN et les domaines OST-HTH. L'avantage principal de cette méthode est de pouvoir



utiliser un seul vecteur d'entrée pour pouvoir mobiliser et cloner le fragment d'intérêt dans plusieurs vecteurs de destination.

La première étape du clonage Gateway® est la réaction BP. La réaction permet la création du vecteur d'entrée (pENTRY) à partir d'un vecteur donneur (pDONR206 ou pDONR221), contenant les sites attP<sub>1</sub> et attP<sub>2</sub>, et du fragment d'intérêt bordé des sites de recombinaison attB<sub>1</sub> et attB<sub>2</sub>. Les sites attB sont ajoutés lors de la PCR. Cette réaction de recombinaison est catalysée par l'enzyme BP clonase. L'enzyme LR catalyse la deuxième étape du clonage Gateway® qui consiste en la création d'un vecteur d'expression (pEXP) contenant la séquence d'intérêt par recombinaison des sites attL et attR présent respectivement dans le vecteur d'entrée et dans le vecteur de destination (pDEST).

Le mélange réactionnel pour la BP est le suivant : 150 ng de pDONR, 150 ng du produit PCR purifié et 1 µL de mélange enzymatique BP clonase II (Invitrogen) dans un volume final de 10 µL. La réaction de recombinaison se fait à 25°C pendant 1h. La réaction est stoppée par ajout de 2 µg de Protéinase K et incubation à 37°C pendant 10 min.

Le mélange réactionnel pour la LR est le suivant : 50 à 150 ng d'un vecteur d'entrée pENTRY, 150 ng de pDEST et 1 µL de mélange enzymatique LR clonase II (Invitrogen) dans un volume final de 10 µL. La réaction de recombinaison se fait à 25°C pendant 1h et est ensuite stoppée avec l'ajout de 2 µg de Protéinase K et incubation de 10 min à 37°C.

Après chaque réaction de recombinaison (BP ou LR), 5 µL de milieu réactionnel sont utilisés pour transformer 50 µL des bactéries chimio-compétentes TO10F'.

### **Transformation des bactéries compétentes par choc thermique**

Les bactéries compétentes TOP10F' sont ajoutées à 5 µL de réactions BP/LR, ou à 5 µL de réaction de ligation et incubées sur glace pendant 30 minutes. Le choc thermique est réalisé pendant 30 secondes à 42°C et les bactéries sont placées 1 minute sur glace. 1 mL de milieu de culture LB (Lysogeny Broth : bacto-tryptone 1% ; extrait de levure 0,5% ; NaCl 0,5%) est ajouté avant d'incuber les bactéries 1 heure à 37°C. Les bactéries sont ensuite centrifugées 3 min à 11000g, le surnageant éliminé et le culot est repris dans 100 µL de milieu LB et étalé sur des boîtes de Petri contenant du milieu LB agar (bacto-tryptone 1% ; extrait de levure 0,5% ; NaCl 0,5% ; agar 1%) supplémenté avec l'antibiotique de sélection de la souche bactérienne et du plasmide introduit.

## 1.25 METHODES RELATIVES A L'ARN

### 1.25.1 TRANSCRIPTION *IN VITRO*

Les ARN simple brin de 400 nt utilisés comme substrats pour tester l'activité enzymatique de DAE ont été produits par transcription *in vitro* avec le kit RiboMAX™ Marge Scale RNA production system (Promega). Le mélange réactionnel, fait à température ambiante (T amb) avec un volume final de 20 µL, contient : 1/5 (v/v) de buffer T7 5x, 2,5 U de RNaseOUT™ (ThermoFischer Scientific), 1,5 µL de rNTP (ATP, UTP, CTP, GTP ; 10 mM/chaque), 0,5 µL de UTP (0,5 mM), 20 µCi de α-32P UTP, 200 ng d'ADN et 2 µL de mélange Enzymatique T7. Le mélange est incubé à 37°C pendant 3h. 2U de RQ1 RNases-Free DNase (Promega) sont ensuite ajoutés et le mélange est incubé à 37°C pendant 30 min. Ensuite 40 µL de TRI-Reagent® (Sigma Aldrich) et 8 µL de chloroforme sont ajoutés. Le mélange est vortexé 15 s et incubé à T amb 15 min. Après une centrifugation à 16000g à 4°C pendant 15 min, le surnageant est prélevé et purifié par chromatographie d'exclusion sur une résine Sephadex G-50. L'éluât est récupéré et 40 µg de glycogène ainsi que 0,8 volumes d'Isopropanol 100% sont ajoutés. Ensuite le mélange est vortexé pendant 5 s et centrifugé à 16000g, 4°C pendant 15 min. Le surnageant est éliminé et le culot est lavé deux fois avec 500 µL d'éthanol 100% (v/v). Après le dernier lavage l'éthanol est éliminé et le culot est laissé sécher pendant 10 min. 1 µL d'ARN est chargé avec 1/2 (v/v) de tampon de charge 2X (formamide 95% (v/v), bleu de bromophénol 0,025% (p/v), xylène cyanol FF 0,025% (p/v), EDTA 5 mM, SDS 0,025% (p/v), pH 8.5) sur un gel 8% de polyacrylamide-Urée 7M (10,5g UREA, 5mL 40% acrylamide (19:1), 1,25 mL de TBE 10x, 250 µL de APS 10%, 25 µL de TEMED, volume final 25 mL). La migration se fait à 100V pendant 1h dans du tampon TBE 0,5x. Les ARN sont ensuite quantifiés avec un Scintillateur.

### 1.25.2 EXTRACTION D'ARN

100 mg de matériel végétal sont mis dans un tube Eppendorf préalablement refroidi à l'azote liquide et contenant des billes de verre de diamètre 4mm (Carl Roth). Les échantillons sont broyés 8 s au Silamat® S6 (Ivoclar Vivadent), à une vitesse moyenne de 4500 mouvements/min. 800 µL de TRI-Reagent® sont ajoutés avant un second broyage de 20 s au Silamat®. Un volume de 160 µL de chloroforme est ajouté et les échantillons sont vortexés afin de les homogénéiser puis centrifugés 15 minutes à 4°C à 18000g. Cette étape permet la séparation des échantillons en trois phases : la phase supérieure aqueuse contient les ARN, la phase intermédiaire contient principalement l'ADN et la phase inférieure phénolique contient la majorité des protéines. La phase aqueuse est prélevée et transférée dans un nouveau tube. Les ARN sont précipités par l'ajout de 400 µL d'isopropanol 100% (v/v) suivi d'une homogénéisation par plusieurs inversions du tube. Après une incubation de 10 min à T amb et une centrifugation de 10 min 4°C à 18000g, les culots sont lavés deux fois à l'éthanol à 75%

(v/v) puis repris dans 100 µL d'eau ultra pure (Milli-Q, système de filtration Millipore). Ensuite 100 µL d'un mélange Phénol : Chloroforme : Alcool isoamylique, pH 4,5 (25 :24 :1,v/v) sont ajoutés. Après agitation, les échantillons sont centrifugés à 4°C pendant 15 minutes à 18000g. La phase aqueuse est prélevée et les ARN sont précipités par l'ajout de 1/10 (v/v) de NaAC pH 5,2, 0,5 µg de glycogène et 2,5/1 (v/v) d'éthanol 100%. La précipitation se fait à -80°C pendant la nuit. Ensuite le mélange est centrifugé 30 min à 4°C à 18000g. Les culots sont lavés deux fois à l'éthanol 75% (v/v) puis repris dans 30 à 50 µL d'eau Milli-Q. Les ARN sont ensuite dosés par spectrophotométrie au NanoDrop. La qualité est contrôlée en déposant 500ng d'ARN extraits additionnés de tampon de charge 2x sur un gel d'agarose 1% (p/v) après une dénaturation de 5 minutes à 65°C.

### **1.25.3 QPCR**

Tout d'abord 2 µg d'ARN purifié ont été incubés 30 min à 37°C avec 1U de DNase, RNase free (Thermo-Fischer Scientific) en présence de 4 U de RNase OUT®. La réaction a été stoppée avec l'ajout de 20 mM d'EDTA et incubation à 65°C pendant 10 min.

Ensuite les ADNc ont été synthétisés à partir de 2 µg d'ARN avec l'enzyme SuperScript IV (Invitrogen) avec un mélange d'oligo d(T) et d'hexamères avec une séquence aléatoire. Le mélange réactionnel est composé de : 0,5 mM de chaque dNTP, 2,5 µL d'oligo-d(T)<sub>17</sub>, 0,4 µL d'hexamères random (Promega) dans un volume final de 13 µL. Les ADNc ont été quantifiés avec le kit SYBR Green qPCR (Roche) et des amorces spécifiques aux gènes analysés. La PCR a été faite à la plateforme d'analyse de l'expression génique à l'IBMP avec le LightCycler® 480 Real-Time PCR System (Roche) dans une plaque de 384 puits avec les conditions suivantes : dénaturation initiale de 10 min à 95°C ensuite 45 cycles de dénaturation à 95°C pendant 15 s, hybridation à 60°C pendant 20 s et élongation à 72°C pendant 40 s. Les niveaux d'expression ont été calculés avec la méthode des  $\Delta\Delta C_t$ , normalisés avec ceux des gènes de référence ACTIN2 (AT3G18780) et TIP41 (AT4G34270).

### **1.25.4 CHROMATOGRAPHIE EN COUCHE MINCE (CCM)**

#### **Maqueur UDP**

Pour produire le marqueur UDP tout d'abord un oligonucléotide ADN avec une extrémité ARN en 5' (/5App/CTGACNNNNNNNNNNNNNNNTGGAATTCTCGGGTGCCAAGGC/iBiodT//iBiodT//3ddC/) a été marqué radioactivement en 5'. Le mélange réactionnel est le suivant : 1,5 µCi de  $\alpha$ -32P UTP, 2 µL de buffer NEB 2 10x, 2U de Cid1 (NEB), 5 U de RNase OUT® (Thermo-Fischer Scientific), 0,5 µM d'oligo ARN dans un volume final de 20 µL. Le mélange est incubé 15 min à 25°C, la réaction est ensuite stoppée par incubation à 75°C pendant 5 min. L'oligonucléotide marqué ainsi produit est

purifié sur colonne Sephadex® G50. L'UDP a été produit par incubation à 37°C pendant 30 min de 5 µL d'oligonucléotide marqué avec 1 µL de PNPase (Sigma-Aldrich) dans un volume final de 15 µL.

### **Marqueur UMP**

Pour le test d'activité enzymatique de XRN1 utilisé comme contrôle pour la CCM l'ARNsb de 400 nt radiomarké internement avec du  $\alpha$ -<sup>32</sup>P UTP a été déphosphorylé en 5' par incubation 30 min à 37°C avec 1U de RppH (NewEngland BioLabs) en présence d'1/10 de buffer 3 NEB10x dans un volume final de 10 µL. L'ARN a été ensuite incubé avec 2U de XRN1 (NewEngland BioLabs) pendant 1 h à 37°C.

### **Marqueur Phosphate (P)**

Pour produire le marqueur P le même oligonucléotide utilisé pour le marqueur UDP a été marqué en 5' avec du <sup>32</sup>P. Le mélange réactionnel est composé de : 2,5 µCi de  $\gamma$ -<sup>32</sup>P dATP, 1,5 µL de Buffer A PNK 10x (Thermo-Fischer Scientific), 10 U de T<sub>4</sub> polynucléotide kinase (Thermo-Fischer Scientific). La réaction se fait pendant 30 min à 37°C. L'oligonucléotide marqué est purifié sur colonne Sephadex® G50, 10 µL sont ensuite utilisés comme substrat pour 1U de phosphatase alcaline thermosensible FASTAP (Thermo-Fischer Scientific) dans 1,5 µL de buffer FASTAP 10x dans un volume final de 15 µL. Le mélange est incubé 30 min à 37°C.

### **Marqueur UTP**

Le marqueur UTP a été produit par dilution 1/10000 du stock de  $\alpha$ -<sup>32</sup>P UTP.

### **Migration sur plaque CCM**

2 µL des produits de réaction de DAE ainsi que des marqueurs ont été déposés sur plaque de cellulose (TLC PEI Cellulose F, Merck) carrée (10x10 cm), laissée en contact pendant 30 min dans un tampon contenant 0,5 M de LiCl et 1 M d'acide formique.

## **1.25.5 NORTHERN BLOT**

### **Électrophorèse sur gel de d'agarose au formaldéhyde**

Pour la détection de la GFP dans les patchs infiltrés de *N. benthamiana*, entre 2 et 5 µg d'ARN dans dilués du formamide 50% sont chargés sur gel d'agarose 1% avec le tampon de charge contenant : 1x HEPES, 10% (v/v) de formaldéhyde, 4% (v/v) de formamide, 0.025% bleu de bromophénol. La migration est réalisée dans une solution 1x HEPES pendant 1 heure à 100V sous une hotte chimique.

### **Transfert sur membrane**

Après migration, le gel est lavé 1 fois pendant 5 minutes dans de l'eau MilliQ, puis 5 minutes dans de la solution 10x SSC (3M de chlorure de sodium et 0,3M de citrate de sodium, pH 7.0). Les ARN sont transférés par capillarité sur une membrane de nylon chargée positivement Hybond N+ de Amersham (GE Healthcare). Pour le transfert, un sandwich constitué de 2 papiers Whatman imbibés de tampon 10x SSC, suivi du gel d'agarose, puis de la membrane Hybond N+, et enfin encore 2 papiers de Whatman déposés sur 3 éponges carrées imbibées dans un récipient contenant du tampon 10x SSC. Du papier absorbant est déposé sur le montage pour permettre la diffusion du tampon suivi d'une plaque en métal avec un poids d'environ 1 kg.

### **Coloration de la membrane**

Après transfert, la membrane est incubée 5 minutes dans une solution de SSC 2x et déposée sur Whatman. Les ARN sont fixés sur la membrane par rayonnement UV à 100 x 1200 µjoules pendant 30 s, 2 fois, avec un appareil Stratalinker (Stratagene). La membrane est ensuite lavée 5 minutes avec de l'eau ultra pure et colorée dans 0,02% de bleu de méthylène, 0,3 M d'acétate de sodium pH 5,2. La membrane ainsi colorée est scannée. Avant de procéder à l'hybridation, la membrane est décolorée dans 0,2x SSC, 1% SDS.

### **Hybridation avec sonde radioactive**

La membrane est hybridée 30 minutes avec tampon d'hybridation (PerfectHyb®) à 65°C. La sonde radioactive est rajoutée dans la solution d'hybridation et incubée sur la nuit à 65°C sous rotation. Ensuite la membrane est lavée au moins 3 fois dans du 2x SSC, 0,1% SDS pendant 20 min avant d'être exposée avec un écran PhosphorImager ou à un film autoradiographique à T amb.

### **1.25.6 ANALYSES RNASEQ**

Les ARN purifiés à partir de plantules de 14 jours cultivées in vitro ont été quantifiés au fluorimètre Qubit (Invitrogen) et la qualité des ARN a été analysée avec le système Bioanalyzer 2100 (Agilent) disponible à la plateforme Analyse de l'expression génique à l'IBMP. Seuls les ARN ayant un RIN (RNA Integrity Number) > 6,8 ont été utilisés pour la préparation des bibliothèques. Les bibliothèques ont été préparées à partir de 2 µg d'ARN totaux avec le kit SENSE mRNA-Seq Library Prep Kit V2 for Illumina Platforms (Lexogen) qui permet la production de banques d'ADNc à partir d'ARN poly(A) +. Les banques ont été quantifiées au fluorimètre Qubit et leur qualité a été testée avec le système Bioanalyzer. Les banques ont été séquencées en single read (50 cycles) à la plateforme de séquençage GenomEast (IGBMC, Strasbourg) sur un séquenceur Illumina HiSeq 4000. Les fichiers

fastq ont été générés grâce aux logiciels RTA (v2.7.3) et bcl2fastq (v2.17.1.14). Les séquences ont ensuite été mappées contre le génome d'Arabidopsis (TAIR10) avec Hisat2 (v2.1.0) en utilisant les paramètres par défaut à l'exception de la taille maximum des introns autorisée qui a été réduite à 2000 bp. Les alignements ainsi obtenus ont été annotés grâce au logiciel HTseq count (v0.10.0) en se basant sur le fichier d'annotation Araport11\_GFF3\_genes\_transposons.201606.gff. Afin d'identifier les gènes différentiellement exprimés entre les échantillons, une analyse statistique a été réalisée sur R (v.3.4.2) grâce au package DEseq2 (v1.18.1). Le modèle statistique dans ce package est basé sur une distribution binomiale négative implémenté (Love et al., 2014). Les p-values ont été ajustées avec la méthode de Benjamini et Hochberg. Les gènes avec une p-value ajustée  $< 0.05$  et un Log Fold-change  $> 1$  ont été considérés comme variant de manière significative.

## 1.26 METHODES RELATIVES AUX PROTEINES

### 1.26.1 EXTRACTION DES PROTEINES

Une aliquote de matériel végétal d'*A. thaliana* ou de *N. benthamiana* est mis dans un tube Eppendorf préalablement refroidi à l'azote liquide et contenant des billes de verre de diamètre 4mm (Carl Roth). Les échantillons sont broyés une première fois 8 secondes au Silamat®, à une vitesse moyenne de 4500 mouvements par min puis immédiatement remis dans l'azote liquide. Les échantillons sont broyés une seconde fois, pendant 20 secondes, en présence de 100 µL de tampon d'extraction SDS-urée (Tris pH 6,8 62,5 mM; urée 4M; SDS 3% (p/v); glycérol 10% (p/v); bleu de bromophénol 0,01%) et ensuite chauffés pendant 5 min à 85°C. Le broyat obtenu est centrifugé 5 min à 16000g pour éliminer les débris cellulaires. Les échantillons ainsi obtenus sont analysés sur gel de polyacrylamide et western blot.

### 1.26.2 WESTERN BLOT

#### SDS-PAGE

Les échantillons contenant les protéines à analyser sont chargés sur SDS-PAGE afin de séparer les protéines selon leur poids moléculaire. Les protéines migrent d'abord dans le gel de concentration (0,6 M Tris pH 6,8, 0,5% (p/v) SDS, 4% (p/v) 37,5/1 acrylamide/bisacrylamide) ensuite dans le gel de séparation (1,1 M Tris pH 8,8, 0,3% (p/v) SDS, 10% (p/v) 37,5/1 acrylamide/bisacrylamide). La migration a été faite dans une cellule Mini-PROTEAN® (Bio-Rad) à 100 V pendant 1 h 30 min dans un tampon Tris-Glycine-SDS (25 mM Tris, 192 mM glycine, 0,1% (p/v) SDS, pH 8,6).

#### Transfert sur membrane PVDF

Après migration, les protéines présents dans le gel sont transférées sur une membrane PVDF 0,45 µm (Immobilion-P, Merck Millipore) par transfère liquide avec le système Mini Trans-blot® (Bio-Rad) après activation de la membrane dans 100% (v/v) éthanol. Le transfert a lieu à 4°C pendant 1 h 30 min à 80 V dans un tampon Tris glycine (25 mM Tris, 192 mM glycine, 40% (v/v) éthanol).

#### Immunodetection

Après le transfert, la membrane est saturée pendant 30 min dans 5% (p/v) de lait écrémé dilué dans du Tris-Buffered Saline-Tween 20 (20 mM Tris-HCl (pH 7,4); 150 mM NaCl; 0,2% Tween-20 (v/v)). La membrane est ensuite incubée sur la nuit à 4°C avec l'anticorps primaire antiGFP dilué 1/30000 (gracieusement fourni par David Gilmer) dans 2% de lait écrémé dilué au TBST. Ensuite la membrane est lavée 5 fois pendant 5 min dans du TBST et incubée avec l'anticorps secondaire GAR couplé à la HRP (Horseadish Peroxidase) dilué au 1/10000 dans 2% de lait écrémé dilué avec le TBST.

Enfin la membrane est lavée par 5 lavages consécutifs de 5 minutes avant d'être révélée. La détection des protéines est réalisée par chimiluminescence grâce à l'ajout du substrat Lumi-Light Western Blotting Substrate (Roche). La luminescence est détectée par film autoradiographique ou avec l'appareil FUSION FX.

### **Coloration de la membrane au bleu de Coomassie**

La membrane a été colorée avec une solution contenant le colorant bleu de Coomassie R-250 (40% (v/v) éthanol, 10% (v/v) acide acétique, 0,25g/L de Coomassie R-250) pendant 30 min/1 h à T amb. La membrane est ensuite décolorée pendant plusieurs heures dans une solution contenant 40% (v/v) d'éthanol et 10% (v/v) d'acide acétique.

## **1.26.3 IMMUNOPRECIPITATION ET SPECTROMETRIE DE MASSE**

### **Immunoprecipitation**

Pour l'IP classique 0,9 g de bouquets floraux fermés issu de plantes sauvages et de lignées transformées sont broyés en chambre froide avec mortier et pilon préalablement refroidi à 4°C dans 4,5 mL de tampon de lyse (150 mM NaCl ; 50 mM Tris-HCl pH 8 ; 1% de TritonX-100) en présence d'inhibiteurs de protéases (Complete, EDTA/free, Roche). Pour les IP avec crosslink 0,3 g de bouquet floraux fermés ont été broyés pendant 10 min à 4°C comme décrit précédemment dans du tampon de lyse contenant 50 mM NaCl, 50 mM Tris, inhibiteurs de protéases, 1% TritonX-100 et 240 µL de formaldéhyde 16%. Le crosslink a été stoppé par ajout de 250 µL de glycine 2 M. Pour l'IP classique c'est à ce moment que les lysats obtenus sont divisés en trois réplicats techniques, pour l'IP crosslink c'est seulement après l'incubation avec les billes magnétiques, avant de charger sur les colonnes que les échantillons sont séparés en deux réplicats d'affinité. Dans les deux cas les lysats, après broyage, sont clarifiés par deux étapes de centrifugation successives à 13500g, 10 min à 4°C. Un volume de 60 µL, l'Input, est prélevé et mélangé avec le tampon de charge Laemmli 4x (0,2 M Tris-HCl pH 6,8, 2% (p/v) SDS, 40% (v/v) glycérol, 20% (v/v) β-mercaptoéthanol, 0,02% (p/v) bleu de bromophénol). Ensuite 50 µL de billes magnétiques (Miltenyi) couplé à l'anticorps antiGFP sont ajouté aux lysats clarifiés. Le mélange est incubé pendant 30 min sous rotation à 4°C. Les lysats sont ensuite chargés, 200 µL par 200 µL sur des colonnes µMACS préalablement équilibrées avec 200 µL de tampon de lyse et accrochées sur un support aimanté. 60 µL du troisième ou quatrième dépôt, le « flow-through », sont prélevés et mélangé avec du tampon Laemmli 4x. Quatre lavages successifs sont effectués avec 200 µL de tampon de lavage (20 mM Tris-HCl pH 7,5, Triton 0,1%). Pour finir, l'élution des protéines retenues sur la colonne est réalisée à température ambiante en chargeant 100 µL de



tampon d'éluion (fourni avec le kit) préalablement chauffé à 95°C. Les inputs, flow-through et les éluâtes ont été analysés par western blot.

### **Analyse des échantillons par spectrométrie de masse**

Les échantillons d'éluion ont été analysés par spectrométrie de masse à la Plateforme de Protéomique de l'Institut de Biologie Moléculaire et Cellulaire de Strasbourg (IBMC). Les protéines, digérées avec la trypsine, ont été analysées par nanoLC-MS/MS sur un spectromètre de masse QExactive+ couplé à l'EASY-nanoLC-1000 (Thermo-Fisher Scientific). Les données ont été analysées par comparaison avec la banque de données TAIR10 en utilisant le logiciel Mascot (v2.5; Matrix Science, London, UK) et Proline (v1.4) (<http://proline.profiroteomics.fr/>). Le nombre total de spectres détectés a été utilisé pour quantifier la quantité de chaque protéine dans chaque expérience d'immunoprécipitation. Afin d'identifier les protéines enrichies de manière significative des analyses statistiques ont été réalisées sous R (v3.4.2) en utilisant le package IPInquiry (v3.1) développé par Hélène Zuber. Le test statistique est basé sur le modèle négatif binomial implémenté dans les packages edgeR et msmsTest. Les données ont été normalisées en utilisant la méthode « médiane des ratios » similaire à celle du package DEseq2 (Love et al., 2014). Pour chaque protéine, une p-value ajustée (procédure de correction de Benjamini-Hochberg) et un fold-change ont été calculés. Les volcanoplots, représentant le log<sub>2</sub> (foldchange) et le log<sub>10</sub> (p-value ajustée) respectivement sur les axes des abscisses et des ordonnées ont été tracés sur R en utilisant le package ggplot2 (v3.1.1).

## **1.26.4 PRODUCTION ET PURIFICATION DES PROTEINES RECOMBINANTES**

### **Expression des protéines recombinantes**

La production des protéines est réalisée dans la souche bactérienne d'expression BL21 d'Escherichia coli transformé avec le plasmide pMGWA contenant la séquence de DAE, D55N ou D153N. 5ml de pré-culture contenant 1% de glucose et 100 µg/ml d'ampicilline, ont été incubés pendant la nuit à 37°C. 2ml de la pré-culture ont ensuite été ajoutés à 200ml de LB supplémenté avec 100 µg/ml d'ampicilline et 1% de glucose. Cette culture a été incubée à 37°C jusqu'à avoir une DO<sub>600nm</sub> de 0,8. Ensuite les cultures ont été mises 30 min à 4°C. L'induction a été faite avec 1 mM d'IPTG pendant 3h à 20°C. Après induction, la culture a été centrifugée 20min à 4000g, les surnageants éliminés et les culots lavés une première fois dans 50ml de PBS 1x (140mM NaCl ; 2,7mM KCl ; 10mM Na<sub>2</sub>HPO<sub>4</sub> ; 1,8mM KH<sub>2</sub>PO<sub>4</sub>), puis une deuxième fois dans 20ml de PBS 1x. Après avoir centrifugé 15min à 4000g, tous les surnageants sont soigneusement éliminés avant de congeler instantanément les culots dans de l'azote liquide. Les culots bactériens sont ensuite conservés à -80°C.

### **Purification des protéines recombinantes**

Les culots bactériens sont resuspendus dans le tampon de lyse. Le volume de tampon de lyse est ajusté afin d'avoir une  $DO_{600nm}$  d'environ 20. Les parois bactériennes sont lysées avec 2 passages à la presse française et les échantillons sont centrifugés à 35000g pendant 30 minutes à 4°C pour éliminer les débris cellulaires. Pour la purification par le système HPLC Äkta (GE Healthcare), les lysats clarifiés ont été donnés à Nicolas Baumberger et Laurence Hergott de la plateforme de production et purification des protéines de l'IBMP. Pour la purification en batch le kit Proteus 1-Step Batch Plus Spin Columns a été utilisé (Generon). 335  $\mu$ L de la suspension d'amylose 60% (New England BioLabs Inc) est lavée 2 fois avec 15 ml de tampon de lyse (20mM MOPS pH 7,2 ; 250mM KCl ; 15% glycérol (v/v) ; 0,2mM DTT ; 0,1% Tween 20 (v/v) ; 1x d'inhibiteur de protéase (Complete, EDTA/free, Roche)). La résine est ajoutée au lysat bactérien et le mélange est incubé 3h à 4°C sous rotation. Après incubation le mélange est introduit dans une colonne de purification par gravité, et la résine est lavée à 4 reprises avec 20 mL de tampon de lavage (20mM MOPS pH 7,2 ; 250mM KCl ; 15% glycérol (v/v) ; 0,1% Tween 20 (v/v)). Les échantillons sont élués par 2 éluations consécutives avec 3 mL de tampon d'éluion (20mM MOPS pH 7,2 ; 150mM KCl ; 15% glycérol (v/v) ; 0,1% Tween 20 (v/v) 10 mM maltose monohydrate). Une étape ultérieure de filtration sur gel a été faite à la plateforme de production et purification des protéines de l'IBMP sur le système Äkta. Des aliquotes des protéines ainsi purifiées ont été congelés instantanément avec de l'azote liquide et conservés à -80°C.

### **Coloration du gel au SYPRO® Ruby**

Le SYPRO® Ruby (Thermo-Fischer Scientific) est un colorant fluorescent capable de se lier par des interactions non-covalentes aux protéines. Ce colorant a une haute sensibilité et il a été utilisé pour la quantification de protéines recombinantes purifiées. Le gel est coloré dans 50ml de SYPRO® Ruby pendant la nuit à T amb sous agitation. Le gel est ensuite décoloré dans 100ml de solution de décoloration composée de 10% (v/v) de méthanol et de 7% (v/v) d'acide acétique pendant 30 min à l'obscurité. Le SYPRO® Ruby a deux pics d'excitation, à 450nm et à 280nm, et un pic d'émission à 618nm. Les images de gel sont visualisées et enregistrées avec un système Chemidoc. Afin de quantifier les protéines purifiées, une courbe d'étalonnage a été établie en quantifiant les signaux de fluorescence d'une gamme de protéine BSA (Bovine Serum Albumin) de concentrations connues en utilisant ImageJ.

### 1.26.5 TEST D'ACTIVITE ENZYMATIQUE

La protéine purifiée (MBP-DAE, MBP-D55N, MBP-D153N) à une concentration de 100 nM a été incubée avec 2 nM des substrats ARN ou ADN radiomarqués dans un tampon de réaction composé de : 300 mM Tris-HCl pH8, 400 mM NaCl, 45 mM MgCl<sub>2</sub> et 200 µg/ml de BSA pendant 1 h dans un bain marie à 25°C. 5 µL ont été prélevés à chaque instant t (t<sub>0</sub>, t<sub>15</sub>, t<sub>30</sub>, t<sub>45</sub>, t<sub>60</sub>). La réaction est stoppée par ajout de 5 µL de tampon de charge 2X (formamide 95% (v/v), bleu de bromophénol 0,025% (p/v), xylène cyanol FF 0,025% (p/v), EDTA 5 mM, SDS 0,025% (p/v), pH 8.5). Les produits de la réaction sont chauffés à 70°C pendant 3 min, transférés sur glace et ensuite chargés sur gel dénaturant 8% polyacrylamide-Urée 7M (10,5g UREA, 5mL 40% acrylamide (19:1), 2,5 mL de TBE 10x, 250 µL de APS 10%, 25 µL de TEMED, volume final 25 mL). La migration se fait pendant 1h à 100 V dans du TBE 1X.

### 1.26.6 DOUBLE-HYBRIDE DE LEVURE

Une pre-culture de milieu YPD (Clontech) a été faite à partir de la souche de levure AH109 qui a poussé pendant 7 jours sur un milieu YPD Agar (Clontech) et incubée sur la nuit à 30°C sous agitation ; Ensuite 300 mL de milieu YPD ont été inoculés avec la preculture jusqu'à une DO<sub>600</sub> de 0,3 et ensuite incubée 4h à 30°C sous agitation ; à une DO<sub>600</sub> = 0,5 la culture est centrifugée 5 min à 1000 à T amb. Le culot est resuspendu dans 1,5 mL de solution TE/LiAc stérile (1 mM Tris-HCl, 0,1 mM EDTA pH 8, 0,1 M LiAc). Ensuite 0,1 µg de plasmide avec le domaine d'activation (pGADT7), 0,1 µg de plasmide contenant le domaine de liaison à l'ADN (pGBT9) et 0,1 mg de sperme de saumon sont mélangés avec 100 µL de cellules de levure compétente et avec 600 µL de tampon TE/LiAc/PEG stérile ( 1 mM Tris-HCl, 0,1 mM EDTA pH 8, 0,1 M LiAc, 50% (p/v) PEG PM(3500)). Après 30 min d'incubation à 30°C sous agitation, 70 µL de DMSO sont ajoutés et les levures sont incubées à 42°C pendant 15 min (choc thermique) puis elles sont refroidies sur glace. Les cellules sont centrifugées à 10000g pendant 30 s et les culots sont resuspendus dans 500 µL de buffer TE stérile (1 mM Tris-HCl, 0,1 mM EDTA pH8). Ensuite 100 µL sont déposés sur boîte contenant un milieu de sélection : SD - Leu -Trp, SD -Leu -Trp, -His, SD - Leu -Trp -His - Ade. Les boîtes sont incubées à 30°C pendant 7 jours maximum.

Les colonies obtenues sont ensuite inoculées dans 2 mL de milieu SD -Leu -Trp et incubées à 30°C sous agitation. La culture est centrifugée 2 min à 2000g et le culot resuspendu dans 1 mL d'eau stérile. Le volume est ensuite ajusté pour obtenir une DO<sub>600</sub> = 1,5. Ensuite 10 µL des dilutions 10<sup>-1</sup>, 10<sup>-2</sup>, 10<sup>-3</sup> sont déposés sur boîtes sur milieu agar SD -Leu -Trp, SD -Leu -Trp, -His, SD - Leu -Trp -His - Ade, toujours en présence de 5 mM de 3-AT (3-amino-1,2,4-triazole) pour éviter l'autoactivation. Les boîtes sont incubées à 30°C pendant 3- 10 jours avant d'être prises en photo.

## 1.27 METHODES RELATIVES A LA MANIPULATION DES PLANTES

### 1.27.1 AGROINFILTRATION DES FEUILLES DE *N.BENTHAMIANA*

#### Transformation d'*A.tumefaciens* par choc thermique

Les bactéries GV3101 compétentes (50 µL) sont ajoutées à 100ng de plasmides purifiés et incubées pendant 20 minutes sur glace. Ensuite elles sont incubées 5 minutes à 37°C pour le choc thermique. Après incubation de 2 minutes sur glace, 1 mL de milieu de culture LB est ajouté avant d'incuber les bactéries 2 heures à 28°. Les bactéries sont ensuite centrifugées 3 min à 11000g, le surnageant éliminé et le culot repris dans 100 µL de milieu LB et étalé sur boîte de Pétri contenant du milieu LB agar supplémenté avec la gentamycine, la rifampicine et avec l'antibiotique de sélection du plasmide introduit, à 100 µg/mL chacune. Les bactéries étalées sont ensuite incubées à 28°C pendant 3 jours. Une colonie isolée a été ensuite prélevée et inoculée dans 3 mL de LB liquide avec la gentamycine, la rifampicine et avec l'antibiotique de sélection du plasmide introduit, à 100 µg/mL chacune. Après incubation à 28°C pendant une nuit, ou plus, jusqu'à saturation, 1 mL de culture est ensuite mélangé 1 :1 (v/v) avec du glycérol 50% et gardée à -80°C (« glycerol stock »).

#### Culture et agroinfiltration

50 µL de glycérol stock sont inoculés dans 10 mL de LB liquide avec les antibiotiques. Après incubation à 28°C pendant 24h la culture est centrifugé 10 min à 4000g à T amb et resuspendue dans le tampon d'infiltration (10 mM de MgCl<sub>2</sub>, 100 µM d'acétosyringone). L'absorbance (DO) à 600 nm est ensuite mesuré et le volume de la culture est ajusté pour avoir une DO = 1. Les cultures co-infiltrées sont mélangé avec un ratio 1 :1. Cette préparation est ensuite infiltrée dans le coté abaxial d'une feuille de *N.Benthamiana*, qui a été légèrement blessée avec une aiguille.

### 1.27.2 TRANSFORMATION STABLE D'*A. THALIANA* PAR TREMPAGE FLORAL

Cette méthode a été utilisé pour obtenir les lignées RFP-DAE/YFP-DCP1 et RFP-D153N/YFP-DCP1. Dans ce cas les agrobactéries ont été cultivées et traitée comme indiqué dans le paragraphe précédent mais le culot a été resuspendu dans du milieu d'infiltration composé de : 5% sucrose (p/v), 0.005% de Silvet L-77 (v/v) et 200 µM d'acétosyringone. La suspension a été incubée 1 heure à T amb. Les hampes florales ont ensuite été trempées entièrement dans le milieu d'infiltration pendant 1 min. Ensuite les plantes ont été transférées en logette, 2 jours à l'obscurité, ensuite dans des conditions de jours long jusqu'au jaunissement des siliques et la maturation des graines. Les graines To récoltées, ont été semées et sélectionnées *in vitro* dans un milieu MS solide contenant 25 µg/mL de hygromycine et 250 mg/mL du bactériostatique cefotaxime.

### **1.27.3 CULTURE DES PLANTULES *IN VITRO***

La surface des graines est stérilisée par incubation dans une solution de désinfection (éthanol 70%, Triton X-100 0,1%) pendant 10 min sous agitation douce. Après un rinçage à l'éthanol 100% pendant 5 min les graines sont mises à sécher stérilement sous hotte à flux laminaire. Elles sont ensuite réparties sur milieu solide MS (MS0255 (Duchefa) pH 5,7, agarose 1% (p/v), sucrose 1% (p/v)) à l'aide d'un cure-dent.

### **1.28 MICROSCOPIE CONFOCALE**

Des microscopes Zeiss à balayage laser (Laser Scanning Microscope LSM 700 ou 780), à la plateforme de Microscopie et d'Imagerie de l'IBMP ont été utilisés avec une acquisition des images de type « line by line ». Une macro ImageJ créée par Jeome Mutterer (IBMP) a été utilisée pour la quantification des colocalisations.

### **1.29 ARBRE PHYLOGENETIQUE**

L'arbre phylogénétique a été construit sur la plateforme <https://www.phylogeny.fr/>. Les séquences ont été recherchées avec l'outil tBLASTn. Les alignements multiples ont été faits avec MUSCLE (v3.8.31). Après l'alignement les positions avec un gap ont été éliminées. L'arbre phylogénétique a été construit avec la méthode du maximum de vraisemblance implémentée dans le logiciel PhyML (v3.1/3.0 aLRT, approximate likelihood-ratio test). Les valeurs de bootstrap sont calculées sur 100 itérations. La représentation graphique de l'arbre a été faite avec TreeDyn (v198.3).

## 1.30 LISTE AMORCES

|  |   |                       |
|--|---|-----------------------|
| Transcription in vitro                     |   |                       |
| T7GFP <sub>1 400</sub> F                   | TAATACGACTCACTATAGGGAGTAAAGGAGAAGAAGCTTTTCACT                                 |                       |
| T7GFP <sub>4 400</sub> R                   | TAATACGACTCACTATAGGGTTCCGTCCTCCTTGAATCGA                                      |                       |
| 3G GFP <sub>1 400</sub> F                  | GGGAGTAAAGGAGAAGAAGCTTTTCACT  |                       |
| 3G GFP <sub>4 400</sub> R                  | GGGTTCCGTCCTCCTTGAATCGA   |                       |
| Clonage domaine DAE                        |   |                       |
| F (attB <sub>1</sub> ) TEV NYN             | GGGGACAAGTTTGTACAAAAAAGCAGGCTGGGAAAACCTGTATTTTCAGGGCATGATACAAAACGCTATGTCAAGTT |                       |
| F attB <sub>1</sub> TEV OST<br>HTH         | GGGGACAAGTTTGTACAAAAAAGCAGGCTGGGAAAACCTGTATTTTCAGGGCCAGGCGAGTCTATGGCATGT      |                       |
| R attB <sub>2</sub> NYN                    | GGGGACCACTTTGTACAAGAAAGCTGGGTGAAAACCTGCGAAACCATCAAGG                          |                       |
| R attB <sub>2</sub> OST HTH                | GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAAACCGGGCTGATGAGTT                            |                       |
| Genotypage                                 |   |                       |
| SALK_127441-RP                             | GCGACTATCGTTCCTCGTATCG  | <i>dae-3 et dae-2</i> |
| SALK_127441-LP                             | GCAAGGAGCTTACATTGCTTG   |                       |
| SALK_009495-LP                             | GGAGAATTGTCCCGTTCCTAG   | <i>dae-1</i>          |
| SALK_009495-RP                             | TCCTCCATTGGTAGTTACCCC   |                       |
| Thethering-Overlap PCR                     |   |                       |
| PCR <sub>1</sub>                           |   |                       |
| B-GFP_ATG-F                                | GAGGGTTAATATGAAGACTAATCTTTTTC   |                       |
| GFP_STOP-R                                 | TTAGAGTTCGTCATGTTTG   |                       |
| BoxB F                                     | CGCGTGCTGAAGAACGAGCAG   |                       |
| G-BoxB-R                                   | TAGTCTTCATATTAACCCTCACTAAAGGGAAGC   |                       |
| GFP_ATG-F                                  | ATGAAGACTAATCTTTTTC   |                       |
| B-GFP_STOP-R                               | GTTGGACTTATTAGAGTTCGTCATGTTTG   |                       |
| P-BoxB-F                                   | CGAACTCTAACGCGTGCTGAAGAACGAGCAG   |                       |
| BoxB R                                     | ATTAACCCTCACTAAAGGGAAGC   |                       |
| lambdaN-F                                  | ATGGACGCACAAACACGACG  |                       |
| R-HA-R                                     | AGGAGGCCATAGCGTAATCTGGAACATCATA   |                       |
| H-RFP-F                                    | AGATTACGCTATGGCCTCCTCCGAGGACGTC   |                       |
| DAE-R                                      | TTAAACCGGGCTGATGAGTT  |                       |
| lambdaN-F                                  | ATGGACGCACAAACACGACG  |                       |
| E-HA-R                                     | TTTGTATCATAGCGTAATCTGGAACATCATA   |                       |
| H-DAE-F                                    | AGATTACGCTATGATACAAAACGCTATGTC  |                       |
| DAE-R                                      | TTAAACCGGGCTGATGAGTT  |                       |
| PCR <sub>2</sub>                           |   |                       |
| AttB <sub>1</sub> -BoxB-F                  | ggggacaagtttgtacaaaaaagcaggctCGCGTGCTGAAGAACGAGCAG                            |                       |
| AttB <sub>2</sub> -GFP-R                   | GGGGACCACTTTGTACAAGAAAGCTGGGTTTAGAGTTCGTCATGTTTG                              |                       |
| AttB <sub>1</sub> -GFP <sub>5</sub> -ATG-F | ggggacaagtttgtacaaaaaagcaggctATGAAGACTAATCTTTTTC                              |                       |

|                               |   |                        |
|-------------------------------|---|------------------------|
| AttB2-BoxB-R                  | GGGGACCACCTTTGTACAAGAAAGCTGGGTTATTAACCCTCACTAAAGGAAGC | GFP with BoxB in 5'UTR |
| AttB1-lambdaN-F               | ggggacaagtttgtaaaaaagcaggctATGGACGCACAAACACGACG       |                        |
| AttB2-DAE-R                   | GGGGACCACCTTTGTACAAGAAAGCTGGGTTTTAAACCGGGCTGATGAGTT   |                        |
| lambdaN-F1                    | AGGTGTCCACTCCCAGTTC                                   |                        |
| R-HA-R                        | AGGAGGCCATAGCGTAATCTGGAACATCATA                       | GFP with BoxB in 3'UTR |
| E-HA-R                        | TTTGTATCATAGCGTAATCTGGAACATCATA                       |                        |
| lambdaN-F2                    | GGCTAGAGTACTTAATACGACTCACT                            |                        |
| R-HA-R                        | AGGAGGCCATAGCGTAATCTGGAACATCATA                       |                        |
| E-HA-R                        | TTTGTATCATAGCGTAATCTGGAACATCATA                       | lambdaN fusion RFP-DAE |
| AttB1-lambdaN-F1              | ggggacaagtttgtaaaaaagcaggctAGGTGTCCACTCCCAGTTC        |                        |
| AttB1-lambdaN-F2              | ggggacaagtttgtaaaaaagcaggctGGCTAGAGTACTTAATACGACTCACT |                        |
| QPCR                          |   |                        |
| DAE F2                        | TCAGGTACCTTGAGTCCAC                                   |                        |
| DAE R2                        | GGCCACCCATAATCCGATA                                   |                        |
| DAE F1                        | TCCCGTTCCTAGCGATGTAC                                  |                        |
| DAE R1                        | CGCTTTATCCGACGCATCTT                                  |                        |
| ACT2q F1                      | TTGACCAAGCAGCATGAA                                    |                        |
| ACT2 R1                       | CCGATCCAGACACTGTAATTCCTT                              |                        |
| TIP41 F1                      | GTGAAAAGTGTGGAGAGAAGCAA                               |                        |
| TIP41 R1                      | TCAACTGGATACCCTTTGCA                                  |                        |
| <b>QPCR-Validation RNAseq</b> |   |                        |
| 5G26270q-F                    | CTGTACCCGTTCCGTGCTT                                   |                        |
| 5G26270q-R                    | TGGCCTGTGAAACATATCCA                                  |                        |
| 3G42090q-F                    | CGAAGCCCTCCTCCTAAATC                                  |                        |
| 3G42090q-R                    | CTTCATAGCTGCATCCACCA                                  |                        |
| 1G73120q-F                    | AGAAGAGTGCCTCGCATGAT                                  |                        |
| 1G73120q-R                    | ATTCCCGTACCTTCATGTGG                                  |                        |
| 5G10760q-F                    | CTGAGAGTTGTTCCGCTTCC                                  |                        |
| 5G10760q-R                    | GTTTTACCGCAACCAAAGT                                   |                        |
| 1G14930q-F                    | ACACTTGGGTTTTTCCATCG                                  |                        |
| 1G14930q-R                    | TCCGGGAAGAGATTGTTTTG                                  |                        |
| 3G02670q-F                    | TGATCTCCACCTCACACCA                                   |                        |
| 3G02670q-R                    | TGACTGCAAAGTGGAGAACG                                  |                        |
| 1G68350q-F                    | AGGGTCAAACCTCCTCAGCA                                  |                        |
| 1G68350q-R                    | ATACCGGCCGATACACTGAG                                  |                        |
| 2G06995q-F                    | AACCGGCCTATACAAAACC                                   |                        |
| 2G06995q-R                    | CTGAGGCCTGAGGGTAAACA                                  |                        |
| 5G47000q-F                    | GGATCCCGTGCTGATAAAGA                                  |                        |
| 5G47000q-R                    | AGACGCTAAAAGCCCTAGCC                                  |                        |
| 5G50915q-F                    | CCCTACACCTCACTCCTCCA                                  |                        |

|                            |                             |
|----------------------------|-----------------------------|
| 5G50915q-R                 | CTGGTCGCAGGAAGAAGAAC        |
| 1G67105q-F                 | <b>ACCGGTCTTTTCTCCGTTTT</b> |
| 1G67105q-R                 | AATCGGAGGTGACGAAGATG        |
| 3G42090_2q-F               | GGTTGTTGAATTTGCTCACGTC      |
| 3G42090_2q-R               | CCATCTCCTCTTCTTGCTGC        |
| 4G10250q-F                 | AGGAGCTTTGTTGCTTGAA         |
| 4G10250q-R                 | GCTTGTGTCCCTCTCGAGTC        |
| 2G05914q-F                 | ACCTCACTTCGCTTCTCAGC        |
| 2G05914q-R                 | ACATCTCCGGTGATTGGAAG        |
| 1G04533q-F                 | CTGGCCCTATCACCATCACT        |
| 1G04533q-R                 | CAGGGGAGTCAATGAAAAGC        |
| 1G20070q-F                 | TTTCCTGAAGGATTCGATGG        |
| 1G20070q-R                 | TTCCCAGCAGCTTCTACGTT        |
| 1G53540q-F                 | GAGGTTGAGGATGGCAACAT        |
| 1G53540q-R                 | GAGGTTGAGGATGGCAACAT        |
| 2G20880q-F                 | AACTCCCGTGCATCAGAAAC        |
| 2G20880q-R                 | TCTGGAATTGCATTGAACCA        |
| 3G17110q-F                 | GTAGAGGGTCCGGTTCAGGT        |
| 3G17110q-R                 | CCGTTGTTTCCTCCTCCATA        |
| 4G28140q-F                 | GAGATCCGTAAGCCACGAAG        |
| 4G28140q-R                 | AAATTCAGTGTGCGCTGTG         |
| 1G52690q-F                 | ACTTTAGGGCTTCGCACTGA        |
| 1G52690q-R                 | ATCGGCATAACCTCACGAAC        |
| 3G02480q-F                 | AAGACCGCATGAACAAAAG         |
| 3G02480q-R                 | AATGGACGCAAGGAAACAAC        |
| <b>Sonde Northern blot</b> |                             |
| GFP1                       | AGTAAAGGAGAAGAACTTTTCACT    |
| GFP4                       | TTCCGTCCTCCTTGAAATCGA       |



# BIBLIOGRAPHIE

- Achsel, T., Brahms, H., Kastner, B., Bachi, A., Wilm, M., and Lührmann, R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J.* **18**: 5789–802.
- Ahmed, I. et al. (2014). Mutations in DCPS and EDC3 in autosomal recessive intellectual disability indicate a crucial role for mRNA decapping in neurodevelopment. *Hum. Mol. Genet.* **24**: 3172–3180.
- Aizer, A. and Shav-Tal, Y. (2008). Intracellular trafficking and dynamics of P bodies. *Prion* **2**: 131–134.
- De Alba, A.E.M., Moreno, A.B., Gabriel, M., Mallory, A.C., Christ, A., Bounon, R., Balzergue, S., Aubourg, S., Gautheret, D., Crespi, M.D., Vaucheret, H., and Maizel, A. (2015). In plants, decapping prevents RDR6-dependent production of small interfering RNAs from endogenous mRNAs. *Nucleic Acids Res.* **43**: 2902–2913.
- Alberti, S. and Dormann, D. (2019). Liquid–Liquid Phase Separation in Disease. *Annu. Rev. Genet.* **53**.
- Allen, E. and Howell, M.D. (2010). miRNAs in the biogenesis of trans-acting siRNAs in higher plants. *Semin. Cell Dev. Biol.* **21**: 798–804.
- De Almeida, C., Scheer, H., Gobert, A., Fileccia, V., Martinelli, F., Zuber, H., and Gagliardi, D. (2018). RNA uridylation and decay in plants. *Philos. Trans. R. Soc. B Biol. Sci.* **373**.
- De Almeida, C., Scheer, H., Zuber, H., and Gagliardi, D. (2017). RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. *Wiley Interdiscip. Rev. RNA* **9**.
- Anantharaman, V. and Aravind, L. (2006). The NYN domains: Novel predicted RNAses with a PIN domain-like fold. *RNA Biol.* **3**: 18–27.
- Anderson, P. and Kedersha, N. (2002a). Stressful initiations. *J. Cell Sci.* **115**: 3227–3234.
- Anderson, P. and Kedersha, N. (2002b). Visibly stressed: The role of eIF2, TIA-1, and stress granules in protein translation. *Cell Stress Chaperones* **7**: 213–221.
- Ayache, J., Bénard, M., Ernout-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.
- Basbouss-Serhal, I., Pateyron, S., Cochet, F., Leymarie, J., and Bailly, C. (2017). 5' to 3' mRNA Decay Contributes to the Regulation of Arabidopsis Seed Germination by Dormancy. *Plant Physiol.* **173**: 1709–1723.
- Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* **20**: 1885–1898.
- Berndt, H., Harnisch, C., Rammelt, C., Stöhr, N., Zirkel, A., Dohm, J.C., Himmelbauer, H., Tavanez, J.-P., Hüttelmaier, S., and Wahle, E. (2012). Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. *RNA* **18**: 958–72.

- Beta, R.A.A. and Balatsos, N.A.A.** (2018). Tales around the clock: Poly(A) tails in circadian gene expression. *Wiley Interdiscip. Rev. RNA* **9**: 1–14.
- Bicknell, A.A. and Ricci, E.P.** (2017). When mRNA translation meets decay. *Biochem. Soc. Trans.* **45**: 339–351.
- Bird, J.G. et al.** (2016). The mechanism of RNA 5' capping with NAD<sup>+</sup>, NADH and desphospho-CoA. *Nature* **535**: 444–447.
- 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.
- Bloch, D.B., Nobre, R.A., Bernstein, G.A., and Yang, W.-H.** (2011). Identification and characterization of protein interactions in the mammalian mRNA processing body using a novel two-hybrid assay. *Exp. Cell Res.* **317**: 2183–2199.
- Boccaro, M., Sarazin, A., Thiébeauld, O., Jay, F., Voinnet, O., Navarro, L., and Colot, V.** (2014). The Arabidopsis miR472-RDR6 silencing pathway modulates PAMP- and effector-triggered immunity through the post-transcriptional control of disease resistance genes. *PLoS Pathog.* **10**: e1003883.
- 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.
- Bologna, N.G. and Voinnet, O.** (2014). The Diversity, Biogenesis, and Activities of Endogenous Silencing Small RNAs in *Arabidopsis*. *Annu. Rev. Plant Biol.* **65**: 473–503.
- Bouchoucha, A., Waltz, F., Bonnard, G., Arrivé, M., Hammann, P., Kuhn, L., Schelcher, C., Zuber, H., Gobert, A., and Giegé, P.** (2019). Determination of protein-only RNase P interactome in Arabidopsis mitochondria and chloroplasts identifies a complex between PRORP1 and another NYN domain nuclease. *Plant J.*
- Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Séraphin, B.** (2000). A Sm-like protein complex that participates in mRNA degradation. *EMBO J.* **19**: 1661–71.
- Branscheid, A., Marchais, A., Schott, G., Lange, H., Gagliardi, D., Andersen, S.U., Voinnet, O., and Brodersen, P.** (2015). SKI2 mediates degradation of RISC 5'-cleavage fragments and prevents secondary siRNA production from miRNA targets in Arabidopsis. *Nucleic Acids Res.* **43**: 10975–10988.
- Braun, J.E., Tritschler, F., Haas, G., Igreja, C., Truffault, V., Weichenrieder, O., and Izaurralde, E.** (2010). The C-terminal  $\alpha$ - $\alpha$  superhelix of Pat is required for mRNA decapping in metazoa. *EMBO J.* **29**: 2368–2380.
- Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., and Voinnet, O.** (2008). Widespread Translational Inhibition by Plant miRNAs and siRNAs. *Science* (80-.). **320**: 1185–1190.
- Buchan, J.R. and Parker, R.** (2009). Eukaryotic Stress Granules : The Ins and Out of Translation What are Stress Granules ? *Mol. Cell* **36**: 932.
- Cao, G.-Y., Li, M.-Z., Wang, H., Shi, L.-Y., and Su, Y.Q.** (2018). Interference with the C-terminal structure of

- MARF1 causes defective oocyte meiotic division and female infertility in mice. *J. Biomed. Res.* **32**: 58–67.
- Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Gilbert, K.B., Montgomery, T.A., Nguyen, T., Cuperus, J.T., and Carrington, J.C.** (2012). Functional analysis of three *Arabidopsis* argonautes using slicer-defective mutants. *Plant Cell* **24**: 3613–3629.
- Castellano, L.A. and Bazzini, A.A.** (2017). Poly (A) tails : longer is not always better. *Nat. Publ. Gr.* **24**: 1010–1011.
- Causier, B., Li, Z., De Smet, R., Lloyd, J.P.B., Van De Peer, Y., and Davies, B.** (2017). Conservation of Nonsense-Mediated mRNA Decay Complex Components Throughout Eukaryotic Evolution. *Sci. Rep.* **7**.
- Cevher, M.A., Zhang, X., Fernandez, S., Kim, S., Baquero, J., Nilsson, P., Lee, S., Virtanen, A., and Kleiman, F.E.** (2010). Nuclear deadenylation/polyadenylation factors regulate 3' processing in response to DNA damage. *EMBO J.* **29**: 1674–87.
- Chantarachot, T. and Bailey-Serres, J.** (2018). Polysomes, stress granules, and processing bodies: A dynamic triumvirate controlling cytoplasmic mRNA fate and function. *Plant Physiol.* **176**: 254–269.
- Charenton, C., Taverniti, V., Gaudon-Plesse, C., Back, R., Séraphin, B., and Graille, M.** (2016). Structure of the active form of Dcp1–Dcp2 decapping enzyme bound to m7GDP and its Edc3 activator. *Nat. Struct. Mol. Biol.* **23**: 982–986.
- Chen, X.-J., Zhang, X.-H., Hu, L.-D., Zhang, J.-Q., Jiang, Y., Yang, Y., and Yan, Y.-B.** (2016). DsCaf1 is involved in environmental stress response of *Dunaliella salina*. *Int. J. Biol. Macromol.* **82**: 369–74.
- Chen, Y.G., Kowtoniuk, W.E., Agarwal, I., Shen, Y., and Liu, D.R.** (2009). LC/MS analysis of cellular RNA reveals NAD-linked RNA. *Nat. Chem. Biol.* **5**: 879–881.
- Chiba, Y., Johnson, M.A., Lidder, P., Vogel, J.T., van Erp, H., and Green, P.J.** (2004). AtPARN is an essential poly(A) ribonuclease in *Arabidopsis*. *Gene* **328**: 95–102.
- Chicois, C., Scheer, H., Garcia, S., Zuber, H., Mutterer, J., Chicher, J., Hammann, P., Gagliardi, D., and Garcia, D.** (2018). The UPF1 interactome reveals interaction networks between RNA degradation and translation repression factors in *Arabidopsis*. *Plant J.* **96**: 119–132.
- Cho, H., Kim, K.M., Han, S., Choe, J., Park, S.G., Choi, S.S., and Kim, Y.K.** (2012). Staufen1-Mediated mRNA Decay Functions in Adipogenesis. *Mol. Cell* **46**: 495–506.
- 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.
- Cohen, L.S., Mikhli, C., Jiao, X., Kiledjian, M., Kunkel, G., and Davis, R.E.** (2005). Dcp2 Decaps m<sup>2,2,7</sup>GpppN-Capped RNAs, and Its Activity Is Sequence and Context Dependent. *Mol. Cell. Biol.* **25**: 8779–8791.
- 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.
- Conti, E. and Izaurralde, E.** (2005). Nonsense-mediated mRNA decay: Molecular insights and mechanistic

- variations across species. *Curr. Opin. Cell Biol.*
- Coots, R.A., Liu, X.M., Mao, Y., Dong, L., Zhou, J., Wan, J., Zhang, X., and Qian, S.B.** (2017). m6A Facilitates eIF4F-Independent mRNA Translation. *Mol. Cell* **68**: 504-514.e7.
- Darling, A.L., Liu, Y., Oldfield, C.J., and Uversky, V.N.** (2018). Intrinsically Disordered Proteome of Human Membrane-Less Organelles. *Proteomics* **18**: 1-12.
- Decker, C.J. and Parker, R.** (1993). A turnover pathway for both stable and unstable mRNAs in yeast: Evidence for a requirement for deadenylation. *Genes Dev.* **7**: 1632-1643.
- Dehecq, M., Decourty, L., Namane, A., Proux, C., Kanaan, J., Le Hir, H., Jacquier, A., and Saveanu, C.** (2018). Nonsense-mediated mRNA decay involves two distinct Upf1-bound complexes. *EMBO J.* **37**.
- Deshmukh, M. V., Jones, B.N., Quang-Dang, D.-U., Flinders, J., Floor, S.N., Kim, C., Jemielity, J., Kalek, M., Darzynkiewicz, E., and Gross, J.D.** (2008). mRNA Decapping Is Promoted by an RNA-Binding Channel in Dcp2. *Mol. Cell* **29**: 324-336.
- Deyholos, M.K., Cavaness, G.F., Hall, B., King, E., Punwani, J., Van Norman, J., and Sieburth, L.E.** (2003). VARICOSE, a WD-domain protein, is required for leaf blade development. *Development* **130**: 6577-88.
- Van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E., and Séraphin, B.** (2002). Human Dcp2: A catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J.* **21**: 6915-6924.
- Dunckley, T. and Parker, R.** (1999). The DCP2 protein is required for mRNA decapping in *Saccharomyces cerevisiae* and contains a functional MutT motif. *EMBO J.* **18**: 5411-22.
- Dunster, K., Lai, F.P.L., and Sentry, J.W.** (2005). Limkain b1, a novel human autoantigen localized to a subset of ABCD3 and PXF marked peroxisomes. *Clin. Exp. Immunol.* **140**: 556-563.
- Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H.** (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *TL - 16. Nat. Struct. Mol. Biol.* **16 VN-r**: 49-55.
- Edwards, K., Johnstone, C., and Thompson, C.** (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* **19**: 1349.
- Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E.** (2007a). P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. *Mol. Cell. Biol.* **27**: 3970-3981.
- Eulalio, A., Rehwinkel, J., Stricker, M., Huntzinger, E., Yang, S.F., Doerks, T., Dorner, S., Bork, P., Boutros, M., and Izaurralde, E.** (2007b). Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing. *Genes Dev.* **21**: 2558-2570.
- Fan, S.-J., Marchand, V., and Ephrussi, A.** (2011). *Drosophila* Ge-1 promotes P body formation and oskar mRNA localization. *PLoS One* **6**: e20612.
- Fejes-Toth, K., Sotirova, V., Sachidanandam, R., Assaf, G., Hannon, G.J., Kapranov, P., Foissac, S., Willingham, A.T., Duttagupta, R., Dumais, E., and Gingeras, T.R.** (2009). Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. *Nature* **457**: 1028-1032.
- 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.
- Fernandes, R., Nogueira, G., da Costa, P.J., Pinto, F., and Romão, L.** (2019). Nonsense-Mediated mRNA Decay in Development, Stress and Cancer. In *Advances in Experimental Medicine and Biology* (Springer New York LLC), pp. 41–83.
- Floor, S.N., Jones, B.N., Hernandez, G.A., and Gross, J.D.** (2010). A split active site couples cap recognition by Dcp2 to activation. *Nat. Struct. Mol. Biol.* **17**: 1096–101.
- Fortes, P., Kufel, J., Fornerod, M., Polycarpou-Schwarz, M., Lafontaine, D., Tollervey, D., and Mattaj, I.W.** (1999). Genetic and Physical Interactions Involving the Yeast Nuclear Cap-Binding Complex. *Mol. Cell. Biol.* **19**: 6543–6553.
- 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**: 279–90.
- Gao, M., Fritz, D.T., Ford, L.P., and Wilusz, J.** (2000). Interaction between a poly(A)-specific ribonuclease and the 5' cap influences mRNA deadenylation rates in vitro. *Mol. Cell* **5**: 479–488.
- Garcia, D.** (2008). A miRacle in plant development: Role of microRNAs in cell differentiation and patterning. *Semin. Cell Dev. Biol.* **19**: 586–595.
- 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.
- 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**: 1046–1048.
- Gobert, A., Bruggeman, M., and Giegé, P.** (2019). Involvement of PIN-like domain nucleases in tRNA processing and translation regulation. *IUBMB Life*.
- Gobert, A., Gutmann, B., Taschner, A., Göringer, M., Holzmann, J., Hartmann, R.K., Rossmann, W., and Giegé, P.** (2010). A single *Arabidopsis* organellar protein has RNase P activity. *Nat. Struct. Mol. Biol.* **17**: 740–744.
- Godwin, A.R., Kojima, S., Green, C.B., and Wilusz, J.** (2013). Kiss your tail goodbye: the role of PARN, Nocturnin, and Angel deadenylases in mRNA biology. *Biochim. Biophys. Acta* **1829**: 571–9.
- Goeres, D.C., Van Norman, J.M., Zhang, W., Fauver, N.A., Spencer, M. Lou, and Sieburth, L.E.** (2007). Components of the *Arabidopsis* mRNA decapping complex are required for early seedling development. *Plant Cell* **19**: 1549–64.
- Gonatopoulos-Pournatzis, T. and Cowling, V.H.** (2014). Cap-binding complex (CBC). *Biochem. J.* **457**: 231–242.
- Gong, C. and Maquat, L.E.** (2011). LncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu element. *Nature* **470**: 284–290.
- Gowravaram, M., Schwarz, J., Khilji, S.K., Urlaub, H., and Chakrabarti, S.** (2019). Insights into the

- assembly and architecture of a Staufen-mediated mRNA decay (SMD)-competent mRNP. *Nat. Commun.* **10**.
- Gregory, B.D., O'Malley, R.C., Lister, R., Urich, M.A., Tonti-Filippini, J., Chen, H., Millar, A.H., and Ecker, J.R.** (2008). A Link between RNA Metabolism and Silencing Affecting Arabidopsis Development. *Dev. Cell* **14**: 854–866.
- Gu, K., Mok, L., and Chong, M.M.W.** (2018). Regulating gene expression in animals through RNA endonucleolytic cleavage. *Heliyon* **4**: e00908.
- Gy, I., Gascioli, V., Laressergues, D., Morel, J.B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C.** (2007). Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *Plant Cell* **19**: 3451–3461.
- Haas, G., Braun, J.E., Igraja, C., Tritschler, F., Nishihara, T., and Izaurrealde, E.** (2010). HPat provides a link between deadenylation and decapping in metazoa. *J. Cell Biol.* **189**: 289–302.
- Habacher, C. and Ciosk, R.** (2017a). ZC3H12A/MCP1P/Regnase-1-related endonucleases: An evolutionary perspective on molecular mechanisms and biological functions. *BioEssays* **39**: 1–6.
- Hanson, G. and Collier, J.** (2018). Translation and Protein Quality Control: Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* **19**: 20–30.
- Harrison, A.F. and Shorter, J.** (2017). RNA-binding proteins with prion-like domains in health and disease. *Biochem. J.* **474**: 1417–1438.
- Hayashi, M., Nanba, C., Saito, M., Kondo, M., Takeda, A., Watanabe, Y., and Nishimura, M.** (2012). Loss of XRN4 Function Can Trigger Cosuppression in a Sequence-Dependent Manner. *Plant Cell Physiol.* **53**: 1310–1321.
- He, F. and Jacobson, A.** (2015). Control of mRNA decapping by positive and negative regulatory elements in the Dcp2 C-terminal domain. *RNA* **21**: 1633–47.
- He, F. and Jacobson, A.** (2001). Upf1p, Nmd2p, and Upf3p Regulate the Decapping and Exonucleolytic Degradation of both Nonsense-Containing mRNAs and Wild-Type mRNAs. *Mol. Cell. Biol.* **21**: 1515–1530.
- 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**: 1–9.
- Hopkins, K.C., McLane, L.M., Maqbool, T., Panda, D., Gordesky-Gold, B., and Cherry, S.** (2013). A genome-wide RNAi screen reveals that mRNA decapping restricts bunyaviral replication by limiting the pools of dcp2-accessible targets for cap-snatching. *Genes Dev.* **27**: 1511–1525.
- Horvathova, I., Voigt, F., Kotrys, A. V., Zhan, Y., Artus-Revel, C.G., Eglinger, J., Stadler, M.B., Giorgetti, L., and Chao, J.A.** (2017). The Dynamics of mRNA Turnover Revealed by Single-Molecule Imaging in Single Cells. *Mol. Cell* **68**: 615-625.e9.
- Howell, M.D., Fahlgren, N., Chapman, E.J., Cumbie, J.S., Sullivan, C.M., Givan, S.A., Kasschau, K.D., and Carrington, J.C.** (2007). Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-

- LIKE<sub>4</sub> pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell* **19**: 926–942.
- Hubstenberger, A. et al.** (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Mol. Cell* **68**: 144-157.e5.
- Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E.** (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **14**: 2609–2617.
- Iwasaki, S., Takeda, A., Motose, H., and Watanabe, Y.** (2007). Characterization of Arabidopsis decapping proteins AtDCP1 and AtDCP2, which are essential for post-embryonic development. *FEBS Lett.*
- Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R.** (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* **164**: 487–498.
- Jang, G.-J.J., Yang, J.-Y.Y., Hsieh, H.-L.L., and Wu, S.-H.H.** (2019). Processing bodies control the selective translation for optimal development of Arabidopsis young seedlings. *Proc. Natl. Acad. Sci. U. S. A.* **116**: 6451–6456.
- Jeong, H.-J., Kim, Y.J., Kim, S.H., Kim, Y.-H., Lee, I.-J., 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.
- Jiang, H.Y., Yang, W.H., Gulick, T., Bloch, K.D., and Bloch, D.B.** (2005). Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. *Rna* **11**: 1795–1802.
- Jiao, X., Doamekpor, S.K., Bird, J.G., Nickels, B.E., Tong, L., Hart, R.P., and Kiledjian, M.** (2017). 5' End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding. *Cell* **168**: 1015-1027.e10.
- Jinek, M., Eulalio, A., Lingel, A., Helms, S., Conti, E., and Izaurralde, E.** (2008). The C-terminal region of Ge-1 presents conserved structural features required for P-body localization. *RNA* **14**: 1991–8.
- Julius, C. and Yuzenkova, Y.** (2017). Bacterial RNA polymerase caps RNA with various cofactors and cell wall precursors. *Nucleic Acids Res.* **45**: 8282–8290.
- Julius, C. and Yuzenkova, Y.** (2019). Noncanonical RNA-capping: Discovery, mechanism, and physiological role debate (Blackwell Publishing Ltd).
- Kamenska, A., Simpson, C., Vindry, C., Broomhead, H., Bénard, M., Ernoult-Lange, M., Lee, B.P., Harries, L.W., Weil, D., and Standart, N.** (2016). The DDX6–4E-T interaction mediates translational repression and P-body assembly. *Nucleic Acids Res.* **44**: 6318–6334.
- Kanemitsu, Y., Fujitani, M., Fujita, Y., Zhang, S., Su, Y.Q., Kawahara, Y., and Yamashita, T.** (2017). The RNA-binding protein MARF1 promotes cortical neurogenesis through its RNase activity domain. *Sci. Rep.* **7**: 1–9.
- Kashima, I., Jonas, S., Jayachandran, U., Buchwald, G., Conti, E., Lupas, A.N., and Izaurralde, E.** (2010). SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. *Genes Dev.* **24**: 2440–2450.
- 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. *Proc. Natl. Acad. Sci. U. S. A.* **97**: 13985–13990.
- Katagiri, F., Thilmony, R., and He, S.Y.** (2002). The *Arabidopsis thaliana*-*Pseudomonas syringae* Interaction. *Arab. B.* **1**: e0039.
- Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fitzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P.** (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol.* **169**: 871–884.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E.** (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**: 845–858.
- 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**: 1585–95.
- Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R., Parker, R., 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 Resource The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Mol. Cell*: 1–13.
- Kierzkowski, D., Kmieciak, M., Piontek, P., Wojtaszek, P., Szweykowska-Kulinska, Z., and Jarmolowski, A.** (2009). The *Arabidopsis* CBP20 targets the cap-binding complex to the nucleus, and is stabilized by CBP80. *Plant J.* **59**: 814–25.
- Kiledjian, M.** (2018). Eukaryotic RNA 5'-End NAD<sup>+</sup> Capping and DeNADding. *Trends Cell Biol.* **28**: 454–464.
- Kim, Y.K. and Maquat, L.E.** (2019). UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. *RNA* **25**: 407–422.
- Kosmacz, M., Gorka, M., Schmidt, S., Luzarowski, M., Moreno, J.C., Szlachetko, J., Leniak, E., Sokolowska, E.M., Sofroni, K., Schnittger, A., and Skirycz, A.** (2019). Protein and metabolite composition of *Arabidopsis* stress granules. *New Phytol.* **222**: 1420–1433.
- Kramer, S. and McLennan, A.G.** (2019). The complex enzymology of mRNA decapping: Enzymes of four classes cleave pyrophosphate bonds. *Wiley Interdiscip. Rev. RNA* **10**: 1–17.
- Kurihara, Y.** (2017). Activity and roles of *Arabidopsis thaliana* XRN family exoribonucleases in noncoding RNA pathways. *J. Plant Res.* **130**: 25–31.
- Kurosaki, T. and Maquat, L.E.** (2013). Rules that govern UPF1 binding to mRNA 3' UTRs. *Proc. Natl. Acad. Sci. U. S. A.* **110**: 3357–3362.
- Kurosaki, T., Popp, M.W., and Maquat, L.E.** (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat. Rev. Mol. Cell Biol.* **20**: 406–420.
- Kwasnik, A., Wang, V.Y.-F., Krzyszton, M., Gozdek, A., Zakrzewska-Placzek, M., Stepniak, K., Poznanski, J., Tong, L., and Kufel, J.** (2019). *Arabidopsis* DXO1 links RNA turnover and chloroplast function independently of its enzymatic activity. *Nucleic Acids Res.* **47**: 4751–4764.
- Kwon, Y.J., Park, M.J., Kim, S.G., Baldwin, I.T., and Park, C.M.** (2014). Alternative splicing and nonsense-



- mediated decay of circadian clock genes under environmental stress conditions in Arabidopsis. *BMC Plant Biol.* **14**.
- De la Parra, C., Ernlund, A., Alard, A., Ruggles, K., Ueberheide, B., and Schneider, R.J.** (2018). A widespread alternate form of cap-dependent mRNA translation initiation. *Nat. Commun.* **9**.
- Łabno, A., Tomecki, R., and Dziembowski, A.** (2016). Cytoplasmic RNA decay pathways - Enzymes and mechanisms. *Biochim. Biophys. Acta - Mol. Cell Res.* **1863**: 3125–3147.
- LaGrandeur, T.E. and Parker, R.** (1998). Isolation and characterization of Dcp1p, the yeast mRNA decapping enzyme. *EMBO J.* **17**: 1487–96.
- Lai, T., Cho, H., Liu, Z., Bowler, M.W., Piao, S., Parker, R., Kim, Y.K., and Song, H.** (2012). Structural basis of the PNR2-mediated link between mRNA surveillance and decapping. *Structure* **20**: 2025–2037.
- Lange, H. et al.** (2019). RST1 and RIPR connect the cytosolic RNA exosome to the Ski complex in Arabidopsis. *Nat. Commun.* **10**.
- Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Rättsch, G., and Weigel, D.** (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 8795–800.
- Lewis, J.D., Izaurralde, E., Jarmolowski, A., McGuigan, C., and Mattaj, I.W.** (1996). A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev.* **10**: 1683–98.
- Li, B., Wu, H., and Guo, H.** (2018). Plant mRNA decay: extended roles and potential determinants. *Curr. Opin. Plant Biol.* **45**: 178–184.
- Li, T. et al.** (2015). Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *EMBO J.* **34**: 1630–1647.
- Liang, W., Li, C., Liu, F., Jiang, H., Li, S., Sun, J., Wu, X., and Li, C.** (2009). The Arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. *Cell Res.* **19**: 307–16.
- Lim, J., Ha, M., Chang, H., Kwon, S.C., Simanshu, D.K., Patel, D.J., and Kim, V.N.** (2014). Uridylation by TUT4 and TUT7 marks mRNA for degradation. *Cell* **159**: 1365–1376.
- Lima, S.A., Chipman, L.B., Nicholson, A.L., Chen, Y.H., Yee, B.A., Yeo, G.W., Collier, J., and Pasquinelli, A.E.** (2017). Short poly(A) tails are a conserved feature of highly expressed genes. *Nat. Struct. Mol. Biol.* **24**: 1057–1063.
- Lin, M. Der, Jiao, X., Grima, D., Newbury, S.F., Kiledjian, M., and Chou, T. Bin** (2008). Drosophila processing bodies in oogenesis. *Dev. Biol.* **322**: 276–288.
- Lindsay, A.J. and McCaffrey, M.W.** (2011). Myosin Va is required for P body but not stress granule formation. *J. Biol. Chem.* **286**: 11519–11528.
- Lloyd, J.P.B. and Davies, B.** (2013). SMG1 is an ancient nonsense-mediated mRNA decay effector. *Plant J.* **76**: 800–10.
- Lloyd, J.P.B. and Garcia, D.** (2018). The evolution and diversity of the nonsense-mediated mRNA decay pathway [ version 2 ; referees : 4 approved ] Referee Status :: 1–27.

- Lobel, J.H., Tibble, R.W., and Gross, J.D.** (2019). Pat1 activates late steps in mRNA decay by multiple mechanisms. *bioRxiv*: 594168.
- Love, M.I., Huber, W., and Anders, S.** (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**.
- Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A.** (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. *EMBO J.* **32**: 1855–1868.
- Lukhovitskaya, N. and Ryabova, L.A.** (2019). Cauliflower mosaic virus transactivator protein (TAV) can suppress nonsense-mediated decay by targeting VARICOSE, a scaffold protein of the decapping complex. *Sci. Rep.* **9**: 7042.
- Lv, H., Zhu, Y., Qiu, Y., Niu, L., Teng, M., and Li, X.** (2015). Structural analysis of Dis3L2, an exosome-independent exonuclease from *Schizosaccharomyces pombe*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**: 1284–1294.
- Lykke-Andersen, J.** (2002). Identification of a Human Decapping Complex Associated with hUpf Proteins in Nonsense-Mediated Decay. *Mol. Cell. Biol.* **22**: 8114–8121.
- Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M.G., and Arraiano, C.M.** (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* **32**: 1842–1854.
- Mangus, D.A., Evans, M.C., Agrin, N.S., Smith, M., Gongidi, P., and Jacobson, A.** (2004). Positive and Negative Regulation of Poly(A) Nuclease. *Mol. Cell. Biol.* **24**: 5521–5533.
- 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**.
- Maragozidis, P., Papanastasi, E., Scutelnic, D., Totomi, A., Kokkori, I., Zarogiannis, S.G., Kerenidi, T., Gourgoulianis, K.I., and Balatsos, N.A.A.** (2015). Poly(A)-specific ribonuclease and Nocturnin in squamous cell lung cancer: Prognostic value and impact on gene expression. *Mol. Cancer* **14**: 187.
- Marnef, A., Sommerville, J., and Ladomery, M.R.** (2009). RAP55: insights into an evolutionarily conserved protein family. *Int. J. Biochem. Cell Biol.* **41**: 977–81.
- Marshall, A.N., Han, J., Kim, M., and Van Hoof, A.** (2018). Conservation of mRNA quality control factor Ski7 and its diversification through changes in alternative splicing and gene duplication. *Proc. Natl. Acad. Sci. U. S. A.* **115**: E6808–E6816.
- Martínez, J., Ren, Y.G., Nilsson, P., Ehrenberg, M., and Virtanen, A.** (2001). The mRNA Cap Structure Stimulates Rate of Poly(A) Removal and Amplifies Processivity of Degradation. *J. Biol. Chem.* **276**: 27923–27929.
- 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.
- Matheny, T., Rao, B., and Parker, R.** (2019). Transcriptome-wide comparison of stress granules and P-bodies

- reveals that translation plays a major role in RNA partitioning. *Mol. Cell. Biol.*
- McLennan, A.G.** (2006). The Nudix hydrolase superfamily. *Cell. Mol. Life Sci.* **63**: 123–143.
- Mérai, Z., Benkovics, A.H., Nyikó, T., Debreczeny, M., Hiripi, L., Kerényi, Z., Kondorosi, É., and Silhavy, D.** (2013). The late steps of plant nonsense-mediated mRNA decay. *Plant J.* **73**: 50–62.
- Merret, R., Descombin, J., Juan, Y. ting, Favory, J.J., Carpentier, M.C., Chaparro, C., Charng, Y. yung, Deragon, J.M., and Bousquet-Antonelli, C.** (2013). XRN<sub>4</sub> and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. *Cell Rep.* **5**: 1279–1293.
- Mino, T., Murakawa, Y., Fukao, A., Vandebon, A., Wessels, H., and Ori, D.** (2015). Regnase-1 and Roquin Regulate a Common Element in Inflammatory mRNAs by Spatiotemporally Distinct Mechanisms. *Cell* **161**: 1058–1073.
- Moon, D.H., Segal, M., Boyraz, B., Guinan, E., Hofmann, I., Cahan, P., Tai, A.K., and Agarwal, S.** (2015). Poly(A)-specific ribonuclease (PARN) mediates 3'-end maturation of the telomerase RNA component. *Nat. Genet.* **47**: 1482–8.
- Morgan, M., Much, C., Digiacomio, M., Azzi, C., Ivanova, I., Vitsios, D.M., Pistolic, J., Collier, P., Moreira, P.N., Benes, V., Enright, A.J., and Carroll, D.O.** (2017). the mammalian maternal transcriptome. *Nat. Publ. Gr.*
- Mugridge, J.S., Collier, J., and Gross, J.D.** (2018a). Structural and molecular mechanisms for the control of eukaryotic 5'-3' mRNA decay. *Nat. Struct. Mol. Biol.* **25**: 1077–1085.
- Mugridge, J.S., Tibble, R.W., Ziemniak, M., Jemielity, J., and Gross, J.D.** (2018b). Structure of the activated Edc1-Dcp1-Dcp2-Edc3 mRNA decapping complex with substrate analog poised for catalysis. *Nat. Commun.* **9**: 1152.
- Muhlrad, D. and Parker, R.** (2005). The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p. *EMBO J.* **24**: 1033–45.
- Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J.** (2013). XRN 5'→3' exoribonucleases: Structure, mechanisms and functions. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1829**: 590–603.
- Nagarajan, V.K., Kukulich, P.M., von Hagel, B., and Green, P.J.** (2019). RNA degradomes reveal substrates and importance for dark and nitrogen stress responses of Arabidopsis XRN<sub>4</sub>. *Nucleic Acids Res.* **47**: 9216–9230.
- Nakamura, A., Amikura, R., Hanyu, K., and Kobayashi, S.** (2001). Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis. *Development* **128**: 3233–3242.
- 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**.
- Nicholson, A.L. and Pasquinelli, A.E.** (2019). Tales of Detailed Poly(A) Tails. *Trends Cell Biol.* **29**: 191–200.

- Nicholson, P., Gkratsou, A., Josi, C., Colombo, M., and Mühlemann, O.** (2018). Dissecting the functions of SMG5, SMG7, and PNRC2 in nonsense-mediated mRNA decay of human cells. *RNA* **24**: 557–573.
- Nicholson, P., Josi, C., Kurosawa, H., Yamashita, A., and Mühlemann, O.** (2014). A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic Acids Res.* **42**: 9217–9235.
- Nishimura, N., Okamoto, M., Narusaka, M., Yasuda, M., Nakashita, H., Shinozaki, K., Narusaka, Y., and Hirayama, T.** (2009). ABA Hypersensitive Germination2-1 Causes the Activation of Both Abscisic Acid and Salicylic Acid Responses in Arabidopsis. *Plant Cell Physiol.* **50**: 2112–2122.
- Nishimura, T., Fakim, H., Brandmann, T., Youn, J.Y., Gingras, A.C., Jinek, M., and Fabian, M.R.** (2018). Human MARF1 is an endoribonuclease that interacts with the DCP1:2 decapping complex and degrades target mRNAs. *Nucleic Acids Res.* **46**: 12008–12021.
- Oeffinger, M. and Zenklusen, D.** (2019). The Biology of mRNA: Structure and Function.
- Ohtani, M. and Wachter, A.** (2019). NMD-Based Gene Regulation - A Strategy for Fitness Enhancement in Plants? *Plant Cell Physiol.* **60**: 1953–1960.
- Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S.** (2012). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Res.* **40**: 1251–1266.
- Olmedo, G., Guo, H., Gregory, B.D., Nourizadeh, S.D., Aguilar-Henonin, L., Li, H., An, F., Guzman, P., and Ecker, J.R.** (2006). ETHYLENE-INSENSITIVE5 encodes a 5'→3' exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 13286–93.
- Otsuka, Y., Kedersha, N.L., and Schoenberg, D.R.** (2009). Identification of a Cytoplasmic Complex That Adds a Cap onto 5'-Monophosphate RNA. *Mol. Cell. Biol.* **29**: 2155–2167.
- Ottens, F. and Gehring, N.H.** (2016). Physiological and pathophysiological role of nonsense-mediated mRNA decay. *Pflugers Arch. Eur. J. Physiol.* **468**: 1013–1028.
- Ozgur, S., Chekulaeva, M., and Stoecklin, G.** (2010). Human Pat1b Connects Deadenylation with mRNA Decapping and Controls the Assembly of Processing Bodies. *Mol. Cell. Biol.* **30**: 4308–4323.
- Pabis, M., Neufeld, N., Steiner, M.C., Bojic, T., Shav-Tal, Y., and Neugebauer, K.M.** (2013). The nuclear cap-binding complex interacts with the U4/U6-U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA* **19**: 1054–63.
- Pan, S., Li, K. en, Huang, W., Zhong, H., Wu, H., Wang, Y., Zhang, H., Cai, Z., Guo, H., Chen, X., and Xia, Y.** (2019). Arabidopsis DXO1 possesses deNADding and exonuclease activities and its mutation affects defense-related and photosynthetic gene expression. *J. Integr. Plant Biol.*
- Paquette, D.R., Tibble, R.W., Daifuku, T.S., and Gross, J.D.** (2018). Control of mRNA decapping by autoinhibition. *Nucleic Acids Res.* **46**: 6318–6329.
- Parker, R. and Sheth, U.** (2007). P Bodies and the Control of mRNA Translation and Degradation. *Mol. Cell* **25**: 635–646.
- Parker, R. and Song, H.** (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat. Struct. Mol.*

Biol. **11**: 121–127.

- Pavlopoulou, A., Vlachakis, D., Balatsos, N.A.A., and Kossida, S.** (2013). A comprehensive phylogenetic analysis of deadenylases. *Evol. Bioinforma.* **2013**: 491–497.
- Pelechano, V., Wei, W., and Steinmetz, L.M.** (2015). Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell* **161**: 1400–12.
- Le Pen, J. et al.** (2018). Terminal uridylyltransferases target RNA viruses as part of the innate immune system. *Nat. Struct. Mol. Biol.* **25**: 778–786.
- Perea-Resa, C., Carrasco-López, C., Catalá, R., Turečková, V., Novak, O., Zhang, W., Sieburth, L., Jiménez-Gómez, J.M., and Salinas, J.** (2015). The LSM1-7 complex differentially regulates arabidopsis tolerance to abiotic stress conditions by promoting selective mRNA decapping. *Plant Cell* **28**: 505–520.
- Potuschak, T., Vansiri, A., Binder, B.M., Lechner, E., Vierstra, R.D., and Genschik, P.** (2006). The exoribonuclease XRN4 is a component of the ethylene response pathway in Arabidopsis. *Plant Cell* **18**: 3047–57.
- Poulsen, C., Vaucheret, H., and Brodersen, P.** (2013). Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *Plant Cell* **25**: 22–37.
- Radhakrishnan, A., Chen, Y.-H., Martin, S., Alhusaini, N., Green, R., and Collier, J.** (2016). The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. *Cell*: 1–11.
- Radó-Trilla, N. and Albà, M.** (2012). Dissecting the role of low-complexity regions in the evolution of vertebrate proteins. *BMC Evol. Biol.* **12**.
- Rajgor, D., Hanley, J.G., and Shanahan, C.M.** (2016). Identification of novel nesprin-1 binding partners and cytoplasmic matrin-3 in processing bodies. *Mol. Biol. Cell* **27**: 3894–3902.
- Ramanathan, A., Robb, G.B., and Chan, S.-H.** (2016). mRNA capping: biological functions and applications. *Nucleic Acids Res.* **44**: 7511–7526.
- Rao, S., Amorim, R., Niu, M., Breton, Y., Tremblay, M.J., and Mouland, A.J.** (2019). Host mRNA decay proteins influence HIV-1 replication and viral gene expression in primary monocyte-derived macrophages. *Retrovirology* **16**.
- Rayson, S., Ashworth, M., de Torres Zabala, M., Grant, M., and Davies, B.** (2012). The salicylic acid dependent and independent effects of NMD in plants. *Plant Signal. Behav.* **7**: 1434–7.
- Ren, G., Chen, X., and Yu, B.** (2012). Uridylation of miRNAs by HEN1 SUPPRESSOR1 in arabidopsis. *Curr. Biol.* **22**: 695–700.
- Ries, R.J., Zaccara, S., Klein, P., Olarerin-George, A., Namkoong, S., Pickering, B.F., Patil, D.P., Kwak, H., Lee, J.H., and Jaffrey, S.R.** (2019). m6A enhances the phase separation potential of mRNA. *Nature*.
- Rissland, O.S. and Norbury, C.J.** (2009). Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. *Nat. Struct. Mol. Biol.* **16**: 616–623.
- Roegiers, F. and Jan, Y.N.** (2000). Staufen: A common component of mRNA transport in oocytes and neurons? *Trends Cell Biol.* **10**: 220–224.
- Sarowar, S., Oh, H.W., Cho, H.S., Baek, K.-H., Seong, E.S., Joung, Y.H., Choi, G.J., Lee, S., and Choi, D.**

- (2007). Capsicum annuum CCR4-associated factor CaCAF<sub>1</sub> is necessary for plant development and defence response. *Plant J.* **51**: 792–802.
- Schäfer, I.B., Yamashita, M., Schuller, J.M., Schüssler, S., Reichelt, P., Strauss, M., and Conti, E.** (2019). Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase. *Cell* **177**: 1619–1631.e21.
- Schmidt, S.A., Foley, P.L., Jeong, D.H., Rymarquis, L.A., Doyle, F., Tenenbaum, S.A., Belasco, J.G., and Green, P.J.** (2015). Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. *Nucleic Acids Res.* **43**: 309–323.
- Schoenberg, D.R., Baron-Benhamou, J., Gehring, N.H., Kulozik, A.E., and Hentze, M.W.** (2004). Using the  $\lambda$ N Peptide to Tether Proteins to RNAs. In *mRNA Processing and Metabolism* (Humana Press), pp. 135–154.
- Shatkin, A.J. and Manley, J.L.** (2000). The ends of the affair: Capping and polyadenylation. *Nat. Struct. Biol.* **7**: 838–842.
- Shaul, O.** (2015). Unique Aspects of Plant Nonsense-mediated mRNA Decay. *Trends Plant Sci.* **xx**: 1–13.
- She, M., Decker, C.J., Chen, N., Tumati, S., Parker, R., and Song, H.** (2006). Crystal structure and functional analysis of Dcp2p from *Schizosaccharomyces pombe*. *Nat. Struct. Mol. Biol.* **13**: 63–70.
- Sheth, U. and Parker, R.** (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* (80-. ). **300**: 805–808.
- Shirokikh, N.E. and Preiss, T.** (2018). Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. *Wiley Interdiscip. Rev. RNA* **9**: 1–45.
- Shivaprasad, P. V., Chen, H.M., Patel, K., Bond, D.M., Santos, B.A.C.M., and Baulcombe, D.C.** (2012). A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell* **24**: 859–874.
- Shukla, S., Bjerke, G.A., Muhlrad, D., Yi, R., and Parker, R.** (2019). The RNase PARN Controls the Levels of Specific miRNAs that Contribute to p53 Regulation. *Mol. Cell* **73**: 1204–1216.e4.
- Shukla, S. and Parker, R.** (2017). PARN Modulates Y RNA Stability and Its 3'-End Formation. *Mol. Cell. Biol.* **37**.
- Shukla, S., Schmidt, J.C., Goldfarb, K.C., Cech, T.R., and Parker, R.** (2016). Inhibition of telomerase RNA decay rescues telomerase deficiency caused by dyskerin or PARN defects. *Nat. Struct. Mol. Biol.* **23**: 286–92.
- Sikorska, N., Zuber, H., Gobert, A., Lange, H., and Gagliardi, D.** (2017). RNA degradation by the plant RNA exosome involves both phosphorolytic and hydrolytic activities. *Nat. Commun.* **8**.
- Son, A., Park, J.E., and Kim, V.N.** (2018). PARN and TOE1 Constitute a 3' End Maturation Module for Nuclear Non-coding RNAs. *Cell Rep.* **23**: 888–898.
- Song, M.G. and Kiledjian, M.** (2007). 3' terminal oligo U-tract-mediated stimulation of decapping. *RNA* **13**: 2356–2365.
- Sorenson, R.S., Deshotel, M.J., Johnson, K., Adler, F.R., and Sieburth, L.E.** (2018). Arabidopsis mRNA

- decay landscape arises from specialized RNA decay substrates, decapping-mediated feedback, and redundancy. *Proc. Natl. Acad. Sci. U. S. A.* **115**: E1485–E1494.
- Souret, F.F., Kastenmayer, J.P., and Green, P.J.** (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. *Mol. Cell* **15**: 173–183.
- St Johnston, D., Beuchle, D., and Nüsslein-Volhard, C.** (1991). *staufen*, a gene required to localize maternal RNAs in the Drosophila egg. *Cell* **66**: 51–63.
- Standart, N. and Weil, D.** (2018). P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. *Trends Genet.* **34**: 612–626.
- Di Stefano, B. et al.** (2019). The RNA Helicase DDX6 Controls Cellular Plasticity by Modulating P-Body Homeostasis. *Cell Stem Cell* **25**: 622–638.e13.
- 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**: 1879–92.
- Steiger, M., Carr-Schmid, A., Schwartz, D.C., Kiledjian, M., and Parker, R.** (2003). Analysis of recombinant yeast decapping enzyme. *RNA* **9**: 231–8.
- Stoecklin, G. and Kedersha, N.** (2013). Relationship of GW/P-Bodies with Stress Granules. In, pp. 197–211.
- 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., Sugiura, K., Sun, F., Pendola, J.K., Cox, G.A., Handel, M.A., Schimenti, J.C., and Eppig, J.J.** (2012a). MARF1 regulates essential oogenic processes in mice. *Science* (80-. ). **335**: 1496–1499.
- Su, Y.Q., Sun, F., Handel, M.A., Schimentic, J.C., and Eppig, J.J.** (2012b). Meiosis arrest female 1 (MARF1) has nuage-like function in mammalian oocytes. *Proc. Natl. Acad. Sci. U. S. A.* **109**: 18653–18660.
- Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P.** (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* **508**: 66–71.
- Sureshkumar, S., Dent, C., Seleznev, A., Tasset, C., and Balasubramanian, S.** (2016). Nonsense-mediated mRNA decay modulates FLM-dependent thermosensory flowering response in Arabidopsis. *Nat. Plants* **2**.
- Suzuki, H.I., Arase, M., Matsuyama, H., Choi, Y.L., Ueno, T., Mano, H., Sugimoto, K., and Miyazono, K.** (2011). MCP1P1 ribonuclease antagonizes dicer and terminates microRNA biogenesis through precursor microRNA degradation. *Mol. Cell* **44**: 424–436.
- Suzuki, Y., Arai, T., Green, P.J., Yamaguchi, J., and Chiba, Y.** (2015). AtCCR4a and AtCCR4b are Involved in Determining the Poly(A) Length of Granule-bound starch synthase 1 Transcript and Modulating Sucrose and Starch Metabolism in Arabidopsis thaliana. *Plant Cell Physiol.* **56**: 863–74.
- Swisher, K.D. and Parker, R.** (2011). Interactions between Upf1 and the Decapping Factors Edc3 and Pat1 in *Saccharomyces cerevisiae*. *PLoS One* **6**: e26547.
- Szádeczky-Kardoss, I., Csorba, T., Auber, A., Schamberger, A., Nyikó, T., Taller, J., Orbán, T.I., Burgyán, J., and Silhavy, D.** (2018). The nonstop decay and the RNA silencing systems operate cooperatively in

- plants. *Nucleic Acids Res.* **46**: 4632–4648.
- 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.
- Tharun, S., He, W., Mayes, A.E., Lennertz, P., Beggs, J.D., and Parker, R.** (2000). Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* **404**: 515–518.
- Tharun, S. and Parker, R.** (2001). Targeting an mRNA for decapping: Displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. *Mol. Cell* **8**: 1075–1083.
- Thran, M., Link, K., and Sonnewald, U.** (2012). The Arabidopsis *DCP2* gene is required for proper mRNA turnover and prevents transgene silencing in Arabidopsis. *Plant J.* **72**: 368–377.
- Tomecki, R. and Dziembowski, A.** (2010). Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism. *RNA* **16**: 1692–1724.
- Tritschler, F., Braun, J.E., Motz, C., Igreja, C., Haas, G., Truffault, V., Izaurrealde, E., and Weichenrieder, O.** (2009). DCP1 forms asymmetric trimers to assemble into active mRNA decapping complexes in metazoa. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 21591–6.
- Trotman, J.B. and Schoenberg, D.R.** (2019). A recap of RNA recapping. *Wiley Interdiscip. Rev. RNA* **10**.
- Tsai, H.Y., Chen, C.C.G., Conte, D., Moresco, J.J., Chaves, D.A., Mitani, S., Yates, J.R., Tsai, M.D., and Mello, C.C.** (2015). A ribonuclease coordinates siRNA amplification and mRNA Cleavage during NAI. *Cell* **160**: 407–419.
- Tseng, C.-K., Wang, H.-F., Schroeder, M.R., and Baumann, P.** (2018). The H/ACA complex disrupts triplex in hTR precursor to permit processing by RRP6 and PARN. *Nat. Commun.* **9**: 5430.
- Tsuzuki, M., Motomura, K., Kumakura, N., and Takeda, A.** (2017). Interconnections between mRNA degradation and RDR-dependent siRNA production in mRNA turnover in plants. *J. Plant Res.* **130**: 211–226.
- Tu, B. et al.** (2015). Distinct and Cooperative Activities of HESO1 and URT1 Nucleotidyl Transferases in MicroRNA Turnover in Arabidopsis. *PLOS Genet.* **11**: e1005119.
- Tuck, A.C., Rankova, A., Arpat, A.B., Liechti, L.A., Hess, D., Iesmantavicius, V., Castelo-Szekely, V., Gatfield, D., and Bühler, M.** (2020). Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. *Mol. Cell* **0**: 1–15.
- Ustianenko, D., Hrossova, D., Potesil, D., Chalupnikova, K., Hrazdilova, K., Pachernik, J., Cetkovska, K., Uldrijan, S., Zdrahal, Z., and Vanacova, A.** (2013). Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* **19**: 1632–1638.
- Ustianenko, D., Pasulka, J., Feketova, Z., Bednarik, L., Zigackova, D., Fortova, A., Zavolan, M., and Vanacova, S.** (2016). TUT-DIS 3 L 2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs.: 1–13.
- Vazquez, F., Gascioli, V., Crété, P., and Vaucheret, H.** (2004). The Nuclear dsRNA Binding Protein HYL1 Is Required for MicroRNA Accumulation and Plant Development, but Not Posttranscriptional Transgene Silencing. *Curr. Biol.* **14**: 346–351.



- 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.
- Vindry, C., Marnef, A., Broomhead, H., Twyffels, L., Ozgur, S., Stoecklin, G., Llorian, M., Smith, C.W., Mata, J., Weil, D., and Standart, N.** (2017). Dual RNA Processing Roles of Pat1b via Cytoplasmic Lsm1-7 and Nuclear Lsm2-8 Complexes. *Cell Rep.* **20**: 1187–1200.
- Vogel, F., Hofius, D., Paulus, K.E., Jungkunz, I., and Sonnewald, U.** (2011). The second face of a known player: Arabidopsis silencing suppressor AtXRN4 acts organ-specifically. *New Phytol.* **189**: 484–493.
- Walley, J.W., Kelley, D.R., Nestorova, G., Hirschberg, D.L., and Dehesh, K.** (2010). Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. *Plant Physiol.* **152**: 866–75.
- Walters, R.W., Matheny, T., Mizoue, L.S., Rao, B.S., Muhlrud, D., and Parker, R.** (2017). Identification of NAD<sup>+</sup> capped mRNAs in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **114**: 480–485.
- Wang, Z., Jiao, X., Carr-Schmid, A., and Kiledjian, M.** (2002). The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **99**: 12663–12668.
- Wawer, I., Golisz, A., Sulkowska, A., Kawa, D., Kulik, A., and Kufel, J.** (2018). mRNA Decapping and 5'-3' Decay Contribute to the Regulation of ABA Signaling in Arabidopsis thaliana. *Front. Plant Sci.* **9**: 312.
- Webster, M.W., Chen, Y.H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Collier, J., and Passmore, L.A.** (2018). mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Mol. Cell* **70**: 1089-1100.e8.
- Webster, M.W., Stowell, J.A.W., Tang, T.T.L., and Passmore, L.A.** (2017). Analysis of mRNA deadenylation by multi-protein complexes. *Methods* **126**: 95–104.
- Wolf, J. and Passmore, L.A.** (2014). mRNA deadenylation by Pan2-Pan3. *Biochem. Soc. Trans.* **42**: 184–187.
- Wolozin, B.** (2012). Regulated protein aggregation: Stress granules and neurodegeneration. *Mol. Neurodegener.* **7**.
- Wurm, J.P., Holdermann, I., Overbeck, J.H., Mayer, P.H.O., and Sprangers, R.** (2017). Changes in conformational equilibria regulate the activity of the Dcp2 decapping enzyme. *Proc. Natl. Acad. Sci.* **114**: 6034–6039.
- Wurm, J.P. and Sprangers, R.** (2019). Dcp2: an mRNA decapping enzyme that adopts many different shapes and forms. *Curr. Opin. Struct. Biol.* **59**: 115–123.
- Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S.A., and Carrington, J.C.** (2005). Expression of Arabidopsis MIRNA genes. *Plant Physiol.* **138**: 2145–2154.
- Xin, X.-F. and He, S.Y.** (2013). *Pseudomonas syringae* pv. tomato DC3000: A Model Pathogen for Probing Disease Susceptibility and Hormone Signaling in Plants. *Annu. Rev. Phytopathol.* **51**: 473–498.
- 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**: 3270–9.
- Xu, J. and Chua, N.-H.** (2012). Dehydration stress activates Arabidopsis MPK6 to signal DCP1

- phosphorylation. *EMBO J.* **31**: 1975–84.
- Xu, J., Peng, W., Sun, Y., Wang, X., Xu, Y., Li, X., Gao, G., and Rao, Z.** (2012). Structural study of MCPIP1 N-terminal conserved domain reveals a PIN-like RNase. *40*: 6957–6965.
- Xu, J., Yang, J.-Y., Niu, Q.-W., and Chua, N.-H.** (2006). Arabidopsis DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. *Plant Cell* **18**: 3386–98.
- Yamashita, A.** (2013). Role of SMG-1-mediated Upf1 phosphorylation in mammalian nonsense-mediated mRNA decay. *Genes to Cells* **18**: 161–175.
- Yamashita, A., Chang, T.-C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.-Y.A., and Shyu, A.-B.** (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat. Struct. Mol. Biol.* **12**: 1054–63.
- Yao, Q. et al.** (2018). Ribonuclease activity of MARF1 controls oocyte RNA homeostasis and genome integrity in mice. *Proc. Natl. Acad. Sci. U. S. A.* **115**: 11250–11255.
- Yi, H., Park, J., Ha, M., Lim, J., Chang, H., and Kim, V.N.** (2018). PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. *Mol. Cell* **70**: 1081-1088.e5.
- Yokogawa, M., Tsushima, T., Noda, N.N., and Kumeta, H.** (2016). Structural basis for the regulation of enzymatic activity of Regnase-1 by domain-domain interactions. *Nat. Publ. Gr.*: 1–10.
- Youn, J.Y., Dyakov, B.J.A., Zhang, J., Knight, J.D.R., Vernon, R.M., Forman-Kay, J.D., and Gingras, A.C.** (2019). Properties of Stress Granule and P-Body Proteomes. *Mol. Cell.*
- Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X.** (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science (80- )*. **307**: 932–935.
- Yu, X., Li, B., Jang, G.J., Jiang, S., Jiang, D., Jang, J.C., Wu, S.H., Shan, L., and He, P.** (2019). Orchestration of Processing Body Dynamics and mRNA Decay in Arabidopsis Immunity. *Cell Rep.* **28**: 2194-2205.e6.
- Yu, X., Willmann, M.R., Anderson, S.J., and Gregory, B.D.** (2016). Genome-Wide Mapping of Uncapped and Cleaved Transcripts Reveals a Role for the Nuclear mRNA Cap-Binding Complex in Co-translational RNA Decay in Arabidopsis. *Plant Cell: tpc.00456.2016.*
- Zeidan, O., He, F., Zhang, F., Zhang, H., Jacobson, A., and Hinnebusch, A.G.** (2018). Conserved mRNA-granule component Scd6 targets Dhh1 to repress translation initiation and activates Dcp2-mediated mRNA decay in vivo. *PLoS Genet.* **14**.
- Zhai, J. et al.** (2011). MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* **25**: 2540–2553.
- Zhang, H., Xia, R., and Walbot, V.** (2015a). Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Curr. Opin. Plant Biol.* **27**: 84–90.
- Zhang, H., Zhong, H., Zhang, S., Shao, X., Ni, M., Cai, Z., Chen, X., and Xia, Y.** (2019). NAD tagSeq reveals that NAD<sup>+</sup>-capped RNAs are mostly produced from a large number of protein-coding genes in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **116**: 12072–12077.
- Zhang, W., Murphy, C., and Sieburth, L.E.** (2010). Conserved RNaseIII domain protein functions in

- cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 15981–15985.
- Zhang, X., Zhu, Y., Liu, X., Hong, X., Xu, Y., Zhu, P., Shen, Y., Wu, H., Ji, Y., Wen, X., Zhong, C., Zhao, Q., Wong, Y., Lu, J., Guo, H.** (2015a). Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. *Science* (80-. ). **348**: 120–123.
- Zhang, X., Devany, E., Murphy, M.R., Glazman, G., Persaud, M., and Kleiman, F.E.** (2015b). PARN deadenylase is involved in miRNA-dependent degradation of TP53 mRNA in mammalian cells. *Nucleic Acids Res.* **43**: 10925–38.
- Zhang, X. and Guo, H.** (2017). mRNA decay in plants: both quantity and quality matter. *Curr. Opin. Plant Biol.* **35**: 138–144.
- Zhang, Z., Hu, F., Sung, M.W., Shu, C., Castillo-González, C., Koiwa, H., Tang, G., Dickman, M., Li, P., and Zhang, X.** (2017). RISC-interacting clearing 3'- 5' exoribonucleases (RICES) degrade uridylated cleavage fragments to maintain functional RISC in Arabidopsis thaliana. *Elife* **6**.
- Zhu, L., Kandasamy, S.K., Liao, S.E., and Fukunaga, R.** (2018). LOTUS domain protein MARF1 binds CCR4-NOT deadenylase complex to post-transcriptionally regulate gene expression in oocytes. *Nat. Commun.* **9**: 4031.
- Zigáčková, D. and Vaňáčková, Š.** (2018). The role of 3' end uridylation in RNA metabolism and cellular physiology. *Philos. Trans. R. Soc. B Biol. Sci.* **373**.
- Zinder, J.C. and Lima, C.D.** (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes Dev.* **31**: 88–100.
- Zuber, H., Scheer, H., Ferrier, E., Sement, F.M., Mercier, P., Stupfler, B., and Gagliardi, D.** (2016). Uridylation and PABP Cooperate to Repair mRNA Deadened Ends in Arabidopsis. *Cell Rep.* **14**: 2707–2717.

# ANNEXES

## Annexe 1 Liste des protéines enrichies dans l'IP GFP-DAE

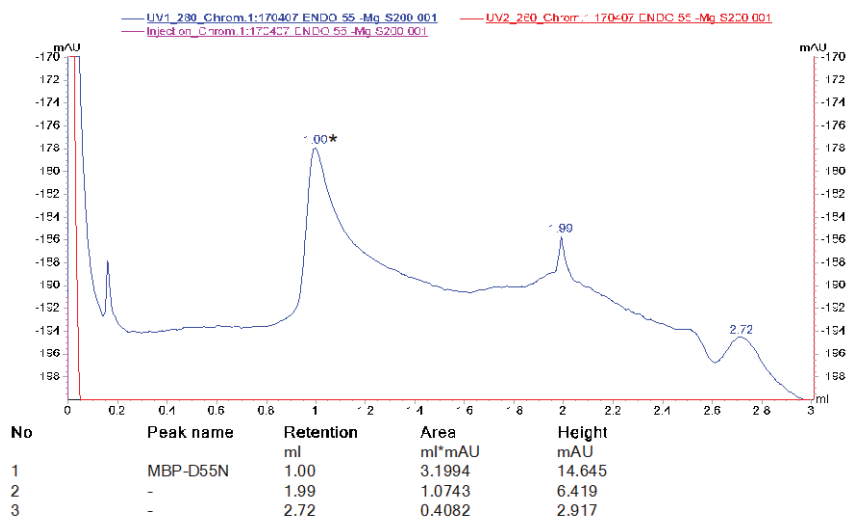
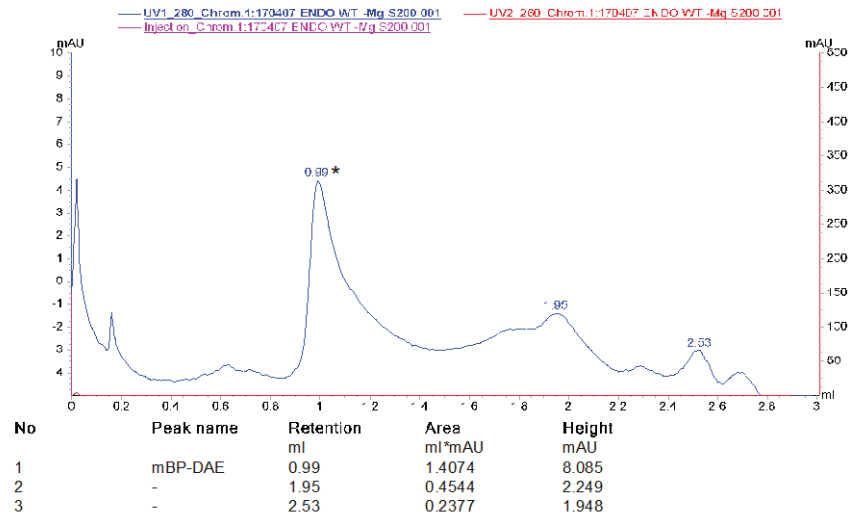
Le tableau contient les protéines avec un LogFC > 1 et avec une adjp (adjusted p-value) < 0,05.

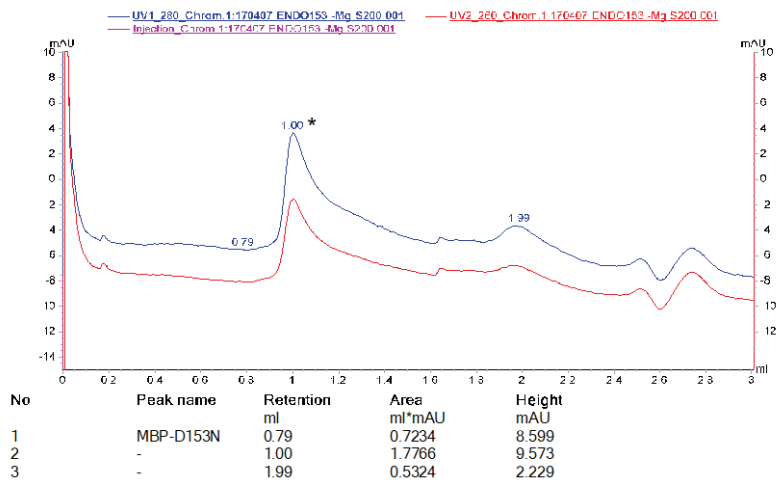
| AGI         | LogFC       | p.value     | adjp        | Nome du gène | Description  | Moyenne du nombre de spectres |       |
|-------------|-------------|-------------|-------------|--------------|--|-------------------------------|-------|
|             |             |             |             |              |  | Colo                          | DAE   |
| AT2G15560.1 | 8,609134361 | 4,84801E-54 | 4,62501E-51 | DAE          | At2g15560<br>[Source:UniProtKB/TrEMBL;Acc:Q6NQ97]  | 0,3                           | 453,4 |
| AT5G02500.1 | 3,255668077 | 1,85031E-23 | 8,82597E-21 | MED37E       | Probable mediator of RNA polymerase II transcription subunit 37e<br>[Source:UniProtKB/Swiss-Prot;Acc:P22953] | 14,1                          | 346,0 |
| AT3G09440.1 | 2,947184435 | 6,56862E-23 | 2,08882E-20 | HSP70-3      | Heat shock protein 70 (Hsp 70) family protein<br>[Source:UniProtKB/TrEMBL;Acc:A0A178V176]                    | 10,8                          | 226,2 |
| AT1G08370.1 | 4,960089907 | 5,5284E-22  | 1,05482E-19 | DCP1         | mRNA-decapping enzyme-like protein<br>[Source:UniProtKB/Swiss-Prot;Acc:Q95JF3]                               | 0,3                           | 38,2  |
| AT5G02490.1 | 2,990410543 | 1,5924E-19  | 2,17022E-17 | MED37D       | Probable mediator of RNA polymerase II transcription subunit 37c<br>[Source:UniProtKB/Swiss-Prot;Acc:P22954] | 10,4                          | 238,0 |
| AT3G12580.1 | 2,616709408 | 3,3949E-17  | 4,04842E-15 | MED37C       | Probable mediator of RNA polymerase II transcription subunit 37c<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LHA8] | 9,9                           | 173,0 |
| AT3G13300.1 | 4,254968565 | 6,43941E-14 | 6,1432E-12  | VCS          | Enhancer of mRNA-decapping protein 4<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LTT8]                             | 0,3                           | 22,8  |
| AT3G62250.1 | 3,114658538 | 2,32722E-09 | 1,00917E-07 | RPS27AC      | Ubiquitin-40S ribosomal protein S27a-3<br>[Source:UniProtKB/Swiss-Prot;Acc:P59233]                           | 1,2                           | 16,6  |
| AT1G16030.1 | 3,447250962 | 1,20269E-07 | 3,37461E-06 | HSP70-5      | Hsp70b<br>[Source:UniProtKB/TrEMBL;Acc:A0A178W9N7]   | 0,9                           | 47,1  |
| AT5G62690.1 | 2,767242496 | 4,51198E-07 | 1,13274E-05 | TUB2         | encodes tubulin beta-2/beta-3 chain The mRNA is cell-to-cell mobile(source:TAIR)                             | 0,6                           | 17,6  |
| AT2G37620.1 | 2,49271366  | 2,48291E-05 | 0,000426677 | ACT1         | Member of the actin gene family. Expressed in mature pollen.   | 0,4                           | 11,4  |
| AT3G49490.1 | 4,501957354 | 4,64759E-05 | 0,000729638 |              | Uncharacterized protein TgC5.90<br>[Source:UniProtKB/TrEMBL;Acc:Q9SCK9]                                      | 0,0                           | 4,4   |
| AT4G24780.1 | 4,394954717 | 4,74188E-05 | 0,000729638 |              | Probable pectate lyase 18<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9C5M8]  | 0,0                           | 2,8   |
| AT1G04820.1 | 2,487107605 | 0,000198211 | 0,002455761 | TOR2         | Encodes an alpha tubulin isoform that is expressed in roots, leaves and flowers.                             | 0,3                           | 8,6   |
| AT3G13290.1 | 2,19803985  | 0,000208476 | 0,002517546 | VCR          | Varicose-related protein [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT9]  | 0,4                           | 6,9   |
| AT3G20820.1 | 1,246037227 | 0,000271069 | 0,003153657 |              | Leucine-rich repeat (LRR) family protein<br>[Source:UniProtKB/TrEMBL;Acc:Q9LT39]                             | 3,3                           | 25,8  |
| AT3G07010.1 | 3,039086636 | 0,000343089 | 0,003850674 |              | Probable pectate lyase 8<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9M8Z8]   | 0,1                           | 4,3   |
| AT2G32080.1 | 3,894238428 | 0,000407916 | 0,004411075 | PUR1         | PUR ALPHA-1<br>[Source:UniProtKB/TrEMBL;Acc:A0A178VXA3]  | 0,0                           | 2,8   |
| AT2G34160.1 | 3,70748881  | 0,001012143 | 0,009374602 |              | Uncharacterized protein At2g34160<br>[Source:UniProtKB/Swiss-Prot;Acc:O22969]                                | 0,0                           | 1,9   |
| AT4G27060.1 | 2,36151503  | 0,001436589 | 0,012689869 | TOR1         | Microtubule-associated protein TORTIFOLIA1<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9To41]                       | 0,8                           | 16,9  |
| AT4G14960.2 | 2,225794047 | 0,001474385 | 0,012786935 | TUBA6        | Tubulin alpha chain<br>[Source:UniProtKB/TrEMBL;Acc:A0A178UXP7]  | 0,3                           | 3,7   |
| AT1G29910.1 | 2,529047195 | 0,002939642 | 0,022616274 | AB180, CAB3  | member of Chlorophyll a/b-binding protein family   | 0,2                           | 6,1   |
| AT4G32020.1 | 3,469093009 | 0,003094078 | 0,023018059 |              | AT4g32020/F10N7_170<br>[Source:UniProtKB/TrEMBL;Acc:O49389]  | 0,0                           | 1,1   |
| AT3G22310.1 | 3,425722701 | 0,00304624  | 0,023018059 | RH9          | DEAD-box ATP-dependent RNA helicase 9, mitochondrial<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LUW6]             | 0,0                           | 1,6   |
| AT3G22330.1 | 2,184861914 | 0,003056272 | 0,023018059 | RH53         | DEAD-box ATP-dependent RNA helicase 53,<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LUW6]                          | 0,3                           | 4,8   |

|             |             |             |             |                 |  |     |      |
|-------------|-------------|-------------|-------------|-----------------|--|-----|------|
|             |             |             |             |                 | mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9LUW5]   |     |      |
| AT5G64260.1 | 1,917617472 | 0,003456427 | 0,024980543 | EXL2            | Protein EXORDIUM-like 2<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9FE06]  | 0,4 | 5,4  |
| AT3G06980.1 | 3,373657848 | 0,00442861  | 0,03152906  | RH50            | DEAD-box ATP-dependent RNA helicase 50<br>[Source:UniProtKB/Swiss-Prot;Acc:Q8GUG7]   | 0,0 | 1,0  |
| AT5G48650.1 | 3,335219385 | 0,004592379 | 0,032452809 |                 | Nuclear transport factor 2 (NTF2) family protein with RNA binding (RRM-RBD-RNP motifs) domain-containing protein<br>[Source:UniProtKB/TrEMBL;Acc:Q84JH2] | 0,0 | 1,4  |
| AT5G19770.1 | 2,061207127 | 0,005199363 | 0,035943422 | TUBULIN ALPHA-3 | tubulin 3 (source:TAIR)  | 0,3 | 7,4  |
| AT2G17870.1 | 3,311155884 | 0,005254809 | 0,036065383 | CSP3            | Cold shock domain-containing protein 3<br>[Source:UniProtKB/Swiss-Prot;Acc:Q94C69]   | 0,0 | 1,3  |
| AT3G58510.1 | 1,61913196  | 0,005443772 | 0,037095419 | RH11            | DEAD-box ATP-dependent RNA helicase 11<br>[Source:UniProtKB/Swiss-Prot;Acc:Q8LA13]   | 1,6 | 10,1 |
| AT1G02080.1 | 3,264308018 | 0,006034074 | 0,040538781 |                 | Transcription regulator<br>[Source:UniProtKB/TrEMBL;Acc:F4HVV6]  | 0,0 | 1,0  |
| AT2G27030.3 | 3,170478674 | 0,006807418 | 0,044178752 | CAM5            | encodes a calmodulin that has higher affinity to kinesin-like calmodulin binding motor protein than CAM4 or CAM6. The mRNA is cell-to-cell mobile.       | 0,0 | 2,2  |
| AT5G57110.1 | 3,205371094 | 0,007220093 | 0,04592384  | ACA8            | Calcium-transporting ATPase<br>[Source:UniProtKB/TrEMBL;Acc:QoWV19]  | 0,0 | 1,1  |
| AT3G53420.1 | 2,273030158 | 0,00755443  | 0,047727986 | PIP2-1          | PIP2A<br>[Source:UniProtKB/TrEMBL;Acc:AoA178V8P4]  | 0,2 | 3,4  |
| AT1G72610.1 | 3,257028783 | 0,007862008 | 0,04902193  | GLP1            | Germin-like protein subfamily 3 member 1<br>[Source:UniProtKB/Swiss-Prot;Acc:P94040]   | 0,0 | 1,3  |

**Annexe 2** Analyse par gel filtration de l'état d'agrégation de MBP-DAE, MBP-D55N et MBP-D153N

Colonne : Superdex200 5/150 GE  
 Buffer : 20mM MOPS pH 7,2 ; 150mM KCl ; 15% glycérol (v/v) ; 0,1% Tween 20 (v/v)  
 50ul injectés  
 Flux 0,1ml/min







**Annexe 3** Analyse par spectrométrie de masse des éluâtes contenant les extraits protéiques MBP-DAE, MBP-D55N et MBP-D153N.

| Accession                | Description  | Spectral count |
|--------------------------|--|----------------|
| <a href="#">MBP-D55N</a> |  | 1195           |
| MALE_ECO57               | Maltose-binding periplasmic protein OS=Escherichia coli O157:H7 GN=malE PE=1 SV=1                                | 103            |
| DNAK_ECO24               | Chaperone protein DnaK OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=dnaK PE=3 SV=1                    | 11             |
| CH601_ECOK1              | 60 kDa chaperonin 1 OS=Escherichia coli O1:K1 / APEC GN=groL1 PE=3 SV=1  | 9              |
| EFTU1_ECO24              | Elongation factor Tu 1 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1                    | 7              |
| DEAD_ECO57               | ATP-dependent RNA helicase DeaD OS=Escherichia coli O157:H7 GN=deaD PE=3 SV=3                                    | 7              |
| RPOC_ECO24               | DNA-directed RNA polymerase subunit beta' OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpoC PE=1 SV=1 | 5              |
| ATPB_ECO24               | ATP synthase subunit beta OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=atpD PE=3 SV=1                 | 5              |
| OMP_FCOLI                | Outer membrane protein F OS=Escherichia coli (strain K12) GN=ompF PE=1 SV=1                                      | 4              |
| NUOCD_ECOL5              | NADH-quinone oxidoreductase subunit C/D OS=Escherichia coli O6:K15:H31 (strain 536 / UPEC) GN=nuoC PE=3 SV=1     | 4              |
| DNAJ_ECO24               | Chaperone protein DnaJ OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=dnaJ PE=3 SV=1                    | 4              |
| DEGP_ECO57               | Periplasmic serine endoprotease DegP OS=Escherichia coli O157:H7 GN=degP PE=3 SV=1                               | 3              |
| FTSH_ECO57               | ATP-dependent zinc metalloprotease FtsH OS=Escherichia coli O157:H7 GN=ftsH PE=3 SV=1                            | 3              |
| DHG_ECOLI                | Quinoprotein glucose dehydrogenase OS=Escherichia coli (strain K12) GN=gcd PE=1 SV=3                             | 3              |
| HFLC_ECO57               | Modulator of FtsH protease HflC OS=Escherichia coli O157:H7 GN=hflC PE=3 SV=1                                    | 3              |
| ODP1_ECO57               | Pyruvate dehydrogenase E1 component OS=Escherichia coli O157:H7 GN=aceE PE=1 SV=2                                | 3              |
| NUOG_ECO57               | NADH-quinone oxidoreductase subunit G OS=Escherichia coli O157:H7 GN=nuoG PE=3 SV=3                              | 3              |
| FEPA_ECOLI               | Ferrienterobactin receptor OS=Escherichia coli (strain K12) GN=fepA PE=1 SV=2                                    | 3              |
| ODO1_ECO57               | 2-oxoglutarate dehydrogenase E1 component OS=Escherichia coli O157:H7 GN=sucA PE=3 SV=1                          | 3              |
| SDHA_ECO57               | Succinate dehydrogenase flavoprotein subunit OS=Escherichia coli O157:H7 GN=sdhA PE=3 SV=1                       | 3              |
| SLYB_ECO57               | Outer membrane lipoprotein SlyB OS=Escherichia coli O157:H7 GN=slyB PE=3 SV=1                                    | 3              |
| RL5_ECO24                | 50S ribosomal protein L5 OS=Escherichia coli O139:H28  | 3              |

|            |   |   |
|------------|---|---|
|            | (strain E24377A / ETEC) GN=rplE PE=3 SV=1   |   |
| RODZ_ECO24 | Cytoskeleton protein RodZ OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rodZ PE=3 SV=1                           | 3 |
| CAN_ECOLI  | Carbonic anhydrase 2 OS=Escherichia coli (strain K12)<br>GN=can PE=1 SV=1   | 3 |
| ACRA_ECO57 | Multidrug efflux pump subunit AcrA OS=Escherichia coli<br>O157:H7 GN=acrA PE=3 SV=1   | 2 |
| SDAC_ECO57 | Serine transporter OS=Escherichia coli O157:H7 GN=sdaC<br>PE=3 SV=1   | 2 |
| ATZN_ECOLI | Lead, cadmium, zinc and mercury-transporting ATPase<br>OS=Escherichia coli (strain K12) GN=zntA PE=1 SV=1                     | 2 |
| BLUF_ECOLI | Blue light- and temperature-regulated antirepressor BluF<br>OS=Escherichia coli (strain K12) GN=bluF PE=1 SV=1                | 2 |
| YGIB_ECO24 | UPFo441 protein YgiB OS=Escherichia coli O139:H28 (strain<br>E24377A / ETEC) GN=ygiB PE=3 SV=1                                | 2 |
| DLDH_ECO57 | Dihydrolipoyl dehydrogenase OS=Escherichia coli O157:H7<br>GN=lpdA PE=3 SV=2  | 2 |
| RS10_ECO24 | 30S ribosomal protein S10 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsJ PE=3 SV=1                           | 2 |
| ATPA_ECO24 | ATP synthase subunit alpha OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=atpA PE=3 SV=1                          | 2 |
| RS3_ECO24  | 30S ribosomal protein S3 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsC PE=3 SV=1                            | 2 |
| LPP_ECO57  | Major outer membrane lipoprotein OS=Escherichia coli<br>O157:H7 GN=lpp PE=3 SV=1  | 2 |
| MENB_ECOL6 | 1,4-dihydroxy-2-naphthoyl-CoA synthase OS=Escherichia<br>coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=menB<br>PE=3 SV=1 | 2 |
| RL4_ECO24  | 50S ribosomal protein L4 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rplD PE=3 SV=1                            | 2 |
| KPRS_ECO57 | Ribose-phosphate pyrophosphokinase OS=Escherichia coli<br>O157:H7 GN=prs PE=3 SV=2  | 2 |
| PTND_ECOLI | Mannose permease IID component OS=Escherichia coli<br>(strain K12) GN=manZ PE=1 SV=2  | 2 |
| RS5_ECO24  | 30S ribosomal protein S5 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsE PE=3 SV=1                            | 2 |
| SECF_ECO57 | Protein translocase subunit SecF OS=Escherichia coli<br>O157:H7 GN=secF PE=3 SV=1   | 2 |
| YIAF_ECO57 | Uncharacterized protein YiaF OS=Escherichia coli O157:H7<br>GN=yiaF PE=4 SV=2   | 2 |
| YHCB_ECO57 | Putative cytochrome d ubiquinol oxidase subunit 3<br>OS=Escherichia coli O157:H7 GN=yhcB PE=3 SV=2                            | 2 |
| RPOA_ECO24 | DNA-directed RNA polymerase subunit alpha<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=rpoA PE=1 SV=1        | 2 |
| METH_ECOLI | Methionine synthase OS=Escherichia coli (strain K12)<br>GN=metH PE=1 SV=5   | 2 |
| HFLK_ECO57 | Protein HflK OS=Escherichia coli O157:H7 GN=hflK PE=3<br>SV=1   | 2 |

|             |  |   |
|-------------|--|---|
| CLPP_ECO24  | ATP-dependent Clp protease proteolytic subunit<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=clpP PE=3 SV=1      | 2 |
| RS12_ECO24  | 30S ribosomal protein S12 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsL PE=3 SV=1                              | 2 |
| BAMA_ECO24  | Outer membrane protein assembly factor BamA<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=bamA PE=3 SV=1         | 1 |
| LPTD_ECO57  | LPS-assembly protein LptD OS=Escherichia coli O157:H7<br>GN=lptD PE=3 SV=1   | 1 |
| GLRX4_ECO57 | Glutaredoxin-4 OS=Escherichia coli O157:H7 GN=grxD PE=3<br>SV=1  | 1 |
| ATMA_ECO57  | Magnesium-transporting ATPase, P-type 1 OS=Escherichia<br>coli O157:H7 GN=mgtA PE=3 SV=1   | 1 |
| ALAA_ECOL6  | Glutamate-pyruvate aminotransferase AlaA OS=Escherichia<br>coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=alaA<br>PE=3 SV=1  | 1 |
| ROXA_ECOLI  | 50S ribosomal protein L16 3-hydroxylase OS=Escherichia coli<br>(strain K12) GN=roxA PE=1 SV=2                                    | 1 |
| EFG_ECO24   | Elongation factor G OS=Escherichia coli O139:H28 (strain<br>E24377A / ETEC) GN=fusA PE=3 SV=1                                    | 1 |
| YDIH_ECO57  | Uncharacterized protein YdiH OS=Escherichia coli O157:H7<br>GN=ydiH PE=4 SV=2  | 1 |
| OMPA_ECO57  | Outer membrane protein A OS=Escherichia coli O157:H7<br>GN=ompA PE=3 SV=1  | 1 |
| CLPX_ECO24  | ATP-dependent Clp protease ATP-binding subunit ClpX<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=clpX PE=3 SV=1 | 1 |
| TIG_ECO24   | Trigger factor OS=Escherichia coli O139:H28 (strain E24377A<br>/ ETEC) GN=tig PE=3 SV=1  | 1 |
| ADIA_ECOLI  | Biodegradative arginine decarboxylase OS=Escherichia coli<br>(strain K12) GN=adiA PE=1 SV=1                                      | 1 |
| RS11_ECO24  | 30S ribosomal protein S11 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsK PE=3 SV=1                              | 1 |
| ZIPA_ECO24  | Cell division protein ZipA OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=zipA PE=3 SV=1                             | 1 |
| RL1_ECO24   | 50S ribosomal protein L1 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rplA PE=3 SV=1                               | 1 |
| YCJF_ECO24  | UPF0283 membrane protein YcjF OS=Escherichia coli<br>O139:H28 (strain E24377A / ETEC) GN=ycjF PE=3 SV=1                          | 1 |
| TYPA_ECO27  | GTP-binding protein TypA/BipA OS=Escherichia coli<br>O127:H6 (strain E2348/69 / EPEC) GN=typA PE=3 SV=1                          | 1 |
| GATZ_ECO24  | D-tagatose-1,6-bisphosphate aldolase subunit GatZ<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=gatZ PE=3 SV=1   | 1 |
| LACI_ECOLI  | Lactose operon repressor OS=Escherichia coli (strain K12)<br>GN=lacl PE=1 SV=3   | 1 |
| FDOG_ECOLI  | Formate dehydrogenase-O major subunit OS=Escherichia<br>coli (strain K12) GN=fdoG PE=1 SV=5                                      | 1 |

|             |  |   |
|-------------|--|---|
| BLAT_ECOLX  | Beta-lactamase TEM OS=Escherichia coli GN=bla PE=1 SV=1  | 1 |
| METK_ECO24  | S-adenosylmethionine synthase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=metK PE=3 SV=1                   | 1 |
| BAMB_ECOLI  | Outer membrane protein assembly factor BamB OS=Escherichia coli (strain K12) GN=bamB PE=1 SV=1                         | 1 |
| LAST_ECOLI  | Uncharacterized tRNA/rRNA methyltransferase LasT OS=Escherichia coli (strain K12) GN=lasT PE=3 SV=2                    | 1 |
| DNAB_ECO57  | Replicative DNA helicase OS=Escherichia coli O157:H7 GN=dnaB PE=3 SV=1   | 1 |
| RS7_ECO24   | 30S ribosomal protein S7 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsG PE=3 SV=1                        | 1 |
| PPID_ECOL6  | Peptidyl-prolyl cis-trans isomerase D OS=Escherichia coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=ppiD PE=3 SV=1 | 1 |
| BGAL_ECO24  | Beta-galactosidase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=lacZ PE=3 SV=1                              | 1 |
| ADHE_ECO57  | Aldehyde-alcohol dehydrogenase OS=Escherichia coli O157:H7 GN=adhE PE=3 SV=2   | 1 |
| MUKB_ECO24  | Chromosome partition protein MukB OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=mukB PE=3 SV=1               | 1 |
| CLPB_ECO57  | Chaperone protein ClpB OS=Escherichia coli O157:H7 GN=clpB PE=3 SV=1   | 1 |
| RL17_ECO24  | 50S ribosomal protein L17 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplQ PE=3 SV=1                       | 1 |
| RS4_ECO24   | 30S ribosomal protein S4 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsD PE=3 SV=1                        | 1 |
| WZZB3_ECOLX | Chain length determinant protein OS=Escherichia coli GN=wzzB PE=3 SV=1   | 1 |
| ACNB_ECOLI  | Aconitate hydratase B OS=Escherichia coli (strain K12) GN=acnB PE=1 SV=3   | 1 |
| HNS_ECO57   | DNA-binding protein H-NS OS=Escherichia coli O157:H7 GN=hns PE=3 SV=2  | 1 |
| USG_ECOLI   | USG-1 protein OS=Escherichia coli (strain K12) GN=usg PE=1 SV=1  | 1 |
| PYRG_ECO24  | CTP synthase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=pyrG PE=3 SV=1                                    | 1 |
| YBED_ECO24  | UPF0250 protein YbeD OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=ybeD PE=3 SV=1                            | 1 |
| RS6_ECO24   | 30S ribosomal protein S6 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsF PE=3 SV=1                        | 1 |
| YIAD_ECOLI  | Probable lipoprotein YiaD OS=Escherichia coli (strain K12) GN=yiaD PE=1 SV=2   | 1 |
| SECD_ECO57  | Protein translocase subunit SecD OS=Escherichia coli O157:H7 GN=secD PE=3 SV=1   | 1 |
| ACRB_ECOLI  | Multidrug efflux pump subunit AcrB OS=Escherichia coli (strain K12) GN=acrB PE=1 SV=1                                  | 1 |
| SLYD_ECO57  | FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS=Escherichia coli O157:H7 GN=slyD PE=3 SV=1                       | 1 |
| RS2_ECO24   | 30S ribosomal protein S2 OS=Escherichia coli O139:H28  | 1 |

|            |  |   |
|------------|--|---|
|            | (strain E24377A / ETEC) GN=rpsB PE=3 SV=1  |   |
| FLGK_ECOLI | Flagellar hook-associated protein 1 OS=Escherichia coli (strain K12) GN=flgK PE=1 SV=3   | 1 |
| RPOB_ECO24 | DNA-directed RNA polymerase subunit beta OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpoB PE=1 SV=1                              | 1 |
| MLAF_ECO57 | Probable phospholipid import ATP-binding protein MlaF OS=Escherichia coli O157:H7 GN=mlaF PE=3 SV=1  | 1 |
| DHNA_ECOLI | NADH dehydrogenase OS=Escherichia coli (strain K12) GN=ndh PE=1 SV=2   | 1 |
| RPOE_ECO57 | ECF RNA polymerase sigma-E factor OS=Escherichia coli O157:H7 GN=rpoE PE=3 SV=1  | 1 |
| LPTB_ECO57 | Lipopolysaccharide export system ATP-binding protein LptB OS=Escherichia coli O157:H7 GN=lptB PE=3 SV=2                                      | 1 |
| RL2_ECO24  | 50S ribosomal protein L2 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplB PE=3 SV=1  | 1 |
| NUOF_ECOLI | NADH-quinone oxidoreductase subunit F OS=Escherichia coli (strain K12) GN=nuoF PE=1 SV=3   | 1 |
| SOHB_ECOLI | Probable protease SohB OS=Escherichia coli (strain K12) GN=sohB PE=1 SV=1  | 1 |
| SYT_ECO24  | Threonine--tRNA ligase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=thrS PE=3 SV=1  | 1 |
| DBHA_ECO57 | DNA-binding protein HU-alpha OS=Escherichia coli O157:H7 GN=hupA PE=3 SV=1   | 1 |
| CLPA_ECO57 | ATP-dependent Clp protease ATP-binding subunit ClpA OS=Escherichia coli O157:H7 GN=clpA PE=3 SV=1  | 1 |
| ARNA_ECO24 | Bifunctional polymyxin resistance protein ArnA OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=arnA PE=3 SV=1                        | 1 |
| RL10_ECO24 | 50S ribosomal protein L10 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplJ PE=3 SV=1   | 1 |
| RL11_ECO24 | 50S ribosomal protein L11 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplK PE=3 SV=1   | 1 |
| CRA_ECO57  | Catabolite repressor/activator OS=Escherichia coli O157:H7 GN=cra PE=3 SV=1  | 1 |
| FTSZ_ECO57 | Cell division protein FtsZ OS=Escherichia coli O157:H7 GN=ftsZ PE=3 SV=1   | 1 |
| METQ_ECO57 | D-methionine-binding lipoprotein MetQ OS=Escherichia coli O157:H7 GN=metQ PE=3 SV=1  | 1 |
| ODP2_ECOLI | Dihydrolypoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex OS=Escherichia coli (strain K12) GN=aceF PE=1 SV=3 | 1 |
| FTSK_ECO57 | DNA translocase FtsK OS=Escherichia coli O157:H7 GN=ftsK PE=3 SV=1   | 1 |
| IF2_ECO24  | Translation initiation factor IF-2 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=infB PE=3 SV=1                                    | 1 |
| FTSN_ECOLI | Cell division protein FtsN OS=Escherichia coli (strain K12) GN=ftsN PE=1 SV=3  | 1 |
| HSLU_ECO24 | ATP-dependent protease ATPase subunit HslU OS=Escherichia coli O139:H28 (strain E24377A / ETEC)  | 1 |

|            |   |   |
|------------|---|---|
|            | GN=hsIU PE=3 SV=1   |   |
| NDPA_ECO24 | Nucleoid-associated protein YejK OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=yejK PE=3 SV=1 | 1 |

| Accession                | Description  | Spectral count |
|--------------------------|--|----------------|
| <a href="#">MBP-D55N</a> |  | 1195           |
| MALE_ECO57               | Maltose-binding periplasmic protein OS=Escherichia coli O157:H7 GN=malE PE=1 SV=1                                | 103            |
| DNAK_ECO24               | Chaperone protein DnaK OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=dnaK PE=3 SV=1                    | 11             |
| CH601_ECOK1              | 60 kDa chaperonin 1 OS=Escherichia coli O1:K1 / APEC GN=groL1 PE=3 SV=1  | 9              |
| EFTU1_ECO24              | Elongation factor Tu 1 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1                    | 7              |
| DEAD_ECO57               | ATP-dependent RNA helicase DeaD OS=Escherichia coli O157:H7 GN=deaD PE=3 SV=3                                    | 7              |
| RPOC_ECO24               | DNA-directed RNA polymerase subunit beta' OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpoC PE=1 SV=1 | 5              |
| ATPB_ECO24               | ATP synthase subunit beta OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=atpD PE=3 SV=1                 | 5              |
| OMP_F_ECOLI              | Outer membrane protein F OS=Escherichia coli (strain K12) GN=ompF PE=1 SV=1                                      | 4              |
| NUOCD_ECOL5              | NADH-quinone oxidoreductase subunit C/D OS=Escherichia coli O6:K15:H31 (strain 536 / UPEC) GN=nuoC PE=3 SV=1     | 4              |
| DNAJ_ECO24               | Chaperone protein DnaJ OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=dnaJ PE=3 SV=1                    | 4              |
| DEGP_ECO57               | Periplasmic serine endoprotease DegP OS=Escherichia coli O157:H7 GN=degP PE=3 SV=1                               | 3              |
| FTSH_ECO57               | ATP-dependent zinc metalloprotease FtsH OS=Escherichia coli O157:H7 GN=ftsH PE=3 SV=1                            | 3              |
| DHG_ECOLI                | Quinoprotein glucose dehydrogenase OS=Escherichia coli (strain K12) GN=gcd PE=1 SV=3                             | 3              |
| HFLC_ECO57               | Modulator of FtsH protease HflC OS=Escherichia coli O157:H7 GN=hflC PE=3 SV=1                                    | 3              |
| ODP1_ECO57               | Pyruvate dehydrogenase E1 component OS=Escherichia coli O157:H7 GN=aceE PE=1 SV=2                                | 3              |
| NUOG_ECO57               | NADH-quinone oxidoreductase subunit G OS=Escherichia coli O157:H7 GN=nuoG PE=3 SV=3                              | 3              |
| FEPA_ECOLI               | Ferrienterobactin receptor OS=Escherichia coli (strain K12) GN=fepA PE=1 SV=2                                    | 3              |
| ODO1_ECO57               | 2-oxoglutarate dehydrogenase E1 component OS=Escherichia coli O157:H7 GN=sucA PE=3 SV=1                          | 3              |
| SDHA_ECO57               | Succinate dehydrogenase flavoprotein subunit OS=Escherichia coli O157:H7 GN=sdhA PE=3 SV=1                       | 3              |
| SLYB_ECO57               | Outer membrane lipoprotein SlyB OS=Escherichia coli O157:H7 GN=slyB PE=3 SV=1                                    | 3              |
| RL5_ECO24                | 50S ribosomal protein L5 OS=Escherichia coli O139:H28  | 3              |

|            |   |   |
|------------|---|---|
|            | (strain E24377A / ETEC) GN=rplE PE=3 SV=1   |   |
| RODZ_ECO24 | Cytoskeleton protein RodZ OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rodZ PE=3 SV=1                           | 3 |
| CAN_ECOLI  | Carbonic anhydrase 2 OS=Escherichia coli (strain K12)<br>GN=can PE=1 SV=1   | 3 |
| ACRA_ECO57 | Multidrug efflux pump subunit AcrA OS=Escherichia coli<br>O157:H7 GN=acrA PE=3 SV=1   | 2 |
| SDAC_ECO57 | Serine transporter OS=Escherichia coli O157:H7 GN=sdaC<br>PE=3 SV=1   | 2 |
| ATZN_ECOLI | Lead, cadmium, zinc and mercury-transporting ATPase<br>OS=Escherichia coli (strain K12) GN=zntA PE=1 SV=1                     | 2 |
| BLUF_ECOLI | Blue light- and temperature-regulated antirepressor BluF<br>OS=Escherichia coli (strain K12) GN=bluF PE=1 SV=1                | 2 |
| YGIB_ECO24 | UPFo441 protein YgiB OS=Escherichia coli O139:H28 (strain<br>E24377A / ETEC) GN=ygiB PE=3 SV=1                                | 2 |
| DLDH_ECO57 | Dihydrolipoyl dehydrogenase OS=Escherichia coli O157:H7<br>GN=lpdA PE=3 SV=2  | 2 |
| RS10_ECO24 | 30S ribosomal protein S10 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsJ PE=3 SV=1                           | 2 |
| ATPA_ECO24 | ATP synthase subunit alpha OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=atpA PE=3 SV=1                          | 2 |
| RS3_ECO24  | 30S ribosomal protein S3 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsC PE=3 SV=1                            | 2 |
| LPP_ECO57  | Major outer membrane lipoprotein OS=Escherichia coli<br>O157:H7 GN=lpp PE=3 SV=1  | 2 |
| MENB_ECOL6 | 1,4-dihydroxy-2-naphthoyl-CoA synthase OS=Escherichia<br>coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=menB<br>PE=3 SV=1 | 2 |
| RL4_ECO24  | 50S ribosomal protein L4 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rplD PE=3 SV=1                            | 2 |
| KPRS_ECO57 | Ribose-phosphate pyrophosphokinase OS=Escherichia coli<br>O157:H7 GN=prs PE=3 SV=2  | 2 |
| PTND_ECOLI | Mannose permease IID component OS=Escherichia coli<br>(strain K12) GN=manZ PE=1 SV=2  | 2 |
| RS5_ECO24  | 30S ribosomal protein S5 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsE PE=3 SV=1                            | 2 |
| SECF_ECO57 | Protein translocase subunit SecF OS=Escherichia coli<br>O157:H7 GN=secF PE=3 SV=1   | 2 |
| YIAF_ECO57 | Uncharacterized protein YiaF OS=Escherichia coli O157:H7<br>GN=yiaF PE=4 SV=2   | 2 |
| YHCB_ECO57 | Putative cytochrome d ubiquinol oxidase subunit 3<br>OS=Escherichia coli O157:H7 GN=yhcB PE=3 SV=2                            | 2 |
| RPOA_ECO24 | DNA-directed RNA polymerase subunit alpha<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=rpoA PE=1 SV=1        | 2 |
| METH_ECOLI | Methionine synthase OS=Escherichia coli (strain K12)<br>GN=metH PE=1 SV=5   | 2 |
| HFLK_ECO57 | Protein HflK OS=Escherichia coli O157:H7 GN=hflK PE=3<br>SV=1   | 2 |

|             |  |   |
|-------------|--|---|
| CLPP_ECO24  | ATP-dependent Clp protease proteolytic subunit<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=clpP PE=3 SV=1      | 2 |
| RS12_ECO24  | 30S ribosomal protein S12 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsL PE=3 SV=1                              | 2 |
| BAMA_ECO24  | Outer membrane protein assembly factor BamA<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=bamA PE=3 SV=1         | 1 |
| LPTD_ECO57  | LPS-assembly protein LptD OS=Escherichia coli O157:H7<br>GN=lptD PE=3 SV=1   | 1 |
| GLRX4_ECO57 | Glutaredoxin-4 OS=Escherichia coli O157:H7 GN=grxD PE=3<br>SV=1  | 1 |
| ATMA_ECO57  | Magnesium-transporting ATPase, P-type 1 OS=Escherichia<br>coli O157:H7 GN=mgtA PE=3 SV=1   | 1 |
| ALAA_ECOL6  | Glutamate-pyruvate aminotransferase AlaA OS=Escherichia<br>coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=alaA<br>PE=3 SV=1  | 1 |
| ROXA_ECOLI  | 50S ribosomal protein L16 3-hydroxylase OS=Escherichia coli<br>(strain K12) GN=roxA PE=1 SV=2                                    | 1 |
| EFG_ECO24   | Elongation factor G OS=Escherichia coli O139:H28 (strain<br>E24377A / ETEC) GN=fusA PE=3 SV=1                                    | 1 |
| YDIH_ECO57  | Uncharacterized protein YdiH OS=Escherichia coli O157:H7<br>GN=ydiH PE=4 SV=2  | 1 |
| OMPA_ECO57  | Outer membrane protein A OS=Escherichia coli O157:H7<br>GN=ompA PE=3 SV=1  | 1 |
| CLPX_ECO24  | ATP-dependent Clp protease ATP-binding subunit ClpX<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=clpX PE=3 SV=1 | 1 |
| TIG_ECO24   | Trigger factor OS=Escherichia coli O139:H28 (strain E24377A<br>/ ETEC) GN=tig PE=3 SV=1  | 1 |
| ADIA_ECOLI  | Biodegradative arginine decarboxylase OS=Escherichia coli<br>(strain K12) GN=adiA PE=1 SV=1                                      | 1 |
| RS11_ECO24  | 30S ribosomal protein S11 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsK PE=3 SV=1                              | 1 |
| ZIPA_ECO24  | Cell division protein ZipA OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=zipA PE=3 SV=1                             | 1 |
| RL1_ECO24   | 50S ribosomal protein L1 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rplA PE=3 SV=1                               | 1 |
| YCJF_ECO24  | UPF0283 membrane protein YcjF OS=Escherichia coli<br>O139:H28 (strain E24377A / ETEC) GN=ycjF PE=3 SV=1                          | 1 |
| TYPA_ECO27  | GTP-binding protein TypA/BipA OS=Escherichia coli<br>O127:H6 (strain E2348/69 / EPEC) GN=typA PE=3 SV=1                          | 1 |
| GATZ_ECO24  | D-tagatose-1,6-bisphosphate aldolase subunit GatZ<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=gatZ PE=3 SV=1   | 1 |
| LACI_ECOLI  | Lactose operon repressor OS=Escherichia coli (strain K12)<br>GN=lacl PE=1 SV=3   | 1 |
| FDOG_ECOLI  | Formate dehydrogenase-O major subunit OS=Escherichia<br>coli (strain K12) GN=fdoG PE=1 SV=5                                      | 1 |



|             |  |   |
|-------------|--|---|
| BLAT_ECOLX  | Beta-lactamase TEM OS=Escherichia coli GN=bla PE=1 SV=1  | 1 |
| METK_ECO24  | S-adenosylmethionine synthase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=metK PE=3 SV=1                   | 1 |
| BAMB_ECOLI  | Outer membrane protein assembly factor BamB OS=Escherichia coli (strain K12) GN=bamB PE=1 SV=1                         | 1 |
| LAST_ECOLI  | Uncharacterized tRNA/rRNA methyltransferase LasT OS=Escherichia coli (strain K12) GN=lasT PE=3 SV=2                    | 1 |
| DNAB_ECO57  | Replicative DNA helicase OS=Escherichia coli O157:H7 GN=dnaB PE=3 SV=1   | 1 |
| RS7_ECO24   | 30S ribosomal protein S7 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsG PE=3 SV=1                        | 1 |
| PPID_ECOL6  | Peptidyl-prolyl cis-trans isomerase D OS=Escherichia coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=ppiD PE=3 SV=1 | 1 |
| BGAL_ECO24  | Beta-galactosidase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=lacZ PE=3 SV=1                              | 1 |
| ADHE_ECO57  | Aldehyde-alcohol dehydrogenase OS=Escherichia coli O157:H7 GN=adhE PE=3 SV=2   | 1 |
| MUKB_ECO24  | Chromosome partition protein MukB OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=mukB PE=3 SV=1               | 1 |
| CLPB_ECO57  | Chaperone protein ClpB OS=Escherichia coli O157:H7 GN=clpB PE=3 SV=1   | 1 |
| RL17_ECO24  | 50S ribosomal protein L17 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplQ PE=3 SV=1                       | 1 |
| RS4_ECO24   | 30S ribosomal protein S4 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsD PE=3 SV=1                        | 1 |
| WZZB3_ECOLX | Chain length determinant protein OS=Escherichia coli GN=wzzB PE=3 SV=1   | 1 |
| ACNB_ECOLI  | Aconitate hydratase B OS=Escherichia coli (strain K12) GN=acnB PE=1 SV=3   | 1 |
| HNS_ECO57   | DNA-binding protein H-NS OS=Escherichia coli O157:H7 GN=hns PE=3 SV=2  | 1 |
| USG_ECOLI   | USG-1 protein OS=Escherichia coli (strain K12) GN=usg PE=1 SV=1  | 1 |
| PYRG_ECO24  | CTP synthase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=pyrG PE=3 SV=1                                    | 1 |
| YBED_ECO24  | UPF0250 protein YbeD OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=ybeD PE=3 SV=1                            | 1 |
| RS6_ECO24   | 30S ribosomal protein S6 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsF PE=3 SV=1                        | 1 |
| YIAD_ECOLI  | Probable lipoprotein YiaD OS=Escherichia coli (strain K12) GN=yiaD PE=1 SV=2   | 1 |
| SECD_ECO57  | Protein translocase subunit SecD OS=Escherichia coli O157:H7 GN=secD PE=3 SV=1   | 1 |
| ACRB_ECOLI  | Multidrug efflux pump subunit AcrB OS=Escherichia coli (strain K12) GN=acrB PE=1 SV=1                                  | 1 |
| SLYD_ECO57  | FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS=Escherichia coli O157:H7 GN=slyD PE=3 SV=1                       | 1 |
| RS2_ECO24   | 30S ribosomal protein S2 OS=Escherichia coli O139:H28  | 1 |

|            |  |   |
|------------|--|---|
|            | (strain E24377A / ETEC) GN=rpsB PE=3 SV=1  |   |
| FLGK_ECOLI | Flagellar hook-associated protein 1 OS=Escherichia coli (strain K12) GN=flgK PE=1 SV=3   | 1 |
| RPOB_ECO24 | DNA-directed RNA polymerase subunit beta OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpoB PE=1 SV=1                              | 1 |
| MLAF_ECO57 | Probable phospholipid import ATP-binding protein MlaF OS=Escherichia coli O157:H7 GN=mlaF PE=3 SV=1  | 1 |
| DHNA_ECOLI | NADH dehydrogenase OS=Escherichia coli (strain K12) GN=ndh PE=1 SV=2   | 1 |
| RPOE_ECO57 | ECF RNA polymerase sigma-E factor OS=Escherichia coli O157:H7 GN=rpoE PE=3 SV=1  | 1 |
| LPTB_ECO57 | Lipopolysaccharide export system ATP-binding protein LptB OS=Escherichia coli O157:H7 GN=lptB PE=3 SV=2                                      | 1 |
| RL2_ECO24  | 50S ribosomal protein L2 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplB PE=3 SV=1  | 1 |
| NUOF_ECOLI | NADH-quinone oxidoreductase subunit F OS=Escherichia coli (strain K12) GN=nuoF PE=1 SV=3   | 1 |
| SOHB_ECOLI | Probable protease SohB OS=Escherichia coli (strain K12) GN=sohB PE=1 SV=1  | 1 |
| SYT_ECO24  | Threonine--tRNA ligase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=thrS PE=3 SV=1  | 1 |
| DBHA_ECO57 | DNA-binding protein HU-alpha OS=Escherichia coli O157:H7 GN=hupA PE=3 SV=1   | 1 |
| CLPA_ECO57 | ATP-dependent Clp protease ATP-binding subunit ClpA OS=Escherichia coli O157:H7 GN=clpA PE=3 SV=1  | 1 |
| ARNA_ECO24 | Bifunctional polymyxin resistance protein ArnA OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=arnA PE=3 SV=1                        | 1 |
| RL10_ECO24 | 50S ribosomal protein L10 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplJ PE=3 SV=1   | 1 |
| RL11_ECO24 | 50S ribosomal protein L11 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplK PE=3 SV=1   | 1 |
| CRA_ECO57  | Catabolite repressor/activator OS=Escherichia coli O157:H7 GN=cra PE=3 SV=1  | 1 |
| FTSZ_ECO57 | Cell division protein FtsZ OS=Escherichia coli O157:H7 GN=ftsZ PE=3 SV=1   | 1 |
| METQ_ECO57 | D-methionine-binding lipoprotein MetQ OS=Escherichia coli O157:H7 GN=metQ PE=3 SV=1  | 1 |
| ODP2_ECOLI | Dihydrolypoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex OS=Escherichia coli (strain K12) GN=aceF PE=1 SV=3 | 1 |
| FTSK_ECO57 | DNA translocase FtsK OS=Escherichia coli O157:H7 GN=ftsK PE=3 SV=1   | 1 |
| IF2_ECO24  | Translation initiation factor IF-2 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=infB PE=3 SV=1                                    | 1 |
| FTSN_ECOLI | Cell division protein FtsN OS=Escherichia coli (strain K12) GN=ftsN PE=1 SV=3  | 1 |
| HSLU_ECO24 | ATP-dependent protease ATPase subunit HslU OS=Escherichia coli O139:H28 (strain E24377A / ETEC)  | 1 |

|            |   |   |
|------------|---|---|
|            | GN=hslU PE=3 SV=1   |   |
| NDPA_ECO24 | Nucleoid-associated protein YejK OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=yejK PE=3 SV=1 | 1 |

| Accession                 | Description  | Spectral count |
|---------------------------|--|----------------|
| <a href="#">MBP-D153N</a> |  | 416            |
| MALE_ECO57                | Maltose-binding periplasmic protein OS=Escherichia coli O157:H7 GN=malE PE=1 SV=1  | 93             |
| DNAK_ECO24                | Chaperone protein DnaK OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=dnaK PE=3 SV=1  | 18             |
| CH601_ECOK1               | 60 kDa chaperonin 1 OS=Escherichia coli O1:K1 / APEC GN=groL1 PE=3 SV=1  | 17             |
| DEGP_ECO57                | Periplasmic serine endoprotease DegP OS=Escherichia coli O157:H7 GN=degP PE=3 SV=1   | 9              |
| DNAJ_ECO24                | Chaperone protein DnaJ OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=dnaJ PE=3 SV=1  | 9              |
| OMP_F_ECOLI               | Outer membrane protein F OS=Escherichia coli (strain K12) GN=ompF PE=1 SV=1  | 8              |
| ODP1_ECO57                | Pyruvate dehydrogenase E1 component OS=Escherichia coli O157:H7 GN=aceE PE=1 SV=2  | 8              |
| FEPA_ECOLI                | Ferrienterobactin receptor OS=Escherichia coli (strain K12) GN=fepA PE=1 SV=2  | 6              |
| RS5_ECO24                 | 30S ribosomal protein S5 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsE PE=3 SV=1  | 6              |
| NUOCD_ECOL5               | NADH-quinone oxidoreductase subunit C/D OS=Escherichia coli O6:K15:H31 (strain 536 / UPEC) GN=nuoC PE=3 SV=1                                 | 5              |
| OMPA_ECO57                | Outer membrane protein A OS=Escherichia coli O157:H7 GN=ompA PE=3 SV=1   | 5              |
| ATPB_ECO24                | ATP synthase subunit beta OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=atpD PE=3 SV=1   | 5              |
| EFTU1_ECO24               | Elongation factor Tu 1 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1  | 4              |
| ODP2_ECOLI                | Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex OS=Escherichia coli (strain K12) GN=aceF PE=1 SV=3 | 4              |
| RPOC_ECO24                | DNA-directed RNA polymerase subunit beta' OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpoC PE=1 SV=1                             | 4              |
| BAMA_ECO24                | Outer membrane protein assembly factor BamA OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=bamA PE=3 SV=1                           | 4              |
| RL1_ECO24                 | 50S ribosomal protein L1 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rplA PE=3 SV=1  | 3              |
| FTSN_ECOLI                | Cell division protein FtsN OS=Escherichia coli (strain K12) GN=ftsN PE=1 SV=3  | 3              |
| LPP_ECO57                 | Major outer membrane lipoprotein OS=Escherichia coli O157:H7 GN=lpp PE=3 SV=1  | 3              |
| RS12_ECO24                | 30S ribosomal protein S12 OS=Escherichia coli O139:H28   | 3              |

|             |   |   |
|-------------|---|---|
|             | (strain E24377A / ETEC) GN=rpsL PE=3 SV=1   |   |
| LPTD_ECO57  | LPS-assembly protein LptD OS=Escherichia coli O157:H7<br>GN=lptD PE=3 SV=1  | 2 |
| ARNA_ECO24  | Bifunctional polymyxin resistance protein ArnA<br>OS=Escherichia coli O139:H28 (strain E24377A / ETEC)<br>GN=arnA PE=3 SV=1 | 2 |
| DLDH_ECO57  | Dihydrolipoyl dehydrogenase OS=Escherichia coli O157:H7<br>GN=lpdA PE=3 SV=2  | 2 |
| RS8_ECO24   | 30S ribosomal protein S8 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsH PE=3 SV=1                          | 2 |
| FDOG_ECOLI  | Formate dehydrogenase-O major subunit OS=Escherichia<br>coli (strain K12) GN=fdoG PE=1 SV=5                                 | 2 |
| EPTC_ECOLI  | Phosphoethanolamine transferase EptC OS=Escherichia coli<br>(strain K12) GN=eptC PE=1 SV=1                                  | 2 |
| RS10_ECO24  | 30S ribosomal protein S10 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsJ PE=3 SV=1                         | 2 |
| RL2_ECO24   | 50S ribosomal protein L2 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpLB PE=3 SV=1                          | 2 |
| NFUA_ECO24  | Fe/S biogenesis protein NfuA OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=nfuA PE=3 SV=1                      | 2 |
| OMPX_ECO57  | Outer membrane protein X OS=Escherichia coli O157:H7<br>GN=ompX PE=3 SV=1   | 2 |
| HFLK_ECO57  | Protein HflK OS=Escherichia coli O157:H7 GN=hflK PE=3<br>SV=1   | 1 |
| YHCB_ECO57  | Putative cytochrome d ubiquinol oxidase subunit 3<br>OS=Escherichia coli O157:H7 GN=yhcB PE=3 SV=2                          | 1 |
| RL11_ECO24  | 50S ribosomal protein L11 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpLK PE=3 SV=1                         | 1 |
| TYP_A_ECO27 | GTP-binding protein TypA/BipA OS=Escherichia coli<br>O127:H6 (strain E2348/69 / EPEC) GN=typA PE=3 SV=1                     | 1 |
| FTSK_ECO57  | DNA translocase FtsK OS=Escherichia coli O157:H7 GN=ftsK<br>PE=3 SV=1   | 1 |
| BGAL_ECO24  | Beta-galactosidase OS=Escherichia coli O139:H28 (strain<br>E24377A / ETEC) GN=lacZ PE=3 SV=1                                | 1 |
| ACNB_ECOLI  | Aconitate hydratase B OS=Escherichia coli (strain K12)<br>GN=acnB PE=1 SV=3   | 1 |
| RS6_ECO24   | 30S ribosomal protein S6 OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rpsF PE=3 SV=1                          | 1 |
| PNP_ECO24   | Polyribonucleotide nucleotidyltransferase OS=Escherichia<br>coli O139:H28 (strain E24377A / ETEC) GN=pnp PE=3 SV=1          | 1 |
| RODZ_ECO24  | Cytoskeleton protein RodZ OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=rodZ PE=3 SV=1                         | 1 |
| DHG_ECOLI   | Quinoprotein glucose dehydrogenase OS=Escherichia coli<br>(strain K12) GN=gcd PE=1 SV=3                                     | 1 |
| SYT_ECO24   | Threonine--tRNA ligase OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=thrS PE=3 SV=1                            | 1 |
| ISCS_ECO24  | Cysteine desulfurase IscS OS=Escherichia coli O139:H28<br>(strain E24377A / ETEC) GN=iscS PE=3 SV=1                         | 1 |
| SECF_ECO57  | Protein translocase subunit SecF OS=Escherichia coli  | 1 |

|            |  |   |
|------------|--|---|
|            | O157:H7 GN=secF PE=3 SV=1  |   |
| ATPA_ECO24 | ATP synthase subunit alpha OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=atpA PE=3 SV=1                          | 1 |
| RPOA_ECO24 | DNA-directed RNA polymerase subunit alpha OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpoA PE=1 SV=1           | 1 |
| NUOB_ECO24 |  | 1 |
| NUOF_ECOLI | NADH-quinone oxidoreductase subunit F OS=Escherichia coli (strain K12) GN=nuoF PE=1 SV=3                                   | 1 |
| TOLC_ECOLI | Outer membrane protein TolC OS=Escherichia coli (strain K12) GN=tolC PE=1 SV=3   | 1 |
| RPOD_ECOLI | RNA polymerase sigma factor RpoD OS=Escherichia coli (strain K12) GN=rpoD PE=1 SV=2  | 1 |
| HFLC_ECO57 | Modulator of FtsH protease HflC OS=Escherichia coli O157:H7 GN=hflC PE=3 SV=1  | 1 |
| PTND_ECO57 | Mannose permease IID component OS=Escherichia coli O157:H7 GN=manZ PE=3 SV=1   | 1 |
| PPID_EC0L6 | Peptidyl-prolyl cis-trans isomerase D OS=Escherichia coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=ppiD PE=3 SV=1     | 1 |
| METH_EC0LI | Methionine synthase OS=Escherichia coli (strain K12) GN=metH PE=1 SV=5   | 1 |
| FLGK_EC0LI | Flagellar hook-associated protein 1 OS=Escherichia coli (strain K12) GN=flgK PE=1 SV=3                                     | 1 |
| RS3_ECO24  | 30S ribosomal protein S3 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsC PE=3 SV=1                            | 1 |
| KPRS_ECO57 | Ribose-phosphate pyrophosphokinase OS=Escherichia coli O157:H7 GN=prs PE=3 SV=2  | 1 |
| CAN_EC0LI  | Carbonic anhydrase 2 OS=Escherichia coli (strain K12) GN=can PE=1 SV=1   | 1 |
| RL13_ECO24 |  | 1 |
| IBPB_ECO24 | Small heat shock protein IbpB OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=ibpB PE=3 SV=1                       | 1 |
| CLPX_ECO24 | ATP-dependent Clp protease ATP-binding subunit ClpX OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=clpX PE=3 SV=1 | 1 |
| IF2_ECO24  | Translation initiation factor IF-2 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=infB PE=3 SV=1                  | 1 |
| RS9_ECO24  | 30S ribosomal protein S9 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsI PE=3 SV=1                            | 1 |
| GNTU_ECO57 |  | 1 |
| EFG_ECO24  | Elongation factor G OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=fusA PE=3 SV=1                                 | 1 |
| IBPA_ECO24 | Small heat shock protein IbpA OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=ibpA PE=3 SV=1                       | 1 |
| YGHW_ECO57 | Uncharacterized protein YghW OS=Escherichia coli O157:H7 GN=yghW PE=4 SV=1   | 1 |
| MBHS_EC0L6 | Hydrogenase-1 small chain OS=Escherichia coli O6:H1  | 1 |

|            |  |   |
|------------|--|---|
|            | (strain CFT073 / ATCC 700928 / UPEC) GN=hyaA PE=3 SV=1   |   |
| GYRB_ECO57 | DNA gyrase subunit B OS=Escherichia coli O157:H7 GN=gyrB PE=3 SV=2   | 1 |
| CYDA_ECOL6 | Cytochrome bd-I ubiquinol oxidase subunit 1 OS=Escherichia coli O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=cydA PE=3 SV=1 | 1 |
| MBHT_ECO57 | Hydrogenase-2 small chain OS=Escherichia coli O157:H7 GN=hybO PE=3 SV=1  | 1 |
| ADIA_ECOLI | Biodegradative arginine decarboxylase OS=Escherichia coli (strain K12) GN=adiA PE=1 SV=1                                     | 1 |
| YFHM_ECOLI | Uncharacterized lipoprotein YfhM OS=Escherichia coli (strain K12) GN=yfhM PE=1 SV=1  | 1 |
| YGIQ_ECO57 | UPF0313 protein YgiQ OS=Escherichia coli O157:H7 GN=ygiQ PE=3 SV=1   | 1 |
| RL17_ECO24 | 50S ribosomal protein L17 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpIQ PE=3 SV=1                             | 1 |
| RS11_ECO24 | 30S ribosomal protein S11 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsK PE=3 SV=1                             | 1 |
| KDSB_ECO24 | 3-deoxy-manno-octulosonate cytidyltransferase OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=kdsB PE=3 SV=1         | 1 |
| SUFB_ECOLI | FeS cluster assembly protein SufB OS=Escherichia coli (strain K12) GN=sufB PE=1 SV=2   | 1 |
| LAST_ECOLI | Uncharacterized tRNA/rRNA methyltransferase LasT OS=Escherichia coli (strain K12) GN=lasT PE=3 SV=2                          | 1 |
| NUOG_ECO57 | NADH-quinone oxidoreductase subunit G OS=Escherichia coli O157:H7 GN=nuoG PE=3 SV=3  | 1 |
| PTKA_ECO57 | Galactitol-specific phosphotransferase enzyme IIA component OS=Escherichia coli O157:H7 GN=gatA PE=3 SV=1                    | 1 |
| RL14_ECO24 | 50S ribosomal protein L14 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpIN PE=3 SV=1                             | 1 |
| SSPB_ECO57 | Stringent starvation protein B OS=Escherichia coli O157:H7 GN=sspB PE=3 SV=1   | 1 |
| LACI_ECOLI | Lactose operon repressor OS=Escherichia coli (strain K12) GN=lacl PE=1 SV=3  | 1 |
| ADHE_ECO57 | Aldehyde-alcohol dehydrogenase OS=Escherichia coli O157:H7 GN=adhE PE=3 SV=2   | 1 |
| RS15_ECO24 | 30S ribosomal protein S15 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=rpsO PE=3 SV=1                             | 1 |
| INAA_ECOLI |  | 1 |
| LEPA_ECO24 | Elongation factor 4 OS=Escherichia coli O139:H28 (strain E24377A / ETEC) GN=lepA PE=3 SV=1                                   | 1 |
| G6PD_ECO57 | Glucose-6-phosphate 1-dehydrogenase OS=Escherichia coli O157:H7 GN=zwf PE=3 SV=1   | 1 |

#### Annexe 4 Liste des protéines enrichies dans l'IP YFP-DCP1

Le tableau contient les protéines avec un LogFC > 1 et avec une adjp (adjusted p-value) > 0,05.

| AGI         | LogFC      | p.value    | adjp       | Nom du gène | Description  | Moyenne du nombre de spectres |       |
|-------------|------------|------------|------------|-------------|--|-------------------------------|-------|
|             |            |            |            |             |  | Colo                          | DCP1  |
| AT1G08370.1 | 11,11444   | 4,2997E-26 | 7,1159E-23 | DCP1        | mRNA-decapping enzyme-like protein<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9SJF3]                       | 0,0                           | 402,8 |
| AT3G13300.1 | 8,01482099 | 1,1292E-13 | 9,3438E-11 | VCS         | Enhancer of mRNA-decapping protein 4<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LTT8]                     | 0,0                           | 47,5  |
| AT3G49490.1 | 7,84426976 | 3,7475E-13 | 2,0674E-10 |             | Uncharacterized protein TgC5.90<br>[Source:UniProtKB/TrEMBL;Acc:Q9SCK9]                              | 0,0                           | 42,3  |
| AT2G45810.1 | 6,73560031 | 8,4408E-09 | 3,4924E-06 | RH6         | DEAD-box ATP-dependent RNA helicase 6<br>[Source:UniProtKB/Swiss-Prot;Acc:Q94BV4]                    | 0,0                           | 18,0  |
| AT4G00660.2 | 6,78367511 | 1,7E-08    | 5,627E-06  | RH8         | DEAD-box ATP-dependent RNA helicase 8<br>[Source:UniProtKB/Swiss-Prot;Acc:Q8RXX6]                    | 0,0                           | 17,0  |
| AT3G13290.1 | 6,47610499 | 2,6373E-08 | 7,2744E-06 | VCR         | Varicose-related protein [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT9]                                    | 0,0                           | 16,8  |
| AT3G61240.1 | 6,36738704 | 5,9859E-08 | 1,4152E-05 | RH12        | DEAD-box ATP-dependent RNA helicase 12<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9M2E0]                   | 0,0                           | 13,5  |
| AT2G15560.1 | 6,26350827 | 8,7997E-08 | 1,8204E-05 | DAE         | At2g15560<br>[Source:UniProtKB/TrEMBL;Acc:Q6NQ97]  | 0,0                           | 12,7  |
| AT1G07310.1 | 5,790747   | 1,3933E-06 | 0,00025621 |             | Calcium-dependent lipid-binding (CaLB domain) family protein<br>[Source:UniProtKB/TrEMBL;Acc:Q9LNVo] | 0,0                           | 9,8   |
| AT1G20620.1 | 3,95590039 | 5,6145E-06 | 0,00092921 | CAT3        | catalase 3 [Source:TAIR;Acc:AT1G20620]   | 0,3                           | 14,0  |
| AT3G06650.1 | 3,30167717 | 1,0901E-05 | 0,0015503  | ACLB-1      | ACLB-1<br>[Source:UniProtKB/TrEMBL;Acc:A0A178V742]   | 1,0                           | 20,5  |
| AT5G49460.1 | 3,21394891 | 1,1241E-05 | 0,0015503  | ACLB-2      | ATP-citrate synthase beta chain protein 2<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9FGX1]                | 1,0                           | 20,3  |
| AT3G08530.1 | 3,20436565 | 4,5362E-05 | 0,00577499 | CHC2        | Clathrin heavy chain 2 [Source:UniProtKB/Swiss-Prot;Acc:QoWLB5]                                      | 0,8                           | 20,0  |
| AT3G11130.1 | 2,99307852 | 6,3418E-05 | 0,0069971  | CHC1        | Clathrin heavy chain 1 [Source:UniProtKB/Swiss-Prot;Acc:QoWJN6]                                      | 1,0                           | 21,3  |
| AT3G12400.1 | 4,70169694 | 0,00054415 | 0,04573346 | ELC         | Ubiquitin-conjugating enzyme/RWD-like protein<br>[Source:UniProtKB/TrEMBL;Acc:A0A19LRG3]             | 0,0                           | 4,0   |
| AT3G58570.1 | 2,6561434  | 0,00055267 | 0,04573346 | RH52        | DEAD-box ATP-dependent RNA helicase 52<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9M2F9]                   | 0,7                           | 10,8  |

## Annexe 5 Liste des protéines enrichies dans l'IP GFP-DCP2

Le tableau contient les protéines avec un LogFC > 1 et avec une adjp (adjusted p-value) < 0,05.

| AGI         | LogFC      | p.value    | adjp       | Nom du gène | Description  | Moyenne du nombre de spectres |       |
|-------------|------------|------------|------------|-------------|--|-------------------------------|-------|
|             |            |            |            |             |  | Colo                          | DCP2  |
| AT5G13570.2 | 11,3135173 | 2,1874E-18 | 3,6202E-15 | DCP2        | decapping 2 [Source:TAIR;Acc:AT5G13570]  | 0,0                           | 586,7 |
| AT3G13300.1 | 10,7581628 | 1,4581E-16 | 1,2066E-13 | VCS         | Enhancer of mRNA-decapping protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT8]        | 0,0                           | 395,3 |
| AT3G13290.1 | 10,0409649 | 2,7653E-14 | 1,5255E-11 | VCR         | Varicose-related protein [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT9]                    | 0,0                           | 241,0 |
| AT1G06000.1 | 6,71294467 | 5,4452E-06 | 0,00060079 | UGT8gC1     | Glycosyltransferase (Fragment) [Source:UniProtKB/TrEMBL;Acc:W8PVE7]                  | 0,0                           | 25,7  |
| AT4G13940.1 | 5,53369072 | 0,00047706 | 0,02924229 | SAHH1       | Adenosylhomocysteinase 1 [Source:UniProtKB/Swiss-Prot;Acc:O23255]                    | 0,0                           | 15,0  |
| AT3G13470.1 | 5,47949767 | 0,00056659 | 0,03233479 | CPN60B2     | Chaperonin 60 subunit beta 2, chloroplastic [Source:UniProtKB/Swiss-Prot;Acc:Q9LJE4] | 0,0                           | 15,3  |
| AT5G56010.1 | 5,46421967 | 0,00060586 | 0,03314739 | HSP90-3     | Heat shock protein 90-3 [Source:UniProtKB/Swiss-Prot;Acc:P51818]                     | 0,0                           | 15,0  |
| AT1G64740.1 | 5,45442785 | 0,00062089 | 0,03314739 | TUBA1       | Tubulin alpha-1 chain [Source:UniProtKB/Swiss-Prot;Acc:P11139]                       | 0,0                           | 15,0  |
| AT1G79920.1 | 3,76769226 | 0,00107123 | 0,04665482 |             | Heat shock protein 70 (Hsp 70) family protein [Source:TAIR;Acc:AT1G79920]            | 0,2                           | 17,3  |



## Annexe 6 Liste des protéines enrichies dans l'IP GFP-DAE avec crosslink

Le tableau contient les protéines avec un LogFC > 1 et avec une adjp (adjusted p-value) < 0,05.

| AGI         | LogFC      | LR         | p.value    | adjp       | Nom du gène | Description  | Moyenne du nombre de spectres |        |
|-------------|------------|------------|------------|------------|-------------|--|-------------------------------|--------|
|             |            |            |            |            |             |  | Colo                          | DAE    |
| AT2G15560.1 | 10,8357558 | 181,189708 | 2,6648E-41 | 3,6694E-38 | DAE         | At2g15560<br>[Source:UniProtKB/TrEMBL;Acc:Q6NQ97]  | 0                             | 309,5  |
| AT5G56010.1 | 9,42066895 | 102,567591 | 4,169E-24  | 1,4352E-21 | HSP90-3     | Heat shock protein 90-3<br>[Source:UniProtKB/Swiss-Prot;Acc:P51818]  | 0                             | 104    |
| AT5G56030.2 | 9,3641733  | 100,419193 | 1,2333E-23 | 3,3964E-21 | HSP81-2     | Heat shock protein 81-2<br>[Source:UniProtKB/TrEMBL;Acc:F4K6B6]  | 0                             | 100,25 |
| AT5G56000.1 | 9,0240477  | 87,1332503 | 1,0145E-20 | 2,3283E-18 | HSP90-4     | Hsp81.4<br>[Source:UniProtKB/TrEMBL;Acc:AoA178UQ52]  | 0                             | 76,75  |
| AT1G08370.1 | 8,40798651 | 68,7762381 | 1,1029E-16 | 1,8984E-14 | DCP1        | mRNA-decapping enzyme-like protein<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9SJF3]                               | 0                             | 59,25  |
| AT3G13300.1 | 8,25085751 | 64,9640421 | 7,6277E-16 | 1,167E-13  | VCS         | Enhancer of mRNA-decapping protein 4<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LTT8]                             | 0                             | 60,5   |
| AT5G02500.1 | 2,53681867 | 55,9402496 | 7,4707E-14 | 1,0287E-11 | MED37E      | Probable mediator of RNA polymerase II transcription subunit 37e<br>[Source:UniProtKB/Swiss-Prot;Acc:P22953] | 10,75                         | 257,75 |
| AT4G27060.1 | 7,62222892 | 48,4145324 | 3,45E-12   | 3,3933E-10 | TOR1        | Microtubule-associated protein TORTIFOLIA1<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9To41]                       | 0                             | 37,25  |
| AT5G47010.1 | 7,58039995 | 47,1380083 | 6,616E-12  | 6,0735E-10 | UPF1        | Regulator of nonsense transcripts 1 homolog<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9FJR0]                      | 0                             | 44     |
| AT5G02490.1 | 2,34086707 | 44,3127657 | 2,7989E-11 | 2,2671E-09 | MED37D      | Probable mediator of RNA polymerase II transcription subunit 37c<br>[Source:UniProtKB/Swiss-Prot;Acc:P22954] | 8,25                          | 175,5  |
| AT3G09440.1 | 2,29575108 | 41,4889637 | 1,1854E-10 | 9,0682E-09 | HSP70-3     | Heat shock protein 70 (Hsp 70) family protein<br>[Source:UniProtKB/TrEMBL;Acc:AoA178V176]                    | 8                             | 161    |
| AT3G12580.1 | 2,1194734  | 33,8464449 | 5,9638E-09 | 3,8489E-07 | MED37C      | Probable mediator of RNA polymerase II transcription subunit 37c<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LHA8] | 7                             | 127,5  |
| AT4G00660.2 | 6,96286205 | 33,1731308 | 8,4307E-09 | 5,0475E-07 | RH8         | DEAD-box ATP-dependent RNA helicase 8<br>[Source:UniProtKB/Swiss-Prot;Acc:Q8RXK6]                            | 0                             | 26,75  |
| AT3G13290.1 | 6,81099145 | 30,2055852 | 3,8859E-08 | 2,058E-06  | VCR         | Varicose-related protein<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LTT9]   | 0                             | 24     |
| AT2G45810.1 | 3,40629773 | 24,4755557 | 7,5259E-07 | 3,3429E-05 | RH6         | DEAD-box ATP-dependent RNA helicase 6<br>[Source:UniProtKB/Swiss-Prot;Acc:Q94BV4]                            | 0,5                           | 27     |
| AT5G52640.1 | 6,47034902 | 24,1342048 | 8,985E-07  | 3,8664E-05 | HSP90-1     | Heat shock protein 90-1<br>[Source:UniProtKB/Swiss-Prot;Acc:P27323]  | 0                             | 14,75  |
| AT2G18960.1 | 3,84266904 | 21,2689301 | 3,9915E-06 | 0,00015704 | AHA1        | Plasma membrane ATPase<br>[Source:UniProtKB/TrEMBL;Acc:AoA178VSN5]   | 0,25                          | 20     |
| AT4G30190.2 | 3,65591102 | 18,0786603 | 2,1196E-05 | 0,00081076 | HA2         | Plasma membrane ATPase<br>[Source:UniProtKB/TrEMBL;Acc:F4JPJ7]   | 0,25                          | 17     |

|             |            |            |            |            |       |  |      |       |
|-------------|------------|------------|------------|------------|-------|--|------|-------|
| AT5G62670.1 | 5,96176396 | 17,1390464 | 3,4741E-05 | 0,00129292 | AHA11 | ATPase 11, plasma membrane-type [Source:UniProtKB/Swiss-Prot;Acc:Q9LV11]                       | 0    | 11,75 |
| AT3G61240.1 | 2,91817094 | 16,2912911 | 5,4313E-05 | 0,00196813 | RH12  | DEAD-box ATP-dependent RNA helicase 12 [Source:UniProtKB/Swiss-Prot;Acc:Q9M2E0]                | 0,5  | 20,5  |
| AT2G35110.1 | 5,82317417 | 15,7901186 | 7,0771E-05 | 0,0024363  | NAP1  | Protein NAP1 [Source:UniProtKB/Swiss-Prot;Acc:Q5S2C4]  | 0    | 12,5  |
| AT2G28790.1 | 5,77366992 | 15,3255637 | 9,0484E-05 | 0,00296657 |       | Pathogenesis-related thaumatin superfamily protein [Source:TAIR;Acc:AT2G28790]                 | 0    | 12,75 |
| AT1G20620.1 | 5,71473342 | 14,6558529 | 0,00012903 | 0,00413207 | CAT3  | catalase 3 [Source:TAIR;Acc:AT1G20620]   | 0    | 11,75 |
| AT3G08530.1 | 1,50817039 | 12,9777618 | 0,00031521 | 0,00885812 | CHC2  | Clathrin heavy chain 2 [Source:UniProtKB/Swiss-Prot;Acc:QoWLB5]                                | 3,5  | 56    |
| AT5G22060.1 | 5,48778977 | 12,459698  | 0,00041583 | 0,01145188 | ATJ2  | At5g22060 [Source:UniProtKB/TrEMBL;Acc:QoV7U1]   | 0    | 9,5   |
| AT5G18410.1 | 5,47130439 | 12,2944605 | 0,0004543  | 0,01203034 | PIR   | transcription activators [Source:TAIR;Acc:AT5G18410]   | 0    | 8,5   |
| AT1G48480.1 | 5,4252707  | 11,9303369 | 0,00055227 | 0,01382688 | RKL1  | Probable inactive receptor kinase At1g48480 [Source:UniProtKB/Swiss-Prot;Acc:Q9LP77]           | 0    | 9     |
| AT2G21390.1 | 5,40392595 | 11,7926805 | 0,00059464 | 0,01462179 |       | Coatomer subunit alpha-2 [Source:UniProtKB/Swiss-Prot;Acc:Q9JTG9]                              | 0    | 9,5   |
| AT3G11130.1 | 1,41432113 | 11,7572452 | 0,00060607 | 0,01464134 | CHC1  | Clathrin heavy chain 1 [Source:UniProtKB/Swiss-Prot;Acc:QoWNJ6]                                | 3,75 | 55,75 |
| AT5G57350.1 | 5,39870825 | 11,7014798 | 0,0006245  | 0,01482659 | AHA3  | Plasma membrane ATPase [Source:UniProtKB/TrEMBL;Acc:AoA178UHY0]                                | 0    | 8     |
| AT5G58960.1 | 5,27782977 | 10,8228505 | 0,00100255 | 0,02226632 | GIL1  | Plant protein of unknown function (DUF641) [Source:TAIR;Acc:AT5G58960]                         | 0    | 7,75  |
| AT3G47950.1 | 5,18367271 | 10,1427997 | 0,00144866 | 0,0306893  | AHA4  | Plasma membrane ATPase [Source:UniProtKB/TrEMBL;Acc:AoA178VFO4]                                | 0    | 7     |
| AT5G67630.1 | 4,98969243 | 8,90062948 | 0,00285072 | 0,05528798 |       | RuvB-like helicase [Source:UniProtKB/TrEMBL;Acc:Q9FJW0]  | 0    | 6,75  |
| AT1G14830.1 | 2,0202285  | 8,80107067 | 0,00301054 | 0,05757654 | DRP1C | DRP1C [Source:UniProtKB/TrEMBL;Acc:AoA178W4V6]   | 0,75 | 17    |
| AT3G22330.1 | 4,81357814 | 7,8788294  | 0,00500168 | 0,08829897 | RH53  | DEAD-box ATP-dependent RNA helicase 53, mitochondrial [Source:UniProtKB/Swiss-Prot;Acc:Q9LUW5] | 0    | 6,25  |

## Annexe 7 Liste des protéines enrichies dans l'IP YFP-DCP1 avec crosslink

Le tableau contient les protéines avec un LogFC > 1 et avec une adjp (adjusted p-value) < 0,05.

| AGI         | LogFC      | p.value    | adjp       | Nom du gène   | Description   | Moyenne du nombre de spectres |        |
|-------------|------------|------------|------------|---------------|---|-------------------------------|--------|
|             |            |            |            |               |   | Colo                          | DCP1   |
| AT1G08370.1 | 10,3382696 | 2,828E-110 | 3,894E-107 | <b>DCP1</b>   | mRNA-decapping enzyme-like protein [Source:UniProtKB/Swiss-Prot;Acc:Q9S5JF3]                      | 0                             | 269,75 |
| AT3G13300.1 | 9,1422105  | 4,422E-49  | 1,5223E-46 | <b>VCS</b>    | Enhancer of mRNA-decapping protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT8]                     | 0                             | 112,25 |
| AT4G00660.2 | 8,45793263 | 2,3537E-31 | 6,4822E-29 | <b>RH8</b>    | DEAD-box ATP-dependent RNA helicase 8 [Source:UniProtKB/Swiss-Prot;Acc:Q8RXK6]                    | 0                             | 74,25  |
| AT2G45810.1 | 5,22977626 | 1,5863E-29 | 3,6406E-27 | <b>RH6</b>    | DEAD-box ATP-dependent RNA helicase 6 [Source:UniProtKB/Swiss-Prot;Acc:Q94BV4]                    | 0,5                           | 77,25  |
| AT3G13290.1 | 8,24615201 | 3,5675E-27 | 6,1405E-25 | <b>VCR</b>    | Varicose-related protein [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT9]                                 | 0                             | 60,75  |
| AT3G61240.1 | 4,5487228  | 6,1811E-18 | 8,5114E-16 | <b>RH12</b>   | DEAD-box ATP-dependent RNA helicase 12 [Source:UniProtKB/Swiss-Prot;Acc:Q9M2E0]                   | 0,5                           | 48     |
| AT3G08530.1 | 2,48971979 | 1,4551E-16 | 1,8216E-14 | <b>CHC2</b>   | Clathrin heavy chain 2 [Source:UniProtKB/Swiss-Prot;Acc:QoWLB5]                                   | 3,5                           | 65,75  |
| AT3G11130.1 | 2,38147859 | 9,244E-16  | 1,0607E-13 | <b>CHC1</b>   | Clathrin heavy chain 1 [Source:UniProtKB/Swiss-Prot;Acc:QoWLN6]                                   | 3,75                          | 66     |
| AT5G47010.1 | 7,12879936 | 2,3824E-13 | 2,0504E-11 | <b>UPF1</b>   | Regulator of nonsense transcripts 1 homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9FJR0]              | 0                             | 28,5   |
| AT3G49490.1 | 7,10183345 | 3,4245E-13 | 2,7738E-11 |               | Uncharacterized protein T9C5.90 [Source:UniProtKB/TrEMBL;Acc:Q9SCK9]                              | 0                             | 29,75  |
| AT2G15560.1 | 7,02749807 | 1,3387E-12 | 1,0241E-10 | <b>DAE</b>    | At2g15560 [Source:UniProtKB/TrEMBL;Acc:Q6NQ97]  | 0                             | 28,25  |
| AT2G35110.1 | 6,69299388 | 3,8047E-10 | 2,7574E-08 | <b>NAP1</b>   | Protein NAP1 [Source:UniProtKB/Swiss-Prot;Acc:Q5S2C4]   | 0                             | 18,75  |
| AT1G54490.1 | 6,51847831 | 3,4215E-09 | 2,3557E-07 | <b>XRN4</b>   | XRN4 [Source:UniProtKB/TrEMBL;Acc:AoA384KRW9]   | 0                             | 18     |
| AT1G07310.1 | 6,011479   | 6,7444E-07 | 4,2214E-05 |               | Calcium-dependent lipid-binding (CaLB domain) family protein [Source:UniProtKB/TrEMBL;Acc:Q9LNV0] | 0                             | 14     |
| AT5G58960.1 | 5,94525512 | 1,3807E-06 | 8,2661E-05 | <b>GIL1</b>   | Plant protein of unknown function (DUF641) [Source:TAIR;Acc:AT5G58960]                            | 0                             | 11,75  |
| AT3G06650.1 | 3,77031187 | 2,9855E-06 | 0,00015812 | <b>ACLB-1</b> | ACLB-1 [Source:UniProtKB/TrEMBL;Acc:AoA178V742]   | 0,25                          | 16,5   |
| AT5G49460.1 | 3,75365447 | 3,4673E-06 | 0,00017683 | <b>ACLB-2</b> | ATP-citrate synthase beta chain protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9FGX1]                | 0,25                          | 16,5   |
| AT5G18410.1 | 5,73590022 | 7,5549E-06 | 0,00037154 | <b>PIR</b>    | transcription activators [Source:TAIR;Acc:AT5G18410]  | 0                             | 9,75   |
| AT4G30190.2 | 3,68097283 | 9,6976E-06 | 0,00046047 | <b>HA2</b>    | Plasma membrane ATPase [Source:UniProtKB/TrEMBL;Acc:F4JPJ7]                                       | 0,25                          | 11     |
| AT2G18960.1 | 3,64826927 | 1,2974E-05 | 0,00059549 | <b>AHA1</b>   | Plasma membrane ATPase [Source:UniProtKB/TrEMBL;Acc:AoA178VSN5]                                   | 0,25                          | 10,75  |
| AT1G20620.1 | 5,65197453 | 1,404E-05  | 0,00062366 | <b>CAT3</b>   | catalase 3 [Source:TAIR;Acc:AT1G20620]  | 0                             | 9      |
| AT5G62670.1 | 5,60399729 | 2,1557E-05 | 0,00092318 | <b>AHA11</b>  | ATPase 11, plasma membrane-type [Source:UniProtKB/Swiss-Prot;Acc:Q9LV11]                          | 0                             | 7,5    |
| AT1G79090.1 | 5,4697204  | 4,3524E-05 | 0,00168719 | <b>PAT1</b>   | Protein PAT1 homolog  | 0                             | 8,75   |

|             |            |            |            |             |  |   |      |
|-------------|------------|------------|------------|-------------|--|---|------|
|             |            |            |            |             | [Source:UniProtKB/Swiss-Prot;Acc:QoWPK4]   |   |      |
| AT3G12400.1 | 5,36243286 | 8,2163E-05 | 0,0030578  | ELC         | Ubiquitin-conjugating enzyme/RWD-like protein<br>[Source:UniProtKB/TrEMBL;Acc:AoA19LRG3]                 | 0 | 8,25 |
| AT2G28790.1 | 5,22555755 | 0,00017963 | 0,00618378 |             | Pathogenesis-related thaumatin superfamily<br>protein [Source:TAIR;Acc:AT2G28790]                        | 0 | 7,25 |
| AT3G12915.1 | 5,17599552 | 0,00024247 | 0,0079494  |             | Ribosomal protein S5/Elongation factor G/III/V<br>family protein;(source:Araport11)                      | 0 | 6,5  |
| AT5G13570.2 | 5,14725335 | 0,00028321 | 0,00906931 | <b>DCP2</b> | decapping 2 [Source:TAIR;Acc:AT5G13570]  | 0 | 6,25 |
| AT5G53620.1 | 5,11117933 | 0,00033426 | 0,01010864 |             | At5g53620<br>[Source:UniProtKB/TrEMBL;Acc:Q93XY1]  | 0 | 6,25 |
| AT5G56010.1 | 5,10441462 | 0,00033769 | 0,01010864 | HSP90-3     | Heat shock protein 90-3<br>[Source:UniProtKB/Swiss-Prot;Acc:P51818]                                      | 0 | 6,5  |
| AT5G56000.1 | 5,03626435 | 0,00046593 | 0,01336625 | HSP90-4     | Hsp81.4<br>[Source:UniProtKB/TrEMBL;Acc:AoA178UQ52]  | 0 | 6,25 |
| AT5G53330.1 | 4,99001691 | 0,000556   | 0,01531225 |             | Proline-rich cell wall protein-like<br>[Source:UniProtKB/TrEMBL;Acc:Q8LG11]                              | 0 | 6,5  |
| AT2G21390.1 | 4,97955808 | 0,00063034 | 0,01701926 |             | Coatomer subunit alpha-2<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9SJT9]                                     | 0 | 5,5  |
| AT5G56030.2 | 4,96473449 | 0,0006435  | 0,01704033 | HSP81-2     | Heat shock protein 81-2<br>[Source:UniProtKB/TrEMBL;Acc:F4K6B6]  | 0 | 6    |
| AT3G53120.1 | 4,93142035 | 0,00075259 | 0,01955322 | VPS37-1     | Vacuolar protein-sorting-associated protein 37<br>homolog 1 [Source:UniProtKB/Swiss-<br>Prot;Acc:Q9SCP9] | 0 | 5,75 |
| AT5G19400.1 | 4,87816118 | 0,00100927 | 0,02396156 | <b>SMG7</b> | Protein SMG7 [Source:UniProtKB/Swiss-<br>Prot;Acc:A9QM73]  | 0 | 4,75 |
| AT5G57350.1 | 4,72194638 | 0,00191806 | 0,04280202 | AHA3        | Plasma membrane ATPase<br>[Source:UniProtKB/TrEMBL;Acc:AoA178UHYo]                                       | 0 | 4    |
| AT1G48480.1 | 4,71323361 | 0,00193793 | 0,04280202 | RKL1        | Probable inactive receptor kinase At1g48480<br>[Source:UniProtKB/Swiss-Prot;Acc:Q9LP77]                  | 0 | 4,25 |
| AT2G46225.2 | 4,70443691 | 0,00195826 | 0,04280202 | ABIL1       | ABI-1-like 1<br>[Source:UniProtKB/TrEMBL;Acc:B3H5V3]   | 0 | 4,5  |

## Annexe 8 Liste des protéines enrichies dans l'IP GFP-DCP2 avec crosslink

Le tableau contient les protéines avec un LogFC > 1 et avec une adjp (adjusted p-value) < 0,05.

| AGI         | LogFC      | p.value    | adjp       | Nom du gène | Description  | Moyenne du nombre de spectres |        |
|-------------|------------|------------|------------|-------------|--|-------------------------------|--------|
|             |            |            |            |             |  | Colo                          | DCP2   |
| AT5G13570.2 | 10,9017318 | 1,918E-143 | 2,642E-140 | <b>DCP2</b> | decapping 2 [Source:TAIR;Acc:AT5G13570]  | 0                             | 365    |
| AT3G13300.1 | 9,80905484 | 4,5084E-74 | 3,104E-71  | <b>VCS</b>  | Enhancer of mRNA-decapping protein 4 [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT8]        | 0                             | 172,5  |
| AT3G13290.1 | 9,20064755 | 4,1988E-52 | 1,4454E-49 | <b>VCR</b>  | Varicose-related protein [Source:UniProtKB/Swiss-Prot;Acc:Q9LTT9]                    | 0                             | 113,25 |
| AT1G79920.1 | 7,13768263 | 1,5937E-14 | 2,4384E-12 | HSP70       | Heat shock protein 70 (Hsp 70) family protein [Source:TAIR;Acc:AT1G79920]            | 0                             | 23,5   |
| AT1G79930.1 | 6,97799911 | 3,6739E-13 | 5,059E-11  | HSP70-14    | Heat shock 70 kDa protein 14 [Source:UniProtKB/Swiss-Prot;Acc:Q9S7Co]                | 0                             | 21,25  |
| AT4G00660.2 | 6,92117295 | 7,4766E-13 | 9,3593E-11 | RH8         | DEAD-box ATP-dependent RNA helicase 8 [Source:UniProtKB/Swiss-Prot;Acc:Q8RXK6]       | 0                             | 23,5   |
| AT1G54490.1 | 6,46332598 | 9,8039E-10 | 9,6428E-08 | <b>XRN4</b> | XRN4 [Source:UniProtKB/TrEMBL;Acc:AoA384KRW9]  | 0                             | 17,25  |
| AT2G45810.1 | 3,64630279 | 5,251E-09  | 4,8204E-07 | RH6         | DEAD-box ATP-dependent RNA helicase 6 [Source:UniProtKB/Swiss-Prot;Acc:Q94BV4]       | 0,5                           | 20     |
| AT3G61240.1 | 3,59064739 | 1,173E-08  | 1,0095E-06 | RH12        | DEAD-box ATP-dependent RNA helicase 12 [Source:UniProtKB/Swiss-Prot;Acc:Q9M2E0]      | 0,5                           | 19,75  |
| AT1G08370.1 | 6,19613015 | 2,7736E-08 | 2,1218E-06 | <b>DCP1</b> | mRNA-decapping enzyme-like protein [Source:UniProtKB/Swiss-Prot;Acc:Q9SJF3]          | 0                             | 13,75  |
| AT5G56010.1 | 5,86056224 | 8,0476E-07 | 5,2769E-05 | HSP90-3     | Heat shock protein 90-3 [Source:UniProtKB/Swiss-Prot;Acc:P51818]                     | 0                             | 10,75  |
| AT5G56030.2 | 5,81205601 | 1,1526E-06 | 6,9004E-05 | HSP81-2     | Heat shock protein 81-2 [Source:UniProtKB/TrEMBL;Acc:F4K6B6]                         | 0                             | 11     |
| AT5G56000.1 | 5,64540207 | 4,7933E-06 | 0,00026401 | HSP90-4     | Hsp81.4 [Source:UniProtKB/TrEMBL;Acc:AoA178UQ52]                                     | 0                             | 9,25   |
| AT5G47010.1 | 5,53325588 | 1,1711E-05 | 0,00057594 | <b>UPF1</b> | Regulator of nonsense transcripts 1 homolog [Source:UniProtKB/Swiss-Prot;Acc:Q9FJR0] | 0                             | 8      |
| AT1G77680.1 | 5,28562027 | 5,9369E-05 | 0,00272505 | <b>SOV</b>  | Inactive exonuclease DIS3L2 [Source:UniProtKB/Swiss-Prot;Acc:QoWPN0]                 | 0                             | 6,75   |
| AT3G12915.1 | 5,13414674 | 0,00014692 | 0,00652593 |             | Ribosomal protein S5/Elongation factor G/III/V family protein;(source:Araport11)     | 0                             | 5,75   |
| AT3G52930.1 | 5,00798411 | 0,00025555 | 0,01099657 | FBA8        | Fructose-bisphosphate aldolase [Source:UniProtKB/TrEMBL;Acc:AoA178V8L4]              | 0                             | 6,25   |
| AT1G79090.1 | 4,95835249 | 0,00036632 | 0,01528534 | <b>PAT1</b> | Protein PAT1 homolog [Source:UniProtKB/Swiss-Prot;Acc:QoWPK4]                        | 0                             | 5      |
| AT2G28790.1 | 4,90344605 | 0,00044189 | 0,01690241 |             | Pathogenesis-related thaumatin superfamily protein [Source:TAIR;Acc:AT2G28790]       | 0                             | 5,5    |
| AT1G20620.1 | 4,83607692 | 0,0006367  | 0,02307197 | CAT3        | catalase 3 [Source:TAIR;Acc:AT1G20620]   | 0                             | 4,75   |
| AT1G56070.1 | 1,21525287 | 0,0007022  | 0,02417315 | LOS1        | Elongation factor 2 [Source:UniProtKB/Swiss-Prot;Acc:Q9ASR1]                         | 3,75                          | 26     |

## Annexe 9 Listes de gènes issues des comparaison entre les loci perturbés chez vcs-6 et DAE OE ou D153N OE

**(A) Gènes communs entre vcs-6 DOWN et DAE OE UP**

| AGI       | Description  |
|-----------|--|
| AT2G43050 | Probable pectinesterase/pectinesterase inhibitor 16 [Source:UniProtKB/Swiss-Prot |

**(B) Gènes communs entre vcs-6 DOWN et D153N OE UP**

| AGI       | Description   |
|-----------|---|
| AT1G04660 | Glycine-rich protein [Source:UniProtKB/TrEMBL   |
| AT1G04800 | At1g04800 [Source:UniProtKB/TrEMBL  |
| AT1G05540 | T25N20.19 [Source:UniProtKB/TrEMBL  |
| AT1G07485 | Uncharacterized protein At1g07490 [Source:UniProtKB/TrEMBL                                  |
| AT1G11850 | F12F1.31 protein [Source:UniProtKB/TrEMBL   |
| AT1G18990 | Probable myosin-binding protein 5 [Source:UniProtKB/Swiss-Prot                              |
| AT1G22590 | AGAMOUS-like 87 [Source:UniProtKB/TrEMBL  |
| AT1G24577 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL                                   |
| AT1G26911 | NA  |
| AT1G30282 | other RNA [Source:TAIR  |
| AT1G31690 | Amine oxidase [Source:UniProtKB/TrEMBL  |
| AT1G62975 | Transcription factor bHLH125 [Source:UniProtKB/Swiss-Prot                                   |
| AT1G66100 | Probable thionin-2.4 [Source:UniProtKB/Swiss-Prot   |
| AT1G69320 | CLE10 [Source:UniProtKB/TrEMBL  |
| AT1G76530 | Protein PIN-LIKES 4 [Source:UniProtKB/Swiss-Prot  |
| AT2G28740 | Histone H4 [Source:UniProtKB/Swiss-Prot   |
| AT2G32690 | Glycine-rich protein 23 [Source:UniProtKB/Swiss-Prot  |
| AT2G33810 | Squamosa promoter-binding-like protein 3 [Source:UniProtKB/Swiss-Prot                       |
| AT2G40100 | Chlorophyll a-b binding protein, chloroplastic [Source:UniProtKB/TrEMBL                     |
| AT2G40475 | Expressed protein [Source:UniProtKB/TrEMBL  |
| AT3G03776 | Hydroxyproline-rich glycoprotein family protein [Source:UniProtKB/TrEMBL                    |
| AT3G10185 | Gibberellin-regulated family protein [Source:TAIR   |
| AT3G14900 | Uncharacterized protein At3g14900 [Source:UniProtKB/TrEMBL                                  |
| AT3G26235 | unknown protein   |
| AT3G26960 | Pollen Ole e 1 allergen and extensin family protein [Source:UniProtKB/TrEMBL                |
| AT3G29030 | Expansin [Source:UniProtKB/TrEMBL   |
| AT3G29780 | Protein RALF-like 27 [Source:UniProtKB/Swiss-Prot   |
| AT3G58070 | GIS [Source:UniProtKB/TrEMBL  |
| AT4G19430 | At4g19430 [Source:UniProtKB/TrEMBL  |
| AT4G29020 | Putative glycine-rich protein 5 [Source:UniProtKB/TrEMBL                                    |
| AT5G15160 | Transcription factor PRE2 [Source:UniProtKB/Swiss-Prot                                      |
| AT5G22430 | Pollen Ole e 1 allergen and extensin family protein [Source:UniProtKB/TrEMBL                |
| AT5G22580 | Stress-response A/B barrel domain-containing protein At5g22580 [Source:UniProtKB/Swiss-Prot |

|           |   |
|-----------|---|
| AT5G25840 | DUF1677 family protein (DUF1677) [Source:UniProtKB/TrEMBL                         |
| AT5G39210 | Protein CHLORORESPIRATORY REDUCTION 7, chloroplastic [Source:UniProtKB/Swiss-Prot |
| AT5G58310 | Methylesterase 18 [Source:UniProtKB/Swiss-Prot                                    |
| AT5G60910 | Agamous-like MADS-box protein AGL8 [Source:UniProtKB/Swiss-Prot                   |
| AT5G67070 | Protein RALF-like 34 [Source:UniProtKB/Swiss-Prot                                 |

### (C) Gènes communs entre vcs-6 DOWN et DAE OE DOWN

| AGI       | Description  |
|-----------|--|
| AT3G55710 | UDP-glycosyltransferase 7F2 [Source:UniProtKB/Swiss-Prot |
| AT4G15660 | Monothiol glutaredoxin-S8 [Source:UniProtKB/Swiss-Prot   |

### (D) Gènes communs entre vcs-6 DOWN et D153N OE DOWN

| AGI       | Description  |
|-----------|--|
| AT1G01600 | Cytochrome P <sub>450</sub> 86A <sub>4</sub> [Source:UniProtKB/Swiss-Prot                            |
| AT1G02380 | T6A9.7 protein [Source:UniProtKB/TrEMBL  |
| AT1G02630 | Equilibrative nucleotide transporter 8 [Source:UniProtKB/Swiss-Prot                                  |
| AT1G05700 | Leucine-rich repeat transmembrane protein kinase protein [Source:TAIR                                |
| AT1G05805 | Transcription factor bHLH128 [Source:UniProtKB/Swiss-Prot  |
| AT1G06100 | Delta-9 desaturase-like 2 protein [Source:UniProtKB/Swiss-Prot                                       |
| AT1G06160 | Ethylene-responsive transcription factor ERF094 [Source:UniProtKB/Swiss-Prot                         |
| AT1G06360 | Fatty acid desaturase family protein [Source:TAIR  |
| AT1G06680 | Oxygen-evolving enhancer protein 2-1, chloroplastic [Source:UniProtKB/Swiss-Prot                     |
| AT1G08090 | High-affinity nitrate transporter 2.1 [Source:UniProtKB/Swiss-Prot                                   |
| AT1G10000 | Ribonuclease H-like superfamily protein [Source:UniProtKB/TrEMBL                                     |
| AT1G10020 | At1g10020 [Source:UniProtKB/TrEMBL   |
| AT1G10760 | Alpha-glucan water dikinase 1, chloroplastic [Source:UniProtKB/Swiss-Prot                            |
| AT1G11303 | G-type lectin S-receptor-like serine/threonine-protein kinase At1g11303 [Source:UniProtKB/Swiss-Prot |
| AT1G11350 | G-type lectin S-receptor-like serine/threonine-protein kinase SD1-13 [Source:UniProtKB/Swiss-Prot    |
| AT1G12570 | At1g12570 [Source:UniProtKB/TrEMBL   |
| AT1G12710 | phloem protein 2-A12 [Source:TAIR  |
| AT1G12730 | GPI transamidase subunit PIG-U [Source:UniProtKB/TrEMBL  |
| AT1G13100 | Cytochrome P <sub>450</sub> 71B29 [Source:UniProtKB/Swiss-Prot                                       |
| AT1G14920 | DELLA protein GAI [Source:UniProtKB/Swiss-Prot   |
| AT1G17920 | Uncharacterized protein At1g17920 (Fragment) [Source:UniProtKB/TrEMBL                                |
| AT1G17990 | Putative 12-oxophytodienoate reductase-like protein 2B [Source:UniProtKB/Swiss-Prot                  |
| AT1G18000 | UNC93-like protein 1 [Source:UniProtKB/Swiss-Prot  |

|           |  |
|-----------|--|
| AT1G18020 | Putative 12-oxophytodienoate reductase-like protein 2B [Source:UniProtKB/Swiss-Prot                                |
| AT1G18620 | unknown protein  |
| AT1G20390 | NA   |
| AT1G21130 | Indole glucosinolate O-methyltransferase 4 [Source:UniProtKB/Swiss-Prot  |
| AT1G21270 | Wall-associated receptor kinase 2 [Source:UniProtKB/Swiss-Prot   |
| AT1G22570 | Protein NRT1/ PTR FAMILY 5.15 [Source:UniProtKB/Swiss-Prot   |
| AT1G22900 | Dirigent protein 11 [Source:UniProtKB/Swiss-Prot   |
| AT1G27285 | NA   |
| AT1G27460 | Protein NPGR1 [Source:UniProtKB/Swiss-Prot   |
| AT1G29020 | Calcium-binding EF-hand family protein [Source:TAIR  |
| AT1G29690 | MACPF domain-containing protein CAD1 [Source:UniProtKB/Swiss-Prot  |
| AT1G29720 | Probable LRR receptor-like serine/threonine-protein kinase At1g29720 [Source:UniProtKB/Swiss-Prot                  |
| AT1G29910 | Chlorophyll a-b binding protein 3, chloroplastic [Source:UniProtKB/Swiss-Prot                                      |
| AT1G29930 | Chlorophyll a-b binding protein 1, chloroplastic [Source:UniProtKB/Swiss-Prot                                      |
| AT1G33240 | GT-2-like 1 [Source:TAIR   |
| AT1G44050 | Cysteine/Histidine-rich C1 domain family protein [Source:UniProtKB/TrEMBL  |
| AT1G49830 | Basic helix-loop-helix (BHLH) DNA-binding superfamily protein [Source:UniProtKB/TrEMBL                             |
| AT1G50890 | TORTIFOLIA1-like protein 1 [Source:UniProtKB/Swiss-Prot  |
| AT1G51940 | LysM domain receptor-like kinase 3 [Source:UniProtKB/Swiss-Prot  |
| AT1G53430 | Probable LRR receptor-like serine/threonine-protein kinase At1g53430 [Source:UniProtKB/Swiss-Prot                  |
| AT1G53440 | Probable LRR receptor-like serine/threonine-protein kinase At1g53440 [Source:UniProtKB/Swiss-Prot                  |
| AT1G55450 | S-adenosyl-L-methionine-dependent methyltransferases superfamily protein [Source:UniProtKB/TrEMBL                  |
| AT1G56510 | Disease resistance protein ADR2 [Source:UniProtKB/Swiss-Prot   |
| AT1G56540 | Disease resistance protein (TIR-NBS-LRR class) family [Source:UniProtKB/TrEMBL                                     |
| AT1G59870 | ABC transporter G family member 36 [Source:UniProtKB/Swiss-Prot  |
| AT1G59910 | Formin-like protein 7 [Source:UniProtKB/Swiss-Prot   |
| AT1G60095 | Mannose-binding lectin superfamily protein [Source:TAIR  |
| AT1G60460 | unknown protein  |
| AT1G61520 | Chlorophyll a-b binding protein, chloroplastic [Source:UniProtKB/TrEMBL  |
| AT1G62500 | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein [Source:UniProtKB/TrEMBL |
| AT1G63710 | Cytochrome P450 86A7 [Source:UniProtKB/Swiss-Prot  |
| AT1G65790 | Receptor-like serine/threonine-protein kinase SD1-7 [Source:UniProtKB/Swiss-Prot                                   |
| AT1G66060 | Uncharacterized protein F15E12.9 [Source:UniProtKB/TrEMBL  |
| AT1G68520 | Zinc finger protein CONSTANS-LIKE 6 [Source:UniProtKB/Swiss-Prot   |
| AT1G70090 | Probable galacturonosyltransferase-like g [Source:UniProtKB/Swiss-Prot   |
| AT1G70520 | Cysteine-rich receptor-like protein kinase 2 [Source:UniProtKB/Swiss-Prot  |
| AT1G70740 | At1g70740 [Source:UniProtKB/TrEMBL   |
| AT1G70820 | At1g70820 [Source:UniProtKB/TrEMBL   |
| AT1G71970 | At1g71970/F17M19_12 [Source:UniProtKB/TrEMBL   |
| AT1G72930 | Toll/interleukin-1 receptor-like protein [Source:UniProtKB/Swiss-Prot  |
| AT1G74810 | Putative boron transporter 5 [Source:UniProtKB/Swiss-Prot  |



|           |   |
|-----------|---|
| AT1G75960 | Probable acyl-activating enzyme 8 [Source:UniProtKB/Swiss-Prot  |
| AT1G76110 | High mobility group B protein 9 [Source:UniProtKB/Swiss-Prot  |
| AT1G77210 | Sugar transport protein 14 [Source:UniProtKB/Swiss-Prot   |
| AT1G79840 | HD-ZIP IV family of homeobox-leucine zipper protein with lipid-binding START domain-containing protein [Source:UniProtKB/TrEMBL |
| AT2G02070 | Protein indeterminate-domain 5, chloroplastic [Source:UniProtKB/Swiss-Prot  |
| AT2G02950 | Protein PHYTOCHROME KINASE SUBSTRATE 1 [Source:UniProtKB/Swiss-Prot   |
| AT2G13610 | ABC transporter G family member 5 [Source:UniProtKB/Swiss-Prot  |
| AT2G15040 | NA  |
| AT2G15090 | 3-ketoacyl-CoA synthase 8 [Source:UniProtKB/Swiss-Prot  |
| AT2G15880 | Pollen-specific leucine-rich repeat extensin-like protein 3 [Source:UniProtKB/Swiss-Prot  |
| AT2G18700 | Probable alpha,alpha-trehalose-phosphate synthase [UDP-forming] 11 [Source:UniProtKB/Swiss-Prot                                 |
| AT2G20835 | At2g20835 [Source:UniProtKB/TrEMBL  |
| AT2G23290 | At2g23280 [Source:UniProtKB/TrEMBL  |
| AT2G26910 | PEC1 [Source:UniProtKB/TrEMBL   |
| AT2G28630 | 3-ketoacyl-CoA synthase 12 [Source:UniProtKB/Swiss-Prot   |
| AT2G29290 | Tropinone reductase homolog At2g29290 [Source:UniProtKB/Swiss-Prot  |
| AT2G30432 | TCL1 [Source:UniProtKB/TrEMBL   |
| AT2G34090 | maternal effect embryo arrest 18 [Source:TAIR   |
| AT2G34420 | Chlorophyll a-b binding protein, chloroplastic [Source:UniProtKB/TrEMBL   |
| AT2G34510 | Uncharacterized protein At2g34510 [Source:UniProtKB/TrEMBL  |
| AT2G34960 | CAT5 [Source:UniProtKB/TrEMBL   |
| AT2G35650 | Glucomannan 4-beta-mannosyltransferase 7 [Source:UniProtKB/Swiss-Prot   |
| AT2G36490 | Protein ROS1 [Source:UniProtKB/Swiss-Prot   |
| AT2G37260 | WRKY transcription factor 44 [Source:UniProtKB/Swiss-Prot   |
| AT2G39580 | CONTAINS InterPro DOMAIN/s: Putative zinc-finger domain (InterPro:IPR019607)  |
| AT2G40030 | DNA-directed RNA polymerase V subunit 1 [Source:UniProtKB/Swiss-Prot  |
| AT2G40750 | Probable WRKY transcription factor 54 [Source:UniProtKB/Swiss-Prot  |
| AT2G40960 | Expressed protein [Source:UniProtKB/TrEMBL  |
| AT2G41560 | Calcium-transporting ATPase 4, plasma membrane-type [Source:UniProtKB/Swiss-Prot  |
| AT2G41990 | At2g41990 [Source:UniProtKB/TrEMBL  |
| AT2G42380 | Basic leucine zipper 34 [Source:UniProtKB/Swiss-Prot  |
| AT2G43010 | phytochrome interacting factor 4 [Source:TAIR   |
| AT2G44230 | At2g44230/F4l1.4 [Source:UniProtKB/TrEMBL   |
| AT2G44500 | O-fucosyltransferase 20 [Source:UniProtKB/Swiss-Prot  |
| AT2G44740 | Cyclin-U4-1 [Source:UniProtKB/Swiss-Prot  |
| AT2G45850 | AT-hook motif nuclear-localized protein 9 [Source:UniProtKB/Swiss-Prot  |
| AT3G02130 | LRR receptor-like serine/threonine-protein kinase RPK2 [Source:UniProtKB/Swiss-Prot   |
| AT3G02170 | Protein LONGIFOLIA 2 [Source:UniProtKB/Swiss-Prot   |
| AT3G06880 | Transducin/WD40 repeat-like superfamily protein [Source:UniProtKB/TrEMBL  |
| AT3G06990 | Cysteine/Histidine-rich C1 domain family protein [Source:UniProtKB/TrEMBL   |

|           |  |
|-----------|--|
| AT3G08920 | Rhodanese-like domain-containing protein 10 [Source:UniProtKB/Swiss-Prot                         |
| AT3G08940 | Chlorophyll a-b binding protein CP29.2, chloroplastic [Source:UniProtKB/Swiss-Prot               |
| AT3G11110 | RING-H2 finger protein ATL66 [Source:UniProtKB/Swiss-Prot  |
| AT3G11820 | SYR1 [Source:UniProtKB/TrEMBL  |
| AT3G13000 | AT3G13000 protein [Source:UniProtKB/TrEMBL   |
| AT3G13065 | Protein STRUBBELIG-RECEPTOR FAMILY 4 [Source:UniProtKB/Swiss-Prot                                |
| AT3G16670 | AT3g16670/MGL6_12 [Source:UniProtKB/TrEMBL   |
| AT3G19680 | Uncharacterized protein At3g19680 [Source:UniProtKB/TrEMBL                                       |
| AT3G26490 | BTB/POZ domain-containing protein At3g26490 [Source:UniProtKB/Swiss-Prot                         |
| AT3G28180 | Glycosyltransferase (Fragment) [Source:UniProtKB/TrEMBL  |
| AT3G29320 | Alpha-glucan phosphorylase 1 [Source:UniProtKB/Swiss-Prot  |
| AT3G30340 | WAT1-related protein At3g30340 [Source:UniProtKB/Swiss-Prot                                      |
| AT3G45780 | RPT1 [Source:UniProtKB/TrEMBL  |
| AT3G47470 | Chlorophyll a-b binding protein, chloroplastic [Source:UniProtKB/TrEMBL                          |
| AT3G47620 | Transcription factor TCP14 [Source:UniProtKB/Swiss-Prot  |
| AT3G49500 | RNA-dependent RNA polymerase 6 [Source:UniProtKB/Swiss-Prot                                      |
| AT3G50950 | Disease resistance RPP13-like protein 4 [Source:UniProtKB/Swiss-Prot                             |
| AT3G52550 | At3g52550 [Source:UniProtKB/TrEMBL   |
| AT3G53800 | Fes1B [Source:UniProtKB/TrEMBL   |
| AT3G54460 | F-box protein At3g54460 [Source:UniProtKB/Swiss-Prot   |
| AT3G54810 | GATA transcription factor [Source:UniProtKB/TrEMBL   |
| AT3G55710 | UDP-glycosyltransferase 76F2 [Source:UniProtKB/Swiss-Prot  |
| AT3G56400 | Probable WRKY transcription factor 70 [Source:UniProtKB/Swiss-Prot                               |
| AT3G56410 | Uncharacterized protein T5P19_60 [Source:UniProtKB/TrEMBL  |
| AT3G57300 | DNA helicase INO80-like protein [Source:UniProtKB/TrEMBL   |
| AT3G57470 | Insulin-degrading enzyme-like 2 [Source:UniProtKB/Swiss-Prot                                     |
| AT3G57530 | Calcium-dependent protein kinase 32 [Source:UniProtKB/Swiss-Prot                                 |
| AT3G57780 | Nucleolar-like protein [Source:UniProtKB/TrEMBL  |
| AT3G58120 | BZIP61 [Source:UniProtKB/TrEMBL  |
| AT3G58620 | TTL4 [Source:UniProtKB/TrEMBL  |
| AT3G59080 | AT3g59080/F17J16_130 [Source:UniProtKB/TrEMBL  |
| AT3G59110 | Protein kinase superfamily protein [Source:UniProtKB/TrEMBL                                      |
| AT3G59310 | Solute carrier family 35 protein (DUF914) [Source:UniProtKB/TrEMBL                               |
| AT3G60290 | 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein [Source:UniProtKB/TrEMBL |
| AT3G60320 | Nitrate regulatory gene2 protein [Source:UniProtKB/Swiss-Prot                                    |
| AT3G61470 | Photosystem I chlorophyll a/b-binding protein 2, chloroplastic [Source:UniProtKB/Swiss-Prot      |
| AT3G61840 | Auxin response factor, putative (DUF688) [Source:UniProtKB/TrEMBL                                |
| AT4G01330 | Protein kinase superfamily protein [Source:UniProtKB/TrEMBL                                      |
| AT4G01950 | glycerol-3-phosphate acyltransferase 3 [Source:TAIR  |
| AT4G03110 | RNA-binding protein BRN1 [Source:UniProtKB/Swiss-Prot  |
| AT4G04340 | Protein OSCA1 [Source:UniProtKB/Swiss-Prot   |

|           |  |
|-----------|--|
| AT4G04570 | Cysteine-rich receptor-like protein kinase 40 [Source:UniProtKB/Swiss-Prot   |
| AT4G08115 | NA   |
| AT4G08620 | Sulfate transporter 1.1 [Source:UniProtKB/Swiss-Prot   |
| AT4G08950 | Protein EXORDIUM [Source:UniProtKB/Swiss-Prot  |
| AT4G09160 | Patellin-5 [Source:UniProtKB/Swiss-Prot  |
| AT4G13210 | Putative pectate lyase 14 [Source:UniProtKB/Swiss-Prot   |
| AT4G16857 | Disease resistance protein (TIR-NBS-LRR class) family protein [Source:UniProtKB/TrEMBL                             |
| AT4G16860 | Disease resistance protein RPP4 [Source:UniProtKB/Swiss-Prot   |
| AT4G16870 | NA   |
| AT4G16910 | NA   |
| AT4G16950 | Disease resistance protein RPP5 [Source:UniProtKB/Swiss-Prot   |
| AT4G17460 | Homeobox-leucine zipper protein HAT1 [Source:UniProtKB/Swiss-Prot  |
| AT4G19520 | Probable disease resistance protein At4g19520 [Source:UniProtKB/Swiss-Prot   |
| AT4G19530 | disease resistance protein (TIR-NBS-LRR class) family [Source:TAIR   |
| AT4G22513 | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein [Source:UniProtKB/TrEMBL |
| AT4G22517 | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein [Source:UniProtKB/TrEMBL |
| AT4G22520 | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein [Source:UniProtKB/TrEMBL |
| AT4G23180 | Cysteine-rich receptor-like protein kinase 10 [Source:UniProtKB/Swiss-Prot   |
| AT4G24670 | Tryptophan aminotransferase-related protein 2 [Source:UniProtKB/Swiss-Prot   |
| AT4G25350 | Phosphate transporter PHO1 homolog 4 [Source:UniProtKB/Swiss-Prot  |
| AT4G26530 | Fructose-bisphosphate aldolase [Source:UniProtKB/TrEMBL  |
| AT4G26540 | LRR receptor-like serine/threonine-protein kinase [Source:UniProtKB/Swiss-Prot                                     |
| AT4G27440 | Protochlorophyllide reductase B, chloroplastic [Source:UniProtKB/Swiss-Prot  |
| AT4G27870 | Vacuolar iron transporter (VIT) family protein [Source:TAIR  |
| AT4G28080 | Protein TSS [Source:UniProtKB/Swiss-Prot   |
| AT4G28180 | At4g28180 [Source:UniProtKB/TrEMBL   |
| AT4G30440 | GAE1 [Source:UniProtKB/TrEMBL  |
| AT4G34390 | XLG2 [Source:UniProtKB/TrEMBL  |
| AT4G35470 | Plant intracellular Ras-group-related LRR protein 4 [Source:UniProtKB/Swiss-Prot                                   |
| AT4G36030 | Armadillo repeat only 3 [Source:UniProtKB/TrEMBL   |
| AT4G37080 | Protein of unknown function, DUF547 [Source:TAIR   |
| AT4G38550 | Arabidopsis phospholipase-like protein (PEARL1 4) family [Source:TAIR  |
| AT4G38950 | Kinesin-like protein KIN-7F [Source:UniProtKB/Swiss-Prot   |
| AT4G39770 | Probable trehalose-phosphate phosphatase H [Source:UniProtKB/Swiss-Prot  |
| AT5G03140 | L-type lectin-domain containing receptor kinase VIII.2 [Source:UniProtKB/Swiss-Prot                                |
| AT5G03350 | Lectin-like protein [Source:UniProtKB/Swiss-Prot   |
| AT5G08330 | Transcription factor TCP21 [Source:UniProtKB/Swiss-Prot  |
| AT5G10250 | Phototropic-responsive NPH3 family protein [Source:TAIR  |
| AT5G11250 | Disease resistance protein (TIR-NBS-LRR class) [Source:UniProtKB/TrEMBL  |
| AT5G13950 | unknown protein  |
| AT5G21930 | Copper-transporting ATPase PAA2, chloroplastic [Source:UniProtKB/Swiss-Prot  |

|           |  |
|-----------|--|
| AT5G22940 | FRA8 homolog [Source:TAIR  |
| AT5G23130 | Peptidoglycan-binding LysM domain-containing protein [Source:UniProtKB/TrEMBL                    |
| AT5G25880 | Malic enzyme [Source:UniProtKB/TrEMBL  |
| AT5G26910 | TRM8 [Source:UniProtKB/TrEMBL  |
| AT5G28450 | Chlorophyll a-b binding protein, chloroplastic [Source:UniProtKB/TrEMBL                          |
| AT5G33370 | GDSL esterase/lipase At5g33370 [Source:UniProtKB/Swiss-Prot                                      |
| AT5G34830 | unknown protein  |
| AT5G37450 | Leucine-rich repeat protein kinase family protein [Source:TAIR                                   |
| AT5G42250 | Alcohol dehydrogenase-like 7 [Source:UniProtKB/Swiss-Prot  |
| AT5G43520 | Cysteine/Histidine-rich C1 domain family protein [Source:UniProtKB/TrEMBL                        |
| AT5G45472 | Potential natural antisense gene, locus overlaps with AT5G45470 [Source:TAIR                     |
| AT5G49770 | Probable leucine-rich repeat receptor-like protein kinase At5g49770 [Source:UniProtKB/Swiss-Prot |
| AT5G52600 | Transcription factor MYB82 [Source:UniProtKB/Swiss-Prot  |
| AT5G53500 | Similarity to unknown protein [Source:UniProtKB/TrEMBL   |
| AT5G54270 | Chlorophyll a-b binding protein 3, chloroplastic [Source:UniProtKB/Swiss-Prot                    |
| AT5G62090 | Probable transcriptional regulator SLK2 [Source:UniProtKB/Swiss-Prot                             |
| AT5G66940 | Dof zinc finger protein DOF5.8 [Source:UniProtKB/Swiss-Prot                                      |
| ATCG00480 | ATP synthase subunit beta, chloroplastic [Source:UniProtKB/Swiss-Prot                            |

#### (E) Gènes communs entre vcs-6 UP et DAE OE UP

| AGI       | Description   |
|-----------|---|
| AT1G01355 | Putative endonuclease or glycosyl hydrolase [Source:UniProtKB/TrEMBL                          |
| AT1G01420 | UDP-glucosyl transferase 72B3 [Source:TAIR  |
| AT1G01470 | LSR3 [Source:UniProtKB/TrEMBL   |
| AT1G02470 | Polyketide cyclase/dehydrase and lipid transport superfamily protein [Source:UniProtKB/TrEMBL |
| AT1G02770 | UPF0725 protein At1g02770 [Source:UniProtKB/Swiss-Prot  |
| AT1G03445 | Serine/threonine-protein phosphatase [Source:UniProtKB/TrEMBL                                 |
| AT1G17710 | Inorganic pyrophosphatase 2 [Source:UniProtKB/Swiss-Prot                                      |
| AT1G20070 | T20H2.15 protein [Source:UniProtKB/TrEMBL   |
| AT1G21520 | Uncharacterized protein At1g21520/F24J8_4 [Source:UniProtKB/TrEMBL                            |
| AT1G21525 | NA  |
| AT1G23840 | At1g23840/F5O8_37 [Source:UniProtKB/TrEMBL  |
| AT1G24575 | At1g24575 [Source:UniProtKB/TrEMBL  |
| AT1G32950 | Subtilisin-like protease SBT3.4 [Source:UniProtKB/Swiss-Prot                                  |
| AT1G33960 | P-loop containing nucleoside triphosphate hydrolases superfamily protein [Source:TAIR         |
| AT1G35230 | Classical arabinogalactan protein 5 [Source:UniProtKB/Swiss-Prot                              |
| AT1G43605 |   |
| AT1G45616 | Receptor-like protein 6 [Source:UniProtKB/Swiss-Prot  |
| AT1G54575 | unknown protein   |

|           |  |
|-----------|--|
| AT1G55920 | Serine acetyltransferase 1, chloroplastic [Source:UniProtKB/Swiss-Prot                                   |
| AT1G63990 | Meiotic recombination protein SPO11-2 [Source:UniProtKB/Swiss-Prot                                       |
| AT1G64795 |  |
| AT1G65480 | PEBP (phosphatidylethanolamine-binding protein) family protein [Source:TAIR                              |
| AT1G66920 | LEAF RUST 10 DISEASE-RESISTANCE LOCUS RECEPTOR-LIKE PROTEIN KINASE-like 2.4 [Source:UniProtKB/Swiss-Prot |
| AT1G67810 | SufE-like protein 2, chloroplastic [Source:UniProtKB/Swiss-Prot  |
| AT1G68350 | At1g68350 [Source:UniProtKB/TrEMBL   |
| AT1G69720 | Heme oxygenase 3, chloroplastic [Source:UniProtKB/Swiss-Prot   |
| AT1G73120 | At1g73120 [Source:UniProtKB/TrEMBL   |
| AT1G75920 | GDSL-like Lipase/Acylhydrolase superfamily protein [Source:UniProtKB/TrEMBL                              |
| AT1G76960 | F22K20.6 protein [Source:UniProtKB/TrEMBL  |
| AT2G02820 | Transcription factor MYB88 [Source:UniProtKB/Swiss-Prot  |
| AT2G04040 | Protein DETOXIFICATION [Source:UniProtKB/TrEMBL  |
| AT2G04050 | Protein DETOXIFICATION 3 [Source:UniProtKB/Swiss-Prot  |
| AT2G04460 | NA   |
| AT2G05380 | glycine-rich protein 3 short isoform [Source:TAIR  |
| AT2G05510 | At2g05510 [Source:UniProtKB/TrEMBL   |
| AT2G05914 | Potential natural antisense gene, locus overlaps with AT2G05915 [Source:TAIR                             |
| AT2G25000 | Probable WRKY transcription factor 60 [Source:UniProtKB/Swiss-Prot                                       |
| AT2G25440 | Receptor-like protein 20 [Source:UniProtKB/Swiss-Prot  |
| AT2G26750 | Alpha/beta-Hydrolases superfamily protein [Source:UniProtKB/TrEMBL                                       |
| AT2G32680 | Receptor like protein 23 [Source:UniProtKB/Swiss-Prot  |
| AT2G33080 |  |
| AT2G34490 | Cytochrome P450 710A2 [Source:UniProtKB/Swiss-Prot   |
| AT2G36970 | Glycosyltransferase (Fragment) [Source:UniProtKB/TrEMBL  |
| AT2G37770 | NADPH-dependent aldo-keto reductase, chloroplastic [Source:UniProtKB/Swiss-Prot                          |
| AT2G43570 | CHI [Source:UniProtKB/TrEMBL   |
| AT3G01970 | Probable WRKY transcription factor 45 [Source:UniProtKB/Swiss-Prot                                       |
| AT3G04070 | NAC transcription factor 47 [Source:UniProtKB/Swiss-Prot   |
| AT3G11150 | 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein [Source:UniProtKB/TrEMBL         |
| AT3G13784 | Beta-fructofuranosidase, insoluble isoenzyme CWINV5 [Source:UniProtKB/Swiss-Prot                         |
| AT3G14710 | F-box/FBD/LRR-repeat protein At3g14710 [Source:UniProtKB/Swiss-Prot                                      |
| AT3G17050 | NA   |
| AT3G18610 | Nucleolin 2 [Source:UniProtKB/Swiss-Prot   |
| AT3G21080 | ABC transporter-like protein [Source:UniProtKB/TrEMBL  |
| AT3G25010 | Receptor like protein 41 [Source:UniProtKB/TrEMBL  |
| AT3G25020 | Receptor-like protein 42 [Source:UniProtKB/Swiss-Prot  |
| AT3G26320 | Cytochrome P450 71B36 [Source:UniProtKB/Swiss-Prot   |
| AT3G26840 | Acyltransferase-like protein At3g26840, chloroplastic [Source:UniProtKB/Swiss-Prot                       |
| AT3G28210 | Zinc finger AN1 domain-containing stress-associated protein 12 [Source:UniProtKB/Swiss-Prot              |

|           |  |
|-----------|--|
| AT3G28540 | AAA-ATPase At3g28540 [Source:UniProtKB/Swiss-Prot  |
| AT3G28890 | Receptor-like protein 43 [Source:UniProtKB/Swiss-Prot  |
| AT3G50770 | Probable calcium-binding protein CML41 [Source:UniProtKB/Swiss-Prot                                  |
| AT4G02380 | senescence-associated gene 21 [Source:TAIR   |
| AT4G08870 | ARGAH2 [Source:UniProtKB/TrEMBL  |
| AT4G11890 | Protein kinase superfamily protein [Source:UniProtKB/TrEMBL  |
| AT4G11910 | Protein STAY-GREEN 2, chloroplastic [Source:UniProtKB/Swiss-Prot                                     |
| AT4G13900 | NA   |
| AT4G13920 | Receptor-like protein 50 [Source:UniProtKB/Swiss-Prot  |
| AT4G14020 | At4g14020 [Source:UniProtKB/TrEMBL   |
| AT4G23140 | cysteine-rich RLK (RECEPTOR-like protein kinase) 6 [Source:TAIR                                      |
| AT4G23310 | Putative cysteine-rich receptor-like protein kinase 23 [Source:UniProtKB/Swiss-Prot                  |
| AT4G27150 | SESA2 [Source:UniProtKB/TrEMBL   |
| AT4G27480 | Core-2/l-branching beta-1,6-N-acetylglucosaminyltransferase family protein [Source:UniProtKB/TrEMBL  |
| AT4G31370 | Fasciclin-like arabinogalactan protein 5 [Source:UniProtKB/Swiss-Prot                                |
| AT4G35720 | DUF241 domain protein, putative (DUF241) [Source:UniProtKB/TrEMBL                                    |
| AT4G36700 | Vicilin-like seed storage protein At4g36700 [Source:UniProtKB/Swiss-Prot                             |
| AT4G38560 | Phospholipase like protein [Source:UniProtKB/TrEMBL  |
| AT4G39480 | Cytochrome P450, family g6, subfamily A, polypeptide g [Source:UniProtKB/TrEMBL                      |
| AT5G03545 | At5g03545 [Source:UniProtKB/TrEMBL   |
| AT5G05960 | At5g05960 [Source:UniProtKB/TrEMBL   |
| AT5G07560 | GRP20 [Source:UniProtKB/TrEMBL   |
| AT5G10625 | Flowering-promoting factor 1-like protein 2 [Source:UniProtKB/Swiss-Prot                             |
| AT5G10760 | Aspartyl protease AED1 [Source:UniProtKB/Swiss-Prot  |
| AT5G11300 | Cyclin-A2-2 [Source:UniProtKB/Swiss-Prot   |
| AT5G12030 | 17.7 kDa class II heat shock protein [Source:UniProtKB/Swiss-Prot                                    |
| AT5G15960 | Stress-induced protein KIN1 [Source:UniProtKB/Swiss-Prot   |
| AT5G17860 | Cation/calcium exchanger 1 [Source:UniProtKB/Swiss-Prot  |
| AT5G23950 | Calcium-dependent lipid-binding (CaLB domain) family protein [Source:UniProtKB/TrEMBL                |
| AT5G24080 | G-type lectin S-receptor-like serine/threonine-protein kinase At5g24080 [Source:UniProtKB/Swiss-Prot |
| AT5G24110 | Probable WRKY transcription factor 30 [Source:UniProtKB/Swiss-Prot                                   |
| AT5G24200 | alpha/beta-Hydrolases superfamily protein [Source:TAIR   |
| AT5G25250 | Flotillin-like protein 1 [Source:UniProtKB/Swiss-Prot  |
| AT5G26270 | unknown protein  |
| AT5G47550 | Cysteine proteinase inhibitor 5 [Source:UniProtKB/Swiss-Prot   |
| AT5G47590 | Heat shock protein HSP20/alpha crystallin family [Source:UniProtKB/TrEMBL                            |
| AT5G52390 | PAR1 protein [Source:UniProtKB/TrEMBL  |
| AT5G53030 | Uncharacterized protein At5g53030 [Source:UniProtKB/TrEMBL   |
| AT5G59310 | Non-specific lipid-transfer protein 4 [Source:UniProtKB/Swiss-Prot                                   |
| AT5G60250 | RBR-type E3 ubiquitin transferase [Source:UniProtKB/TrEMBL   |
| AT5G60260 | unknown protein  |

|           |   |
|-----------|---|
| AT5G61430 | NAC domain-containing protein 100 [Source:UniProtKB/Swiss-Prot] |
|-----------|---|

### (F) Gènes communs entre vcs-6 UP et D153N OE UP

| AGI       | Description  |
|-----------|--|
| AT1G01060 | LHY1 [Source:UniProtKB/TrEMBL]   |
| AT1G01100 | 60S acidic ribosomal protein P1-1 [Source:UniProtKB/Swiss-Prot]                                    |
| AT1G01355 | Putative endonuclease or glycosyl hydrolase [Source:UniProtKB/TrEMBL]                              |
| AT1G01470 | LSR3 [Source:UniProtKB/TrEMBL]   |
| AT1G01630 | Polyphosphoinositide binding protein, putative [Source:UniProtKB/TrEMBL]                           |
| AT1G01810 | T1N6.23 [Source:UniProtKB/TrEMBL]  |
| AT1G02770 | UPF0725 protein At1g02770 [Source:UniProtKB/Swiss-Prot]  |
| AT1G02780 | 60S ribosomal protein L19-1 [Source:UniProtKB/Swiss-Prot]  |
| AT1G03106 | At1g03105 [Source:UniProtKB/TrEMBL]  |
| AT1G03445 | Serine/threonine-protein phosphatase [Source:UniProtKB/TrEMBL]                                     |
| AT1G03710 | Cystatin/monellin superfamily protein [Source:UniProtKB/TrEMBL]                                    |
| AT1G03760 | Prefoldin chaperone subunit family protein [Source:UniProtKB/TrEMBL]                               |
| AT1G03770 | RING 1B [Source:UniProtKB/TrEMBL]  |
| AT1G04033 |  |
| AT1G04380 | 1-aminocyclopropane-1-carboxylate oxidase homolog 7 [Source:UniProtKB/Swiss-Prot]                  |
| AT1G04533 |  |
| AT1G05080 | Putative FBD-associated F-box protein At1g05080 [Source:UniProtKB/Swiss-Prot]                      |
| AT1G05400 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL]   |
| AT1G05720 | Selenoprotein family protein [Source:UniProtKB/TrEMBL]   |
| AT1G05730 | FAM136A-like protein (DUF842) [Source:UniProtKB/TrEMBL]  |
| AT1G06473 |  |
| AT1G06537 |  |
| AT1G07367 |  |
| AT1G07400 | 17.8 kDa class I heat shock protein [Source:UniProtKB/Swiss-Prot]                                  |
| AT1G07590 | Pentatricopeptide repeat-containing protein At1g07590, mitochondrial [Source:UniProtKB/Swiss-Prot] |
| AT1G07770 | 40S ribosomal protein S15a-1 [Source:UniProtKB/Swiss-Prot]   |
| AT1G07902 |  |
| AT1G07985 | Expressed protein [Source:UniProtKB/TrEMBL]  |
| AT1G08035 | Cyclin-dependent protein kinase inhibitor SMR16 [Source:UniProtKB/Swiss-Prot]                      |
| AT1G08080 | Alpha carbonic anhydrase 7 [Source:UniProtKB/Swiss-Prot]   |
| AT1G08310 | Alpha/beta-Hydrolases superfamily protein [Source:UniProtKB/TrEMBL]                                |
| AT1G08757 |  |
| AT1G09310 | T31J12.3 protein [Source:UniProtKB/TrEMBL]   |
| AT1G09486 | NA   |
| AT1G09500 | At1g09500/F14J9_16 [Source:UniProtKB/TrEMBL]   |

|           |   |
|-----------|---|
| AT1G09797 |   |
| AT1G09867 |   |
| AT1G10865 | Cytochrome C oxidase assembly factor [Source:UniProtKB/TrEMBL   |
| AT1G10960 | Ferredoxin-1, chloroplastic [Source:UniProtKB/Swiss-Prot  |
| AT1G11175 | other RNA [Source:TAIR  |
| AT1G11230 | T28P6.11 protein [Source:UniProtKB/TrEMBL   |
| AT1G11710 | Pentatricopeptide repeat-containing protein At1g11710, mitochondrial [Source:UniProtKB/Swiss-Prot     |
| AT1G12064 | Transmembrane protein [Source:UniProtKB/TrEMBL  |
| AT1G12080 | At1g12080 [Source:UniProtKB/TrEMBL  |
| AT1G12211 | unknown protein   |
| AT1G12810 | Proline-rich family protein [Source:UniProtKB/TrEMBL  |
| AT1G14200 | At1g14200 [Source:UniProtKB/TrEMBL  |
| AT1G14680 | Early endosome antigen [Source:UniProtKB/TrEMBL   |
| AT1G15230 | At1g15230/F9L1_18 [Source:UniProtKB/TrEMBL  |
| AT1G15270 | F9L1.21 protein [Source:UniProtKB/TrEMBL  |
| AT1G15700 | ATP synthase gamma chain 2, chloroplastic [Source:UniProtKB/Swiss-Prot                                |
| AT1G15760 | At1g15760 [Source:UniProtKB/TrEMBL  |
| AT1G15930 | 40S ribosomal protein S12-1 [Source:UniProtKB/Swiss-Prot  |
| AT1G16489 | other RNA [Source:TAIR  |
| AT1G17530 | Mitochondrial import inner membrane translocase subunit TIM23-1 [Source:UniProtKB/Swiss-Prot          |
| AT1G17710 | Inorganic pyrophosphatase 2 [Source:UniProtKB/Swiss-Prot  |
| AT1G17745 | D-3-phosphoglycerate dehydrogenase [Source:TAIR   |
| AT1G18320 | Mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein [Source:TAIR |
| AT1G18773 | FUNCTIONS IN: molecular_function unknown  |
| AT1G19640 | Jasmonate O-methyltransferase [Source:UniProtKB/Swiss-Prot  |
| AT1G20070 | T20H2.15 protein [Source:UniProtKB/TrEMBL   |
| AT1G20120 | GDSL esterase/lipase At1g20120 [Source:UniProtKB/Swiss-Prot   |
| AT1G20220 | Alba DNA/RNA-binding protein [Source:UniProtKB/TrEMBL   |
| AT1G20610 | Cyclin-B2-3 [Source:UniProtKB/Swiss-Prot  |
| AT1G20795 | Putative F-box protein At1g20795 [Source:UniProtKB/Swiss-Prot   |
| AT1G20930 | Cyclin-dependent kinase B2-2 [Source:UniProtKB/Swiss-Prot   |
| AT1G21190 | Sm-like protein LSM3A [Source:UniProtKB/Swiss-Prot  |
| AT1G21290 | NA  |
| AT1G21520 | Uncharacterized protein At1g21520/F24J8_4 [Source:UniProtKB/TrEMBL                                    |
| AT1G21525 | NA  |
| AT1G21900 | Transmembrane emp24 domain-containing protein p24delta5 [Source:UniProtKB/Swiss-Prot                  |
| AT1G22210 | Probable trehalose-phosphate phosphatase C [Source:UniProtKB/Swiss-Prot                               |
| AT1G22403 | other RNA [Source:TAIR  |
| AT1G23090 | Probable sulfate transporter 3.3 [Source:UniProtKB/Swiss-Prot   |
| AT1G23110 | Fold protein [Source:UniProtKB/TrEMBL   |
| AT1G23750 | F5O8.30 protein [Source:UniProtKB/TrEMBL  |



|           |  |
|-----------|--|
| AT1G23840 | At1g23840/F5O8_37 [Source:UniProtKB/TrEMBL   |
| AT1G24260 | At1g24260 [Source:UniProtKB/TrEMBL   |
| AT1G24575 | At1g24575 [Source:UniProtKB/TrEMBL   |
| AT1G24580 | F21J9.24 [Source:UniProtKB/TrEMBL  |
| AT1G25277 |  |
| AT1G26665 | Mediator complex, subunit Med10 [Source:TAIR   |
| AT1G26770 | Expansin [Source:UniProtKB/TrEMBL  |
| AT1G26910 | 60S ribosomal protein L10-2 [Source:UniProtKB/Swiss-Prot   |
| AT1G27030 | At1g27030/T7N9_9 [Source:UniProtKB/TrEMBL  |
| AT1G27580 | Protein of unknown function (DUF295) [Source:TAIR  |
| AT1G28570 | GDSL esterase/lipase At1g28570 [Source:UniProtKB/Swiss-Prot  |
| AT1G29195 | At1g29190/F28N24_12 [Source:UniProtKB/TrEMBL   |
| AT1G29300 | F28N24.3 protein [Source:UniProtKB/TrEMBL  |
| AT1G29640 | At1g29640 [Source:UniProtKB/TrEMBL   |
| AT1G30070 | SGS domain-containing protein [Source:UniProtKB/TrEMBL   |
| AT1G30100 | Probable 9-cis-epoxycarotenoid dioxygenase NCED5, chloroplastic [Source:UniProtKB/Swiss-Prot                       |
| AT1G30160 | At1g30160 [Source:UniProtKB/TrEMBL   |
| AT1G30170 | Uncharacterized protein T2H7.3 [Source:UniProtKB/TrEMBL  |
| AT1G30880 | At1g30880 [Source:UniProtKB/TrEMBL   |
| AT1G31240 | At1g31240 [Source:UniProtKB/TrEMBL   |
| AT1G31335 | At1g31335 [Source:UniProtKB/TrEMBL   |
| AT1G31900 | NA   |
| AT1G34418 | other RNA [Source:TAIR   |
| AT1G35230 | Classical arabinogalactan protein 5 [Source:UniProtKB/Swiss-Prot   |
| AT1G43160 | Ethylene-responsive transcription factor RAP2-6 [Source:UniProtKB/Swiss-Prot                                       |
| AT1G43170 | 60S ribosomal protein L3-1 [Source:UniProtKB/Swiss-Prot  |
| AT1G43590 | NA   |
| AT1G43605 |  |
| AT1G43997 | NA   |
| AT1G44130 | Eukaryotic aspartyl protease family protein [Source:UniProtKB/TrEMBL   |
| AT1G44414 | At1g44414 [Source:UniProtKB/TrEMBL   |
| AT1G47510 | Type IV inositol polyphosphate 5-phosphatase 11 [Source:UniProtKB/Swiss-Prot                                       |
| AT1G48000 | MYB transcription factor [Source:UniProtKB/TrEMBL  |
| AT1G48400 | F-box/RNI-like/FBD-like domains-containing protein [Source:TAIR  |
| AT1G48750 | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein [Source:UniProtKB/TrEMBL |
| AT1G49140 | NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10-A [Source:UniProtKB/Swiss-Prot                        |
| AT1G49405 | CASP-like protein 5C3 [Source:UniProtKB/Swiss-Prot   |
| AT1G49640 | Probable carboxylesterase 3 [Source:UniProtKB/Swiss-Prot   |
| AT1G50010 | Tubulin alpha-2 chain [Source:UniProtKB/Swiss-Prot   |
| AT1G51650 | ATP synthase subunit epsilon, mitochondrial [Source:UniProtKB/Swiss-Prot   |
| AT1G52450 | F6D8.33 [Source:UniProtKB/TrEMBL   |

|           |  |
|-----------|--|
| AT1G52690 | Late embryogenesis abundant protein 7 [Source:UniProtKB/Swiss-Prot   |
| AT1G52740 | Probable histone H2A variant 3 [Source:UniProtKB/Swiss-Prot  |
| AT1G52827 | At1g52827 [Source:UniProtKB/TrEMBL   |
| AT1G52990 | Thioredoxin family protein [Source:UniProtKB/TrEMBL  |
| AT1G53070 | Lectin-like protein At1g53070 [Source:UniProtKB/Swiss-Prot   |
| AT1G53260 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL  |
| AT1G53460 | At1g53460/T3F20_21 [Source:UniProtKB/TrEMBL  |
| AT1G53540 | 17.6 kDa class I heat shock protein 3 [Source:UniProtKB/Swiss-Prot   |
| AT1G53542 | At1g53545 [Source:UniProtKB/TrEMBL   |
| AT1G54575 | unknown protein  |
| AT1G55152 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL  |
| AT1G55265 | At1g55265 [Source:UniProtKB/TrEMBL   |
| AT1G56045 | 60S ribosomal protein L41 [Source:UniProtKB/Swiss-Prot   |
| AT1G56165 | other RNA [Source:TAIR   |
| AT1G56170 | Nuclear transcription factor Y subunit C-2 [Source:UniProtKB/Swiss-Prot  |
| AT1G56555 | unknown protein  |
| AT1G58160 | Jacalin-related lectin 15 [Source:UniProtKB/Swiss-Prot   |
| AT1G59670 | GSTU15 [Source:UniProtKB/TrEMBL  |
| AT1G59860 | 17.6 kDa class I heat shock protein 1 [Source:UniProtKB/Swiss-Prot   |
| AT1G59920 | F23H11.23 protein [Source:UniProtKB/TrEMBL   |
| AT1G59930 | F23H11.24 protein [Source:UniProtKB/TrEMBL   |
| AT1G60180 | NA   |
| AT1G60970 | Coatomer subunit zeta-1 [Source:UniProtKB/Swiss-Prot   |
| AT1G60989 | Defensin-like protein 249 [Source:UniProtKB/Swiss-Prot   |
| AT1G61255 | At1g61255 [Source:UniProtKB/TrEMBL   |
| AT1G61275 | U12  |
| AT1G61280 | Phosphatidylinositol N-acetylglucosaminyltransferase subunit P [Source:UniProtKB/Swiss-Prot                        |
| AT1G61580 | 60S ribosomal protein L3-2 [Source:UniProtKB/Swiss-Prot  |
| AT1G62370 | F2401.10 [Source:UniProtKB/TrEMBL  |
| AT1G62420 | DUF506 family protein (DUF506) [Source:UniProtKB/TrEMBL  |
| AT1G62480 | At1g62480/T3P18_4 [Source:UniProtKB/TrEMBL   |
| AT1G62510 | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein [Source:UniProtKB/TrEMBL |
| AT1G62690 | unknown protein  |
| AT1G63990 | Meiotic recombination protein SPO11-2 [Source:UniProtKB/Swiss-Prot   |
| AT1G64795 |  |
| AT1G65032 | unknown protein  |
| AT1G65150 | T23K8.6 [Source:UniProtKB/TrEMBL   |
| AT1G65480 | PEBP (phosphatidylethanolamine-binding protein) family protein [Source:TAIR  |
| AT1G65500 | F514.4 protein [Source:UniProtKB/TrEMBL  |
| AT1G65720 | At1g65720/F1E22_13 [Source:UniProtKB/TrEMBL  |
| AT1G66040 | Putative E3 ubiquitin-protein ligase ORTHRUS 4 [Source:UniProtKB/Swiss-Prot  |

|           |   |
|-----------|---|
| AT1G66240 | homolog of anti-oxidant 1 [Source:TAIR  |
| AT1G66580 | 60S ribosomal protein L10-3 [Source:UniProtKB/Swiss-Prot                              |
| AT1G66780 | Protein DETOXIFICATION [Source:UniProtKB/TrEMBL                                       |
| AT1G67265 | DVL3 [Source:UniProtKB/TrEMBL   |
| AT1G67760 | TCP-1/cpn60 chaperonin family protein [Source:UniProtKB/TrEMBL                        |
| AT1G68300 | Adenine nucleotide alpha hydrolases-like superfamily protein [Source:UniProtKB/TrEMBL |
| AT1G68350 | At1g68350 [Source:UniProtKB/TrEMBL  |
| AT1G68568 | other RNA [Source:TAIR  |
| AT1G69252 | other RNA [Source:TAIR  |
| AT1G69485 | Ribosomal L32p protein family [Source:UniProtKB/TrEMBL                                |
| AT1G69620 | 60S ribosomal protein L34-2 [Source:UniProtKB/Swiss-Prot                              |
| AT1G69630 | F-box/LRR-repeat protein 13 [Source:UniProtKB/Swiss-Prot                              |
| AT1G70190 | F2oP5.9 protein [Source:UniProtKB/TrEMBL  |
| AT1G71000 | Chaperone DnaJ-domain superfamily protein [Source:UniProtKB/TrEMBL                    |
| AT1G71890 | Sucrose transport protein SUC5 [Source:UniProtKB/Swiss-Prot                           |
| AT1G72260 | Thionin-2.1 [Source:UniProtKB/Swiss-Prot  |
| AT1G72660 | Developmentally-regulated G-protein 2 [Source:UniProtKB/Swiss-Prot                    |
| AT1G74000 | Protein STRICTOSIDINE SYNTHASE-LIKE 11 [Source:UniProtKB/Swiss-Prot                   |
| AT1G74205 | Potential natural antisense gene, locus overlaps with AT1G74210 [Source:TAIR          |
| AT1G74310 | Chaperone protein ClpB1 [Source:UniProtKB/Swiss-Prot                                  |
| AT1G74870 | At1g74870 [Source:UniProtKB/TrEMBL  |
| AT1G75040 | Pathogenesis-related protein 5 [Source:UniProtKB/Swiss-Prot                           |
| AT1G75335 | Nucleolar-like protein [Source:UniProtKB/TrEMBL                                       |
| AT1G75750 | GASA1 [Source:UniProtKB/TrEMBL  |
| AT1G75900 | GDSL esterase/lipase EXL3 [Source:UniProtKB/Swiss-Prot                                |
| AT1G75920 | GDSL-like Lipase/Acylhydrolase superfamily protein [Source:UniProtKB/TrEMBL           |
| AT1G76540 | CDKB2 [Source:UniProtKB/TrEMBL  |
| AT1G77540 | Acetyltransferase At1g77540 [Source:UniProtKB/Swiss-Prot                              |
| AT1G77680 | Inactive exonuclease DIS3L2 [Source:UniProtKB/Swiss-Prot                              |
| AT1G77710 | Ubiquitin-fold modifier 1 [Source:UniProtKB/TrEMBL                                    |
| AT1G78020 | FCS-Like Zinc finger 6 [Source:UniProtKB/Swiss-Prot                                   |
| AT1G78030 | unknown protein   |
| AT1G78290 | SRK2C [Source:UniProtKB/TrEMBL  |
| AT1G78760 | Putative F-box/FBD/LRR-repeat protein At1g78760 [Source:UniProtKB/Swiss-Prot          |
| AT1G79075 | other RNA [Source:TAIR  |
| AT1G79210 | Proteasome subunit alpha type-2-B [Source:UniProtKB/Swiss-Prot                        |
| AT1G80745 | Transcription factor TFIIIC, tau55-related protein [Source:UniProtKB/TrEMBL           |
| AT2G01150 | RING-H2 finger protein 2B [Source:TAIR  |
| AT2G01175 | At2g01175 [Source:UniProtKB/TrEMBL  |
| AT2G01422 |   |
| AT2G02100 | PDF2.2 [Source:UniProtKB/TrEMBL   |

|           |   |
|-----------|---|
| AT2G02510 | NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 3-A [Source:UniProtKB/Swiss-Prot] |
| AT2G02820 | Transcription factor MYB88 [Source:UniProtKB/Swiss-Prot]                                    |
| AT2G03020 | Heat shock protein HSP20/alpha crystallin family [Source:UniProtKB/TrEMBL]                  |
| AT2G03965 |   |
| AT2G04135 | NA  |
| AT2G04160 | Subtilisin-like protease SBT5.3 [Source:UniProtKB/Swiss-Prot]                               |
| AT2G04380 | Uncharacterized protein At2g04380 [Source:UniProtKB/TrEMBL]                                 |
| AT2G04460 | NA  |
| AT2G04490 | NA  |
| AT2G05380 | glycine-rich protein 3 short isoform [Source:TAIR]  |
| AT2G05510 | At2g05510 [Source:UniProtKB/TrEMBL]   |
| AT2G05914 | Potential natural antisense gene, locus overlaps with AT2G05915 [Source:TAIR]               |
| AT2G05950 | NA  |
| AT2G06845 | Beta-galactosidase related protein [Source:TAIR]  |
| AT2G07042 | other RNA [Source:TAIR]   |
| AT2G07213 | other RNA [Source:TAIR]   |
| AT2G08650 |   |
| AT2G08820 |   |
| AT2G10410 | NA  |
| AT2G10980 | NA  |
| AT2G13431 | NA  |
| AT2G13660 | Uncharacterized protein At2g13660 [Source:UniProtKB/TrEMBL]                                 |
| AT2G14900 | Gibberellin-regulated protein 7 [Source:UniProtKB/Swiss-Prot]                               |
| AT2G15000 | unknown protein   |
| AT2G15128 | other RNA [Source:TAIR]   |
| AT2G15220 | At2g15220/F15A23.4 [Source:UniProtKB/TrEMBL]  |
| AT2G15400 | DNA-directed RNA polymerase II, third largest subunit [Source:UniProtKB/TrEMBL]             |
| AT2G15555 | other RNA [Source:TAIR]   |
| AT2G15940 | NA  |
| AT2G16140 | NA  |
| AT2G17690 | F-box protein At2g17690 [Source:UniProtKB/Swiss-Prot]                                       |
| AT2G18230 | PPa2 [Source:UniProtKB/TrEMBL]  |
| AT2G18490 | C2H2-like zinc finger protein [Source:UniProtKB/TrEMBL]                                     |
| AT2G18660 | EG45-like domain containing protein 2 [Source:UniProtKB/Swiss-Prot]                         |
| AT2G19180 | Expressed protein [Source:UniProtKB/TrEMBL]   |
| AT2G19310 | 18.5 kDa class IV heat shock protein [Source:UniProtKB/Swiss-Prot]                          |
| AT2G19740 | 60S ribosomal protein L31-1 [Source:UniProtKB/Swiss-Prot]                                   |
| AT2G19900 | NADP-dependent malic enzyme 1 [Source:UniProtKB/Swiss-Prot]                                 |
| AT2G20490 | H/ACA ribonucleoprotein complex subunit 3-like protein [Source:UniProtKB/Swiss-Prot]        |
| AT2G20880 | Ethylene-responsive transcription factor ERF053 [Source:UniProtKB/Swiss-Prot]               |
| AT2G21060 | GRP2B [Source:UniProtKB/TrEMBL]   |

|           |   |
|-----------|---|
| AT2G21260 | At2g21260 [Source:UniProtKB/TrEMBL  |
| AT2G21350 | RNA-binding CRS1 / YhbY (CRM) domain protein [Source:UniProtKB/TrEMBL         |
| AT2G21560 | Nucleolar-like protein [Source:UniProtKB/TrEMBL                               |
| AT2G21970 | SEP2 [Source:UniProtKB/TrEMBL   |
| AT2G22710 | NA  |
| AT2G23090 | Uncharacterized protein At2g23090 [Source:UniProtKB/Swiss-Prot                |
| AT2G24030 | At2g24030 [Source:UniProtKB/TrEMBL  |
| AT2G24110 | NA  |
| AT2G24165 | NA  |
| AT2G24780 | Putative uncharacterized protein At2g24780 [Source:UniProtKB/TrEMBL           |
| AT2G24940 | MAPR2 [Source:UniProtKB/TrEMBL  |
| AT2G26150 | Heat stress transcription factor A-2 [Source:UniProtKB/Swiss-Prot             |
| AT2G26160 | F-box protein At2g26160 [Source:UniProtKB/Swiss-Prot                          |
| AT2G26355 | other RNA [Source:TAIR  |
| AT2G26515 |   |
| AT2G26760 | CYCB1 [Source:UniProtKB/TrEMBL  |
| AT2G27402 | At2g27402 [Source:UniProtKB/TrEMBL  |
| AT2G27490 | Dephospho-CoA kinase [Source:UniProtKB/Swiss-Prot                             |
| AT2G27970 | Cyclin-dependent kinases regulatory subunit 2 [Source:UniProtKB/Swiss-Prot    |
| AT2G28105 | Replication factor-A carboxy-terminal domain protein [Source:UniProtKB/TrEMBL |
| AT2G28190 | Superoxide dismutase [Cu-Zn] 2, chloroplastic [Source:UniProtKB/Swiss-Prot    |
| AT2G28570 | At2g28570 [Source:UniProtKB/TrEMBL  |
| AT2G29420 | Glutathione S-transferase U7 [Source:UniProtKB/Swiss-Prot                     |
| AT2G29450 | GSTU5 [Source:UniProtKB/TrEMBL  |
| AT2G29500 | 17.6 kDa class I heat shock protein 2 [Source:UniProtKB/Swiss-Prot            |
| AT2G29570 | Proliferating cell nuclear antigen 2 [Source:UniProtKB/Swiss-Prot             |
| AT2G30000 | PHD finger-like domain-containing protein 5A [Source:UniProtKB/Swiss-Prot     |
| AT2G30060 | Ran-binding protein 1 homolog b [Source:UniProtKB/Swiss-Prot                  |
| AT2G30260 | U2 small nuclear ribonucleoprotein B" [Source:UniProtKB/Swiss-Prot            |
| AT2G30942 | Protein of unknown function (DUF3317) [Source:TAIR                            |
| AT2G30945 |   |
| AT2G31380 | B-box zinc finger protein 25 [Source:UniProtKB/Swiss-Prot                     |
| AT2G31980 | Cysteine proteinase inhibitor [Source:UniProtKB/TrEMBL                        |
| AT2G32190 | Cysteine-rich/transmembrane domain A-like protein [Source:UniProtKB/TrEMBL    |
| AT2G32550 | Cell differentiation, Rcd1-like protein [Source:UniProtKB/TrEMBL              |
| AT2G33080 |   |
| AT2G33175 | unknown protein   |
| AT2G33380 | Probable peroxygenase 3 [Source:UniProtKB/Swiss-Prot                          |
| AT2G33610 | SWI/SNF complex subunit SWI3B [Source:UniProtKB/Swiss-Prot                    |
| AT2G33740 | Protein CutA, chloroplastic [Source:UniProtKB/Swiss-Prot                      |
| AT2G33855 | At2g33855 [Source:UniProtKB/TrEMBL  |

|           |  |
|-----------|--|
| AT2G34585 | At2g34585 [Source:UniProtKB/TrEMBL   |
| AT2G34655 | At2g34655 [Source:UniProtKB/TrEMBL   |
| AT2G35570 | NA   |
| AT2G35760 | CASP-like protein 2B2 [Source:UniProtKB/Swiss-Prot   |
| AT2G36120 | Glycine-rich protein DOT1 [Source:UniProtKB/Swiss-Prot   |
| AT2G36160 | 40S ribosomal protein S14-1 [Source:UniProtKB/Swiss-Prot   |
| AT2G36170 | Ubiquitin-60S ribosomal protein L40-1 [Source:UniProtKB/Swiss-Prot   |
| AT2G36590 | Proline transporter 3 [Source:UniProtKB/Swiss-Prot   |
| AT2G36815 | FUNCTIONS IN: molecular_function unknown   |
| AT2G36830 | Aquaporin TIP1-1 [Source:UniProtKB/Swiss-Prot  |
| AT2G36930 | Expressed protein [Source:UniProtKB/TrEMBL   |
| AT2G37060 | Nuclear transcription factor Y subunit B-8 [Source:UniProtKB/Swiss-Prot  |
| AT2G37180 | Aquaporin PIP2-3 [Source:UniProtKB/Swiss-Prot  |
| AT2G37190 | 60S ribosomal protein L12-1 [Source:UniProtKB/Swiss-Prot   |
| AT2G37410 | Mitochondrial import inner membrane translocase subunit TIM17-2 [Source:UniProtKB/Swiss-Prot                       |
| AT2G37530 | unknown protein  |
| AT2G37770 | NADPH-dependent aldo-keto reductase, chloroplastic [Source:UniProtKB/Swiss-Prot                                    |
| AT2G37870 | Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein [Source:UniProtKB/TrEMBL |
| AT2G38220 | RING/U-box superfamily protein [Source:UniProtKB/TrEMBL  |
| AT2G38530 | Non-specific lipid-transfer protein [Source:UniProtKB/TrEMBL   |
| AT2G38640 | Protein LURP-one-related 8 [Source:UniProtKB/Swiss-Prot  |
| AT2G39705 | RTFL8 [Source:UniProtKB/TrEMBL   |
| AT2G39820 | Eukaryotic translation initiation factor 6-1 [Source:UniProtKB/Swiss-Prot  |
| AT2G40010 | 60S acidic ribosomal protein Po-1 [Source:UniProtKB/Swiss-Prot   |
| AT2G41230 | Protein ORGAN SIZE RELATED 1 [Source:UniProtKB/Swiss-Prot  |
| AT2G41650 | At2g41650 [Source:UniProtKB/TrEMBL   |
| AT2G42720 | F-box/LRR-repeat protein At2g42720 [Source:UniProtKB/Swiss-Prot  |
| AT2G43270 | F-box and associated interaction domains-containing protein [Source:UniProtKB/TrEMBL                               |
| AT2G43535 | Defensin-like protein 196 [Source:UniProtKB/Swiss-Prot   |
| AT2G43580 | Endochitinase At2g43580 [Source:UniProtKB/Swiss-Prot   |
| AT2G43620 | Endochitinase At2g43620 [Source:UniProtKB/Swiss-Prot   |
| AT2G43660 | Carbohydrate-binding X8 domain superfamily protein [Source:UniProtKB/TrEMBL  |
| AT2G43840 | Glycosyltransferase [Source:UniProtKB/TrEMBL   |
| AT2G44170 | NA   |
| AT2G44250 | Protein of Unknown Function (DUF239) [Source:TAIR  |
| AT2G44670 | FCS-Like Zinc finger 3 [Source:UniProtKB/Swiss-Prot  |
| AT2G44995 | Unknown gene [Source:TAIR  |
| AT2G45023 | other RNA [Source:TAIR   |
| AT2G45135 | RING/U-box superfamily protein [Source:UniProtKB/TrEMBL  |
| AT2G46690 | Auxin-responsive protein SAUR32 [Source:UniProtKB/Swiss-Prot   |
| AT2G46970 | Transcription factor PIL1 [Source:UniProtKB/Swiss-Prot   |

|           |   |
|-----------|---|
| AT2G47110 | UBQ6 [Source:UniProtKB/TrEMBL   |
| AT2G47380 | Probable cytochrome c oxidase subunit 5C-1 [Source:UniProtKB/Swiss-Prot                   |
| AT2G47770 | TSPO [Source:UniProtKB/TrEMBL   |
| AT2G48150 | Probable glutathione peroxidase 4 [Source:UniProtKB/Swiss-Prot                            |
| AT3G01130 | ATP synthase E chain [Source:UniProtKB/TrEMBL   |
| AT3G01330 | E2F transcription factor-like E2FF [Source:UniProtKB/Swiss-Prot                           |
| AT3G01600 | NAC domain containing protein 44 [Source:UniProtKB/TrEMBL                                 |
| AT3G01620 | Beta-1,4-N-acetylglucosaminyltransferase family protein [Source:UniProtKB/TrEMBL          |
| AT3G02080 | 40S ribosomal protein S19-1 [Source:UniProtKB/Swiss-Prot                                  |
| AT3G02480 | AT3g02480/F16B3_11 [Source:UniProtKB/TrEMBL   |
| AT3G02620 | Plant stearyl-acyl-carrier-protein desaturase family protein [Source:UniProtKB/TrEMBL     |
| AT3G02640 | F16B3.27 protein [Source:UniProtKB/TrEMBL   |
| AT3G03230 | Alpha/beta-Hydrolases superfamily protein [Source:UniProtKB/TrEMBL                        |
| AT3G03341 | Cold-regulated protein [Source:UniProtKB/TrEMBL   |
| AT3G04000 | NADPH-dependent aldehyde reductase 2, chloroplastic [Source:UniProtKB/Swiss-Prot          |
| AT3G04065 |   |
| AT3G04230 | 40S ribosomal protein S16-2 [Source:UniProtKB/Swiss-Prot                                  |
| AT3G04530 | Phosphoenolpyruvate carboxylase kinase 2 [Source:UniProtKB/Swiss-Prot                     |
| AT3G04660 | F-box/kelch-repeat protein At3g04660 [Source:UniProtKB/Swiss-Prot                         |
| AT3G04860 | AT3g04860/T9J14_19 [Source:UniProtKB/TrEMBL   |
| AT3G05400 | Sugar transporter ERD6-like 12 [Source:UniProtKB/Swiss-Prot                               |
| AT3G05955 |   |
| AT3G06320 | At3g06320 [Source:UniProtKB/TrEMBL  |
| AT3G07590 | Small nuclear ribonucleoprotein SmD1a [Source:UniProtKB/Swiss-Prot                        |
| AT3G07680 | Transmembrane emp24 domain-containing protein p24beta2 [Source:UniProtKB/Swiss-Prot       |
| AT3G07800 | Thymidine kinase a [Source:UniProtKB/Swiss-Prot   |
| AT3G08610 | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 1 [Source:UniProtKB/Swiss-Prot |
| AT3G08730 | Non-specific serine/threonine protein kinase [Source:UniProtKB/TrEMBL                     |
| AT3G08990 | Protein yippee-like At3g08990 [Source:UniProtKB/Swiss-Prot                                |
| AT3G09085 | At3g09085 [Source:UniProtKB/TrEMBL  |
| AT3G09350 | Fes1A [Source:UniProtKB/TrEMBL  |
| AT3G09390 | Metallothionein-like protein 2A [Source:UniProtKB/Swiss-Prot                              |
| AT3G09640 | APX2 [Source:UniProtKB/TrEMBL   |
| AT3G09900 | Ras-related protein RABE1e [Source:UniProtKB/Swiss-Prot                                   |
| AT3G09922 | induced by phosphate starvation1 [Source:TAIR   |
| AT3G10090 | 40S ribosomal protein S28-1 [Source:UniProtKB/Swiss-Prot                                  |
| AT3G10630 | Mannosyltransferase, putative [Source:UniProtKB/TrEMBL                                    |
| AT3G11500 | Putative small nuclear ribonucleoprotein polypeptide G [Source:UniProtKB/TrEMBL           |
| AT3G11520 | Cyclin-B1-3 [Source:UniProtKB/Swiss-Prot  |
| AT3G11750 | Dihydroneopterin aldolase 1 [Source:UniProtKB/Swiss-Prot                                  |
| AT3G11940 | 40S ribosomal protein S5-2 [Source:UniProtKB/Swiss-Prot                                   |

|           |   |
|-----------|---|
| AT3G12050 | Aha1 domain-containing protein [Source:UniProtKB/TrEMBL   |
| AT3G12410 | Polynucleotidyl transferase, ribonuclease H-like superfamily protein [Source:UniProtKB/TrEMBL     |
| AT3G12460 | Polynucleotidyl transferase, ribonuclease H-like superfamily protein [Source:UniProtKB/TrEMBL     |
| AT3G12530 | DNA replication complex GINS protein PSF2 [Source:UniProtKB/Swiss-Prot                            |
| AT3G12580 | Probable mediator of RNA polymerase II transcription subunit 37c [Source:UniProtKB/Swiss-Prot     |
| AT3G12870 | At3g12870 [Source:UniProtKB/TrEMBL  |
| AT3G12981 | NA  |
| AT3G13280 | Putative endonuclease or glycosyl hydrolase [Source:UniProtKB/TrEMBL                              |
| AT3G13580 | Ribosomal protein L30/L7 family protein [Source:UniProtKB/TrEMBL                                  |
| AT3G13680 | F-box/kelch-repeat protein At3g13680 [Source:UniProtKB/Swiss-Prot                                 |
| AT3G13784 | Beta-fructofuranosidase, insoluble isoenzyme CWINV5 [Source:UniProtKB/Swiss-Prot                  |
| AT3G13910 | Protein of unknown function (DUF3511) [Source:TAIR  |
| AT3G14060 | At3g14060 [Source:UniProtKB/TrEMBL  |
| AT3G14600 | 60S ribosomal protein L18a-3 [Source:UniProtKB/Swiss-Prot   |
| AT3G14700 | At3g14700 [Source:UniProtKB/TrEMBL  |
| AT3G14710 | F-box/FBD/LRR-repeat protein At3g14710 [Source:UniProtKB/Swiss-Prot                               |
| AT3G15720 | Probable polygalacturonase At3g15720 [Source:UniProtKB/Swiss-Prot                                 |
| AT3G16470 | Jacalin-related lectin 35 [Source:UniProtKB/Swiss-Prot  |
| AT3G16590 | Putative F-box protein At3g16590 [Source:UniProtKB/Swiss-Prot                                     |
| AT3G16900 | Protein LURP-one-related 13 [Source:UniProtKB/Swiss-Prot  |
| AT3G17050 | NA  |
| AT3G17110 | NA  |
| AT3G17130 | Plant invertase/pectin methylesterase inhibitor superfamily protein [Source:UniProtKB/TrEMBL      |
| AT3G17520 | Late embryogenesis abundant protein (LEA) family protein [Source:UniProtKB/TrEMBL                 |
| AT3G17609 | HYH [Source:UniProtKB/TrEMBL  |
| AT3G17611 | Rhomboid-like protein 14, mitochondrial [Source:UniProtKB/Swiss-Prot                              |
| AT3G17790 | Purple acid phosphatase 17 [Source:UniProtKB/Swiss-Prot   |
| AT3G18320 | F-box and associated interaction domains-containing protein [Source:TAIR                          |
| AT3G18550 | TCP family transcription factor [Source:UniProtKB/TrEMBL  |
| AT3G18720 | F-box/kelch-repeat protein At3g18720 [Source:UniProtKB/Swiss-Prot                                 |
| AT3G19350 | Maternally expressed PAB C-terminal protein [Source:UniProtKB/Swiss-Prot                          |
| AT3G20470 | Glycine-rich protein 5 [Source:UniProtKB/Swiss-Prot   |
| AT3G21010 | NA  |
| AT3G21020 | NA  |
| AT3G21080 | ABC transporter-like protein [Source:UniProtKB/TrEMBL   |
| AT3G21550 | DMP2 [Source:UniProtKB/TrEMBL   |
| AT3G22240 | Cysteine-rich/transmembrane domain PCC1-like protein [Source:UniProtKB/TrEMBL                     |
| AT3G22840 | ELIP1 [Source:UniProtKB/TrEMBL  |
| AT3G23325 | Uncharacterized protein At3g23325 [Source:UniProtKB/Swiss-Prot                                    |
| AT3G24000 | Pentatricopeptide repeat-containing protein At3g24000, mitochondrial [Source:UniProtKB/Swiss-Prot |
| AT3G24100 | Putative uncharacterized protein AT3g24100 [Source:UniProtKB/TrEMBL                               |



|           |   |
|-----------|---|
| AT3G24500 | MBF1C [Source:UniProtKB/TrEMBL  |
| AT3G24640 | Lyase [Source:UniProtKB/TrEMBL  |
| AT3G24927 | NA  |
| AT3G25080 | UPF0725 protein At3g25080 [Source:UniProtKB/Swiss-Prot  |
| AT3G25590 | At3g25590 [Source:UniProtKB/TrEMBL  |
| AT3G25980 | Mitotic spindle checkpoint protein MAD2 [Source:UniProtKB/Swiss-Prot  |
| AT3G26135 | 2-oxoglutarate/Fe(II)-dependent oxygenase family protein [Source:UniProtKB/TrEMBL   |
| AT3G26840 | Acyltransferase-like protein At3g26840, chloroplastic [Source:UniProtKB/Swiss-Prot  |
| AT3G27027 | GPI-anchored-like protein (DUF 3339) [Source:UniProtKB/TrEMBL   |
| AT3G27060 | TSO2 [Source:UniProtKB/TrEMBL   |
| AT3G27340 | At3g27340 [Source:UniProtKB/TrEMBL  |
| AT3G27620 | Ubiquinol oxidase 1c, mitochondrial [Source:UniProtKB/Swiss-Prot  |
| AT3G27940 | LOB domain-containing protein 26 [Source:UniProtKB/Swiss-Prot   |
| AT3G28330 | F-box protein At3g28330 [Source:UniProtKB/Swiss-Prot  |
| AT3G28500 | 60S acidic ribosomal protein P2-3 [Source:UniProtKB/Swiss-Prot  |
| AT3G28611 | NA  |
| AT3G28899 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL   |
| AT3G29631 | Nuclease [Source:UniProtKB/TrEMBL   |
| AT3G29639 | BEST Arabidopsis thaliana protein match is: SPOC domain / Transcription elongation factor S-II protein (TAIR:AT5G11430.1) |
| AT3G30120 | NA  |
| AT3G30122 | NA  |
| AT3G32195 | NA  |
| AT3G42150 | At3g42150 [Source:UniProtKB/TrEMBL  |
| AT3G42658 | NA  |
| AT3G42725 | Putative membrane lipoprotein [Source:UniProtKB/TrEMBL  |
| AT3G42806 | NA  |
| AT3G43580 | At3g43580 [Source:UniProtKB/TrEMBL  |
| AT3G43980 | 40S ribosomal protein S29 [Source:UniProtKB/Swiss-Prot  |
| AT3G44070 | Glycosyl hydrolase family 35 protein [Source:UniProtKB/TrEMBL   |
| AT3G44120 | F-box/kelch-repeat protein At3g44120 [Source:UniProtKB/Swiss-Prot   |
| AT3G44265 | Beta-galactosidase-like protein (Fragment) [Source:UniProtKB/TrEMBL   |
| AT3G44790 | MATH domain and coiled-coil domain-containing protein At3g44790 [Source:UniProtKB/Swiss-Prot                              |
| AT3G46020 | At3g46020 [Source:UniProtKB/TrEMBL  |
| AT3G46060 | RABE1c [Source:UniProtKB/TrEMBL   |
| AT3G46150 | Uncharacterized protein F12M12_120 [Source:UniProtKB/TrEMBL   |
| AT3G46230 | 17.4 kDa class I heat shock protein [Source:UniProtKB/Swiss-Prot  |
| AT3G46487 | NA  |
| AT3G46845 |   |
| AT3G47300 | SelT-like protein [Source:UniProtKB/Swiss-Prot  |
| AT3G47480 | Probable calcium-binding protein CML47 [Source:UniProtKB/Swiss-Prot   |

|           |  |
|-----------|--|
| AT3G47965 | AT3g47961 [Source:UniProtKB/TrEMBL   |
| AT3G48700 | Probable carboxylesterase 13 [Source:UniProtKB/Swiss-Prot                                    |
| AT3G49570 | Protein RESPONSE TO LOW SULFUR 3 [Source:UniProtKB/Swiss-Prot                                |
| AT3G50570 | Hydroxyproline-rich glycoprotein family protein [Source:UniProtKB/TrEMBL                     |
| AT3G50980 | Dehydrin Xero 1 [Source:UniProtKB/Swiss-Prot   |
| AT3G51930 | At3g51930 [Source:UniProtKB/TrEMBL   |
| AT3G52310 | ABC-2 type transporter family protein [Source:TAIR   |
| AT3G52560 | Ubiquitin E2 variant 1D-4 [Source:UniProtKB/TrEMBL   |
| AT3G52590 | Ubiquitin-60S ribosomal protein L40-1 [Source:UniProtKB/Swiss-Prot                           |
| AT3G52730 | Ubiquinol-cytochrome C reductase UQCRX/QCR9-like family protein [Source:UniProtKB/TrEMBL     |
| AT3G53370 | DNA-binding protein S1FA1 [Source:UniProtKB/Swiss-Prot                                       |
| AT3G53890 | 40S ribosomal protein S21-1 [Source:UniProtKB/Swiss-Prot                                     |
| AT3G54160 | RNI-like superfamily protein [Source:UniProtKB/TrEMBL  |
| AT3G54363 | unknown protein  |
| AT3G54366 | Unknown gene [Source:TAIR  |
| AT3G55290 | NAD(P)-binding Rossmann-fold superfamily protein [Source:UniProtKB/TrEMBL                    |
| AT3G56780 | FBD, F-box and Leucine Rich Repeat domains containing protein [Source:UniProtKB/TrEMBL       |
| AT3G56890 | F-box associated ubiquitination effector family protein [Source:UniProtKB/TrEMBL             |
| AT3G57160 | Cysteine-rich TM module stress tolerance protein [Source:UniProtKB/TrEMBL                    |
| AT3G57770 | Protein kinase superfamily protein [Source:UniProtKB/TrEMBL                                  |
| AT3G58270 | MATH domain and coiled-coil domain-containing protein At3g58270 [Source:UniProtKB/Swiss-Prot |
| AT3G58480 | IQ domain-containing protein IQM6 [Source:UniProtKB/Swiss-Prot                               |
| AT3G58800 | Secretion-regulating guanine nucleotide exchange factor [Source:UniProtKB/TrEMBL             |
| AT3G59190 | F-box/RNI-like superfamily protein [Source:UniProtKB/TrEMBL                                  |
| AT3G60010 | SK13 [Source:UniProtKB/TrEMBL  |
| AT3G60710 | Putative FBD-associated F-box protein At3g60710 [Source:UniProtKB/Swiss-Prot                 |
| AT3G60840 | 65-kDa microtubule-associated protein 4 [Source:UniProtKB/Swiss-Prot                         |
| AT3G61110 | 40S ribosomal protein S27 [Source:UniProtKB/TrEMBL   |
| AT3G61890 | HB-12 [Source:UniProtKB/TrEMBL   |
| AT3G61900 | SAUR-like auxin-responsive protein family [Source:UniProtKB/TrEMBL                           |
| AT3G62250 | Ubiquitin-40S ribosomal protein S27a-3 [Source:UniProtKB/Swiss-Prot                          |
| AT3G62460 | AT3g62460/T12C14_160 [Source:UniProtKB/TrEMBL  |
| AT3G62810 | At3g62810 [Source:UniProtKB/TrEMBL   |
| AT4G00810 | AT4G00810 protein [Source:UniProtKB/TrEMBL   |
| AT4G00975 | other RNA [Source:TAIR   |
| AT4G01060 | CAPRICE-like MYB3 [Source:TAIR   |
| AT4G01435 |  |
| AT4G01505 | NA   |
| AT4G01700 | At4g01700 [Source:UniProtKB/TrEMBL   |
| AT4G01985 | unknown protein  |
| AT4G02040 | At4g02040 [Source:UniProtKB/TrEMBL   |

|           |  |
|-----------|--|
| AT4G02230 | 60S ribosomal protein L19-3 [Source:UniProtKB/Swiss-Prot                                 |
| AT4G02465 | unknown protein  |
| AT4G02715 | At4g02715 [Source:UniProtKB/TrEMBL   |
| AT4G02840 |  |
| AT4G03911 | NA   |
| AT4G03960 | Probable tyrosine-protein phosphatase DSP4 [Source:UniProtKB/Swiss-Prot                  |
| AT4G04223 | other RNA [Source:TAIR   |
| AT4G04293 | NA   |
| AT4G05100 | Transcription factor MYB74 [Source:UniProtKB/Swiss-Prot                                  |
| AT4G05295 |  |
| AT4G05380 | P-loop containing nucleoside triphosphate hydrolases superfamily protein [Source:TAIR    |
| AT4G06701 | other RNA [Source:TAIR   |
| AT4G06795 |  |
| AT4G06810 |  |
| AT4G06945 |  |
| AT4G07850 | NA   |
| AT4G07940 | Pre-mRNA-splicing factor CWC22-like protein, putative (DUF3245) [Source:UniProtKB/TrEMBL |
| AT4G08590 | E3 ubiquitin-protein ligase ORTHRUS-LIKE 1 [Source:UniProtKB/Swiss-Prot                  |
| AT4G08685 | SAH7 [Source:UniProtKB/TrEMBL  |
| AT4G08867 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL                                |
| AT4G08869 | Defensin-like protein 210 [Source:UniProtKB/Swiss-Prot                                   |
| AT4G09500 | Glycosyltransferase (Fragment) [Source:UniProtKB/TrEMBL                                  |
| AT4G09800 | 40S ribosomal protein S18 [Source:UniProtKB/Swiss-Prot                                   |
| AT4G09830 | Nuclear receptor family 2 group C protein [Source:UniProtKB/TrEMBL                       |
| AT4G10250 | 22.0 kDa heat shock protein [Source:UniProtKB/Swiss-Prot                                 |
| AT4G10695 | CDC68-like protein [Source:UniProtKB/TrEMBL  |
| AT4G11220 | Reticulon-like protein B2 [Source:UniProtKB/Swiss-Prot                                   |
| AT4G11310 | Probable cysteine protease RDL4 [Source:UniProtKB/Swiss-Prot                             |
| AT4G11910 | Protein STAY-GREEN 2, chloroplastic [Source:UniProtKB/Swiss-Prot                         |
| AT4G12382 | F-box protein At4g12382 [Source:UniProtKB/Swiss-Prot                                     |
| AT4G12400 | Hsp70-Hsp90 organizing protein 3 [Source:UniProtKB/Swiss-Prot                            |
| AT4G12580 | At4g12580 [Source:UniProtKB/TrEMBL   |
| AT4G12890 | At4g12890 [Source:UniProtKB/TrEMBL   |
| AT4G13520 | SMAP1 [Source:UniProtKB/TrEMBL   |
| AT4G13530 | Transmembrane protein [Source:UniProtKB/TrEMBL   |
| AT4G14010 | RALFL32 [Source:UniProtKB/TrEMBL   |
| AT4G14020 | At4g14020 [Source:UniProtKB/TrEMBL   |
| AT4G14110 | COP9 signalosome complex subunit 8 [Source:UniProtKB/Swiss-Prot                          |
| AT4G14145 | At4g14145 [Source:UniProtKB/TrEMBL   |
| AT4G14165 | F-box family protein-like protein [Source:UniProtKB/TrEMBL                               |
| AT4G14548 | other RNA [Source:TAIR   |

|           |  |
|-----------|--|
| AT4G14690 | Early light-induced protein 2, chloroplastic [Source:UniProtKB/Swiss-Prot  |
| AT4G15210 | Beta-amylase 5 [Source:UniProtKB/Swiss-Prot  |
| AT4G15910 | DI21 [Source:UniProtKB/TrEMBL  |
| AT4G16240 | Glycine-rich protein homolog [Source:UniProtKB/TrEMBL  |
| AT4G16447 | At4g16447 [Source:UniProtKB/TrEMBL   |
| AT4G16540 | Heat shock protein HSP20/alpha crystallin family [Source:UniProtKB/TrEMBL  |
| AT4G17486 | DeSI-like protein At4g17486 [Source:UniProtKB/Swiss-Prot   |
| AT4G17670 | FCS-Like Zinc finger 2 [Source:UniProtKB/Swiss-Prot  |
| AT4G17980 | NAC domain containing protein 71 [Source:TAIR  |
| AT4G18330 | Translation elongation factor EF1A/initiation factor IF2gamma family protein [Source:UniProtKB/TrEMBL                              |
| AT4G18590 | Replication protein A 14 kDa subunit B [Source:UniProtKB/Swiss-Prot  |
| AT4G19570 | Chaperone DnaJ-domain superfamily protein [Source:UniProtKB/TrEMBL   |
| AT4G19950 | At4g19950 [Source:UniProtKB/TrEMBL   |
| AT4G20150 | Excitatory amino acid transporter [Source:UniProtKB/TrEMBL   |
| AT4G21105 | At4g21108/At4g21108 [Source:UniProtKB/TrEMBL   |
| AT4G21620 | AT4g21620/F17L22_80 [Source:UniProtKB/TrEMBL   |
| AT4G22960 | Protein of unknown function (DUF544) [Source:TAIR  |
| AT4G23493 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL  |
| AT4G23500 | Pectin lyase-like superfamily protein [Source:UniProtKB/TrEMBL   |
| AT4G23600 | Cystine lyase COR13 [Source:UniProtKB/Swiss-Prot   |
| AT4G23680 | AT4g23680/F9D16_150 [Source:UniProtKB/TrEMBL   |
| AT4G23760 | Cox19-like CHCH family protein [Source:UniProtKB/TrEMBL  |
| AT4G23770 | Uncharacterized protein At4g23770 [Source:UniProtKB/TrEMBL   |
| AT4G24026 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL  |
| AT4G24260 | Endoglucanase 21 [Source:UniProtKB/Swiss-Prot  |
| AT4G24420 | RNA-binding (RRM/RBD/RNP motifs) family protein [Source:UniProtKB/TrEMBL   |
| AT4G24890 | Probable inactive purple acid phosphatase 24 [Source:UniProtKB/Swiss-Prot  |
| AT4G24960 | HVA22D [Source:UniProtKB/TrEMBL  |
| AT4G25700 | Beta-carotene 3-hydroxylase 1, chloroplastic [Source:UniProtKB/Swiss-Prot  |
| AT4G25740 | 40S ribosomal protein S10-1 [Source:UniProtKB/Swiss-Prot   |
| AT4G25780 | CAP (Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1 protein) superfamily protein [Source:UniProtKB/TrEMBL |
| AT4G25910 | NifU-like protein 3, chloroplastic [Source:UniProtKB/Swiss-Prot  |
| AT4G26055 | unknown protein  |
| AT4G26810 | At4g26810 [Source:UniProtKB/TrEMBL   |
| AT4G26960 | unknown protein  |
| AT4G27090 | AT4G27090 protein [Source:UniProtKB/TrEMBL   |
| AT4G27810 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL  |
| AT4G28140 | Ethylene-responsive transcription factor ERF054 [Source:UniProtKB/Swiss-Prot   |
| AT4G28230 | At4g28230 [Source:UniProtKB/TrEMBL   |
| AT4G28280 | LLG3 [Source:UniProtKB/TrEMBL  |

|           |   |
|-----------|---|
| AT4G28290 | At4g28290 [Source:UniProtKB/TrEMBL  |
| AT4G29200 | Beta-galactosidase related protein [Source:UniProtKB/TrEMBL   |
| AT4G29370 | F-box/kelch-repeat protein At4g29370 [Source:UniProtKB/Swiss-Prot                                     |
| AT4G29770 | Target of trans acting-siR480/255. [Source:TAIR   |
| AT4G30010 | AT4g30010/F6G3_40 [Source:UniProtKB/TrEMBL  |
| AT4G30450 | At4g30450 [Source:UniProtKB/TrEMBL  |
| AT4G30460 | Glycine-rich protein [Source:UniProtKB/TrEMBL   |
| AT4G31370 | Fasciclin-like arabinogalactan protein 5 [Source:UniProtKB/Swiss-Prot                                 |
| AT4G31520 | SDA1 family protein [Source:UniProtKB/TrEMBL  |
| AT4G31870 | Putative glutathione peroxidase 7, chloroplastic [Source:UniProtKB/Swiss-Prot                         |
| AT4G32240 | At4g32240 [Source:UniProtKB/TrEMBL  |
| AT4G32490 | Early nodulin-like protein 4 [Source:UniProtKB/TrEMBL   |
| AT4G32830 | AUR1 [Source:UniProtKB/TrEMBL   |
| AT4G33040 | Glutaredoxin-C6 [Source:UniProtKB/Swiss-Prot  |
| AT4G33250 | Eukaryotic translation initiation factor 3 subunit K [Source:UniProtKB/Swiss-Prot                     |
| AT4G33420 | Peroxidase superfamily protein [Source:TAIR   |
| AT4G33467 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL   |
| AT4G34590 | bZIP transcription factor 11 [Source:UniProtKB/Swiss-Prot   |
| AT4G34670 | 40S ribosomal protein S3a [Source:UniProtKB/TrEMBL  |
| AT4G34990 | Transcription factor MYB32 [Source:UniProtKB/Swiss-Prot   |
| AT4G35720 | DUF241 domain protein, putative (DUF241) [Source:UniProtKB/TrEMBL                                     |
| AT4G36880 | CP1 [Source:UniProtKB/TrEMBL  |
| AT4G36900 | Ethylene-responsive transcription factor RAP2-10 [Source:UniProtKB/Swiss-Prot                         |
| AT4G37295 | At4g37295 [Source:UniProtKB/TrEMBL  |
| AT4G37300 | AT4g37300/C7A10_60 [Source:UniProtKB/TrEMBL   |
| AT4G38060 | AT4G38060 protein [Source:UniProtKB/TrEMBL  |
| AT4G38210 | Expansin-A20 [Source:UniProtKB/Swiss-Prot   |
| AT4G38410 | At4g38410 [Source:UniProtKB/TrEMBL  |
| AT4G38495 | Chromatin-remodeling complex subunit [Source:UniProtKB/TrEMBL   |
| AT4G38700 | Dirigent protein 15 [Source:UniProtKB/Swiss-Prot  |
| AT4G39480 | Cytochrome P450, family g6, subfamily A, polypeptide g [Source:UniProtKB/TrEMBL                       |
| AT4G39675 | At4g39675 [Source:UniProtKB/TrEMBL  |
| AT4G39930 | At4g39930 [Source:UniProtKB/TrEMBL  |
| AT4G40085 | other RNA [Source:TAIR  |
| AT5G01180 | Protein NRT1/ PTR FAMILY 8.2 [Source:UniProtKB/Swiss-Prot   |
| AT5G01445 | unknown protein   |
| AT5G01775 |   |
| AT5G02220 | Cyclin-dependent protein kinase inhibitor SMR4 [Source:UniProtKB/Swiss-Prot                           |
| AT5G02490 | Probable mediator of RNA polymerase II transcription subunit 37c [Source:UniProtKB/Swiss-Prot         |
| AT5G02510 | BEST Arabidopsis thaliana protein match is: Ubiquitin system component Cue protein (TAIR:AT5G32440.1) |
| AT5G03010 | At5g03010 [Source:UniProtKB/TrEMBL  |

|           |  |
|-----------|--|
| AT5G03285 | other RNA [Source:TAIR   |
| AT5G03545 | At5g03545 [Source:UniProtKB/TrEMBL   |
| AT5G03850 | 40S ribosomal protein S28-1 [Source:UniProtKB/Swiss-Prot   |
| AT5G03890 | Emb [Source:UniProtKB/TrEMBL   |
| AT5G04000 | unknown protein  |
| AT5G04380 | S-adenosyl-L-methionine-dependent methyltransferases superfamily protein [Source:TAIR            |
| AT5G05420 | Peptidyl-prolyl cis-trans isomerase FKBP15-3 [Source:UniProtKB/Swiss-Prot                        |
| AT5G05950 | At5g05950 [Source:UniProtKB/TrEMBL   |
| AT5G06210 | Small RNA-binding protein 11, chloroplastic [Source:UniProtKB/Swiss-Prot                         |
| AT5G06845 |  |
| AT5G07060 | Zinc finger CCCH domain-containing protein 53 [Source:UniProtKB/Swiss-Prot                       |
| AT5G07330 | At5g07330 [Source:UniProtKB/TrEMBL   |
| AT5G07550 | Oleosin [Source:UniProtKB/TrEMBL   |
| AT5G07560 | GRP20 [Source:UniProtKB/TrEMBL   |
| AT5G07610 | F-box protein At5g07610 [Source:UniProtKB/Swiss-Prot   |
| AT5G08080 | Syntaxin of plants 132 [Source:UniProtKB/TrEMBL  |
| AT5G08260 | Serine carboxypeptidase-like 35 [Source:UniProtKB/Swiss-Prot                                     |
| AT5G09065 |  |
| AT5G09225 | Transmembrane protein [Source:UniProtKB/TrEMBL   |
| AT5G09443 | other RNA [Source:TAIR   |
| AT5G09770 | At5g09770 [Source:UniProtKB/TrEMBL   |
| AT5G10625 | Flowering-promoting factor 1-like protein 2 [Source:UniProtKB/Swiss-Prot                         |
| AT5G11300 | Cyclin-A2-2 [Source:UniProtKB/Swiss-Prot   |
| AT5G11410 | Protein kinase superfamily protein [Source:UniProtKB/TrEMBL                                      |
| AT5G11630 | Uncharacterized protein T22P22_20 [Source:UniProtKB/TrEMBL                                       |
| AT5G12020 | 17.6 kDa class II heat shock protein [Source:UniProtKB/Swiss-Prot                                |
| AT5G12030 | 17.7 kDa class II heat shock protein [Source:UniProtKB/Swiss-Prot                                |
| AT5G12110 | At5g12110 [Source:UniProtKB/TrEMBL   |
| AT5G13170 | Bidirectional sugar transporter SWEET15 [Source:UniProtKB/Swiss-Prot                             |
| AT5G13565 |  |
| AT5G13850 | Nascent polypeptide-associated complex subunit alpha-like protein 3 [Source:UniProtKB/Swiss-Prot |
| AT5G14490 | NAC domain containing protein 85 [Source:TAIR  |
| AT5G14830 | NA   |
| AT5G15200 | 40S ribosomal protein S9-1 [Source:UniProtKB/Swiss-Prot  |
| AT5G15250 | ATP-dependent zinc metalloprotease FTSH 6, chloroplastic [Source:UniProtKB/Swiss-Prot            |
| AT5G15520 | 40S ribosomal protein S19-2 [Source:UniProtKB/Swiss-Prot   |
| AT5G15950 | S-adenosylmethionine decarboxylase proenzyme 2 [Source:UniProtKB/Swiss-Prot                      |
| AT5G15960 | Stress-induced protein KIN1 [Source:UniProtKB/Swiss-Prot   |
| AT5G17360 | At5g17360 [Source:UniProtKB/TrEMBL   |
| AT5G17860 | Cation/calcium exchanger 1 [Source:UniProtKB/Swiss-Prot  |
| AT5G18065 | unknown protein  |

|           |  |
|-----------|--|
| AT5G18380 | 40S ribosomal protein S16-3 [Source:UniProtKB/Swiss-Prot   |
| AT5G18790 | At3g06320 [Source:UniProtKB/TrEMBL   |
| AT5G18800 | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 8-B [Source:UniProtKB/Swiss-Prot          |
| AT5G19875 | At5g19875 [Source:UniProtKB/TrEMBL   |
| AT5G20165 | Protein kish [Source:UniProtKB/TrEMBL  |
| AT5G20290 | 40S ribosomal protein S8 [Source:UniProtKB/TrEMBL  |
| AT5G20500 | Glutaredoxin-C4 [Source:UniProtKB/Swiss-Prot   |
| AT5G20520 | WAV2 [Source:UniProtKB/TrEMBL  |
| AT5G20790 | unknown protein  |
| AT5G21090 | Leucine-rich repeat protein 1 [Source:UniProtKB/Swiss-Prot   |
| AT5G22210 | Aspartyl/glutamyl-tRNA (Asn/Gln) amidotransferase subunit B [Source:UniProtKB/TrEMBL                 |
| AT5G23155 | other RNA [Source:TAIR   |
| AT5G23405 | High mobility group B protein 12 [Source:UniProtKB/Swiss-Prot  |
| AT5G23730 | WD repeat-containing protein RUP2 [Source:UniProtKB/Swiss-Prot                                       |
| AT5G23740 | RPS11-BETA [Source:UniProtKB/TrEMBL  |
| AT5G23760 | At5g23760 [Source:UniProtKB/TrEMBL   |
| AT5G23900 | 60S ribosomal protein L13 [Source:UniProtKB/TrEMBL   |
| AT5G24080 | G-type lectin S-receptor-like serine/threonine-protein kinase At5g24080 [Source:UniProtKB/Swiss-Prot |
| AT5G24170 | Vesicle transport protein [Source:UniProtKB/TrEMBL   |
| AT5G24570 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL  |
| AT5G24660 | Protein RESPONSE TO LOW SULFUR 2 [Source:UniProtKB/Swiss-Prot  |
| AT5G24735 | other RNA [Source:TAIR   |
| AT5G25110 | CBL-interacting serine/threonine-protein kinase 25 [Source:UniProtKB/Swiss-Prot                      |
| AT5G25380 | Cyclin-A2-1 [Source:UniProtKB/Swiss-Prot   |
| AT5G25450 | Cytochrome b-c1 complex subunit 7 [Source:UniProtKB/TrEMBL   |
| AT5G25570 | unknown protein  |
| AT5G26146 | other RNA [Source:TAIR   |
| AT5G26270 | unknown protein  |
| AT5G27345 | NA   |
| AT5G27700 | 40S ribosomal protein S21 [Source:UniProtKB/TrEMBL   |
| AT5G27760 | Hypoxia-responsive family protein [Source:UniProtKB/TrEMBL   |
| AT5G27770 | 60S ribosomal protein L22-3 [Source:UniProtKB/Swiss-Prot   |
| AT5G27850 | 60S ribosomal protein L18-3 [Source:UniProtKB/Swiss-Prot   |
| AT5G27940 | WPP domain-containing protein 3 [Source:UniProtKB/Swiss-Prot   |
| AT5G28010 | Polyketide cyclase/dehydrase and lipid transport superfamily protein [Source:UniProtKB/TrEMBL        |
| AT5G28030 | Bifunctional cystathionine gamma-lyase/cysteine synthase [Source:UniProtKB/Swiss-Prot                |
| AT5G28230 | NA   |
| AT5G28888 | NA   |
| AT5G30360 |  |
| AT5G33251 | NA   |
| AT5G34850 | Purple acid phosphatase [Source:UniProtKB/TrEMBL   |

|           |   |
|-----------|---|
| AT5G34871 | other RNA [Source:TAIR  |
| AT5G35375 | Transmembrane protein [Source:UniProtKB/TrEMBL                                |
| AT5G35490 | mto 1 responding up 1 [Source:TAIR  |
| AT5G35495 | NA  |
| AT5G35510 | unknown protein   |
| AT5G35530 | 40S ribosomal protein S3-3 [Source:UniProtKB/Swiss-Prot                       |
| AT5G37175 | NA  |
| AT5G37260 | Protein REVEILLE 2 [Source:UniProtKB/Swiss-Prot                               |
| AT5G37440 | At5g37440 [Source:UniProtKB/TrEMBL  |
| AT5G37670 | 15.7 kDa heat shock protein, peroxisomal [Source:UniProtKB/Swiss-Prot         |
| AT5G37680 | ADP-ribosylation factor-like protein 8c [Source:UniProtKB/Swiss-Prot          |
| AT5G38386 | Putative F-box/LRR-repeat protein At5g38386 [Source:UniProtKB/Swiss-Prot      |
| AT5G38396 | F-box/LRR-repeat protein At5g38396 [Source:UniProtKB/Swiss-Prot               |
| AT5G38565 | Probable FBD-associated F-box protein At5g38565 [Source:UniProtKB/Swiss-Prot  |
| AT5G39800 | Mitochondrial ribosomal protein L27 [Source:UniProtKB/TrEMBL                  |
| AT5G39950 | Thioredoxin H2 [Source:UniProtKB/Swiss-Prot                                   |
| AT5G40370 | Glutaredoxin family protein [Source:UniProtKB/TrEMBL                          |
| AT5G40730 | Arabinogalactan peptide 24 [Source:UniProtKB/Swiss-Prot                       |
| AT5G40942 | NA  |
| AT5G40960 | At5g40960 [Source:UniProtKB/TrEMBL  |
| AT5G40970 | Protein of unknown function (DUF 3339) [Source:TAIR                           |
| AT5G41090 | At5g41090 [Source:UniProtKB/TrEMBL  |
| AT5G41685 | Mitochondrial import receptor subunit TOM7-1 [Source:UniProtKB/Swiss-Prot     |
| AT5G41830 | RNI-like superfamily protein [Source:TAIR                                     |
| AT5G42330 | At5g42330 [Source:UniProtKB/TrEMBL  |
| AT5G42460 | F-box protein At5g42460 [Source:UniProtKB/Swiss-Prot                          |
| AT5G42680 | At5g42680 [Source:UniProtKB/TrEMBL  |
| AT5G43460 | AT5g43460/MWF20_18 [Source:UniProtKB/TrEMBL                                   |
| AT5G44210 | Ethylene-responsive transcription factor 9 [Source:UniProtKB/Swiss-Prot       |
| AT5G44320 | Eukaryotic translation initiation factor 3 subunit D [Source:UniProtKB/TrEMBL |
| AT5G44490 | FBD-associated F-box protein At5g44490 [Source:UniProtKB/Swiss-Prot           |
| AT5G44568 | Transmembrane protein [Source:UniProtKB/TrEMBL                                |
| AT5G44574 | unknown protein   |
| AT5G44575 | At5g44575 [Source:UniProtKB/TrEMBL  |
| AT5G44578 | unknown protein   |
| AT5G44580 | AT5g44580/K15C23_2 [Source:UniProtKB/TrEMBL                                   |
| AT5G44940 | Putative FBD-associated F-box protein At5g44940 [Source:UniProtKB/Swiss-Prot  |
| AT5G44980 | F-box/FBD/LRR-repeat protein At5g44980 [Source:UniProtKB/Swiss-Prot           |
| AT5G45095 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL                     |
| AT5G45990 | crooked neck protein, putative / cell cycle protein, putative [Source:TAIR    |
| AT5G46160 | 50S ribosomal protein HLP, mitochondrial [Source:UniProtKB/Swiss-Prot         |



|           |  |
|-----------|--|
| AT5G46730 | Glycine-rich protein [Source:UniProtKB/TrEMBL  |
| AT5G47550 | Cysteine proteinase inhibitor 5 [Source:UniProtKB/Swiss-Prot   |
| AT5G47600 | 14.7 kDa heat shock protein [Source:UniProtKB/Swiss-Prot   |
| AT5G47890 | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 2 [Source:UniProtKB/Swiss-Prot            |
| AT5G48180 | Nitrile-specifier protein 5 [Source:UniProtKB/Swiss-Prot   |
| AT5G48510 | BTB/POZ domain-containing protein At5g48510 [Source:UniProtKB/Swiss-Prot                             |
| AT5G48570 | Peptidyl-prolyl cis-trans isomerase FKBP65 [Source:UniProtKB/Swiss-Prot                              |
| AT5G50060 | Plant invertase/pectin methylesterase inhibitor superfamily protein [Source:UniProtKB/TrEMBL         |
| AT5G50260 | KDEL-tailed cysteine endopeptidase CEP1 [Source:UniProtKB/Swiss-Prot                                 |
| AT5G50750 | Probable UDP-arabinopyranose mutase 4 [Source:UniProtKB/Swiss-Prot                                   |
| AT5G50810 | Mitochondrial import inner membrane translocase subunit TIM8 [Source:UniProtKB/Swiss-Prot            |
| AT5G51440 | 23.5 kDa heat shock protein, mitochondrial [Source:UniProtKB/Swiss-Prot                              |
| AT5G52300 | Low-temperature-induced 65 kDa protein [Source:UniProtKB/Swiss-Prot                                  |
| AT5G52640 | Heat shock protein 90-1 [Source:UniProtKB/Swiss-Prot   |
| AT5G52990 | Phytolongin Phyl2.2 [Source:UniProtKB/Swiss-Prot   |
| AT5G53030 | Uncharacterized protein At5g53030 [Source:UniProtKB/TrEMBL   |
| AT5G53048 | other RNA [Source:TAIR   |
| AT5G53190 | Bidirectional sugar transporter SWEET [Source:UniProtKB/TrEMBL                                       |
| AT5G53230 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL  |
| AT5G53240 | Putative uncharacterized protein [Source:UniProtKB/TrEMBL  |
| AT5G53300 | Ubiquitin-conjugating enzyme E2 10 [Source:UniProtKB/Swiss-Prot                                      |
| AT5G53560 | CB5-E [Source:UniProtKB/TrEMBL   |
| AT5G53820 | ABA-inducible protein-like [Source:UniProtKB/TrEMBL  |
| AT5G53902 | U3B  |
| AT5G54390 | PAP-specific phosphatase HAL2-like [Source:UniProtKB/Swiss-Prot                                      |
| AT5G54970 | At5g54970 [Source:UniProtKB/TrEMBL   |
| AT5G55530 | At5g55530 [Source:UniProtKB/TrEMBL   |
| AT5G56510 | Pumilio homolog 12 [Source:UniProtKB/Swiss-Prot  |
| AT5G56795 | MT1B [Source:UniProtKB/TrEMBL  |
| AT5G56810 | Putative F-box/FBD/LRR-repeat protein At5g56810 [Source:UniProtKB/Swiss-Prot                         |
| AT5G56990 | BEST Arabidopsis thaliana protein match is: Cystatin/monellin superfamily protein (TAIR:AT5G56920.1) |
| AT5G57815 | Cytochrome c oxidase subunit [Source:UniProtKB/TrEMBL  |
| AT5G58005 | AT5G58005 protein [Source:UniProtKB/TrEMBL   |
| AT5G58375 | At5g58375 [Source:UniProtKB/TrEMBL   |
| AT5G58570 | Transmembrane protein [Source:UniProtKB/TrEMBL   |
| AT5G58590 | At5g58590 [Source:UniProtKB/TrEMBL   |
| AT5G58710 | Peptidyl-prolyl cis-trans isomerase [Source:UniProtKB/TrEMBL   |
| AT5G59030 | Copper transporter 1 [Source:UniProtKB/Swiss-Prot  |
| AT5G59240 | 40S ribosomal protein S8 [Source:UniProtKB/TrEMBL  |
| AT5G59310 | Non-specific lipid-transfer protein 4 [Source:UniProtKB/Swiss-Prot                                   |
| AT5G59320 | Non-specific lipid-transfer protein 3 [Source:UniProtKB/Swiss-Prot                                   |

|           |   |
|-----------|---|
| AT5G59440 | P-loop containing nucleoside triphosphate hydrolases superfamily protein [Source:UniProtKB/TrEMBL |
| AT5G59480 | Haloacid dehalogenase-like hydrolase (HAD) superfamily protein [Source:UniProtKB/TrEMBL           |
| AT5G59600 | Pentatricopeptide repeat-containing protein At5g59600 [Source:UniProtKB/Swiss-Prot                |
| AT5G59720 | 18.1 kDa class I heat shock protein [Source:UniProtKB/Swiss-Prot                                  |
| AT5G59845 | Gibberellin-regulated protein 10 [Source:UniProtKB/Swiss-Prot                                     |
| AT5G60250 | RBR-type E3 ubiquitin transferase [Source:UniProtKB/TrEMBL  |
| AT5G60260 | unknown protein   |
| AT5G60470 | C2H2 and C2HC zinc fingers superfamily protein [Source:TAIR                                       |
| AT5G60890 | Transcription factor MYB34 [Source:UniProtKB/Swiss-Prot   |
| AT5G61660 | At5g61660 [Source:UniProtKB/TrEMBL  |
| AT5G61880 | Mitochondrial import inner membrane translocase subunit PAM16 like 1 [Source:UniProtKB/Swiss-Prot |
| AT5G62240 | Cell cycle regulated microtubule associated protein [Source:UniProtKB/TrEMBL                      |
| AT5G62575 | SDH7B [Source:UniProtKB/TrEMBL  |
| AT5G63690 | Nucleic acid-binding, OB-fold-like protein [Source:UniProtKB/TrEMBL                               |
| AT5G64140 | RPS28 [Source:UniProtKB/TrEMBL  |
| AT5G64350 | Peptidyl-prolyl cis-trans isomerase FKBP12 [Source:UniProtKB/Swiss-Prot                           |
| AT5G64510 | Tunicamycin induced protein [Source:UniProtKB/TrEMBL  |
| AT5G64590 | Emb [Source:UniProtKB/TrEMBL  |
| AT5G64720 | EC1.5 [Source:UniProtKB/TrEMBL  |
| AT5G65850 | F-box protein At5g65850 [Source:UniProtKB/Swiss-Prot  |
| AT5G66240 | Transducin/WD40 repeat-like superfamily protein [Source:UniProtKB/TrEMBL                          |
| AT5G66400 | RAB18 [Source:UniProtKB/TrEMBL  |
| AT5G67160 | Protein ENHANCED PSEUDOMONAS SUSCEPTIBILITY 1 [Source:UniProtKB/Swiss-Prot                        |
| AT5G67490 | At5g67490 [Source:UniProtKB/TrEMBL  |

**Annexe 10** Liste gènes dérégulés dans l'analyse RNAseq choisis pour la validation par QPCR

| AGI       | Description  |
|-----------|--|
| AT1G73120 | F-box/RNI superfamily protein;(source:Araport11)                               |
| AT5G10760 | Aspartyl protease AED1 [Source:UniProtKB/Swiss-Prot                            |
| AT2G20880 | ATERF53  |
| AT3G17110 | pseudogene   |
| AT4G28140 | DREB subfamily A-6 of ERF/AP2 transcription factor family.                     |
| AT1G52690 | LEA7   |
| AT3G02480 | ABA-RESPONSE PROTEIN, Late embryogenesis abundant protein (LEA) family protein |
| AT4G10250 | ER-localized HSP   |
| AT1G04533 | antisense_long_noncoding_rna   |
| AT5G26270 | transmembrane protein  |
| AT3G42090 | Novel transcript transposon/ALY1_homology                                      |
| AT3G02670 | transmembrane protein  |
| AT1G68350 | cotton fiber protein   |
| AT2G05914 | antisensRNA  |
| AT1G20070 | hypothetical protein   |
| AT5G47000 | Peroxidase superfamily protein down  |
| AT1G67105 | lncRNA down  |
| AT5G50915 | Transcription factor bHLH137 [Source:UniProtKB/Swiss-Prot Down                 |
| AT2G06995 | lncRNA down  |

## Étude du rôle d'une nouvelle endoribonucléase à domaine NYN dans la dégradation des ARN chez *Arabidopsis thaliana*

### Résumé

Les mécanismes de dégradation des ARN ont un rôle majeur afin d'assurer la régulation de l'expression des gènes. Ces phénomènes sont cruciaux lors des transitions développementales et en conditions de stress biotiques et abiotiques. Au cours de ma thèse j'ai caractérisé une nouvelle endoribonucléase à domaine NYN, DAE (DCP1 Associated Endoribonuclease). J'ai montré que DAE interagit directement avec l'activateur de decapping DCP1. L'analyse du transcriptome de plantes dans lesquelles l'expression de DAE est altérée suggère qu'elle agit en parallèle du complexe de decapping. Mes résultats révèlent l'existence d'un nouveau complexe, distinct du complexe de decapping canonique et dont DAE est la sous-unité catalytique. L'étude de la fonction biologique de cette nouvelle voie de dégradation des ARNm permettra de mieux comprendre comment ces mécanismes sont orchestrés chez les plantes afin de répondre à des modifications de leur environnement.

Mots clés : dégradation des ARN, endoribonucléase, domaine NYN, decapping, DCP1, *Arabidopsis thaliana*

### Abstract

RNA degradation pathways have a major role in the regulation of gene expression. These mechanisms are crucial determinants of developmental transitions and in the plant response to biotic and abiotic stresses. During my PhD, I characterized a hitherto unknown NYN domain endoribonuclease, DAE (DCP1 Associated Endoribonuclease). I established that DAE interacts directly with the decapping activator DCP1. The transcriptomic analysis of plants with altered DAE expression, strongly suggest that it acts in parallel of the decapping complex. Overall my results reveal the existence of a new complex distinct from the canonical decapping complex, in which DAE is the catalytic subunit. Studying the biological function of this new complex will improve our understanding of how RNA degradation pathways in plants are orchestrated to respond to environmental stimuli.

Key words: RNA degradation, endoribonuclease, NYN domain, decapping, DCP1, *Arabidopsis thaliana*