

REMERCIEMENTS

Je voudrais tout d'abord remercier les membres de mon jury, Claire Remacle, Christine Lelandais-Brière, Rémy Merret et Dominique Gagliardi d'avoir accepté d'évaluer mon travail de thèse.

Je tiens aussi à remercier ma directrice de thèse, Laurence, pour m'avoir accueillie dans son équipe et donné l'opportunité de vivre cette expérience, mais aussi pour sa disponibilité malgré ses obligations et pour la confiance qu'elle m'a accordée.

Je remercie également tous les membres de l'équipe, passés et présents, pour leurs conseils et leur bonne humeur partagée au laboratoire. Merci à Thalia pour son aide et ses conseils pour les expériences, à Elodie qui m'a appris de nombreuses techniques et la rigueur du travail en laboratoire durant les premières années de ma thèse, à Anne-Marie et Patryk pour leurs conversations scientifiques enrichissantes, à Guillaume qui m'a formée durant les premiers mois de ma thèse et à Christina pour son aide les derniers mois. Un grand merci à Herrade, Mickaële, Lauriane, Jeanne et Valentina avec qui j'ai partagé de très bons moments. Sans vous, ma thèse n'aurait pas été aussi agréable au quotidien. Merci aussi aux stagiaires qui ont animé la vie du laboratoire durant quelques semaines : Romain, Marie, Claire, Lucas.

Je remercie aussi Mathilde, collègue de bureau durant les premières années, pour sa bonne humeur et pour ses conseils.

J'adresse aussi mes remerciements à l'ensemble du personnel de l'IBMP avec qui j'ai pu partager des bons moments de convivialité et de cohésion et pour toutes les réponses à mes questions techniques. Mes remerciements s'adressent aussi plus particulièrement aux personnes que j'ai pu solliciter tout au long de ma thèse notamment l'ensemble des jardiniers qui ont su s'occuper de mes plantes tous les jours, la plateforme P3P, la plateforme de séquençage et plus particulièrement Sandrine, ainsi que Jérôme pour son aide et son expertise en microscopie et enfin Anthony pour son support CRISPR.

Je tiens aussi à remercier ma famille qui m'a entourée, soutenue durant cette épreuve et crue en moi depuis le début de l'aventure.

Et enfin, mes derniers remerciements vont à mon être cher, Alexandre, qui m'a soutenue, aidée et réconfortée quand j'en avais besoin. Merci de m'avoir redonné confiance en ce que je faisais et de m'avoir remotivée durant les moments difficiles. Notre rencontre était la plus belle chose qui me soit arrivée durant ma thèse.

TABLE DES MATIERES

ABRÉVIATIONS.....	1
TABLEAU DES ABBREVIATIONS DES ACIDES AMINES.....	3
INTRODUCTION.....	5
I. Les ARN de transfert	7
1. Structure des ARNt.....	7
2. Les gènes d'ARNt	9
3. Transcription des ARNt.....	11
4. Maturation des ARNt	15
i. Maturation des extrémités 5' et 3'.....	15
ii. Ajout du triplet CCA	17
iii. Épissage des introns	17
iv. Modifications de nucléosides.....	19
5. Contrôle qualité des ARNt et leurs voies de clivages.....	21
i. Nuclear Surveillance Pathway	21
ii. Rapid tRNA Decay Pathway.....	23
6. Fonction principale des ARNt : acteurs de la traduction.....	25
i. Aminoacylation des ARNt.....	25
ii. Le mécanisme de traduction.....	27
7. Autres fonctions des ARNt.....	31
i. Fonctions des ARNt non aminoacylés.....	33
ii. Fonction des ARNt aminoacylés.....	35
II. Les petits ARN non codants, dérivant d'ARNt.....	39
1. Classifications des tRFs	39
i. Les petits ARN non-codants.....	39
ii. Les tRFs : une nouvelle classe de petits ARN non-codants.....	41
iii. Les tRFs chez les plantes	43
2. Biogenèse des tRFs.....	45
i. Chez les procaryotes	45
ii. Chez les eucaryotes.....	47
iii. Chez les plantes.....	47
iv. Impact des modifications post-transcriptionnelles sur la biogenèse des tRFs.....	49
3. Fonctions des tRFs	51
i. Les tRFs dans la régulation de la traduction	51
ii. Les tRFs dans la régulation post-traductionnelle.....	57
iii. Les tRFs dans la régulation de la transcription	57
iv. Les tRFs dans la régulation de l'expression de gènes.....	57
4. Fonctions des tRFs chez les plantes	59
i. Inhibition de la traduction par silencing.....	59
ii. Inhibition de la traduction indépendamment du mécanisme de silencing.....	59
iii. Les tRFs comme facteurs d'adaptation aux stress.....	61
5. Préambule au projet de thèse.....	69
MATÉRIEL & MÉTHODES.....	71
MATERIEL.....	73
I. Matériel végétal.....	73
1. <i>Arabidopsis thaliana</i> pour l'étude des RNS.....	73

2. <i>Arabidopsis thaliana</i> pour la visualisation de P-bodies et granules de stress	73
3. <i>Arabidopsis thaliana</i> pour l'étude de DExH1	73
4. <i>Arabidopsis thaliana</i> écotype Landsberg	73
5. <i>Nicotiana benthamiana</i>	73
II. Souches bactériennes.....	73
1. <i>Escherichia coli</i>	73
2. <i>Agrobacterium tumefaciens</i>	74
III. Vecteurs	74
1. pK11.1R	74
2. pLITMUS 28i.....	74
3. pDONR207	74
4. pB7FWG2.....	75
5. pK7WGF2.....	75
6. pALiCE.....	75
IV. Numéros d'accessions	75
V. Oligonucléotides.....	76
VI. Anticorps.....	76
VII. Outils informatiques	76
1. CRISPOR.....	76
2. MacVector	76
3. ImageJ	76
4. Alignements de séquence.....	77
METHODES.....	79
I. Méthodes relatives aux organismes.....	79
1. Méthodes relatives au matériel bactérien	79
<i>i. Préparation de bactéries compétentes d'Escherichia coli</i>	79
<i>ii. Transformation d'E. coli</i>	79
<i>iii. Préparation d'Agrobacterium tumefaciens compétentes</i>	79
<i>iv. Transformation d'Agrobacterium tumefaciens</i>	80
2. Méthodes relatives au matériel végétal	80
<i>i. Stérilisation de graines d'Arabidopsis</i>	80
<i>ii. Culture d'Arabidopsis</i>	80
<i>iii. Transformation stable d'Arabidopsis par floral dip</i>	81
<i>iv. Sélection de graines fluorescentes</i>	81
<i>v. Agro-infiltration de feuille de N. benthamiana</i>	82
<i>vi. Préparation et transfection de protoplastes d'Arabidopsis</i>	82
II. Méthodes relatives à l'ADN	83
1. Extraction d'ADN génomique d'Arabidopsis.....	83
<i>i. Extraction par broyage</i>	83
<i>ii. Extraction rapide par chauffage</i>	83
2. Clonages.....	84
<i>i. Clonage pour la technique CRISPR/Cas9</i>	84
<i>ii. Clonage Gateway</i>	84
<i>iii. Clonage ALiCE</i>	86
3. Amplification d'ADN par PCR.....	87
4. Reverse Transcription - PCR quantitative (RT-qPCR)	87
5. Électrophorèse sur gel d'agarose.....	87
6. Purification de plasmides	88

7. Séquençage d'ADN	88
<i>i.</i> Séquençage à partir de produits de PCR.....	88
<i>ii.</i> Séquençage d'ADN plasmidique	88
8. Marquage de sondes oligonucléotidiques.....	88
III. Méthodes relatives à l'ARN	89
1. Extraction des ARN de feuilles d' <i>A. thaliana</i>	89
<i>i.</i> Extraction des ARN totaux.....	89
<i>ii.</i> Extraction des petits ARN.....	89
2. Préparation d'ADNc.....	89
<i>i.</i> Traitement DNase.....	89
<i>ii.</i> Rétrotranscription	90
3. Electrophorèse sur gel de polyacrylamide dénaturant.....	90
4. Northern blot.....	90
<i>i.</i> Transfert des ARN sur membrane	90
<i>ii.</i> Hybridation avec une sonde radiomarquée.....	90
5. Préparation des banques de petits ARNs pour sRNA-seq	91
<i>i.</i> Traitement des ARN à la phosphatase et à la polynucléotide kinase (PNK).....	91
<i>ii.</i> Ligation des adaptateurs pour le séquençage	91
<i>iii.</i> Retrotranscription et amplification par PCR	91
<i>iv.</i> Découpe sur gel et purification des fragments	92
<i>v.</i> Séquençage Illumina MiSeq	92
6. Détermination de la présence de structures type G-quadruplex.....	92
<i>i.</i> Électrophorèse sur gel de polyacrylamide non-dénaturant	92
<i>ii.</i> Mesure de la fluorescence – NMM	93
IV. Méthodes relatives aux protéines	93
1. Extraction de protéines	93
2. Électrophorèse sur gel de polyacrylamide dénaturant.....	93
3. Western blot.....	94
<i>i.</i> Transfert des protéines sur membrane.....	94
<i>ii.</i> Coloration des protéines	94
<i>iii.</i> Immunodétection	94
4. Purification de partenaires protéiques de tRFs	94
<i>i.</i> Lyse des protoplastes transfectés et chromatographie d'affinité à la streptavidine.....	94
<i>ii.</i> Identification des protéines par spectrométrie de masse (nanoLC-MS/MS).....	95
<i>iii.</i> Analyse des protéines identifiées par spectrométrie de masse.....	95
ANNEXES	97
Annexe I. Nomenclature des tRFs utilisés	97
Annexe II. Composition des solutions, tampons et milieu	97
Annexe III. Antibiotiques utilisés pour la sélection des bactéries.....	100
Annexe IV. Article "Transfection of small noncoding RNAs into Arabidopsis thaliana protoplasts"	100
Annexe V. Oligonucléotides ARN utilisés pour la transfection de protoplastes d' <i>A. thaliana</i>	119
Annexe VI. Séquences ADN correspondantes aux ARN guides utilisés pour les transformations d' <i>A. thaliana</i> par CRISPR/Cas9	119
Annexe VII. Amorces ADN utilisées pour l'amplification d'ADN par PCR.....	120
Annexe VIII. Amorces ADN utilisées pour l'amplification d'ADN par RT-qPCR.....	124
Annexe IX. Amorces ADN utilisées pour le séquençage	124
Annexe X. Sondes ADN utilisées pour la détection d'ARN par northern blot	125
Annexe XI. Anticorps utilisés pour les western blots.....	125

RÉSULTATS	127
CHAPITRE I : LES RNS A L'ORIGINE DES tRFs CHEZ A. THALIANA	129
I. Mise en contexte	129
1. Description des RNases T2.....	131
i. Structure.....	131
ii. Activité enzymatique.....	131
iii. Évolution classe I/II.....	133
2. Les RNS d' <i>A. thaliana</i> : des RNases T2 à l'origine de la biogenèse des tRFs	133
i. Profils d'expression des RNS.....	133
ii. Localisation subcellulaire des RNS	135
iii. Implication des RNS dans la biogenèse des tRFs.....	137
II. Obtention de lignées <i>rns</i> ^{KO}	137
1. Stratégie employée : CRISPR/Cas9	137
i. Conception des ARN guides <i>in silico</i>	139
ii. Transformation de plantes d' <i>A. thaliana</i> par des agrobactéries avec le plasmide pK11.1R.....	141
2. Résultats des lignées <i>rns</i> ^{KO} obtenues.....	143
i. Résultats de séquençage des mutants obtenus.....	143
ii. Validation des mutants <i>rns</i> ^{KO} par western blot	145
III. Caractérisation des mutants <i>rns</i> ^{KO} – résultats préliminaires	147
1. Analyse phénotypique	147
2. Observations de la biogenèse des tRFs	151
3. Conclusion.....	155
CHAPITRE II : ÉTUDE DE LA FONCTION BIOLOGIQUE DU tRF MODELE ALA CHEZ A. THALIANA	157
I. Mise en contexte	157
1. Pourquoi le tRF-5D Ala comme modèle.....	157
2. Stratégie expérimentale employée pour étudier sa fonction biologique	157
II. Aide à la compréhension des résultats	159
1. Les structures G-quadruplex d'ARN.....	159
i. Description	159
ii. Les G-quadruplex chez les plantes	163
2. Les hélicases à ARN	165
i. Généralités	165
ii. Les hélicases à ARN chez <i>A. thaliana</i>	167
III. Résultats de la localisation subcellulaire et des interactants du tRF Ala.....	167
Article <i>Arabidopsis</i> DEA(D/H)-box RNA helicase interacts with the foci-localized tRF Ala.....	168
Figures de l'article	187
IV. Résultats complémentaires à l'article	225
1. Analyses supplémentaires de DExH1	225
2. Génotypage et vérification des lignées <i>dexh1</i> ^{KO}	229
3. Étude de l'impact de DExH1 sur la biogenèse/stabilité du tRF Ala.....	231
4. Production <i>in vitro</i> de protéines de fusion DExH1-GFP	233
V. Conclusion	233
DISCUSSION & PERSPECTIVES	237
LA BIOGENESE DES tRFs CHEZ A. THALIANA	239
UN PAS VERS L'ETUDE FONCTIONNELLE DU tRF ALA	245
BIBLIOGRAPHIE	253

ABRÉVIATIONS

aa-AMP	Aminoacyl-adénylate	DSB	<i>Double Strand Break</i>
aaRS	Aminoacyl-ARNt synthétase	DTT	Dithiothréitol
ABA	Acide abscissique	dTTP	désoxythymidine triphosphate
ADN	Acide désoxyribonucléique	EDTA	<i>Ethylenediaminetetraacetic acid</i>
ADNc	ADN complémentaire	eGFP	<i>enhanced Green Fluorescent Protein</i>
ADNg	ADN génomique	eIF	<i>Initiation Factor</i>
ADNt	Gène d'ARN de transfert	eRF	<i>Release Factor</i>
ADP	Adénosine diphosphate	FDR	<i>False Discovery Rate</i>
AGO	Argonaute	GFP	<i>Green Fluorescent Protein</i>
ALA	Acide δ-aminolevulinique	GluTR	glutamyl-ARNt reductase
AMP	Adénosine monophosphate	Gm^R	résistance à la gentamycine
Amp^R	Résistance à l'ampicilline	HDR	<i>Homology-Directed Repair</i>
APAF-1	<i>Apoptotic Protease Activating Factor-1</i>	Hyg^R	résistance à l'hygromycine
ARN	Acide ribonucléique	IAN	indole-3-acétonitrile
ARN Pol II	ARN polymérase II	IAOx	indole acétaldoxime
ARN Pol III	ARN polymérase III	IPT	phosphate-isopentenyltransferase
ARNg	ARN guide	KO	<i>knock-out</i>
ARNm	ARN messenger	Ler0	Landsberg
ARNnc	ARN non codant	LTR	<i>Long Terminal Repeat</i>
ARNr	ARN ribosomique	MES	<i>2-(N-morpholino) ethanesulfonic acid</i>
ARNt	ARN de transfert	miARN	microARN
ARNt^{Meté}	ARNt méthionine élongateur	MRPP	<i>Mitochondrial RNase P Protein</i>
ARNt^{Metf}	ARNt methionine initiateur formylé	mtARNr	ARN ribosomique mitochondrial
ARNt^{Meti}	ARNt methionine initiateur	mtRF	tRF d'origine mitochondriale
ATE	arginyl-ARNt transférase	NAA	acide α-naphtaleneacetic
ATP	Adénosine triphosphate	NEP	<i>Nuclear-Encoded Polymerase</i>
Bdp1	<i>B double prime 1</i>	NHEJ	<i>Non-Homologous End Joining</i>
BET	bromure d'éthidium	NIQI	<i>Novel Interesting Quadruplex Interaction motif</i>
Brf1	<i>TFIIB related factor 1</i>	NMM	N-Methyl Mesoporphyrin IX
Col-0	Columbia-0	NSP	<i>Nuclear Surveillance Pathway</i>
CRISPR	<i>Clustered Regularly Interspaced Short Palindromic Repeats associated nucleases</i>	nt	nucléotide
Cter	extrémité C-terminal	Nter	extrémité N-terminal
ctRF	tRF d'origine chloroplastique	ntRF	tRF d'origine nucléaire
dATP	désoxyadénosine triphosphate	p/v	poids à volume
DCL	<i>Dicer-like</i>	PABP	<i>Poly-A Binding Protein</i>
dCTP	désoxycytidine triphosphate	PAM	<i>Protospacer Adjacent Motif</i>
dGTP	désoxyguanosine triphosphate	pb	paire de bases
		PBS	<i>Primer Binding Site</i>

PCR	Réaction en chaîne par polymérase	TBE	Tris Borate EDTA
PEP	<i>Plastid-Encoded Polymerase</i>	TBP	<i>TATA Binding Protein</i>
Pi	phosphate inorganique	TEMED	N, N, N',N'-tétraméthyléthylène diamine
PIC	Pre-Initiation Complex	TFIIIB	<i>Transcription Factor IIIB</i>
PIPES	<i>piperazine-N,N'-bis(2-ethanesulfonic acid)</i>	TFIIIC	<i>Transcription Factor IIIC</i>
PNK	polynucleotide kinase	tRF	<i>tRNA derived Fragments</i>
PPR	<i>Pentatricopeptide Repeat</i>	U	unité
pré-miARN	précurseur de microARN	Ub	Ubiquitine
pri-miARN	microARN primaire	UTR	<i>Untranslated Transcribed Region</i>
PRORP	<i>Proteinaceous RNase P</i>	UV	Ultra-Violets
PTGS	<i>Post-Transcriptional Gene Silencing</i>	VIH	virus de l'immunodéficience humaine
Rat1	<i>Ribonucleic-acid-trafficking protein 1</i>	Ws	Wassilewskija
RBP	<i>RNA-binding Protein</i>	Xrn1	Exoribonuclease 1
RE	reticulum endoplasmique	YFP	<i>Yellow Fluorescent Protein</i>
RFP	<i>Red Fluorescent Protein</i>	ZFN	<i>Zinc-Finger nuclease</i>
RIP	<i>RNA Immunoprecipitation</i>		
RISC	<i>RNA Induced Silencing Complex</i>		
RNase P	Ribonucléase P		
RNP	Ribonucléoprotéines		
RNS	<i>S-like Ribonuclease</i>		
rpm	rotations par minute		
RT-qPCR	<i>Quantitative Reverse Transcription PCR</i>		
RTD	<i>Rapid tRNA Decay</i>		
SDS	Sodium Dodécyl Sulfate		
SDS-PAGE	<i>SDS denaturing Polyacrylamide Gel Electrophoresis</i>		
SF1-6	superfamilles 1-6 des hélicases		
SFLN	<i>Schlafen protein</i>		
siARN	ARN interférent		
Sm^R	résistance à la spectinomycine		
Sp^R	résistance à la spectinomycine		
sRNA-seq	<i>small RNA sequencing</i>		
SSC	<i>Standard Sodium Citrate</i>		
STTM	<i>Short Tandem Target Mimic</i>		
TAE	Tris Acétate EDTA		
TALLEN	<i>Transcription Activator-Like Effector-based Nuclease</i>		

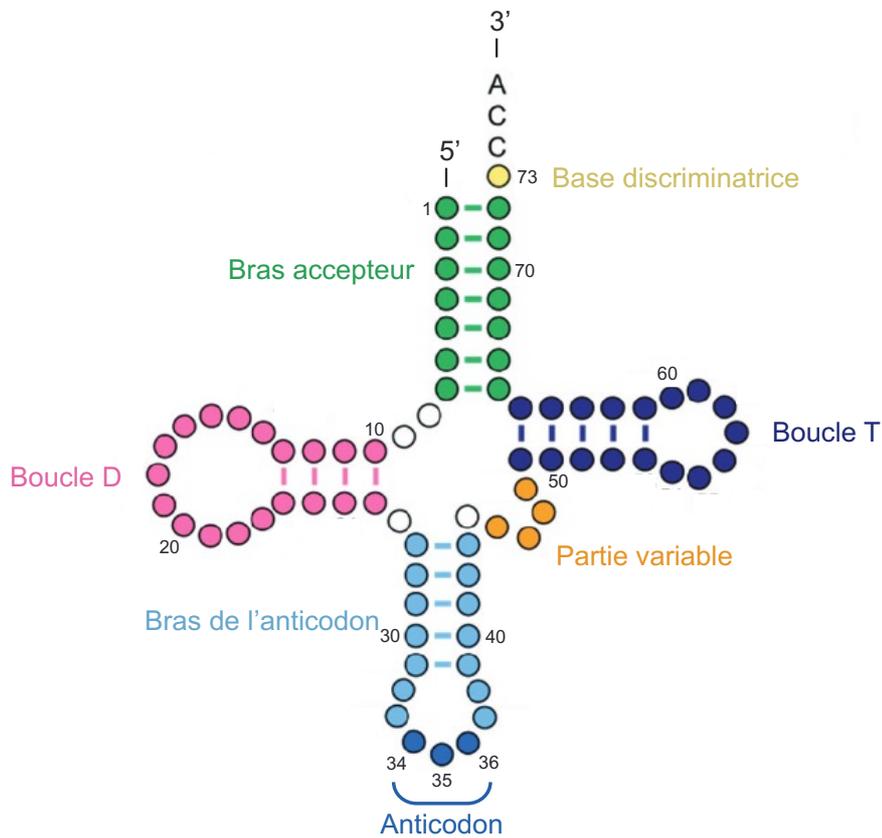
TABLEAU DES ABREVIATIONS DES ACIDES AMINES

Selon le code IUPAC-IUB

A	Ala	Alanine
C	Cys	Cystéine
D	Asp	Acide aspartique
E	Glu	Acide glutamique
F	Phe	Phénylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Méthionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Sérine
T	Thr	Thréonine
V	Val	Valine
W	Trp	Tryptophane
Y	Tyr	Tyrosine

INTRODUCTION

A



B

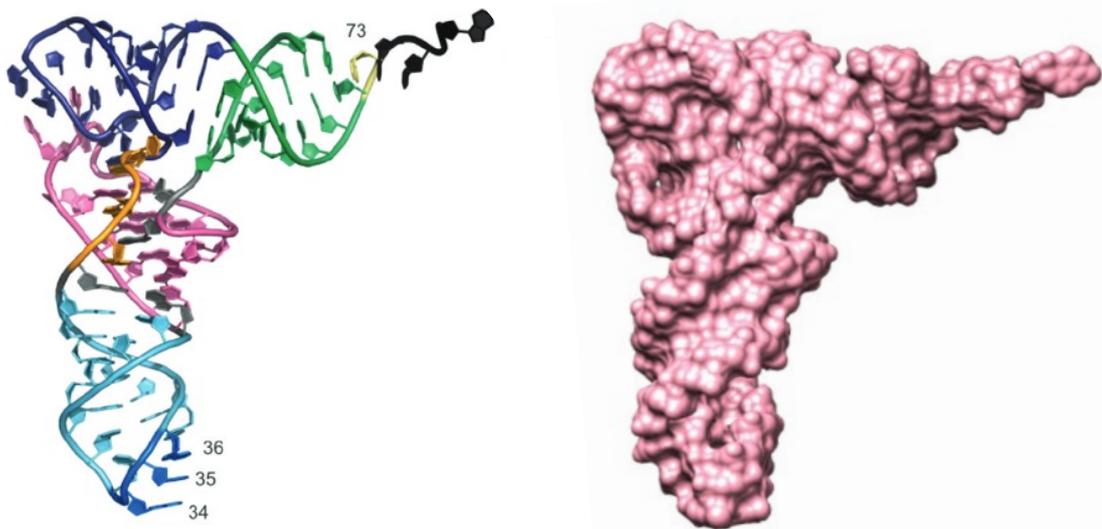


Figure 1 : Structure d'un ARNt. Adaptée de [18]. A : Représentation 2D de la structure secondaire en feuille de trèfle d'un ARNt. En vert sont représentées les bases du bras accepteur, en rose les bases de la boucle D, en bleu les bases du bras de l'anticodon dont en bleu foncé les 3 bases de l'anticodon, en orange les bases de la partie variable, en violet foncé les bases de la boucle T et enfin en jaune la base discriminatrice. Les numéros correspondent aux positions des bases. B : Représentations 3D de la structure tertiaire en L de l'ARNt^{Phe} de levure. Sur la gauche, représentation permettant de visualiser les interactions entre bases. Les couleurs correspondent aux régions décrites en A. Sur la droite, représentation montrant l'encombrement stérique.

INTRODUCTION

I. Les ARN de transfert

Les ARN de transfert (ARNt) sont des macromolécules présentes dans tous les organismes vivants. Ils font partie de la classe des ARN non codants, c'est-à-dire, ceux qui ne seront pas traduits en protéines. L'existence des ARNt a été prédite par Francis Crick en 1955 mais ce n'est que plusieurs années après que les ARNt ont été biochimiquement caractérisés [1]. Tout comme les ARN messagers (ARNm) et les ARN ribosomiques (ARNr), les ARNt sont essentiels au mécanisme de biosynthèse des protéines : la traduction. Ils sont considérés comme des adaptateurs entre l'ARNm qui porte l'information génétique grâce aux codons et la chaîne peptidique naissante au sein du ribosome. Les ARNt, chargés de leur acide aminé (ARNt aminoacylés), reconnaissent par complémentarité de base les codons présents sur l'ARNm, le tout au sein du ribosome. Les ARNt reconnaissent les codons au niveau de leur région dite de l'anticodon. Le code génétique étant dégénéré, plusieurs codons peuvent coder pour un même acide aminé. Uniquement la Méthionine et le Tryptophane sont codés par un seul codon. Par conséquent, il existe des ARNt possédant des anticodons différents mais aminoacylés avec un même acide aminé : ils sont appelés ARNt isoaccepteurs. Les ARNt avec une séquence de l'anticodon identique mais qui possèdent des différences nucléotidiques dans le reste de leur séquence sont appelés ARNt isodécodateurs [2].

1. Structure des ARNt

Les ARNt ont tous une structure semblable, bien conservée chez les eucaryotes. Leur taille varie de 76 à 90 bases, qui en s'appariant entre elles, forment une structure secondaire en tige-boucle. Cette structure secondaire a été pour la première fois élucidée en 1965 à partir de l'ARNt^{Ala} de levure [3]. Ces recherches ont d'ailleurs fait l'objet d'un prix Nobel de médecine en 1968. La structure de l'ARNt s'apparente à une feuille de trèfle (Figure 1 A). Ainsi, de 5' en 3', se trouvent la tige acceptrice de l'acide aminé, la tige-boucle D, la tige-boucle de l'anticodon, la partie variable et la tige-boucle T. La boucle D est ainsi nommée à cause des bases modifiées dihydrouridines (D) conservées aux positions 16 et 17. La boucle T, aussi appelée TΨC, porte ce nom à cause de la présence des bases modifiées thymidine (T), pseudouridine (Ψ) et de la cytidine (C) aux positions 54 à 56. La boucle de l'anticodon permet l'appariement des bases 34, 35 et 36 de l'ARNt aux bases complémentaires situées sur l'ARNm (le codon). La partie variable peut être relativement longue (> 10 bases) notamment chez les eucaryotes pour les ARNt^{Ser} et ARNt^{Leu} ou les ARNt^{Tyr} chez les eubactéries ou dans les organelles [4]. Enfin, le bras accepteur possède en 3' le triplet CCA qui est le site de fixation de l'acide aminé.

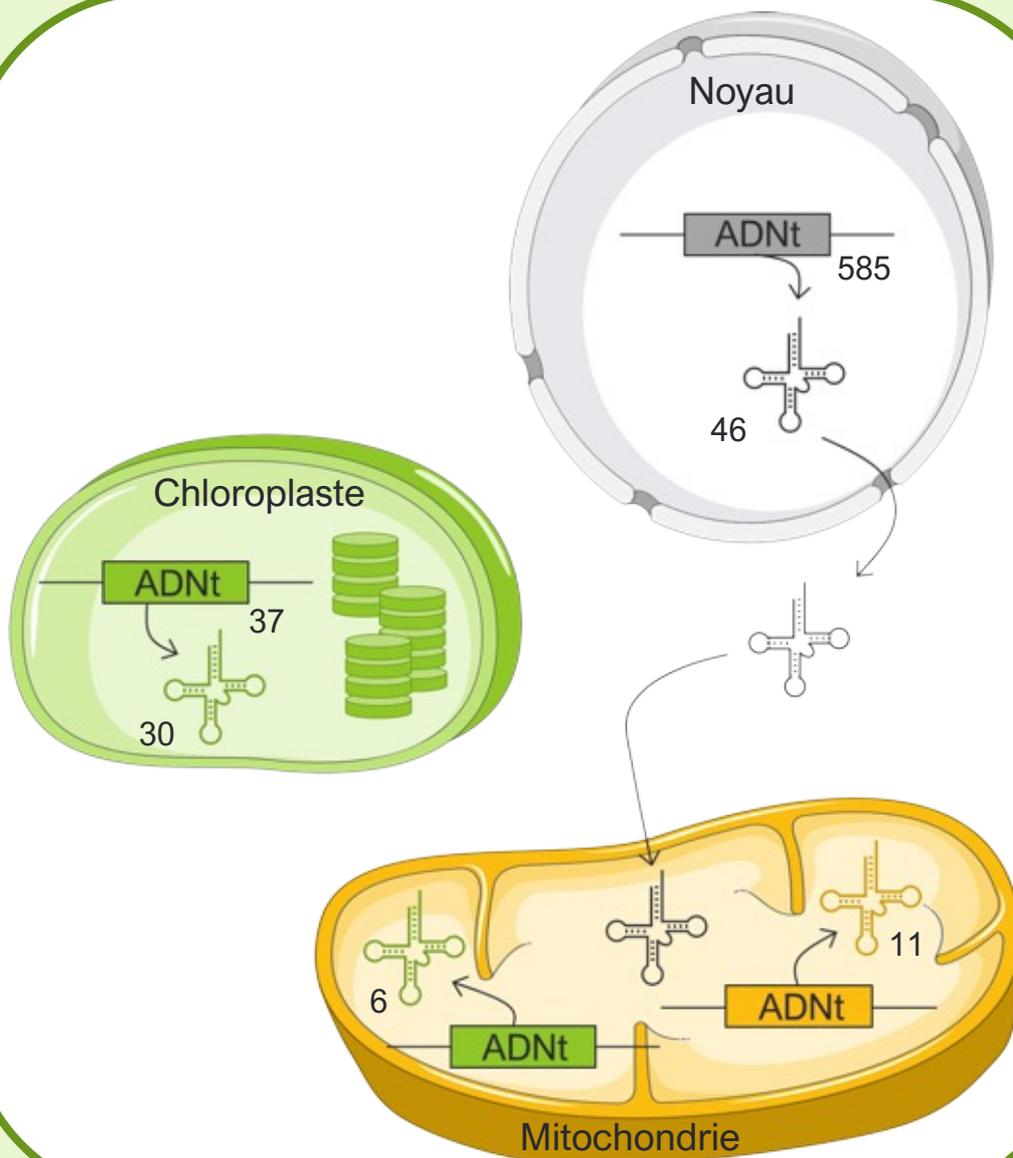


Figure 2 : Schéma de la répartition des gènes d'ARNt au sein de la cellule végétale. Cas du modèle *A. thaliana*. Les ARNt nucléaires, en gris, sont codés par le génome nucléaire puis exportés dans le cytosol. Une partie de cette population d'ARNt est importée dans la mitochondrie. Les ARNt chloroplastiques, en vert, proviennent du génome chloroplastique. Les ARNt mitochondriaux codés par le génome mitochondrial ont deux origines : des gènes provenant de l'alpha-protéobactérie ancestrale à l'origine de la mitochondrie, en jaune, et des gènes d'origine plastidiale transférés dans le génome mitochondrial au cours de l'évolution. Le nombre d'ADNt et d'ARNt isoaccepteurs sont indiqués.

L'ensemble de ces tiges-boucles se replie sur elles-mêmes pour former une structure tridimensionnelle en forme de L (Figure 1 B). Ce sont principalement les boucles D et T qui interagissent entre elles et permettent un repliement de l'ARNt. À l'une des extrémités du L se trouve le bras accepteur qui portera l'acide aminé et à l'autre extrémité se trouve la séquence de l'anticodon, libre, pour s'apparier au codon de l'ARNm. Le bras accepteur est stabilisé par des interactions entre sa tige et la tige T. La boucle de l'anticodon quant à elle interagit avec la tige-boucle D [5].

2. Les gènes d'ARNt

Les ARNt sont codés par de nombreux gènes, appelés ADNt. Le nombre de ces gènes varie selon les organismes. D'une manière générale, on retrouve un plus grand nombre d'ADNt chez les eucaryotes. Par exemple, le génome de la bactérie *Escherichia coli* possède 86 ADNt alors que chez les eucaryotes, on retrouve, entre autres, 275 ADNt chez *Saccharomyces cerevisiae* et 642 chez *Arabidopsis thaliana* [6,7].

Chez les eucaryotes, les ARNt sont codés par des gènes nucléaires mais aussi mitochondriaux. Chez les organismes photosynthétiques, des ADNt sont également codés par le génome plastidial. En effet, la mitochondrie et le chloroplaste proviennent d'endosymbioses d'alpha-protobactérie et de cyanobactérie respectivement et possèdent leur propre génome, relique des génomes bactériens ancestraux. Chez ces deux organites, beaucoup des gènes ont depuis été perdus ou transférés au génome nucléaire au cours de l'évolution. Ainsi, plusieurs ARNt d'origine nucléaires sont importés dans la mitochondrie pour pallier la perte d'ADNt codés par le génome mitochondrial. Ces ARNt sont essentiels pour la lecture des 61 codons sens du code génétique universel et assurer la traduction mitochondriale [8]. Par ailleurs, le génome mitochondrial a aussi acquis, par transfert de fragments d'ADN plastidial, des gènes d'ADNt d'origine chloroplastique. Certains de ces ADNt sont exprimés et ont été nommés *chloroplast-like* [9,10].

Les angiospermes possèdent entre 300 et 2000 ADNt nucléaires [11]. La plante modèle *Arabidopsis thaliana* possède 585 ADNt nucléaires qui codent pour 46 isoaccepteurs (Figure 2). Les ADNt nucléaires sont répartis uniformément sur l'ensemble des chromosomes. Les gènes d'ARNt isoaccepteurs sont répartis de manière aléatoire. Son génome mitochondrial compte 22 ADNt codant 17 ARNt isoaccepteurs. Parmi ces isoaccepteurs, 11 proviennent d'ADNt « natifs » de la mitochondrie et 6 sont des *chloroplast-like*. On observe un manque en ARNt^{Ala}, ARNt^{Arg}, ARNt^{Leu}, ARNt^{Thr} et ARNt^{Val} qui ne sont pas codés par le génome mitochondrial. L'absence de ces ARNt est compensée par l'importation d'ARNt codés par le génome nucléaire [12]. Son génome chloroplastique compte quant à lui 37 ADNt codant pour 30 ARNt isoaccepteur, un nombre suffisant pour assurer la traduction [7,13] [PlantRNA database, <http://seve.ibmp.unistra.fr/plantrna/>].

3. Transcription des ARNt

Les gènes d'ARNt et leur transcription sont bien caractérisés chez la levure *Saccharomyces cerevisiae*. Cet organisme sera utilisé ici comme modèle eucaryotique.

Les gènes d'ARNt contiennent des régions bien conservées qui sont essentielles pour le processus de transcription. Les ADNt possèdent deux régions promotrices internes appelées boîte A et boîte B situées au niveau des séquences codant la boucle D et la boucle T. La distance entre les boîtes A et B varie selon la présence éventuelle d'un intron et la longueur de la partie variable de l'ARNt [14–16]. Une grande proportion de ces gènes possède également en amont des boîtes A et B, une région appelée *TATA box-like*, riche en nucléotides A et T. En aval des boîtes A et B se trouve une région riche en T (généralement de 4 à 15 nucléotides) importante pour la terminaison de la transcription [13,17].

Chez *A. thaliana*, les boîtes A et B ont des séquences consensus : 5'-Tg/aGC/tNNAGTg/tG-3' pour la boîte A et 5'-G/aGTTCGANNC-3' pour la boîte B. La boîte A est située aux positions 8 à 18 et la boîte B se situe entre 30 et 60 nucléotides après la boîte A. La boîte B se situe généralement 13 à 22 nucléotides avant la fin de la séquence codant l'ARNt. On retrouve également la région *TATA box-like* dans les 50 nucléotides en amont de l'ADNt. Chez *A. thaliana* cette séquence est particulièrement riche en AT, à 64 %, comparée à celles d'autres organismes photosynthétiques comme *C. reinhardtii* (seulement 36% riche en AT). Contrairement aux autres eucaryotes, 78% des ADNt d'*A. thaliana* possèdent un triplet CAA parmi les 10 nucléotides en amont des gènes. Ce triplet a été montré comme étant le site d'initiation et de ré-initiation de la traduction chez *Nicotiana tabacum* [17]. Cette séquence est retrouvée chez de nombreuses plantes à fleurs suggérant une conservation du mécanisme d'initiation de la traduction [13]. On retrouve aussi chez les angiospermes (dont *A. thaliana*) la région riche en résidus T en 3' des ADNt. En effet, selon les espèces étudiées, entre 90 à 97% des ADNt possèdent une région riche en T dans les 25 nucléotides en aval de l'ADNt.

Chez la levure *S. cerevisiae*, les ARNt sont transcrits par l'ARN polymérase III (Pol III) au sein du noyau. Cette enzyme transcrit d'ailleurs d'autres ARN non codants comme l'ARNr 5S ou le petit ARN non codant U6 [18,19]. Dans un premier temps, le processus de transcription est pré-initié par le facteur de transcription TFIIIC (Transcription Factor IIIC) qui reconnaît les boîtes A et B des ADNt. L'interaction de TFIIIC avec les boîtes A et B permet par la suite le recrutement d'un second complexe, TFIIIB, constitué de trois sous-unités : la protéine Bdp1 (B double prime 1), la protéine Brf1 (TFIIIB related factor 1) et la protéine TBP (TATA Binding Protein). Le complexe TFIIIB interagit, quant à lui, avec l'extrémité 5' de l'ADNt et notamment avec la *TATA box-like*. L'ensemble TFIIIC-TFIIIB-ADNt permet ensuite le recrutement de la Pol III et l'initiation de la transcription [18,20]. Le processus de transcription s'arrête au niveau de la région riche en T en aval de l'ADNt et donne naissance à un précurseur d'ARNt (pré-ARNt).

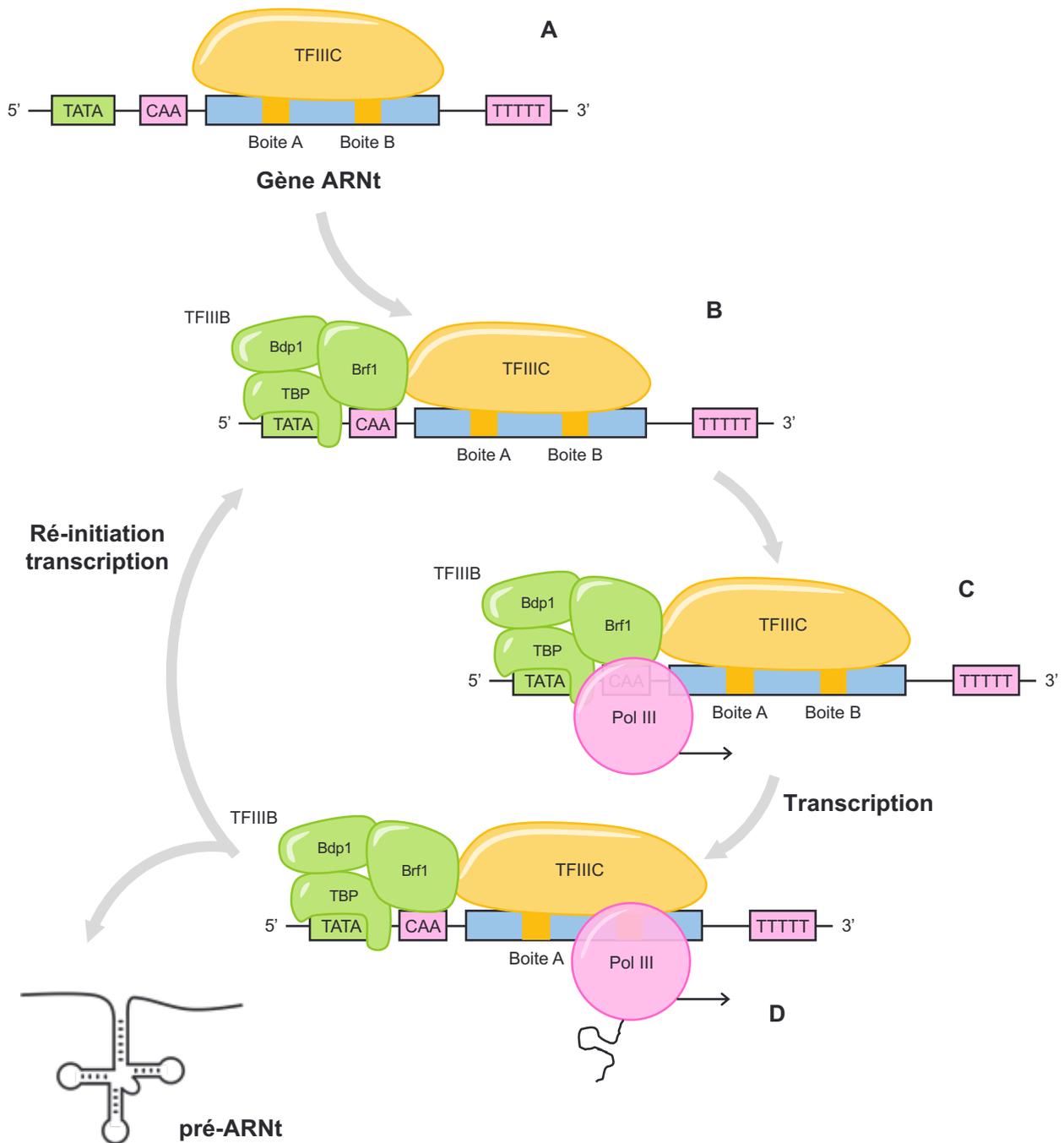


Figure 3 : Schéma du cycle de transcription des pré-ARNT nucléaires chez *A. thaliana*. A : Formation du complexe de pré-initiation TFIIIC. B : recrutement du second complexe TFIIIB composé de Bdp1, Brf1 et TBP. C : Initiation de la transcription par la Pol III. D : Terminaison de la transcription et libération du pré-ARNT. La transcription peut être ré-initié et repartir de l'étape A.

La protéine La reconnaît la queue poly(U) du pré-ARNt et son interaction permet de stabiliser le transcrit avant sa maturation [21]. Ce processus de transcription peut être réinitié plusieurs fois sur le même ADNt, permettant ainsi une forte expression des gènes d'ARNt.

Il a été admis que chez les plantes le mécanisme de transcription nucléaire est proche de celui observé chez la levure. On note cependant comme différence, l'initiation/ré-initiation de la transcription au niveau de triplet CAA en amont de l'ADNt [13,17]. La Figure 3 récapitule les étapes de transcription chez *A. thaliana*.

Chez *A. thaliana*, comme tous les organismes photosynthétiques, la transcription des ARNt s'effectue dans trois compartiments cellulaires : noyau, mitochondries et chloroplastes. Les mécanismes moléculaires sont semblables mais certains composants sont spécifiques du compartiment cellulaire.

Ainsi, chez *A. thaliana*, la transcription des gènes d'ARNt mitochondriaux est réalisée par des ARN polymérase de type phagique, les NEP (Nuclear-Encoded Polymerase). Ces polymérase sont codées par le génome nucléaire puis importées dans la mitochondrie. Aucun facteur de transcription n'a été retrouvé associé aux polymérase NEP lors de la transcription. La transcription, réalisée à partir de promoteurs répartis sur l'ensemble du génome, génère des transcrits polycistroniques.

La transcription dans le chloroplaste est réalisée par les mêmes ARN polymérase NEPs que celles retrouvées dans la mitochondrie, mais aussi par une polymérase PEP (Plastid-Encoded Polymerase) de type bactérien codée par le génome chloroplastique. Chaque polymérase NEP et PEP reconnaît des régions promotrices qui leurs sont spécifiques (pour une revue voir [22]).

A l'issue de la transcription, un pré-ARNt est généré. Il est caractérisé par des extrémités en 5' et 3' appelées *5'leader* et *3'trailer* et certains possèdent encore un intron.

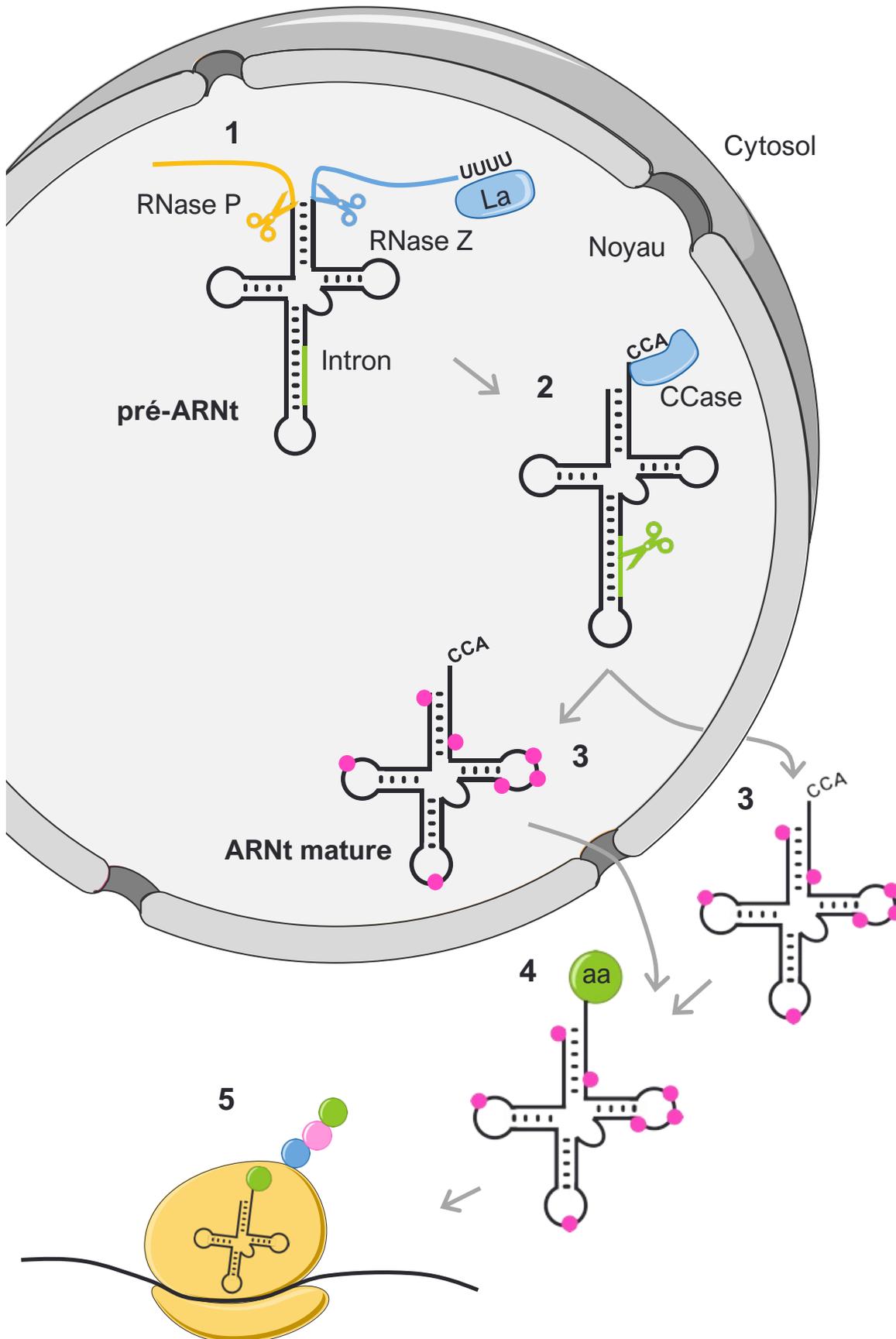


Figure 4 : Schéma récapitulatif des étapes de maturation des pré-ARNt nucléaires chez les eucaryotes. 1 : Maturation des extrémités 5' et 3' des pré-ARNt par les RNases P et Z. 2 : Ajout du triplet CCA en 3' par la CCase et épissage de l'intron (cas des vertébrés où cette étape est réalisée dans le noyau. Chez la levure, cette étape se déroule dans le cytosol à la surface des mitochondries). 3 : Modifications de nucléosides. 4 : Aminoacylation des ARNt matures par des aminoacyl ARNt synthétases. 5 : Prise en charge par les ribosomes.

4. Maturation des ARNt

Après leur transcription, les pré-ARNt sont maturés au cours de plusieurs étapes de modifications post-transcriptionnelles. Ils seront alors prêts pour être aminoacylés et complètement fonctionnels pour la traduction. Les étapes de maturation sont résumées dans la Figure 4.

i. Maturation des extrémités 5' et 3'

Le pré-ARNt tout juste transcrit possède les extrémités 5'*leader* et 3'*trailer*. La première étape de maturation consiste à éliminer les séquences à ces extrémités.

La maturation de l'extrémité 5' des pré-ARNt a lieu au sein du noyau. A l'issue de cette étape, une dizaine de bases en 5' sont éliminées. Elle est effectuée par les ribonucléases P (RNase P) dont il existe deux types. La première à avoir été identifiée est une ribonucléoprotéine (RNP) [23,24]. Elle est présente dans les trois domaines de la vie, à savoir les bactéries, les archées et les eucaryotes. Elle est composée d'une protéine et d'un ARN (ARN P), ce dernier étant essentiel pour son activité catalytique [25]. L'ARN P interagit spécifiquement avec des bases situées dans les bras D et T des pré-ARNt et une partie interagit avec les extrémités 5' et 3' [26]. Une autre catégorie de RNase P, uniquement protéique, a été identifiée dans les mitochondries humaines. L'activité RNase P y est réalisée grâce à un complexe composé des protéines MRPP1, 2 et 3 (*Mitochondrial RNase P Protein*) [27].

Chez *A. thaliana*, l'extrémité 5'*leader* est maturée par trois RNase P orthologues de MRPP3, les protéines PRORP (*Proteinaceous RNase P*). Ces enzymes entièrement protéiques contiennent deux à trois motifs pentatricopeptides (PPR) et un domaine catalytique NYN. Au sein du noyau, ce sont les enzymes PRORP2 et PRORP3 qui maturent les pré-ARNt codés par le génome nucléaire. PRORP1 quant à elle est responsable de la maturation des pré-ARNt dans la mitochondrie et le chloroplaste [28,29].

L'extrémité 3' du pré-ARNt est quant à elle éliminée par la RNase Z. La RNase Z clive après la base discriminatrice en position 73. Elle interagit avec le bras T et la tige acceptrice du pré-ARNt [30].

Chez *A. thaliana*, cette extrémité est éliminée par quatre RNases Z qui sont localisées dans les différents compartiments cellulaires. Ainsi, la RNase AthTrZ^{S1} est localisée dans le cytosol, AthTrZ^{S2} dans les chloroplastes, AthTrZ^{L1} dans les mitochondries et le noyau et AthTrZ^{L2} uniquement dans les mitochondries [30,31].

ii. Ajout du triplet CCA

Pour que l'ARNt puisse être aminoacylé, il doit posséder un triplet de nucléotides CCA à son extrémité 3'. Chez certaines bactéries et quelques archées, les ADNt possèdent une séquence codant le CCA en 3' des ARNt. Le triplet CCA étant déjà présent sur les pré-ARNt, il n'a pas besoin d'être ajouté. En revanche, chez les eucaryotes, aucun ADNt ne code pour le triplet CCA [32]. Il doit donc être ajouté post-transcriptionnellement. Cette étape est réalisée par une enzyme appelée tRNA-nucleotidyl transférase, ou communément CCCase, qui catalyse l'addition du triplet CCA sur l'extrémité 3' maturé par la RNase Z [33].

Chez *A. thaliana*, une seule CCCase a été identifiée. Elle est retrouvée dans le cytosol, la mitochondrie et le chloroplaste. Sa séquence indique également une prédiction de localisation nucléaire mais elle n'a pas pu être confirmée [34].

iii. Épissage des introns

Dans les trois domaines de la vie, un certain nombre de gènes d'ARNt contiennent des introns. Ils doivent être éliminés des pré-ARNt pour permettre une maturation complète en ARNt fonctionnel.

Ce mécanisme a été bien caractérisé pour les ARNt nucléaires de levure. Chez *S. cerevisiae*, plus de 20 % des gènes d'ARNt possèdent un intron (entre les positions 37 et 38) d'une longueur variable entre 14 et 60 nucléotides. Ces introns ont une fonction essentielle dans les processus de maturation des ARNt. Par exemple pour l'ARNt^{Tyr}, il a été montré que la modification de la base U35 de l'anticodon en pseudouridine n'a lieu qu'uniquement lorsque le pré-ARNt contient l'intron. Sans intron, cet ARNt n'est pas modifié et ne peut pas être pris en charge par les voies de contrôle qualité [35].

L'épissage des introns chez la levure a lieu en trois étapes [36]. La première consiste à enlever les introns contenus dans les pré-ARNt. Cette étape est catalysée par des endonucléases spécifiques faisant partie du complexe Sen (Sen2, Sen34, Sen15, Sen54). Il a été montré que Sen 2 et Sen34 sont les sous-unités qui sont responsables du clivage en 5' et 3' de l'intron [37]. Après clivage, on obtient alors un pré-ARNt en deux parties qui doivent être assemblées ensemble : c'est la deuxième étape. La ligation des deux exons du pré-ARNt est catalysée par l'ARNt ligase Trl1. Cette réaction nécessite au préalable une modification des extrémités des exons (2', 3' cyclic phosphate pour l'exon en 5' et une extrémité 5'OH pour l'exon en 3') par différentes enzymes pour permettre à la ligase Trl1 d'assembler les deux exons ensemble [38]. La troisième étape permet d'éliminer le 2' phosphate restant à la jonction de l'épissage par l'enzyme 2'phosphotransférase, Tpt1, conservée chez les eucaryotes. Au sein des eucaryotes, la localisation de l'épissage des introns varie. Chez la levure, l'épissage des introns a lieu dans le cytosol à la surface des mitochondries alors que chez les vertébrés cette étape se déroule dans le noyau [39,40].

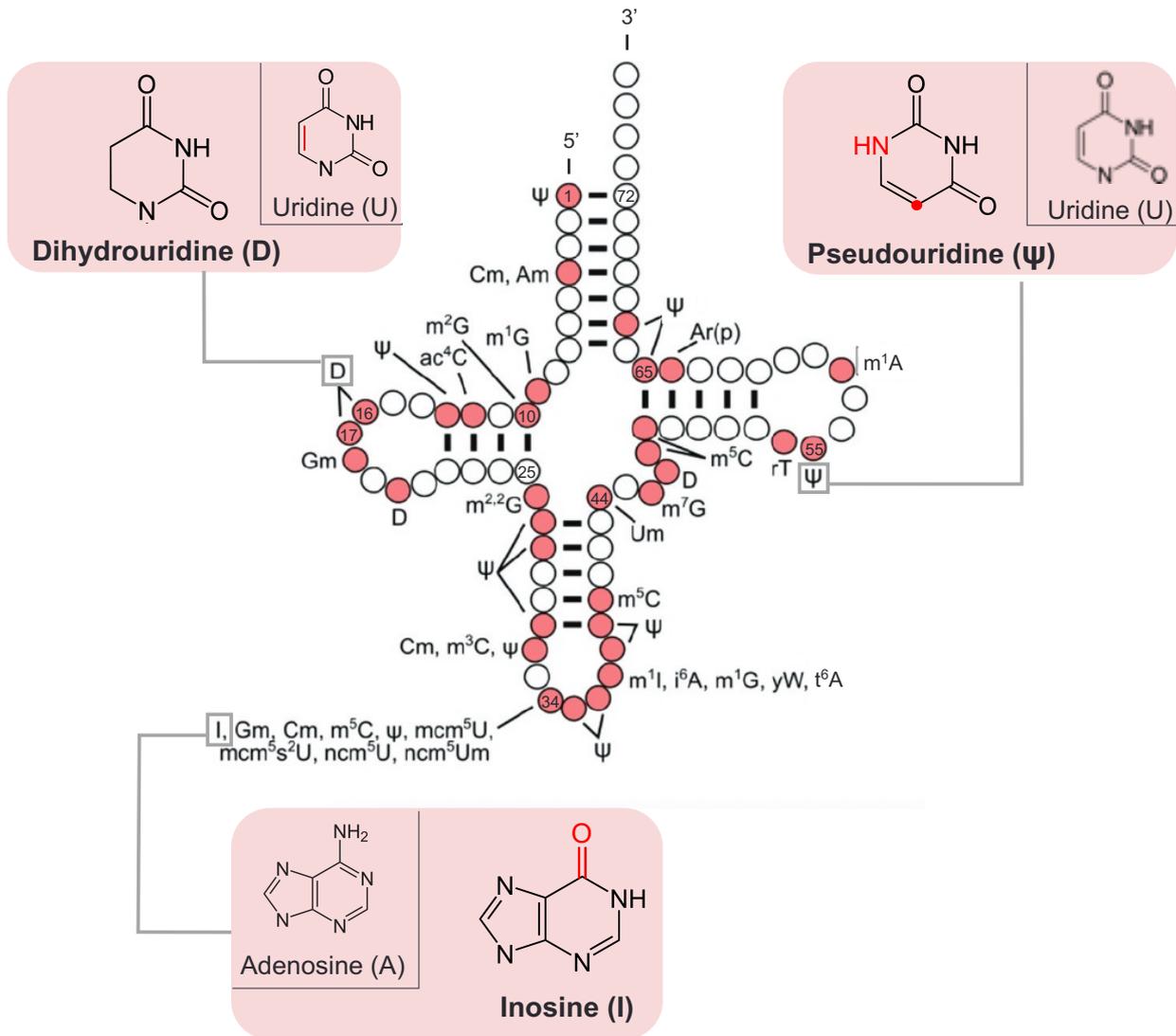


Figure 5 : Représentation d'un ARNt et des bases modifiées chez le modèle *S. cerevisiae*. Adaptée de [18]. Chaque base modifiée est symbolisée par un rond de couleur et les modifications possibles pour cette base sont indiquées. La position des bases est indiquée à l'intérieur des ronds.

Abréviations : pseudouridine (Ψ) ; 2'-O-methylcytidine (Cm) ; 2'-O-methyladenosine (Am) ; 1-methylguanosine (m¹G) ; N2-methylguanosine (m²G) ; N4-acetylcytidine (ac⁴C) ; dihydrouridine (D) ; 2'-O-methylguanosine (Gm) ; N2,N2-dimethylguanosine (m^{2,2}G) ; 3-methylcytidine (m³C) ; inosine (I) ; 5-methylcytidine (m⁵C) ; 5-methoxycarbonyl-methyluridine (mcm⁵U) ; 5-methoxycarbonylmethyl-2-thiouridine (mcm⁵s²U) ; 5-carbamoylmethyluridine (ncm⁵U) ; 5-carbamoylmethyl-2'-O-methyluridine (ncm⁵Um) ; 1-methylinosine (m¹I) ; N6-isopentenyladenosine (i⁶A) ; wybutosine (yW) ; N6-threonylcarbamoyladenosine (t⁶A) ; 2'-O-methyluridine (Um) ; 7-methylguanosine (m⁷G) ; ribothymidine (rT) ; 1-methyladenosine (m¹A) ; 2'-O-ribosyladenosine phosphate (Ar(p)).

Chez les plantes terrestres, deux familles d'ARNt nucléaires possèdent un intron, l'ARNt^{Met} (élongateur) et l'ARNt^{Tyr}, également entre les bases 37 et 38, d'une longueur variable entre 6 et 22 nucléotides [13]. Le mécanisme d'épissage des introns des ARNt nucléaires semble conservé chez les plantes. Chez *A. thaliana*, on retrouve des endonucléases AtSen1 et AtSen2 qui permettent le clivage de l'intron en 5' et 3'. La ligation des deux exons est réalisée par AtRNL et le 2' phosphate résiduel est éliminé par Tpt1 [41].

Chez *A. thaliana*, 6 ARNt chloroplastiques possèdent un intron, plus structuré et de grande taille comparé aux introns d'ARNt nucléaires. Ces introns sont épissés dans le chloroplaste par des protéines à motifs PPR (Pentatricopeptide Repeat) dont OTP51 et THA8 chez *A. thaliana* [13,42–44].

Les ARNt mitochondriaux d'*A. thaliana* ne possèdent pas d'introns et n'ont donc pas besoin de subir cette étape.

iv. Modifications de nucléosides

Les ARNt sont les ARN les plus modifiés post-transcriptionnellement avec environ 12 % de leurs nucléotides portant une modification. Au total, 93 types de modification ont été recensés à ce jour (RNA Modification Database [45]; MODOMICS [46]). Les techniques de spectrométrie de masse et de séquençage à haut débit permettent d'identifier ces modifications mais les limites techniques de ces approches restent encore très présentes [47–49]. Beaucoup de ces modifications sont conservées chez les eucaryotes, bactéries et archées. Des modifications très connues sont la dihydrouridine (D) située dans le bras D et la ribothymidine (rT) dans la boucle T. On retrouve comme type de modification, entre autres, des additions de groupes méthyl aux bases, des acétylations, des modifications d'uridine en pseudouridine (Ψ) ou dihydrouridine (D), ou encore la désamination d'adénine en inosine (I) (Figure 5) [18]. Ces modifications ont lieu tout au long du processus de maturation des pré-ARNt. Elles sont catalysées par de nombreuses enzymes de modifications et ont lieu dans le noyau mais aussi dans le cytosol [50].

Les modifications sont très importantes et jouent un rôle à la fois pour la structure et la stabilité des ARNt mais aussi pour l'aminocyclation et la traduction. Les modifications qui sont importantes pour la structure et la stabilité des ARNt sont plutôt situées sur les bases en dehors de la zone de l'anticodon alors que celles impliquées dans la traduction se situent dans la boucle de l'anticodon.

Ainsi, on retrouve des méthylations de bases à la jonction entre des tiges et des boucles qui empêchent la formation de duplexe d'ARN (la méthylation empêche l'appariement des bases proches) ce qui facilite la formation de la structure tridimensionnelle de l'ARNt. La présence de dihydrouridine (D) dans le bras D améliore la flexibilité de l'ARNt alors que la présence de pseudouridine (Ψ), constitutivement présente dans le bras T, renforce les interactions entre les bases et par conséquent la structure de l'ARNt.

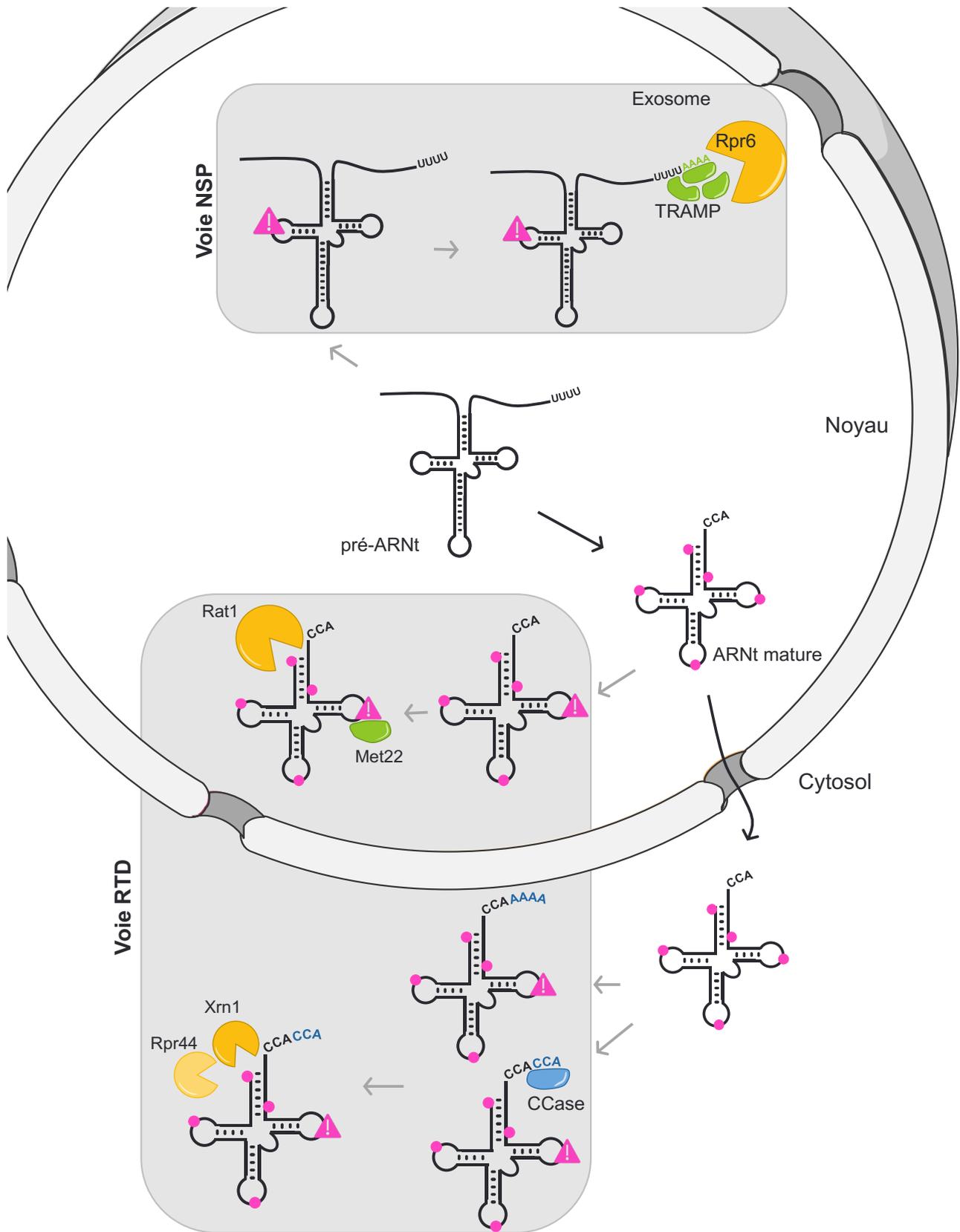


Figure 6 : Schéma récapitulatif des voies de dégradation des ARNt. Lorsqu'un pré-ARNt est aberrant, il est pris en charge par la voie NSP (*Nuclear Surveillance Pathway*) dans le noyau. le complexe TRAMP interagit à son extrémité 3' et ajoute une séquence poly(A) en 3'. L'exosome dégrade le pré-ARNt en 3'. Les ARNt matures aberrants sont pris en charge par la voie de dégradation RTD (*Rapid tRNA Decay*). Dans le noyau, ils sont reconnus par Met22 qui induit une dégradation de l'ARNt mature par Rat1 en 5'. Dans le cytosol, ils sont soit modifiés en 3' par ajout de poly(A) ou par une CCCase qui ajoute un triplet CCA puis ils sont reconnus et dégradés par Xrn1 et Rpr44.

Il a aussi été montré que les modifications hautement conservées du bras T (5-méthylcytidine m^5C , ribothymidine rT ou pseudouridine Ψ) sont essentielles pour l'interaction tridimensionnelle entre le bras D et T [18,51].

Une partie des modifications touche aussi les bases de la boucle de l'anticodon. Les bases aux positions 32, 34, 37 et 38 sont très fréquemment modifiées et ont un rôle important dans la traduction et la croissance de la cellule. Les modifications des bases de l'anticodon peuvent impacter la lecture et le décodage du codon ou l'interaction entre les bases codon-anticodon lors de la traduction au sein du ribosome [18,52]. La modification de la purine en position 37 permet de moduler la structure de la boucle de l'anticodon et stabiliser les interactions entre le codon et l'anticodon [18,53]. Certaines modifications confèrent une identité à l'ARNt. Par exemple, la modification mcm^5s^2U dans la boucle de l'anticodon en position 34 est un déterminant d'identité de l'ARNt^{Lys} chez *E. coli*. L'absence des modifications conférant l'identité à l'ARNt entraîne généralement un déficit de reconnaissance par les ARNt synthétases et une diminution de l'aminocyclation [54]. Dans d'autres cas, ces modifications empêchent l'aminocyclation de l'ARNt avec un acide aminé non apparenté, on les appelle les anti-déterminants [54].

Les modifications des ARNt sont des processus dynamiques. Elles peuvent être régulées en réponse à des stress. Par exemple, en réponse à un traitement au peroxyde d'hydrogène, une augmentation des modifications 2'-O-méthylcytidine (Cm), 5-méthylcytidine (m^5C) et N2,N2-diméthylguanosine (m^2_2G) sur les ARNt de levure *S. cerevisiae* ont été montrés. Les individus n'ayant pas les enzymes responsables de ces modifications sont plus sensibles au stress. Il est ainsi suggéré que ces modifications sont importantes pour stimuler une réponse efficace [55].

5. Contrôle qualité des ARNt et leurs voies de clivages

Les ARNt étant impliqués dans de nombreux processus vitaux pour la cellule, dont la traduction, ils doivent être contrôlés pour éviter la présence d'ARNt dysfonctionnels. Ils subissent donc un contrôle qualité. Tout pré-ARNt ou ARNt hypomodifié et/ou mal mûré de manière plus générale sera éliminé. La dégradation des ARNt se fait via deux voies de contrôle qualité : la *Nuclear Surveillance Pathway* (NSP) et la voie *Rapid tRNA Decay* (RTD) (Figure 6). Ces deux voies sont détaillées dans la revue de Megel et al. [56].

i. *Nuclear Surveillance Pathway*

La voie du *Nuclear Surveillance Pathway* se charge de dégrader en particulier les ARNt hypomodifiés, ayant encore leur intron ou mal mûrés aux extrémités, au sein du noyau. Cette voie décrite principalement chez *S. cerevisiae* est commune aux eucaryotes. Il semble que ce ne soit pas l'absence d'une modification en particulier qui déclenche la dégradation de l'ARNt, mais plus une structure tertiaire globale aberrante.

Dans un premier temps, le pré-ARNt à dégrader va subir une polyadénylation à son extrémité 3' par le complexe TRAMP. Ce complexe est aussi retrouvé dans les voies de dégradation des ARNr et des petits ARN nucléaires (*small nucleolar RNA*). Le complexe TRAMP est constitué d'une polymérase poly(A) (Trf4), d'une hélicase à ARN (Mtr4) et d'une protéine à doigt de zinc (*zinc knuckle*) (Air2). Cette dernière interagit avec une 3' exoribonucléase de l'exosome nucléaire (Rrp6). Après polyadénylation par Trf4, c'est l'exosome qui va dégrader le pré-ARNt par l'extrémité 3' [57].

ii. *Rapid tRNA Decay Pathway*

La voie *Rapid tRNA Decay*, quant à elle, se charge de dégrader les ARNt matures aberrants et non-fonctionnels. C'est aussi chez la levure que cette voie a été bien caractérisée en premier. Il a été observé chez des souches de levure défectueuses pour certaines enzymes de modifications que leurs ARNt étaient dégradés. En conditions normales de croissance, ces souches produisent des ARNt hypomodifiés mais correctement formés et tout à fait fonctionnels. Lorsque les levures sont soumises à un stress thermique à 37°C, ces ARNt hypomodifiés présentent des conformations aberrantes, les rendant non fonctionnels et sont alors dégradés [58,59]. Ces ARNt matures non conformes sont dégradés par la voie *Rapid tRNA Decay* impliquant la protéine Met22 (Methionine-requiring protein 22) et deux 5'-3' exonucléases Rat1 et Xrn1. Cette dernière est localisée dans le cytosol et permet de dégrader les ARNt matures dans ce compartiment cellulaire. En revanche, Rat1 est localisée dans le noyau ce qui suggère que les ARNt peuvent aussi être importés dans le noyau pour y être dégradés [60].

D'autres enzymes, telles que la CCCase, peuvent marquer les ARNt à dégrader. En effet, il a été montré que cette enzyme est capable d'ajouter la séquence CCACCA à l'extrémité 3' des ARNt hypomodifiés ou dont le bras accepteur est défectueux, à condition que l'ARNt possède deux résidus G en 5'. L'extension CCACCA simple brin est détectée et l'ARNt est dégradé par les exonucléases Rrp44 et Xrn1. D'autres ARNt sont marqués pour la dégradation par l'ajout de petites séquences poly(A) après le triplet CCA en 3' [61].

Cependant, il est intéressant de noter que certaines protéines peuvent protéger les ARNt aberrants en interagissant avec eux, entrant alors en compétition avec les exonucléases Rat1 et Xrn1. C'est le cas pour l'ARNt^{Val} qui, en interagissant avec la valyl-ARNt synthétase et le facteur d'élongation eEF1A, est protégé et délivré au ribosome [58].

6. Fonction principale des ARNt : acteurs de la traduction

La fonction principale des ARNt est de participer à la traduction en apportant les acides aminés au ribosome. Pour ce faire, les ARNt doivent auparavant être aminoacylés.

i. Aminoacylation des ARNt

L'étape d' aminoacylation consiste à attacher un acide aminé à l'extrémité CCA en 3' de l'ARNt. Elle est catalysée par les aminoacyl-ARNt synthétases (aaRS). Chaque famille d'ARNt isoaccepteurs spécifiques d'un acide aminé est aminoacylée par une aaRS spécifique. L' aminoacylation se déroule principalement dans le cytosol, en deux étapes. Durant la première, les aaRS activent l'acide aminé en aminoacyl-adénylate (aa-AMP) en présence d'ATP. La deuxième étape permet aux aaRS de transférer l'acide aminé aa-AMP sur l'extrémité acceptrice de l'ARNt (CCA). L'acide aminé aa-AMP est fixé sur l'extrémité 3'-OH de l'adénine par des liaisons ester [62].

Les aaRS sont des acteurs clés pour maintenir la fidélité entre le codon et l'acide aminé correspondant. Des erreurs peuvent conduire à l'incorporation de mauvais acides aminés au sein des chaînes peptidiques et la naissance de protéines mutées. Si ces erreurs ne sont pas détectées par la cellule, il peut en découler des dysfonctionnements de mécanismes biologiques voire la mort cellulaire. Les aaRS ont un système appelé double tamis qui permet de sélectionner les bons acides aminés à transférer sur l'ARNt. Une première sélection a lieu au sein du site actif d' aminoacylation qui exclut les acides aminés en fonction de leur taille (encombrement stérique) ou de leur capacité à interagir. La seconde sélection se déroule dans le site actif hydrolytique qui peut hydrolyser les acides aminés arrivés par erreur [63,64].

En règle générale, chez les archées et les eucaryotes, chaque aaRS est spécifique d'un acide aminé et d'une famille d'ARNt isoaccepteurs. Les aaRS reconnaissent spécifiquement les ARNt grâce à des résidus appelés les éléments déterminant d'identité. Ces résidus sont la plupart du temps de simples nucléotides, des paires de nucléotides ou des motifs particuliers. Pour certains ARNt particuliers ce sont des bases modifiées qui sont élément d'identité comme par exemple les bases Ψ 34, A35 et Ψ 36 de l'anticodon de l'ARNt^{leu} de levure. Pour beaucoup d'ARNt ce sont les bases de l'anticodon ainsi que la base discriminatrice en position 73 (en 5' du triplet CCA du bras accepteur) qui sont les principaux éléments d'identité. Pour une minorité d'ARNt ces déterminants d'identité peuvent se situer sur les autres bases du bras accepteur ou dans les régions D et T (pour une revue, voir [54]). Il existe une particularité pour l'ARNt^{his} auquel une base guanine (G) est ajoutée en première position. Dans ce cas, c'est cette première base qui est essentielle pour la reconnaissance de son aaRS [65].

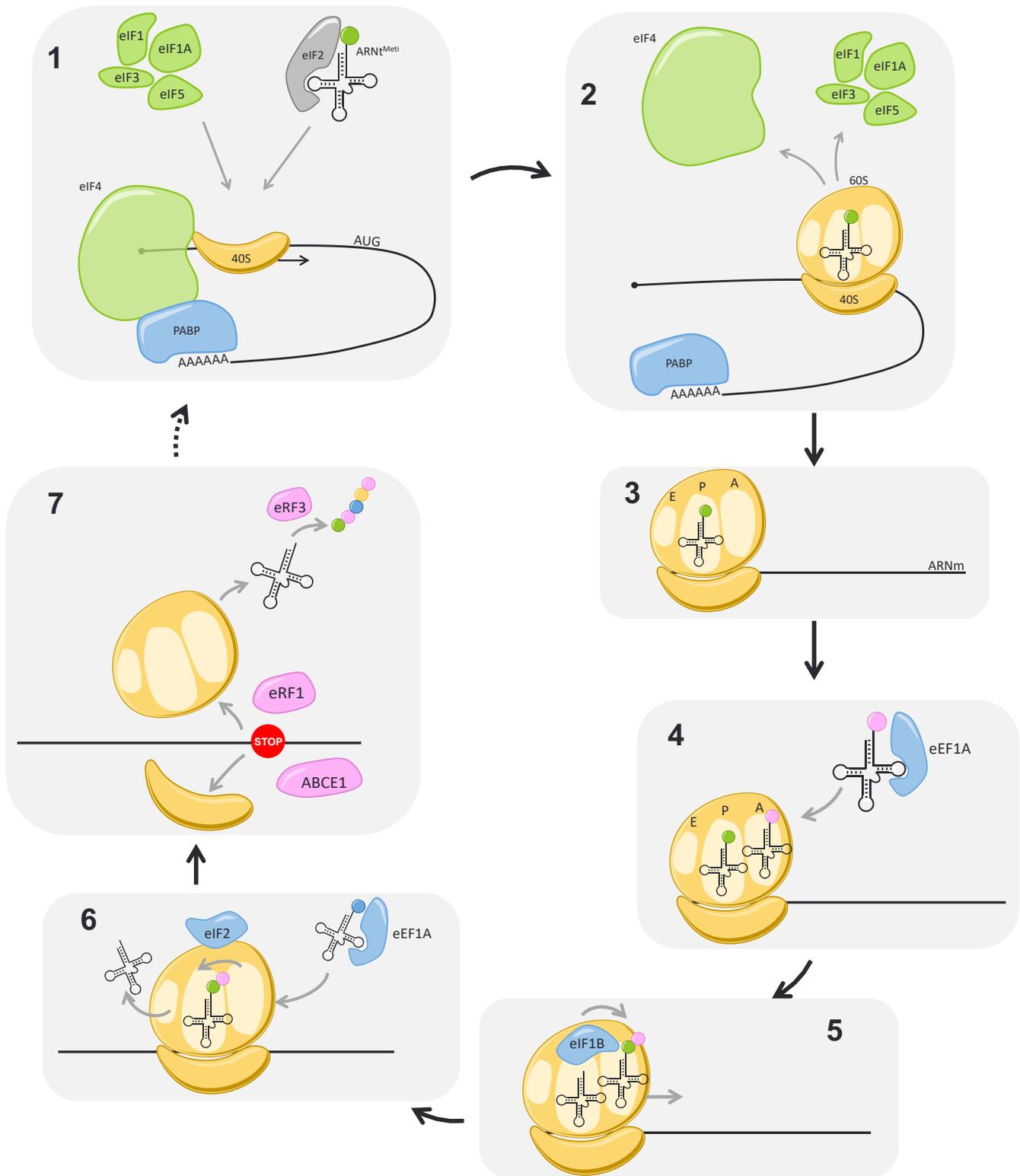


Figure 7 : Schéma récapitulatif des étapes de traduction des ARNm dans le cytosol, chez les eucaryotes. 1 : Initiation. Interaction entre l'ARNm, la petite sous-unité ribosomale 40S, le complexe eIF4 au niveau de la coiffe et PABP sur l'extrémité poly(A). Recrutement des facteurs d'initiation eIF1, eIF1A, eIF3 et eIF5 et de l'ARNt^{Meti} initiateur reconnu par eIF2. 2 : Initiation suite. Au niveau du codon d'initiation AUG, recrutement de la grande sous-unité 60S et libération des facteurs d'initiation. 3 : Elongation. Le premier ARNt est situé dans le site P du ribosome. 4 : Elongation suite. Le facteur d'élongation eEF1A sélectionne les ARNt aminoacylés qui se lient au site A du ribosome. 5 : Elongation suite. Transpeptidation effectuée avec eEF1B. 6 : Elongation suite. Translocation effectuée avec eEF2. Sélection d'un nouvel ARNt aminoacylé par eEF1A. 7 : Terminaison. Arrivés à un codon STOP, les ARNt présents dans le ribosome et la chaîne peptidique sont libérés, les sous-unités ribosomales sont éjectées de l'ARNm avec l'aide de eRF1, eRF3 et ABCE1.

Chez *A. thaliana*, 45 gènes d'aaRS exprimés ont été identifiés dans le génome nucléaire. Environ la moitié de ces aaRS ont un signal d'adressage et une localisation mitochondriale et/ou chloroplastique. Une partie des aaRS est donc importée dans les mitochondries et/ou les chloroplastes pour aminoacyler les ARNt qui sont transcrits dans ces compartiments [66].

ii. Le mécanisme de traduction

La synthèse de protéines consiste à assembler des acides aminés entre eux par des liaisons peptidiques. Elle nécessite une machinerie très complexe impliquant des ribosomes, des ARN, des protéines, des acides aminés et de l'énergie. Autant d'acteurs sont requis pour assurer une haute-fidélité de la séquence protéique par rapport à l'information génétique codée par l'ARNm. Le mécanisme de traduction chez les eucaryotes peut être divisé en quatre étapes : l'initiation, l'élongation, la terminaison et le recyclage des ribosomes. L'ensemble des étapes de la traduction cytosolique chez les eucaryotes est résumé dans la Figure 7.

La traduction dans le cytosol

Chez les procaryotes et les eucaryotes, l'acteur principal de la traduction est le ribosome. Le ribosome eucaryotique est constitué de deux sous-unités : la petite sous-unité 40S et la grande sous-unité 60S. Dans un premier temps, la synthèse protéique commence par l'interaction de la petite sous-unité du ribosome avec les facteurs d'initiation eIF1, eIF1A, eIF3 et eIF5. À ces composants s'ajoutent l'ARNt initiateur (ARNt^{Meti}) reconnu par eIF2. L'ensemble forme le complexe de pré-initiation (PIC). Le complexe d'initiation eIF4 (composé de eIF4A, eIF4B, eIF4E et eIF4G) reconnaît quant à lui l'extrémité 5' de l'ARNm (la coiffe) et va l'amener au complexe PIC. L'extrémité poly(A) en 3' de l'ARNm est protégée des dégradations par la protéine PABP qui reconnaît cette région. PABP et eIF4G interagissent ensemble et de ce mécanisme en découle une circularisation de l'ARNm. L'ensemble de ces complexes scanne l'ARNm jusqu'à ce que l'anticodon de l'ARNt^{Meti} soit correctement positionné par rapport au codon initiateur complémentaire. Le codon AUG est le principal codon initiateur. Il existe cependant d'autres codons moins typiques qui servent aussi d'initiateur, comme par exemple le codon GUG. À ce moment, l'interaction entre l'anticodon et le codon initiateur interrompt le processus de lecture du complexe PIC et permet l'éjection des facteurs d'initiation. La grande sous-unité ribosomale 60S peut alors être recrutée pour former le complexe d'initiation 80S. L'ARNt^{Meti} est lié au ribosome 80S au sein du site P (Peptidyl). Le complexe ribosomal est maintenant compétent pour entrer dans la phase d'élongation de la synthèse protéique. Il est suggéré qu'autant de facteurs d'initiation soient requis pour contrôler minutieusement la synthèse protéique car ce processus est très gourmand en énergie et pourrait avoir de lourdes conséquences s'il est utilisé inutilement [67,68].

Durant la phase d'élongation, les ARNt aminoacylés sont sélectionnés avec l'aide du facteur d'élongation eEF1A grâce à un appariement de bases entre l'anticodon et le codon de l'ARNm, et vont se lier au sein du site A (Aminoacyl) du ribosome. Le ribosome, avec l'aide de eEF1B, catalyse le transfert du groupe méthionyl de l'ARNt^{Meti} au groupe amine de l'ARNt aminoacylé situé dans le site A : c'est la transpeptidation. Ceci permet la formation de la première liaison entre la méthionine et le premier acide aminé de la chaîne peptidique. La chaîne peptidique est maintenant portée par l'ARNt situé dans le site A.

L'ARNm ainsi que l'ARNt situé dans le site A vont se déplacer, ensemble, vers le site P. Cette étape s'appelle la translocation. Le facteur responsable de cette translocation est eEF2. Le codon de l'ARNm situé maintenant dans le site A, libre, va permettre à nouveau l'interaction et la sélection d'un nouvel ARNt aminoacylé grâce à l'appariement des bases entre l'anticodon et le codon. Ce mécanisme de sélection d'ARNt aminoacylé – transpeptidation – translocation va se répéter jusqu'à ce qu'un codon STOP soit lu [67,69].

Lorsque le ribosome lit un codon STOP, aucun ARNt n'interagit au sein du site A. Ce sont les facteurs de libération eRF1 et eRF3 qui vont interagir avec le ribosome au niveau du site A et catalyser l'hydrolyse de la chaîne peptidique dans le site P. Le facteur eRF1 est composé de différents domaines permettant la reconnaissance des codons STOP dans le site A, la libération de la chaîne polypeptidique au niveau du site P et l'interaction avec les facteurs eRF3 et ABCE1. Le facteur eRF3 catalyse la réaction d'hydrolyse de la chaîne peptidique encore liée au dernier ARNt. Ce mécanisme constitue l'étape de terminaison de la synthèse protéique et permet la libération de la protéine. Un dernier facteur protéique, ABCE1, intervient pour faciliter l'éjection de l'ARNm et des derniers ARNt encore présents dans le ribosome, et la dissociation des deux sous-unités du ribosome. Ceci permet un recyclage du ribosome puisque les sous-unités vont être disponibles pour un nouveau cycle de synthèse protéique [67,70].

Chez les plantes, la machinerie de traduction cytosolique est semblable à celle des eucaryotes. On retrouve cependant quelques particularités, notamment à l'étape d'initiation. En plus des facteurs connus chez les autres eucaryotes, chez *A. thaliana* il existe un complexe supplémentaire, eFiso4, qui diffère du complexe canonique eIF4. Chaque complexe semble avoir des affinités pour certains ARNm [71].

La traduction dans les organites

Chez les eucaryotes, la traduction a aussi lieu dans les organites (mitochondries et chloroplastes chez les plantes). Leurs génomes étant originaires d'endosymbioses bactériennes, la traduction est basée sur le modèle bactérien. Les composants sont différents de ceux de la traduction cytosolique. Ici, uniquement les traductions mitochondriale et chloroplastique chez *A. thaliana* vont être présentées.

Chez *A. thaliana*, le génome mitochondrial code pour ses propres ARN ribosomiques (mt-ARNr) et une partie de ses ARNt. Le mitoribosome est proche du ribosome 70S bactérien. Il est constitué d'une grande sous-unité 50S et d'une petite 30S. Cependant, le mitoribosome de plante a divergé au cours de l'évolution. On retrouve une petite sous-unité plus massive que chez les procaryotes et cette sous-unité à la particularité d'avoir une extension particulière. Les ARNr sont globalement plus longs, ils contiennent des insertions qui ne sont pas présentes chez *E. coli*. De plus, des protéines spécifiques du mitoribosome des plantes, les PPR, prennent part à la traduction mitochondriale [72]. Les étapes de la synthèse protéique sont les mêmes que la traduction cytosolique. Cependant il n'existe pas d'homologues pour les facteurs d'initiation. De plus, l'initiation de la traduction a généralement lieu au niveau du codon d'initiation. L'étape d'élongation est similaire à celle effectuée dans le cytosol, impliquant trois facteurs d'élongation (EF-Tu, EF-Ts et EF-G) importés du noyau. Les codons STOP du code génétique universel sont utilisés pour la terminaison de la traduction [73].

Le génome chloroplastique code aussi pour ses propres ARNr. L'initiation de la traduction débute avec la formation d'un complexe composé de la petite sous-unité 30S du ribosome, de l'ARNt^{Metf} initiateur formylé (comme pour la mitochondrie) et de facteurs d'initiation (IF1, IF2 et IF3) qui vont reconnaître le codon initiateur sur l'ARNm. Le mécanisme de reconnaissance du codon initiateur est encore mal compris. Certains ARNm chloroplastiques contiennent la séquence Shine-Dalgarno (issu du système bactérien) qui semblerait être utilisée pour l'initiation de la traduction. Les ARNm ne possédant pas cette séquence particulière semblent être moins structurés, ce qui pourrait permettre au complexe d'initiation de reconnaître directement le codon d'initiation. Le recrutement de la grande sous-unité 50S permet ensuite la formation d'un ribosome 70S fonctionnel pour l'élongation. Trois facteurs d'élongation, homologues de EF-Tu, EF-Ts et EF-G bactériens, sont retrouvés dans le chloroplaste. La traduction se termine lorsque le ribosome atteint l'un des trois codons STOP. Deux facteurs sont impliqués lors de la terminaison : PrfA et PrfB1. Un facteur de libération du ribosome RRF, homologue de RF3 bactérien, existe aussi dans le chloroplaste et permet de dissocier les deux sous-unités 30S et 50S du ribosome et libérer l'ARNm [74].

7. Autres fonctions des ARNt

Bien que les ARNt aient une fonction majeure dans la synthèse protéique, ils sont parfois impliqués dans d'autres processus biologiques essentiels. Qu'ils soient aminoacylés ou non, ils assurent des fonctions variées. Un aperçu chez les eucaryotes des fonctions alternatives que peuvent occuper les ARNt est donné ci-après (Figure 8).

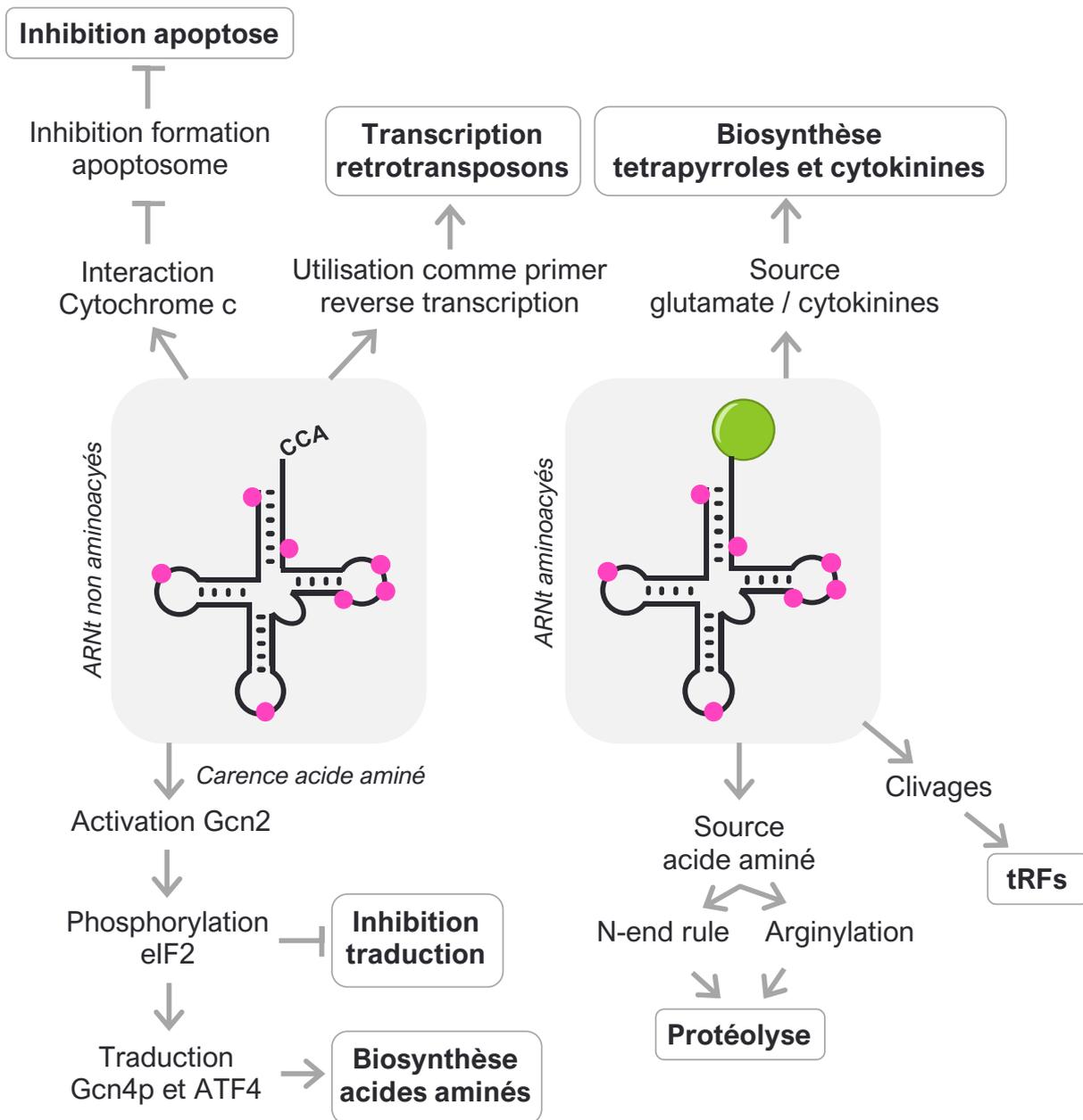


Figure 8 : Schéma récapitulatif des fonctions des ARNt chez les eucaryotes. Les ARNt non aminoacylés (carré gris à gauche) ont une fonction de régulation de l'apoptose et de la transcription de rétrotransposon. Lors de carence en acide aminé ils permettent la régulation globale de la traduction et la biosynthèse d'acides aminés. Les ARNt aminoacylés (carré gris à droite) ont des fonctions dans la synthèse de composés organiques, dans la régulation de la protéolyse et dans la biogenèse de tRFs.

i. Fonctions des ARNt non aminoacylés

Régulation de l'expression des gènes

Les ARNt non aminoacylés peuvent avoir un rôle de régulateur lors de carence en acide aminés. Chez des cellules de levures et de mammifères soumises à une carence en acide aminés, les ARNt non chargés activent la protéine kinase Gcn2p. Il est suggéré que c'est la présence d'ARNt non aminoacylés au niveau du site A du ribosome qui permet l'interaction entre Gcn2p et l'ARNt. Lorsque l'ARNt interagit avec Gcn2p, la kinase change de conformation et devient capable de phosphoryler eIF2. Le facteur d'initiation eIF2 est impliqué lors de l'initiation de la traduction et permet d'amener l'ARNt^{Meti} initiateur au complexe d'initiation de la traduction. Lorsque ce facteur est phosphorylé, il ne peut plus occuper sa fonction de manière aussi efficace qu'en conditions normales. Il en résulte donc une diminution globale de la traduction. Mais la phosphorylation de eIF2 entraîne la traduction de la kinase Gcn4p. Cette kinase agit comme un facteur de transcription et induit l'expression de gènes impliqués dans la synthèse d'acides aminés (dont ATF4). Pour pallier aux carences en acides aminés, les ARNt ont donc le double rôle d'induire le ralentissement de la synthèse protéique et de promouvoir la synthèse d'acides aminés [75,76].

Régulation de l'apoptose

L'apoptose est un mécanisme cellulaire au cours duquel les cellules sont éliminées. La régulation de l'apoptose est essentielle pour l'homéostasie cellulaire puisque ce mécanisme permet de dégrader les cellules endommagées ou nocives. C'est une mort cellulaire programmée. Chez les mammifères, lorsque le signal de l'apoptose est donné, des protéases spécifiques sont activées, les caspases. Elles permettent la dégradation des protéines de la cellule et conduisent à la phagocytose des cellules.

Une des voies de l'apoptose conduit à la perméabilisation des membranes mitochondriales, libérant le cytochrome c de la chaîne respiratoire. Dans le cytosol, le cytochrome c interagit avec le facteur APAF-1 (*apoptotic protease activating factor-1*) pour former ensemble le complexe apoptosome. Ce complexe permet ensuite l'activation des caspases qui induisent l'apoptose. Différents facteurs peuvent réguler la formation de l'apoptosome dont notamment des ARNt. Les ARNt, en interagissant avec le cytochrome c dans le cytosol, créent une compétition avec APAF-1 empêchant la formation de l'apoptosome. Cette interaction semble spécifique à certains ARNt isoaccepteurs, notamment dans des cellules cancéreuses où une accumulation des ARNt^{Arg}, ARNt^{Thr}, ARNt^{Ser} et ARNt^{Tyr} est observée [76,77].

Régulation de la rétrotranscription des éléments transposables

Les rétrotransposons sont des éléments génétiques mobiles qui se transposent dans le génome. Pour leur transposition, certains éléments passent par l'intermédiaire d'ARN notamment les LTR (*Long Terminal Repeat*) rétrotransposons. Ils possèdent des séquences répétées à leurs extrémités nécessaires pour leur rétrotranscription et la séquence PBS d'initiation (*Primer Binding Site*) en 5'. Les ARNt ayant une séquence complémentaire à la PBS sont utilisés comme *primer* pour débiter la rétrotranscription. Par exemple, le virus du VIH, rétrovirus comparable à un rétrotransposon, utilise l'ARNt^{Lys} pour sa rétrotranscription. Les ARNt ont donc un rôle dans la régulation des éléments transposables et dans les infections virales [76].

ii. Fonction des ARNt aminoacylés

Dégradation de protéines de manière ciblée

Pour maintenir l'homéostasie cellulaire, les protéines endommagées, entre autres, doivent être éliminées. Pour cela, il existe des compartiments cellulaires spécifiques, comme les lysosomes ou la vacuole, dans lesquels les protéines sont dégradées de manière non spécifique.

Dans le noyau et le cytosol c'est le protéasome qui est en charge de la dégradation protéique. Cette protéolyse est spécifique, uniquement les protéines marquées sont dégradées. Le protéasome 26S est un complexe de deux sous-unités qui reconnaît, déstructure puis dégrade les protéines marquées.

Les protéines endommagées portent des signaux spécifiques de dégradation appelés les *degrons*. L'ubiquitylation est un mécanisme permettant l'ajout de *degrons* à des protéines spécifiques. Ce mécanisme consiste en la conjugaison d'une chaîne de molécules d'ubiquitines (Ub) sur les résidus Lysines de la protéine. L'ubiquitylation implique les trois enzymes E1 (*Ub-activating enzyme*), E2 (*Ub-conjugating enzyme*) et E3 (*Ub-protein ligase*). La protéine ubiquitylée est reconnue par le protéasome qui assure sa dégradation.

La stabilité d'une protéine peut aussi être déterminée selon la nature de l'acide aminé présent en N-terminal : c'est la voie *N-end rule*. Certains acides aminés sont classés comme résidus stabilisants, d'autres comme déstabilisants. On les appelle les *N-degrons*. Chez les plantes, les acides aminés considérés comme des *N-degrons* sont Arg, Lys, His, Phe, Tyr et Trp. Ils servent de marqueur pour la protéolyse et sont reconnus par une autre classe de E3 ubiquitine ligases appelées les *N-recognins* [77,78]. Chez *A. thaliana*, les *N-recognins* sont des protéines PRT (PROTEOLYSIS). Parmi elles, l'enzyme PRT1 reconnaît les résidus aromatiques (Phe, Tyr, Trp) alors que PRT6 reconnaît spécifiquement les résidus Arg [79]. La voie *N-end rule* semble aussi exister dans le chloroplaste. La *N-recognin* ClpS1-like, homologue de ClpS chez la bactérie, est retrouvée dans ce compartiment [80].

Chez les plantes, la voie *N-end rule* intervient tout au long du cycle de croissance : germination, développement des parties aériennes, sénescence [81]. Les ARNt aminoacylés interviennent dans cette machinerie comme substrat. En effet, les acides aminés portés par les ARNt peuvent servir de substrat aux *N-degrons* et être transférés sur les protéines. Par exemple chez les eucaryotes, les arginyl-ARNt transférases (ATE) transfèrent le résidu Arg des ARNt^{Arg} à l'extrémité N-ter des protéines [76,82].

Synthèse de composés organiques

Les ARNt peuvent permettre la synthèse de composés organiques. Par exemple, chez les plantes, les ARNt^{Glu} sont impliqués dans la biosynthèse de tetrapyrroles (comme la chlorophylle). Cette biosynthèse se déroule dans le stroma des chloroplastes. Dans un premier temps, l'ARNt^{Glu} sert de substrat en glutamate qui est réduit en glutamate-1-semialdéhyde grâce à la glutamyl-ARNt reductase (GluTR). Cet intermédiaire sert dans un deuxième temps de substrat pour être converti en acide δ-aminolevulinique (ALA). Cette molécule ALA est le précurseur universel de la biosynthèse des tetrapyrroles [83].

Les ARNt sont aussi impliqués dans la synthèse d'hormones végétales comme les cytokinines. Les cytokinines sont des phytohormones impliquées dans les processus de croissance et de développement de nombreuses plantes. Les cytokinines sont dérivées de résidus adénines qui portent en plus sur l'atome N⁶ une chaîne dérivée d'isoprène ou une chaîne aromatique. Dans la voie classique de biosynthèse des cytokinines, l'adénosine 5'-phosphate est utilisé durant la première étape comme substrat (ATP, ADP ou AMP) et est catalysé par l'enzyme phosphate-isopentenyltransferase (IPT). S'en suit de nombreuses étapes qui permettent d'aboutir aux cytokinines-isoprènes. Certains ARNt possèdent des bases adénosines liées à un isoprène. Lorsque ces ARNt sont dégradés, ils libèrent ainsi des cytokinines. C'est le cas pour les ARNt^{Cys}, ARNt^{Leu}, ARNt^{Phe}, ARNt^{Ser}, ARNt^{Trp} et ARNt^{Tyr} qui possèdent en position 37 une adénine prénylée [84,85].

A l'origine de la biogenèse de fragments

De nombreuses études ont montré l'existence de petits ARN correspondants à des fragments d'ARNt : les tRFs (*tRNA derived Fragments*). Les données obtenues ces dix dernières années montrent qu'ils ne sont pas uniquement des produits de dégradation issus de clivages aléatoires des ARNt mais bel et bien des ARN fonctionnels appartenant à la classe des petits ARN non-codants.

La seconde partie de cette introduction sera dédiée aux tRFs, à savoir leur classification, leur biogenèse et leurs fonctions avec un intérêt porté plus particulièrement chez les plantes.

II. Les petits ARN non codants, dérivant d'ARNt

1. Classifications des tRFs

i. Les petits ARN non-codants

Parmi les ARN non-codants, on connaît depuis longtemps les ARNr ou encore les ARNt. Depuis, d'autres ARN non-codants ont été identifiés et séparés en deux groupes : les longs ARN non-codants (long ARNnc) d'une taille supérieure à 200 nt et les petits ARN non-codants (petits ARNnc) d'une taille inférieure à 200 nt. Au départ, les études ont montré que ces ARN non-codants pouvaient réguler l'expression des gènes en agissant sur leur transcription ou leur traduction. Par la suite, les mécanismes qui permettent cette régulation ont été mieux caractérisés.

Chez les plantes, les petits ARNnc sont impliqués dans leur développement, la reproduction, la reprogrammation du génome et l'adaptation à leur environnement. Ces ARNnc se trouvent sous forme double brin et sont générés par des ribonucléases de type Dicer (DCL pour *Dicer-Like*) à partir de précurseurs d'ARN. Ils ont une taille entre 20 et 24 nt. Ils sont pris en charge par le complexe de *silencing* RISC (*RNA Induced Silencing Complex*) constitué de protéines Argonaute (AGO). Ce complexe cible des ARN en particulier par complémentarité de séquence entre le petit ARNnc et l'ARNm cible. Selon la nature de l'ARN cible et le type d'AGO présent dans le complexe, ce processus peut engendrer le clivage et la dégradation de l'ARN cible, réprimer sa traduction ou encore servir d'intermédiaire pour recruter d'autres facteurs [86].

Grace à l'évolution des techniques de séquençage à haut débit et d'analyses bio-informatiques, tant en précision qu'en rapidité, le groupe des ARNnc a été mieux caractérisé et de nouvelles classes d'ARN sont entrés dans ce groupe. On retrouve maintenant dans ce groupe les tRFs. Ces petits ARNnc sont retrouvés chez de nombreux organismes dans tous les règnes du vivant. Aujourd'hui dans la littérature, de multiples fonctions sont attribuées aux tRFs et de plus en plus de mécanismes moléculaires dans lesquels ils sont impliqués sont caractérisés. Il est maintenant évident que les tRFs ne doivent plus être considérés comme des produits de dégradation mais bien comme des ARNnc.

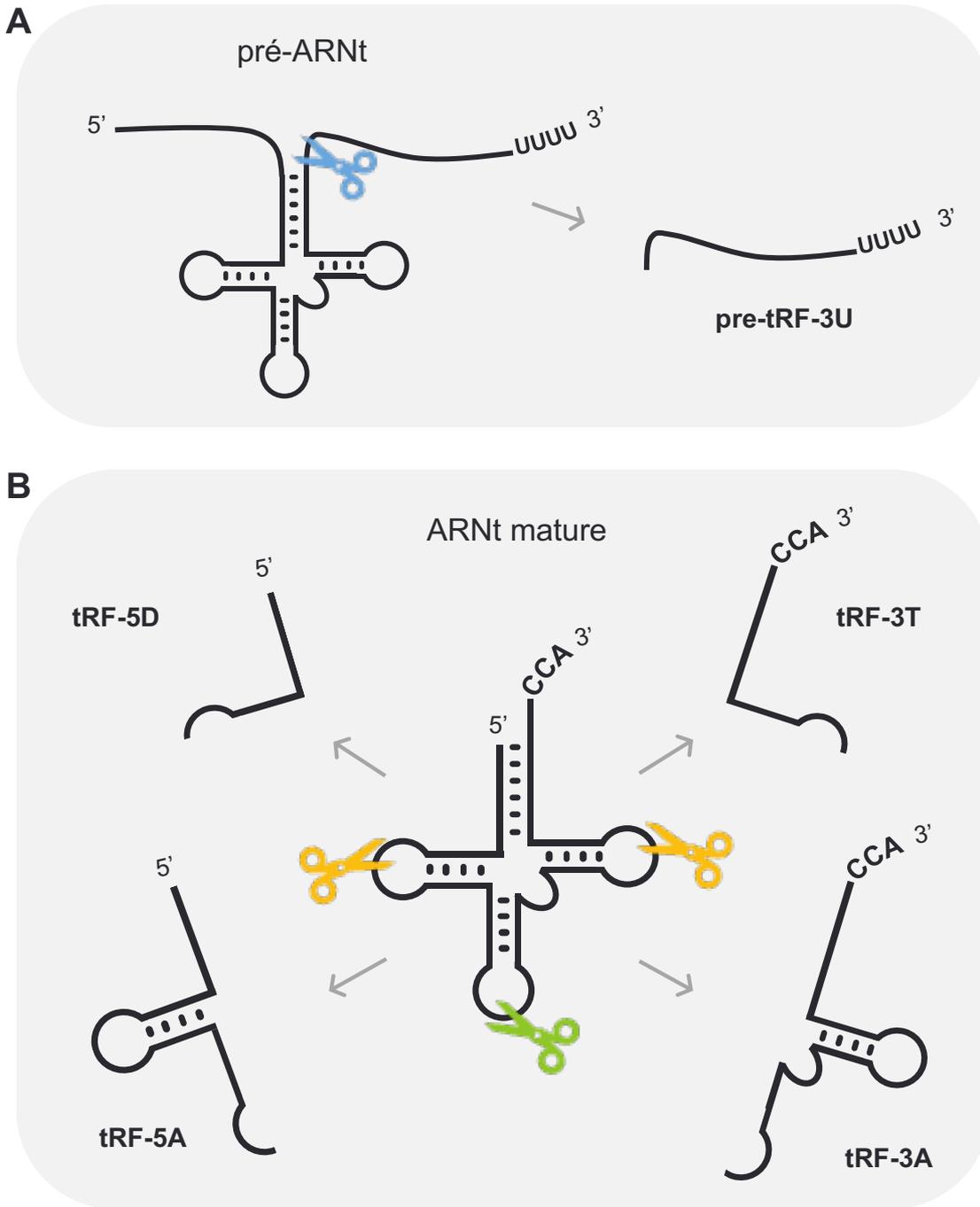


Figure 9 : Schéma récapitulatif des types de tRFs et leur nomenclature. A : pre-tRF-3U générés à partir de clivages de l'extrémité 3'*trailer* de pré-ARNt. B : tRFs générés à partir d'ARNt matures. Les tRF-5D et tRF-3T sont issus de clivages dans la boucle D et T respectivement . Ils sont dits courts et ont une taille entre 15 et 28 nt. Les tRFs-5A et tRF-3A, dits longs, sont issus de clivages dans la boucle de l'anticodon et leur taille varie entre 30 et 40 nt.

ii. Les tRFs : une nouvelle classe de petits ARN non-codants

Les tRFs sont classés en différentes catégories selon la région de l'ARNt dont ils proviennent. Il existe des fragments qui sont issus de clivages endoribonucléolitiques des pré-ARNt et des ARNt matures (Figure 9). Dans la littérature, de nombreuses nomenclatures de tRFs existent. J'ai choisi ici d'utiliser la nomenclature établie par l'équipe dans la revue de Megel et al. [56]. La volonté de la communauté scientifique d'uniformiser la nomenclature des tRFs a abouti récemment à la création d'un outil bio-informatique permettant de nommer les tRFs, tDRnamer (<http://trna.ucsc.edu/tDRnamer/index.html> ; GitHub : UCSC-LoweLab/tDRnamer). Le nom standardisé prend en compte la taille du tRF, la position du fragment par rapport à son ARNt d'origine, la nature de l'ARNt. Un suffixe indique si la séquence du fragment peut correspondre à plusieurs transcrits d'ARNt et le nombre correspondant ainsi que les éventuelles modifications de l'ARNt portées aussi par le tRF. La nomenclature tDRnamer des tRFs utilisés dans mes travaux de thèse est donnée en Annexe I de la partie Matériel & Méthodes.

Lorsque les pré-ARNt sont maturés, les extrémités 5'*leader* et 3'*trailer* sont éliminées. La séquence libérée en 3' est considérée comme un tRF, appelé pre-tRF-3U. Ce fragment a une taille de 15 à 22 nt et contient la séquence poly(U) à son extrémité 3'.

La majorité des tRFs proviennent cependant d'ARNt matures. On distingue alors deux types de tRFs : des fragments longs correspondant à la moitié de l'ARNt et d'autres plus courts correspondant à un quart d'ARNt. Les tRFs longs résultent de clivages dans la région de l'anticodon. Lorsqu'un ARNt est clivé dans la boucle de l'anticodon, il y a alors formation de deux tRFs de 30 à 40 nucléotides, l'un étant la moitié 5' de l'ARNt (tRF-5A) et l'autre étant la moitié 3' de l'ARNt (tRF-3A). Ces longs tRFs sont souvent produits en réponse à des stress, et ce chez de nombreux organismes. Cependant, de plus en plus d'études montrent qu'ils sont aussi retrouvés en absence de stress. Les tRFs courts sont issus des extrémités 5' et 3' des ARNt matures. Ce sont des fragments issus de clivages dans les régions D ou T et sont nommés tRF-5D et tRF-3T respectivement. Ils ont une taille variant de 15 à 28 nt. Les tRF-3T possèdent à leur extrémité 3' le triplet CCA.

La biogenèse de ces tRFs peut aussi entraîner la formation de fragments dits tRFs internes. D'une manière générale, la population de tRFs dans la cellule, tous organismes confondus, ne reflète pas l'abondance des ARNt correspondants. Il est aussi à noter que des fragments plus grands, de 40 à 60 nt, ne sont pas pris en compte dans les banques de petits ARNnc à cause de leur taille. Mais ils ne sont pas à exclure du groupe des tRFs pour autant. Les ARNt étant codés par le génome nucléaire et les génomes d'organites, il est aussi possible de trouver des tRFs provenant d'ARNt d'organites.

iii. Les tRFs chez les plantes

Chez *A. thaliana*, des tRFs issus d'ARNt d'origine nucléaire, chloroplastiques et mitochondriaux ont été retrouvés. L'étude de la population des petits tRFs (< 26 nt) de feuilles d'*A. thaliana* réalisée par [87] montre que 1,9 % des ARNnc sont des tRFs. Les fragments issus d'ARNt matures sont majoritairement représentés puisque seulement 0,2 % sont des pre-tRF-3U issus des pré-ARNt. Les trois quarts des tRFs identifiés proviennent d'ARNt nucléaires (ntRF), environ un quart sont des tRFs d'origine chloroplastique (ctRF) et moins de 1% sont issus d'ARNt mitochondriaux (mtRF).

Il faut savoir qu'il est très difficile d'identifier les tRFs ayant pour origine des ARNt mitochondriaux. En effet, étant donné qu'une partie des ARNt retrouvés dans les mitochondries sont codés par des ADNt *chloroplastic-like* et que d'autres sont importés du noyau, il est difficile de distinguer les mtRF des ctRF et ntRF. Seuls les tRFs issus d'ARNt dits natifs des mitochondries, c'est-à-dire codés par le génome mitochondrial uniquement, peuvent être identifiés de manière certaine. Mais la quantité de ces ARNt, et par conséquent la quantité de ces tRFs générés, reste très minoritaire et leur détection dans les banques de petits ARNnc reposent sur peu de données (peu de *reads*) [87].

Dans l'étude de Cognat et al. [87] la majorité des tRFs identifiés chez *A. thaliana* sont des tRF-5D et tRF-3T, avec une plus grande proportion pour les tRF-5D (environ 55 %) que pour les tRF-3T (environ 30%). Les tRFs d'un ARNt donné ne sont pas forcément représentés par les deux types de tRF (tRF-5D et -3T) en quantité égale. C'est le cas pour l'ARNt^{Ala} (AGC) nucléaire dont le tRF-5D est fortement enrichi alors que le tRF-3T correspondant n'est que très peu présent. Le phénomène inverse existe également, notamment pour l'ARNt^{Asn} (GTT) chloroplastique dont le tRF-3T est fortement enrichi mais pas le tRF-5D. L'ensemble des tRFs identifiés ne dérive principalement que d'une partie des ARNt [87,88]. Seulement 12 des 46 ARNt isoaccepteurs nucléaires et 10 des 30 ARNt isoaccepteurs chloroplastiques sont majoritairement à l'origine de tRF-5D [87].

Parmi les petits tRFs identifiés, ceux de 19 et 20 nucléotides sont les plus abondants. Les tRFs-5D résultent plus majoritairement de clivages dans la boucle D et d'une manière plus modérée dans le bras D. C'est l'inverse pour les tRF-3T où les clivages sont plus abondants dans le bras T que dans la boucle T. Zahra et al. [88] ont analysé des banques issues de 6 plantes angiospermes (*Arabidopsis thaliana*, *Solanum lycopersicum*, *Cicer arietinum*, *Medicago truncatula*, *Oryza sativa* et *Zea mays*) et ont obtenus des résultats similaires. Ils ont même montré que les points de clivages étaient conservés chez ces plantes. Parmi les tRF-3T identifiés, 92% possédaient le triplet CCA en 3'. Ils ont aussi observé que les longs tRFs (tRF-5A et -3A) ne représentent que 10% des tRFs avec aussi cette tendance de tRF-5A plus abondants que les tRF-3A.

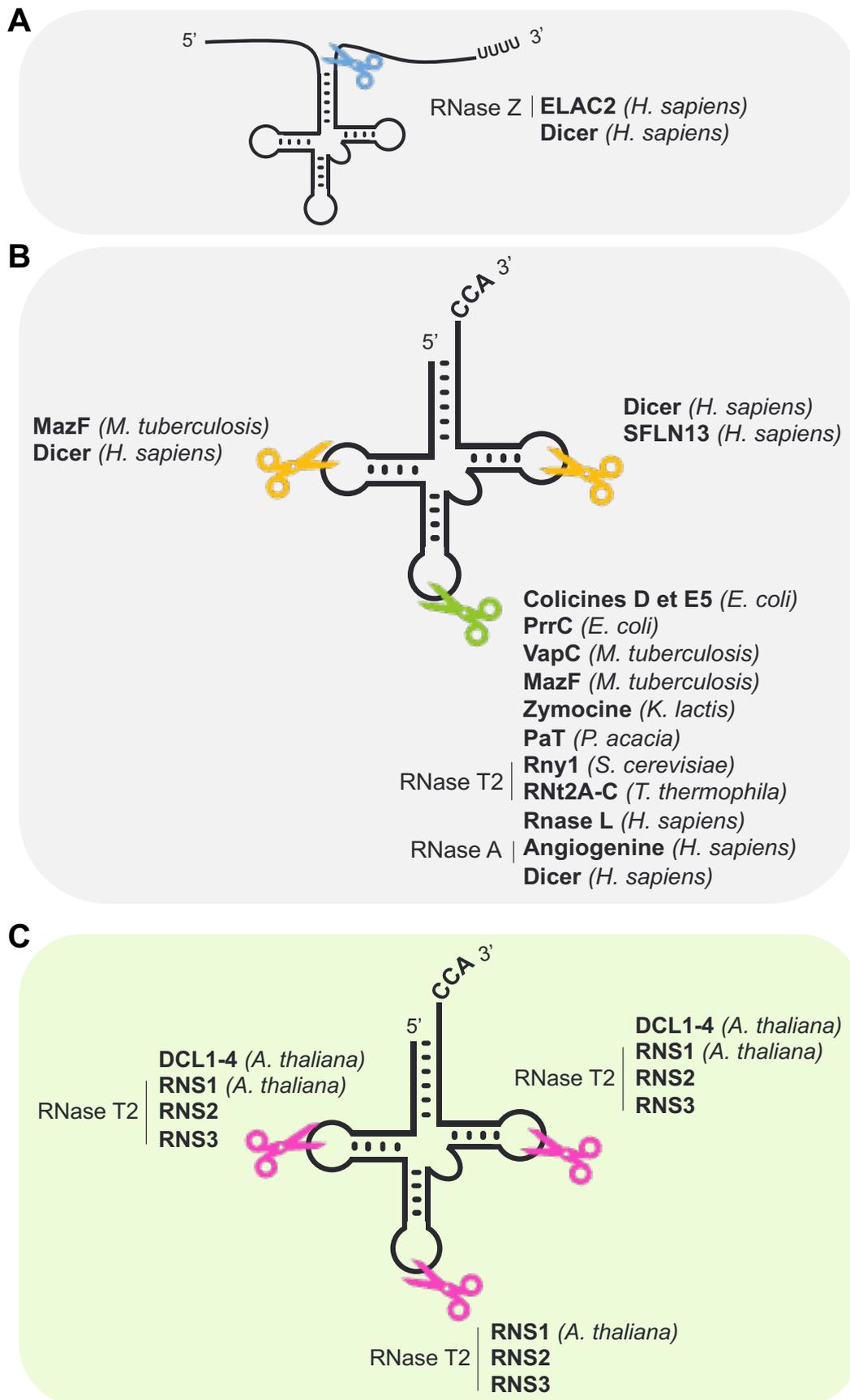


Figure 10 : Schéma récapitulatif des endoribonucléases générant des tRFs. A : Liste d'endoribonucléases générant des tRFs à partir de pré-ARNt. B : Listes d'endoribonucléases générant des tRFs à partir d'ARNt matures. C : Endoribonucléases générant des tRFs à partir d'ARNt mature, chez les plantes.

Au sein de tissus différents (plantules, feuilles et fleurs), la population en tRFs reste relativement similaire. En revanche, la population en tRFs dans les racines présente des différences notables. Ce n'est pas uniquement la population en ctRF qui change (dû à l'absence de chloroplastes dans les racines comparés aux tissus photosynthétiques), mais aussi la population en ntRF [87,88]. L'étude de Cognat et al. [87] montre que globalement, peu de tRF-3T sont retrouvés dans les racines. Concernant les tRF-5D, quatre d'entre eux sont majoritaires et représentent plus de 75% de la population de tRF-5D dans les racines : Arg (ACG), Arg (TCG), Cys (GCA) et Gly (TCC).

En plus d'une disparité de population selon les tissus, on observe aussi des variations des populations de tRFs chez les plantes soumises à des stress. Certains tRFs deviennent plus abondants en cas de stress, d'autres deviennent plus rares. Des tRFs différents sont retrouvés régulés lors de stress chez les différentes plantes étudiées [88]. Par exemple, chez *A. thaliana*, les études de banques de petits ARNnc ont montré que le ntRF-5D Gly (GCC) devient plus abondant lors de différents stress (UV, sécheresse ou froid) [87]. Alors que chez le blé (*Triticum aestivum*), les tRF Glu (CTC), tRF Lys (CTT) et tRF Thr (CGU) sont fortement exprimés lors d'une infection fongique à la fusariose (*Fusarium graminearum*) [89].

2. Biogenèse des tRFs

La biogenèse des tRFs est dépendante d'endoribonucléases capables de cliver les ARNt à des sites particuliers. Certaines sont capables de cliver de l'ARN simple brin (notamment au niveau des boucles), d'autres clivent de l'ARN double brin (au niveau des bras). Comme les tRFs sont retrouvés dans tous les organismes, allant des bactéries aux eucaryotes, les endoribonucléases associées à ces clivages ont été caractérisées chez différents organismes modèles. La Figure 10 récapitule les détails des endonucléases et leur région de clivage donnés ci-après.

i. Chez les procaryotes

Ce sont d'ailleurs en premier chez les procaryotes que des endoribonucléases capables de cliver les ARNt pour générer des tRFs ont été identifiées. La Colicine D et la Colicine E5 chez la bactérie *E. coli* sont capables de cliver certains ARNt particuliers au niveau de la boucle de l'anticodon (ARNt^{Tyr}, ARNt^{His}, ARNt^{Asn}, ARNt^{Asp} pour la Colicine D et uniquement l'ARNt^{Arg} pour la Colicine E5) [90,91]. Une autre endonucléase faisant partie du complexe PrrC, identifiée chez *E. coli*, clive le pré-ARNt^{Lys} au niveau de l'anticodon lors d'une infection par le phage T4 [92]. Chez la bactérie *M. tuberculosis*, plusieurs endonucléases VapC clivent des ARNt spécifiques mais toutes ciblent la boucle de l'anticodon [93]. La toxine MazF est aussi identifiée chez cette bactérie comme étant capable de cliver l'ARNt^{Lys} à la position 35 dans l'anticodon mais aussi l'ARNt^{Pro} à la position 17 dans la boucle D, générant un petit tRF-5D [94].

ii. Chez les eucaryotes

Chez la levure *K. lactis*, la toxine zymocine est capable de cliver l'ARNt^{Glu}, l'ARNt^{Lys} et l'ARNt^{Gln} au niveau de la boucle de l'anticodon. Elle reconnaît uniquement ces ARNt lorsqu'ils possèdent la modification mcm⁵s²U en position 34. La zymocine n'a pas d'activité sur ces ARNt s'ils ne sont pas modifiés [95,96]. Chez l'insecte *P. acacia*, PaT est capable de cliver uniquement l'ARNt^{Gln} dans la boucle de l'anticodon [97].

Chez plusieurs organismes, on retrouve des RNases T2 qui sont à l'origine de la biogenèse des tRFs. Ainsi, chez la levure *S. cerevisiae* et chez le protozoaire *T. thermophila*, Rny1 et les Rnt2A-C, respectivement, sont capables de cliver les ARNt dans la région de l'anticodon [98,99].

Chez l'humain, on retrouve plusieurs types d'endonucléases capables de cliver les ARNt. La RNase Z, ELAC2, clive le pré-ARNt^{Ser} à l'extrémité 3' [100]. La RNase L peut cliver les ARNt^{His}, ARNt^{Pro} et ARNt^{Gln} dans la boucle de l'anticodon [101]. Aussi responsable de clivages dans l'anticodon, l'Angiogenine (ANG) qui est une endonucléase de la famille des RNases A (alors que chez l'animal *R. pipiens*, les RNases A clivent dans le bras D ou la boucle variable [102]) [103,104]. Concernant les petits tRFs, l'endonucléase Dicer est capable de générer ces petits fragments. Il a été montré que Dicer peut cliver l'extrémité 3' du pré-ARNt^{le} [105]. Elle peut aussi cliver dans la boucle D des ARNt^{Gln} et dans la boucle T des ARNt^{Glu} pour former des tRF-5D et -3T [106–108]. Plus récemment, il a aussi été observé que Dicer pouvait être à l'origine de fragments plus grands, probablement des tRF-5A et -3A ou des tRFs internes à partir des ARNt^{Gly}, ARNt^{Glu}, ARNt^{His}, ARNt^{Lys}, ARNt^{Asp} et ARNt^{Asn} [107]. Des protéines de la famille des Schlafen sont aussi capable de cliver les ARNt à l'extrémité 3' ou dans le bras T des ARNt^{Ser} et ARNt^{His} pour SFLN13 [109]. Pour SFLN11, l'activité de clivage des ARNt^{Ser} et ARNt^{Thr} a été montrée mais leurs sites de clivage restent encore à être définis [110].

iii. Chez les plantes

Chez la plante *A. thaliana*, Martinez et al. [111] proposa que des endonucléases *Dicer-like* (DCL) soient responsables de la biogenèse des tRFs. Contrastant avec ce résultats, Alves et al. [112] montra que ce n'était pas les DCL qui génèrent des tRFs mais plutôt une RNase T2. Les résultats de Drouard et al. apportent des informations similaires montrant d'une part que les quatre DCL d'*A. thaliana* ne sont pas responsables de la biogenèse de tRFs et d'autre part que les RNases T2 (RNS1-5), homologues de Rny1 de la levure, sont les acteurs principaux pour la production de tRFs [113]. Le génome nucléaire d'*A. thaliana* code pour 5 RNases T2 : RNS1, RNS2, RNS3, RNS4 et RNS5 (*S-like Ribonuclease*). Il a été montré que les RNS1, 2 et 3 clivent les ARNt dans la boucle de l'anticodon et aussi dans les boucles D et T. Elles sont donc capables de générer autant des petits que des grands tRFs [113].

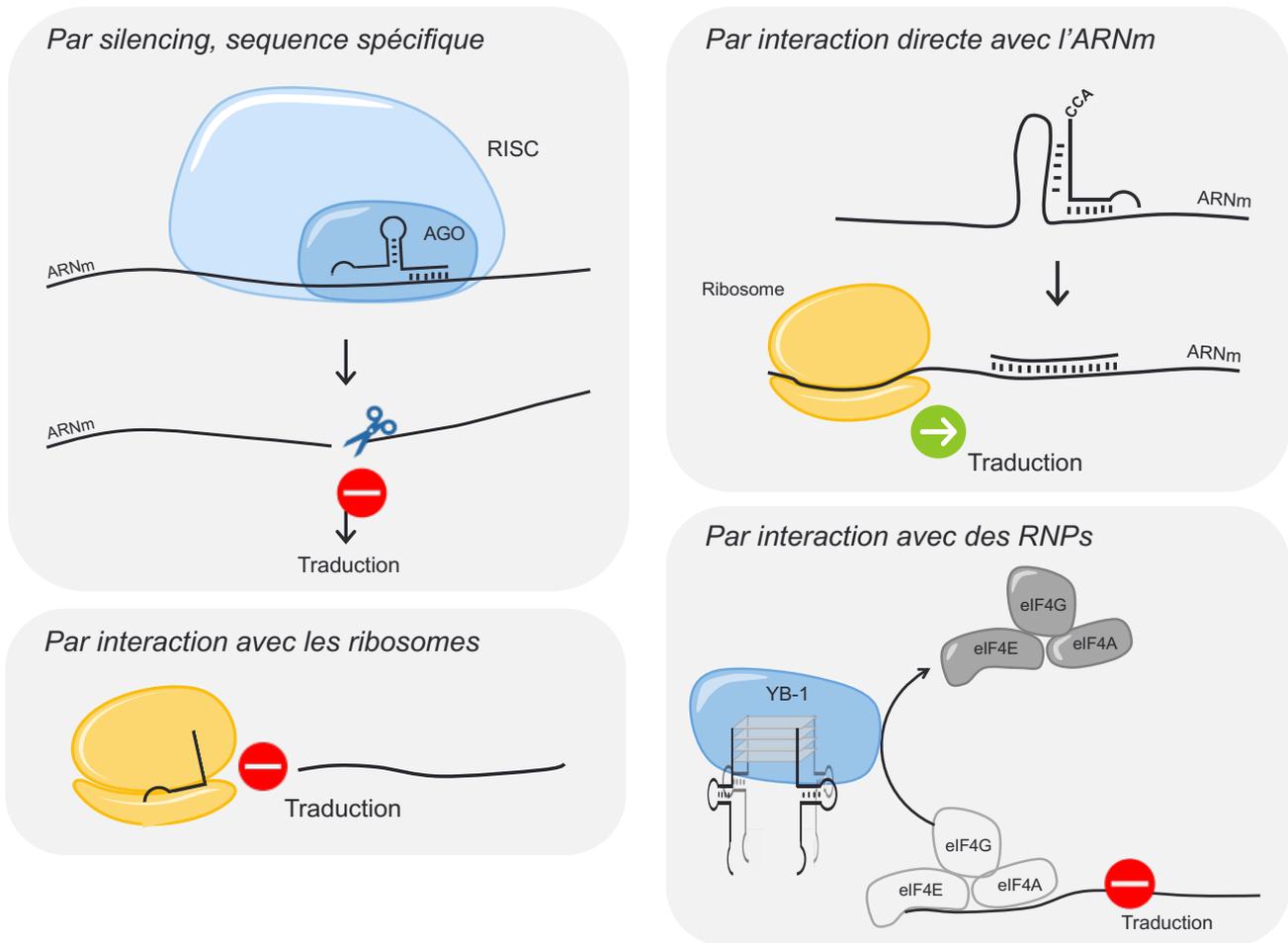
Faisant partie d'un axe d'étude de ma thèse, une description plus complète des RNS chez *A. thaliana*, est donnée dans le Chapitre I de la partie Résultats.

iv. Impact des modifications post-transcriptionnelles sur la biogenèse des tRFs

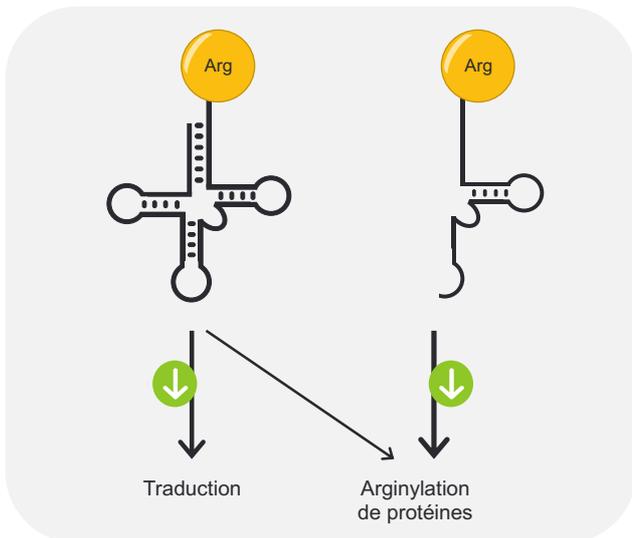
Les modifications des ARNt jouent un rôle important dans la biogenèse des tRFs. En effet, la présence de certaines modifications spécifiques peut empêcher le clivage de l'ARNt et donc la formation de tRFs. C'est le cas par exemple chez la drosophile, en cas de stress thermique, la présence d'une cytosine méthylée en position 38 (m^5C) des ARNt^{Asp}, ARNt^{Gly} et ARNt^{Val} limite la production de tRF-5A [114]. Au contraire, parfois c'est l'absence de modifications qui va améliorer les capacités de fixation et de clivage des endonucléases et donc engendrer une plus forte production de tRFs. Chez l'humain, la diminution du taux de méthylation des cytidines 48 et 49 (m^5C) augmente la production de tRF-5A. Ceci est dû à une meilleure affinité entre l'Angiogenine et l'ARNt non modifié [115].

Chez *A. thaliana*, certaines modifications sont retrouvées très abondantes dans un type d'ARNt donné, comme par exemple la présence de m^1G dans la région D de l'ARNt^{Gly} (GCC) ou de pseudouridine et dihydrouridine dans la région de l'anticodon de l'ARNt^{Glu} (TTC). Globalement, il a été montré que la majorité des modifications d'ARNt se trouvaient dans deux zones, entre les bases 20 et 40 et entre les bases 50 et 60 [88]. Ces positions correspondent aux régions de clivages notamment pour former des tRF-5D, les tRF-5A et -3A et les tRF-3T. Plus généralement, des modifications identiques ont été retrouvées sur des ARNt de six plantes angiospermes différentes (*A. thaliana*, *S. lycopersicum*, *C. arietinum*, *M. truncatula*, *O. sativa* et *Z. mays*) [88]. Certaines modifications sont quant à elles présentent uniquement en cas de stress. C'est le cas notamment de la modification 2'-O-methyladenosine (Am) qui est induite lors de stress salin chez *A. thaliana*, *B. distachyon*, *P. tremula* et *O. sativa* [116]. Pour le moment, aucun lien n'a été prouvé entre les modifications des ARNt et la biogenèse des tRFs chez les plantes. Mais l'abondance de modifications au niveau de sites de clivage des ARNt laisse supposer qu'elles peuvent avoir un impact sur la biogenèse des tRFs.

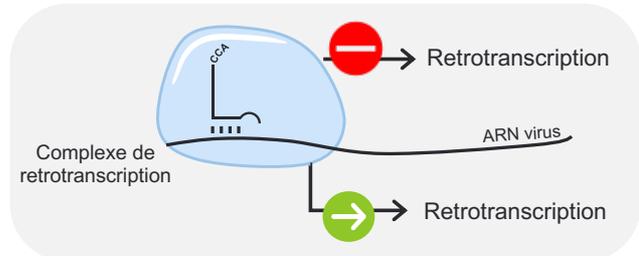
A Régulation de la traduction



B Régulation post-traductionnelle



C Régulation transcriptionnelle



D Régulation expression génique

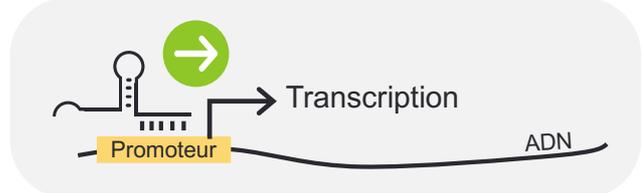


Figure 11 : Récapitulatif des fonctions des tRFs, tous organismes confondus. A : Exemples de fonctions des tRFs dans la régulation de la traduction. B : Exemple d'une fonction de la régulation post-traductionnelle. C : Exemple de fonction de régulations transcriptionnelles. D : Exemple de fonction de régulation de l'expression des gènes.

3. Fonctions des tRFs

Il devient de plus en plus évident que les tRFs peuvent avoir différents mécanismes d'action et fonctions. Cependant les recherches restent difficiles et lentes puisque chaque tRF issu d'un ARNt spécifique peut faire l'objet d'une fonction particulière dans un organisme donné. Il peut donc y avoir un très grand nombre de fonctions à déterminer. Les paragraphes suivants reprennent les grands types de fonction que peuvent avoir les tRFs (tout organisme confondu) et sont résumés dans la Figure 11. Un paragraphe sera dédié aux fonctions connues des tRFs chez les plantes.

i. Les tRFs dans la régulation de la traduction

Inhibition de la traduction par silencing, de manière séquence spécifique

De nombreuses études ont montré que les tRFs sont retrouvés associés aux protéines Argonautes (AGO) et suggèrent leur implication dans le mécanisme de *silencing*. Le mécanisme d'ARN *silencing*, ou ARN interférence, aboutit à une régulation de l'expression des gènes par l'utilisation de petits ARNnc.

Plusieurs voies de *silencing* existent et sont à la fois distinctes et chevauchantes. Pour toutes ces voies, il y a nécessité de produire et/ou maturer l'ARNnc pour former des ARN double brins d'une taille entre 20 et 30 nucléotides. L'un des deux brin, appelé l'ARN guide, est pris en charge par une protéine de la famille des Argonautes (AGO) au sein du complexe RISC (*RNA-induced silencing complex*). Par complémentarité de séquence, le complexe RISC va s'hybrider à l'ARN cible et engendrer soit son clivage et sa dégradation, soit une répression de sa traduction.

Chez les eucaryotes, les deux principaux ARNnc impliqués dans le *silencing* et largement décrits dans la littérature sont les petits ARN interférents (siARN) et les microARN (miARN) (pour une revue, voir [117]). La biogenèse de ces deux types d'ARN est différente. Chez les mammifères, les siARN sont des fragments d'ARN double brins, maturés par les RNases de la famille Dicer. Les miARN sont eux transcrits par l'ARN polymérase II (ARN Pol II) à partir de gènes, en miARN primaires (pri-miARN) au sein du noyau. Ils sont modifiés par Drosha et ses partenaires pour devenir des précurseurs pré-miARN. Ils ont alors une structure tige-boucle. Après export vers le cytosol, ce sont les protéines Dicer qui finissent la maturation en clivant la boucle pour former un ARN double brin. Chez les mammifères, une seule protéine Dicer prend en charge à la fois les siARN et les miARN [118] alors que certaines espèces de drosophiles ont par exemple deux protéines Dicer spécialisées : DCR-1 est spécialisée dans la production de miARN alors que DCR-2 prend en charge les siARN [119].

Les plantes possèdent quant à elle 4 protéines Dicer-like (DCL). La protéine DCL1 joue à la fois le rôle de ses homologues Drosha et celui de Dicer dans la voie de biogenèse des miARN (pour une revue, voir [86]).

Une fois pris en charge par le complexe RISC, l'ARNnc cible l'ARN à réguler par complémentarité de base. Dans le cas des miARN, lorsque le miARN et l'ARN cible présentent une complémentarité imparfaite, ceci conduit à une répression de sa traduction. En revanche, lorsque la complémentarité est parfaite, le complexe RISC induit dans ce cas un clivage de l'ARN cible. Chez les mammifères et la drosophile, les miARN ont le plus souvent une complémentarité imparfaite [117,120]. Dans peu de cas ils présentent une complémentarité parfaite [117,121]. Chez les plantes, cette observation est à l'opposé. Le plus souvent la complémentarité entre le miARN et la cible est quasiment parfaite et entraîne principalement le clivage de l'ARN cible [122–125]. Il est plus rarement observé, lors d'appariement imparfait, une inhibition de la traduction [117,126]. Les siARN s'apparentent aux miARN et permettent le clivage de l'ARN cible par complémentarité quasiment parfaite.

Ainsi, les petits ARNnc sont capables de réguler l'expression des gènes de manière ciblée et post-transcriptionnelle. Ce mécanisme est appelé le *Post-Transcriptional Gene Silencing* (PTGS).

Les études fonctionnelles des tRFs montrent qu'ils pourraient agir comme des miARN ou des siARN et sont pris en charge par les protéines AGO. Le plus souvent, il est observé une diminution de l'ARNm cible, complémentaire au tRF (Figure 11 A) [108,127–129]. Chez les procaryotes, les protéines AGO utilisent non pas des miARN comme guide mais des petits ADN et ciblent non pas des ARNm mais des gènes. Cependant, une étude montre que des petits ARN sont retrouvés associés aux AGO de bactéries, ce qui pourrait laisser envisager une association tRF-AGO [130]. Toutefois, les données chez les procaryotes à ce sujet restent encore très parcellaires. Chez *T. thermophila*, les tRFs sont retrouvés associés aux protéines Piwi (protéines de la famille des AGO). Notamment, un tR-3T est retrouvé dans des complexes dont Twil2 fait partie. Ensemble, ils ciblent les ARNm dont la séquence est complémentaire au tRF. Le complexe induit ensuite le recrutement de l'exonucléase Xrn1 qui va dégrader l'ARNm cible [131]. Chez l'humain aussi, des tRF-5A sont retrouvés en interaction avec des protéines Piwi (notamment Hiwi2) ce qui conduirait à une régulation de la traduction mais le mécanisme n'a pas encore été caractérisé [132].

Régulation de la traduction par interaction directe avec l'ARNm

Dans d'autres cas, des tRFs vont cibler directement des ARNm spécifiques (Figure 11 A). C'est le cas chez l'humain où le tRF-3T issu de l'ARNt^{Leu} s'hybride spécifiquement à un ARNm codant une protéine ribosomale (RPS28, RPS15). Par cette interaction, le tRF change la structure de l'ARNm qui permet sa traductibilité. Dans ce cas, le tRF permet d'augmenter la traduction de protéines essentielles. En absence de tRF, les ARNm codant RPS28 sont moins traduits ce qui va impacter la biogenèse des ribosomes [133].

Régulation de la traduction par interaction avec les ribosomes

Plusieurs études ont montré que les tRFs interagissaient avec les ribosomes (Figure 11 A). En effet, chez l'archée *H. volcanii*, un tRF-5D issu de l'ARNt^{Val} est capable de s'associer à la petite sous-unité 30S du ribosome au niveau du site d'entrée de l'ARNm. Ceci ayant pour conséquence d'empêcher l'ARNm d'entrer dans le ribosome, sa traduction n'a donc pas lieu [134]. Dans des cellules souches de souris, il a été montré que des tRF-5A sont enrichis dans des fractions monosomales et polysomales. Cet enrichissement présente un profil différent selon les cellules étudiées (cellules indifférenciées ou différenciées). Les auteurs suggèrent que les tRFs peuvent interagir avec certains ribosomes (constitués de protéines ribosomales particulières ou portant certaines modifications) et/ou à différents sites du ribosome [135]. Il existe aussi des exemples d'interactions tRF – ribosome mais qui conduisent à une augmentation globale de la traduction. C'est le cas chez le protozoaire *T. brucei* où un tRF-3A issu de l'ARNt^{Thr} est induit en cas de stress. Il est ensuite incorporé dans les ribosomes et cette association permet une meilleure entrée de l'ARNm au sein des ribosomes. La traduction est ainsi améliorée [136].

Régulation de la traduction par interaction avec des ribonucléoprotéines

Les tRFs sont aussi capables de réguler la traduction en interagissant avec des ribonucléoprotéines (RNP) (Figure 11 A). Par exemple, il a été montré dans des cellules humaines que des tRF-5A issus de clivages par ANG des ARNt^{Ala} et ARNt^{Cys} peuvent former des structures G-quadruplex grâce à cinq résidus G qu'ils possèdent en 5'. Ces G-quadruplex sont reconnus par la protéine YB-1 (*Y-box binding protein*), facteur connu pour pouvoir dissocier le facteur d'initiation de la traduction (eIF4F) de l'ARNm. Les résultats montrent que ces tRF-5A sont capables d'interagir avec YB-1 ce qui résulte en la dissociation de eIF4F et de l'ARNm. L'ensemble YB-1 – tRF est ensuite stocké avec l'ensemble du complexe d'initiation de la traduction dans des granules de stress. Ce mécanisme engendre une diminution de la traduction, effet protecteur pour les cellules en cas de stress [137–139].

Les modifications des ARNt lors de leur maturation semblent aussi être un facteur de la régulation de la traduction par interaction avec des RNPs. Par exemple, des tRFs issus de l'extrémité 5' de certains ARNt (notamment ceux possédant des résidu G en 5') portent la modification Ψ (position 8) catalysée par l'enzyme PSU7. L'absence de PSU7 dans des lignées KO peut être reliée à une réduction du niveau de tRF-5D et une augmentation globale de la traduction [140]. Les auteurs présentent un mécanisme où la modification Ψ portée par le tRF permet de mieux séquestrer PABPC, protéine se liant à la séquence poly(A) des ARNm pour initier la traduction. La présence de cette modification conduit à une inhibition de la traduction. Cette étude montre l'importance des modifications des ARNt pour la biogenèse des tRFs et l'impact final que les tRFs peuvent avoir sur la traduction.

ii. Les tRFs dans la régulation post-traductionnelle

Les tRFs ont aussi été retrouvés impliqués dans des mécanismes de régulation post-traductionnels (Figure 11 B). En effet, une modification post-traductionnelle courante est l'arginylation des protéines (ajout d'un acide aminé Arg en N-terminal). Cette modification est, entre autres, utilisée dans la voie *N-end rule* pour la dégradation des protéines (voir I.7.ii). L'arginylation est catalysée par l'enzyme ATE1. Cette enzyme utilise l'arginine chargée sur les ARNt comme substrat pour la transférer sur les protéines. Ce procédé engendre une diminution de la quantité d'ARNt^{Arg} disponible pour la traduction. Une étude récente sur des cellules de souris a montré que les tRF-3A issus des ARNt^{Arg}, et portant eux aussi l'acide aminé Arg, sont aussi source de substrat pour ATE1 [141]. L'utilisation de ces tRFs comme substrat modifie la quantité des ARNt^{Arg} disponibles pour la traduction. La balance entre la traduction et l'arginylation peut donc être modifiée par des tRFs. Aucun lien n'est démontré mais il est suggéré que les tRFs permettent indirectement de réguler la voie *N-end rule* et la protéolyse.

iii. Les tRFs dans la régulation de la transcription

Les tRFs peuvent aussi interagir avec des éléments mobiles comme les virus à ARN ou les retrotransposons (Figure 11 C). Il a été montré qu'en condition normale, l'extrémité 3' des ARNt matures est utilisée par ces éléments comme amorce pour leur retrotranscription. L'équipe de Schorn et al. [142] a montré dans des cellules souches de souris que la transfection de tRF-3T mimant l'extrémité 3' des ARNt matures, crée une compétition et empêche la rétrotranscription de l'ARN viral. Ceci a été aussi observé avec des cellules infectées par le VIH où des tRFs possèdent une complémentarité de séquence avec le site d'initiation (PBS) du VIH [143]. Au contraire, certains virus (comme le *Human T cell leukemia virus type 1*, HTLV-1) utilisent le tRF-3T issu de l'ARNt^{Pro} pour mimer l'ARNt mature et l'utilisent pour initier leur retrotranscription et se répliquer [144].

iv. Les tRFs dans la régulation de l'expression de gènes

Les tRFs pourraient aussi avoir un rôle dans l'hérédité (Figure 11 D). En effet, Chen et al. [145] ont montré chez la souris que des tRF-5A retrouvés dans le sperme avaient une complémentarité de séquence avec certaines régions promotrices. Il semble qu'uniquement les tRFs portant des modifications (méthylations) soient actifs. Cette population de tRFs dans le sperme est transmise à la descendance. Ceci suggère que les tRFs pourraient avoir un rôle dans la régulation de l'expression de certains gènes. Une autre étude a confirmé ce phénomène. Ils ont observé que des souris mâles soumises à un régime pauvre en protéines transmettaient par leur sperme des populations de tRFs différentes à l'embryon. Ces souris accumulent notamment dans leur sperme le tRF-5A issu de l'ARNt^{Gly}. Ce tRF inhibe l'expression de gènes exprimés sous le contrôle du retrotransposon MERVL ce qui conduit à une croissance anormale de la descendance [146].

A Régulation de la traduction

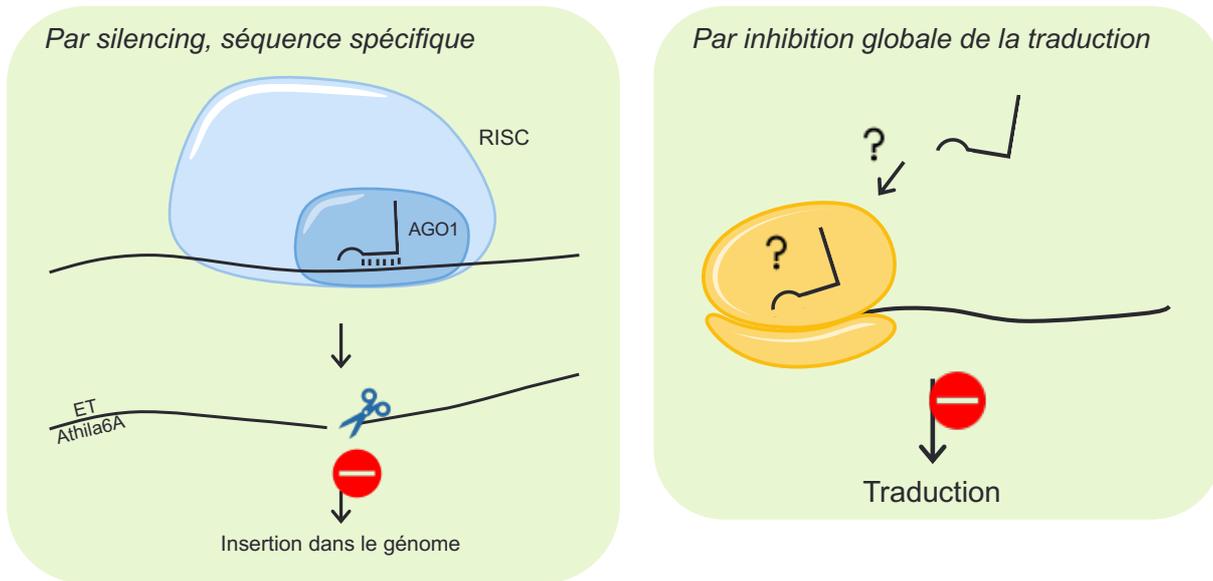


Figure 12 A : Schéma récapitulatif des fonctions des tRFs chez les plantes. Exemples de fonctions des tRFs dans la régulation de la traduction. Inhibition par *silencing* (à gauche) et par association/régulation des ribosomes (à droite).

4. Fonctions des tRFs chez les plantes

Chez les plantes, les données sur les fonctions des tRFs sont encore très parcellaires. Cependant, avec l'évolution des technologies et la ferveur grandissante pour les tRFs, de plus en plus de mécanismes dans lesquels les tRFs sont impliqués sont caractérisés et des fonctions commencent à être déterminées. Les paragraphes ci-dessous détaillent les données parues récemment sur les tRFs chez les plantes. La Figure 12 récapitule uniquement les fonctions chez les plantes.

i. Inhibition de la traduction par silencing

Chez les plantes, les tRFs sont aussi retrouvés associés aux protéines AGO et participent au *silencing* des ARNm [87,111,112,147–150].

L'étude réalisée par Martinez et al. [111] met en évidence une accumulation de tRF-5D dans les grains de pollen d'*A. thaliana*. Les tRFs identifiés font, pour la majorité, 19 nt. Cette accumulation de tRFs est aussi retrouvée dans les organes reproductifs du riz *O. sativa*, du maïs *Z. mays* et de la mousse *P. patens*. Cette équipe montre que les tRFs de 19 nt sont retrouvés enrichis avec AGO1 et émettent l'hypothèse qu'ils pourraient jouer un rôle dans le *silencing* de gènes. L'analyse bio-informatique pour rechercher des ARNm cibles ayant une complémentarité de séquence avec les tRFs montre qu'ils pourraient cibler en grande majorité des ARNm codant des protéines mais aussi pour une part moins importante des éléments transposables. Finalement, les auteurs valident ces prédictions avec le tRF-5D dérivant de l'ARNt^{Met} (CAT) qui est capable de cibler le transcrite de l'élément de la famille Gypsy, Athila6A et de provoquer son clivage. Ici c'est donc un exemple de régulation par *silencing* que peut effectuer un tRF pour contrôler la présence d'éléments transposables (Figure 12 A).

D'autres études, plus fonctionnelles, basées sur le mécanisme de *silencing* sont détaillées dans le paragraphe « Les tRFs comme facteurs d'adaptation aux stress » (II.4.iii).

ii. Inhibition de la traduction indépendamment du mécanisme de silencing

Mon équipe d'accueil a pu montrer que certains tRFs étaient capable d'inhiber la traduction *in vitro* : le tRF-5D issu de l'ARNt^{Ala} (AGC) et le tRF-5D issu de l'ARNt^{Asn} (GTT) [151]. En revanche, les tRFs retrouvés associés à AGO1 dans les précédentes études [87] n'ont aucun impact sur la traduction, suggérant un mécanisme indépendant de celui du *silencing*. D'autant plus qu'il n'y a aucune complémentarité de séquence de ces tRFs avec le transcrite à traduire. Dans cette étude, il a été montré que le tRF-5D Ala de 20 nt est celui qui inhibe le plus la traduction, comparé aux fragments de 18 et 33 nt de ce même ARNt^{Ala}. Le résidu A en position 16 et les résidus G en position 18 et 19 semblent être essentiels pour avoir un effet d'inhibition de la traduction. Finalement, il a été montré que le tRF Ala est capable de s'associer aux polyribosomes actifs mais dans la même mesure que d'autres tRFs, n'expliquant pas un lien entre sa capacité à inhiber la traduction et son interaction avec les ribosomes.

B Facteurs d'adaptation aux stress biotiques

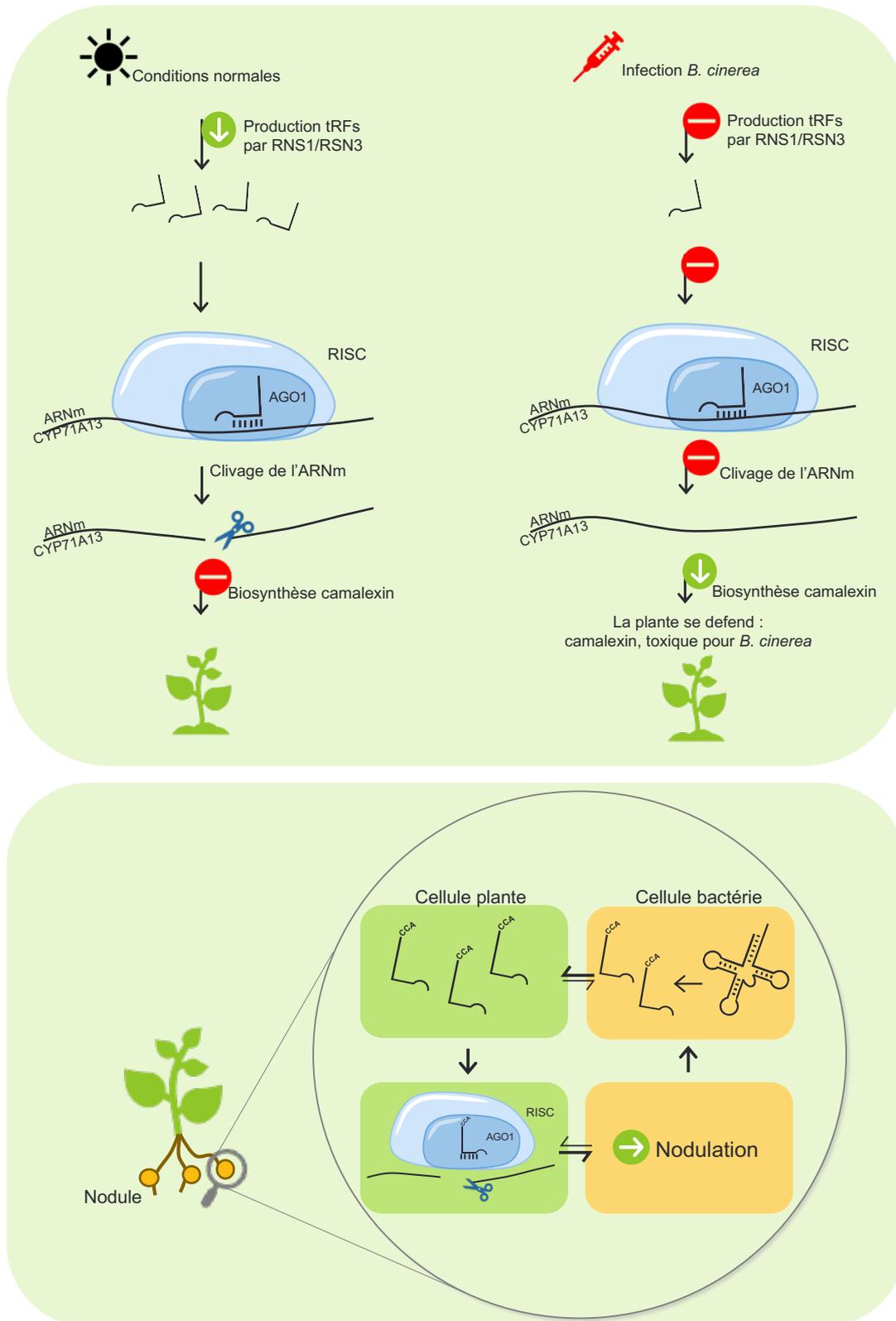


Figure 12 B : Schéma récapitulatif des fonctions des tRFs chez les plantes. Exemples de fonctions des tRFs dans l'adaptation à des stress biotiques. Le premier exemple schématise la fonction de tRFs lors d'infection bactérienne par *B. cinerea*. Le second exemple montre l'implication des tRFs dans la nodulation lors de symbiose avec des rhizobiums.

Il est alors suggéré que le tRF Ala soit impliqué dans un mécanisme de modulation de la traduction plutôt qu'en être un inhibiteur (Figure 12 A).

iii. Les tRFs comme facteurs d'adaptation aux stress

Stress biotiques : interaction hôte-pathogène

Une étude récente de Gu et al. [147] montre que le tRF-5D issu de l'ARNt^{Ala}, clivé par RNS1 et RNS3, est impliqué dans la défense des plantes contre un pathogène. Le mécanisme mis en jeu est celui de *silencing* déjà présenté précédemment. Le tRF-5D Ala est retrouvé enrichi avec AGO1 chez *A. thaliana*. Les résultats montrent qu'il cible l'ARNm du cytochrome P450 71A13 (CYP71A13) dont la séquence est complémentaire à celle du tRF. Ce mécanisme conduit au clivage de cet ARNm. Ce cytochrome est une enzyme qui catalyse la conversion d'indole acetaldoxime (IAOx) en indole-3-acetonitrile (IAN). La molécule d'IAN est un intermédiaire de la voie de biosynthèse de la camalexin (3-thiazol-2'-yl-indole). La camalexin fait partie du groupe des hormones phytoalexines, très abondantes chez les plantes. C'est un composé du métabolite secondaire utilisé par les plantes et qui serait toxique pour les pathogènes [152]. L'équipe de Gu et al. [147] établit alors le mécanisme suivant. En absence d'infection, RNS1 et RNS3 clivent l'ARNt^{Ala} pour former des tRF-5D Ala. Grâce à l'interaction avec AGO1, ces tRFs vont induire le clivage de l'ARNm codant CYP71A13. S'en suit une diminution du taux de camalexin dans l'organisme et le système de défense est réprimé. Lors d'une infection, dans cette étude avec *Botrytis cinerea*, les mécanismes de défense de la plante doivent être activés, donc dé-réprimés. Pour cela, les RNS produisent moins de tRFs et donc l'ARNm de CYP71A13 sera moins clivé. Une plus grande quantité pourra être traduite ce qui va engendrer une biogenèse plus importante de camalexin. La plante se défend donc mieux contre le pathogène (Figure 12 B).

Les résultats de cette étude sont à considérer mais avec précaution. En effet, ici les auteurs montrent que RNS1 et RNS3 sont à l'origine du tRF-5D Ala dans les feuilles d'*A. thaliana*. Les résultats de Megel et al. [113] montrent que RNS2 est aussi impliquée dans la biogenèse de ce tRF, en particulier dans les feuilles où elle est abondante. De plus, le tRF-5D Ala n'a pas été retrouvé associé à AGO1 dans une précédente étude de Cognat et al. [87]. Il reste aussi des questions en suspens concernant la production de tRF et leur régulation.

Un autre exemple de tRFs impliqués lors d'un stress biotique a été observé chez le soja. La condition étudiée est un stress bénéfique pour la plante puisque c'est une symbiose avec les bactéries du sol rhizobium. Cette symbiose provoque la formation d'organes spécifiques au niveau des racines de la plante, les nodules. Le soja tire bénéfice de cette association par l'apport d'azote par les bactéries qui quant à elles profitent des sucres libérés par la plante. L'étude chez le soja montre que des tRFs sont retrouvés dans ces nodules [149]. Ce sont des tRF-3T d'une taille variant entre 18 et 24 nt. Trois tRFs en particulier (issus des ARNt^{Val} (CAC), ARNt^{Gly} (TCC) et ARNt^{Gln} (CTG)) sont eux aussi retrouvés associés à AGO1 et ciblent trois gènes : GmRHD3, GmHAM4 et GmLRX5. Les résultats montrent que ces gènes sont impliqués dans la nodulation. En effet, dans des lignées *KO* pour ces gènes, le nombre de nodulations augmente alors que leur surexpression diminue le nombre de nodules. Cette équipe montre que par leur capacité à interagir avec AGO1, les tRFs ciblent et répriment les gènes qui empêchent la symbiose (Figure 12 B) et conclue que les tRFs sont des régulateurs positifs de la nodulation.

Une étude chez le blé a aussi montré une forte accumulation de tRFs après infection à *Fusarium graminearum*, un champignon pathogène [89]. Le taux de tRFs identifiés est même plus important que le taux de miARN. Ce sont principalement des tRFs de 18 à 21 nt qui s'accumulent lors de l'infection. Chez le blé, 9 RNases T2 sont identifiées, orthologues des RNS de *A. thaliana*. Lors d'une infection par *F. graminearum*, 4 d'entre elles sont fortement induites contrairement aux 5 autres RNases qui ont une expression constante. Ce sont les RNases T2 qui sont à l'origine des tRFs et semblent réguler leur accumulation lors de stress. Parmi les tRFs accumulés lors de l'infection, trois tRFs issus de l'ARNt^{Lys} (CTT) se démarquent des autres par une plus forte accumulation. En se basant sur les données des tRFs chez les plantes, il est probable que ces tRFs puissent réguler l'expression de gènes par *silencing*, comme précédemment. Dans cette étude, ils ont donc regardé, par complémentarité de séquence, les cibles potentielles des tRFs accumulés lors de l'infection. Les résultats de prédiction suggèrent que ces tRFs pourraient cibler des gènes impliqués dans la réplication de l'ADN, la traduction, la dégradation des protéines et la résistance aux maladies et infection par des pathogènes. Pour le moment, ces données n'ont pas encore été vérifiées expérimentalement mais ce serait ici un autre exemple où les tRFs sont impliqués dans la réponse à des attaques de pathogènes chez les plantes.

C Facteurs d'adaptation aux stress abiotiques

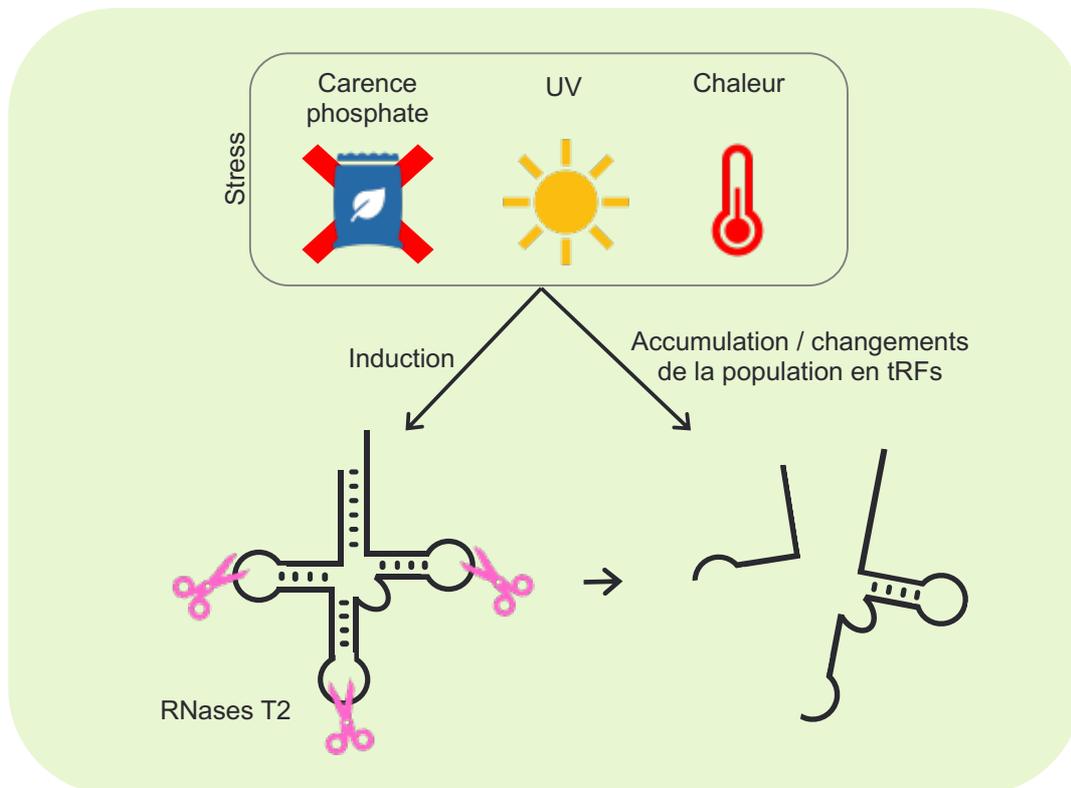


Figure 12 C : Schéma récapitulatif des fonctions des tRFs chez les plantes. Impact des stress abiotiques sur la production et l'accumulation de tRFs.

Une autre étude a mis en évidence un mécanisme où c'est le champignon pathogène, *Fusarium graminearum*, qui produit des tRFs pour réguler l'expression des gènes de la plante hôte [153]. Chez l'orge (*Hordeum vulgare*), trois tRFs produits par le champignon, de manière DCL1/2 dépendante, ont été identifiés parmi les petits ARN présents dans les feuilles infectées. Ce sont des tRFs internes, c'est-à-dire qu'ils ne proviennent ni de l'extrémité 5' ni de l'extrémité 3' des ARNt matures, et ils correspondent plutôt à des fragments partant de la région de l'anticodon et allant jusqu'à la boucle T. Les prédictions et les résultats expérimentaux montrent que ces trois tRFs ciblent des gènes de plantes impliqués dans la défense contre les pathogènes : HvEOL1 est un régulateur négatif putatif de la biosynthèse d'éthylène ; HvBAK1 et HvSERK2 sont des récepteurs à kinase. Ces gènes sont essentiels pour que la plante puisse se défendre contre les pathogènes, comme *F. graminearum*. Les résultats de cette étude montrent d'ailleurs que les champignons produisant moins de tRFs sont moins virulents chez l'orge et le brachypodium (*Brachypodium distachyon*).

Stress abiotiques

Plusieurs analyses réalisées par des équipes différentes indiquent que les populations en tRFs varient en cas de stress abiotiques : certains sont plus fortement accumulés, d'autres deviennent moins abondants (Figure 12 C).

Par exemple, chez *A. thaliana*, 42 tRFs sont dérégulés en cas de stress thermique [88]. La population de tRFs varie non seulement selon le type de stress, mais aussi selon la durée du stress. C'est le cas par exemple chez le maïs (*Z. mays*) où les changements au sein des populations de tRFs sont maximaux au bout de 48h (comparé à 2h de stress). Plus de 1748 tRFs subissent des variations d'expression au bout de 48h dont notamment le tRF-5A de 35 nt issu de l'ARNt^{Gly} (TCC) dont l'expression diminue fortement [88].

Chez la tomate (*Solanum lycopersicum*), après un traitement à l'acide abscissique (ABA, hormone impliquée dans le développement de la plante mais aussi dans leur adaptation à des stress abiotiques), une augmentation de l'expression des tRFs issu de l'ARNt^{Asp} est observée alors que ceux issu de l'ARNt^{Glu} se retrouvent moins exprimés. Plus généralement, en présence d'ABA, l'ensemble des tRFs devient moins abondant. En particulier, le tRF-5D de 20 nt issu de l'ARNt^{Ala} subit une chute drastique [154]. En cas de carence en phosphate chez *A. thaliana* ce sont les tRFs issus des ARNt^{Gly} (TCC) et ARNt^{Asp} (GTC) qui s'accumulent fortement [87,113]. Cette accumulation est d'ailleurs liée à une induction de RNS1 puisque dans des lignées *rns1*^{KO} privées de phosphate, il n'y a pas d'augmentation de la quantité de ces tRFs particuliers [113].

A l'heure actuelle, l'ensemble des données montre que les populations de tRFs varient selon les stress que subit la plante (Figure 12 C). En revanche, les mécanismes moléculaires ne sont pas encore caractérisés. Il est donc impossible à l'heure actuelle de déterminer le rôle exact de ces tRFs dans la réaction / l'adaptation des plantes face aux stress.

Les données concernant les tRFs chez les plantes permettent toutefois de commencer à déterminer leurs fonctions. Ainsi, tout comme chez les autres organismes, les tRFs auraient un rôle important dans la régulation de la traduction, que ce soit de manière globale ou de façon ciblée (ARNm spécifique). Cette régulation de la traduction est d'ailleurs plutôt bien caractérisée lors d'interaction hôte-pathogène. La régulation de la traduction peut être en faveur de la plante hôte qui se défend mieux contre les pathogènes, en faveur du pathogène qui contourne les systèmes de défense de la plante ou en faveur des deux dans le cas de symbiose.

Il est à noter que des tRFs sont retrouvés dans le phloème [155] et qu'ils pourraient être véhiculés dans la plante. Ils serviraient alors de signaux pour indiquer un stress et induire des réponses. C'est d'ailleurs le cas chez l'humain où des tRFs ont été retrouvés dans des vésicules extracellulaires pour être véhiculés entre cellules [156,157]. Les tRFs peuvent aussi être véhiculés sans vésicules comme par exemple dans une étude montrant que des tRF-5A, protégés par des complexes RNP, circulent dans le sang lors de pathologies [158] (pour une revue, voir [159]).

D'une manière générale, les populations de tRFs chez tous les organismes semblent refléter un équilibre particulier qui est modulé très finement selon l'environnement. Les variations de tRFs permettent à l'organisme de s'adapter à son environnement changeant. Un dérèglement de cet équilibre peut aussi conduire à une vulnérabilité de l'organisme face à des agresseurs (bactéries ou virus pathogènes par exemple) et être à l'origine de maladies.

En allant plus loin dans la compréhension des tRFs et de leurs fonctions, l'Homme peut exploiter ces connaissances pour les mettre à profit. Ainsi, une récente étude a montré que des tRFs provenant d'If de Chine (*Taxus chinensis*) sont toxiques pour des cellules cancéreuses humaines. Ils ont montré que ces tRFs sont capables de cibler des gènes impliqués dans le cancer des ovaires chez l'humain [160]. C'est ainsi une piste de thérapie contre les cancers.

5. Préambule au projet de thèse

C'est dans ce contexte de compréhension des tRFs chez les plantes que s'inscrit mon projet de thèse.

Un des deux axes de ma thèse a consisté à poursuivre l'étude des endoribonucléases à l'origine des tRFs chez les plantes. Chez *Arabidopsis*, les études portent essentiellement sur les RNases T2 RNS1, RNS2 et RNS3 car fortement exprimées. Elles sont à l'origine de tRFs courts et longs. Jusqu'à présent, ces résultats ont été obtenus grâce à l'étude de mutants *Knock-Out (KO)* $rns1^{KO}$, $rns2^{KO}$, $rns3^{KO}$ et $rns1rns3^{KO}$. A l'heure actuelle, il n'existe pas de mutant triple *KO* pour les trois principales RNS ($rns1rns2rns3^{KO}$). Ce triple mutant devient donc essentiel pour enrichir les connaissances déjà existantes sur les fonctions des RNS. Mon objectif était donc de réaliser ce mutant triple *KO* et d'apporter de nouvelles informations sur l'expression des RNS et leur rôle dans la biogenèse des tRFs. Les données obtenues sont présentées dans le premier chapitre de la partie Résultats de ce manuscrit.

Le second axe de ma thèse a consisté à déterminer les processus moléculaires impliquant les tRFs. Le choix s'est porté sur un tRF modèle, le tRF-5D de 20 nt issu de l'ARNt^{Ala}. L'équipe a précédemment montré, *in vitro*, que ce tRF est impliqué dans l'inhibition de la traduction. Mon travail a consisté à élucider les mécanismes moléculaires qui conduisent à cette inhibition de traduction. Pour cela, deux approches ont été utilisées pour permettre de mieux comprendre où et comment agit ce tRF : l'étude de sa localisation subcellulaire et l'identification des interactants associés à ce tRF candidat. Ces travaux sont présentés dans le second chapitre de la partie Résultat sous forme d'un article en préparation.

MATÉRIEL & MÉTHODES

MATERIEL

I. Matériel végétal

1. *Arabidopsis thaliana* pour l'étude des RNS

Des plantes de l'écotype Wassilewskija (Ws) et Col-0 ont été utilisées pour l'étude des RNS et la création de mutants *Knock-Out* par CRISPR/Cas9. Les lignées d'insertion de l'écotype Ws, *rns1^{KO}* (FLAG_566A08), *rns3^{KO}* (FLAG_164A04) et *rns4^{KO}* (FLAG_327E02), proviennent du Centre de Stock d'*Arabidopsis* de Versailles (OV-IJPB). Le double mutant *rns1rns3^{KO}* a été obtenu par croisement des lignées *rns1^{KO}* et *rns3^{KO}* [113]. Les lignées d'insertion de l'écotype Col-0, *rns2^{KO}* (SALK_69588) et *rns4^{KO}* (SALK_55304), proviennent du centre *Arabidopsis Biological Resource Center*.

2. *Arabidopsis thaliana* pour la visualisation de P-bodies et granules de stress

Pour la visualisation de granules (P-bodies et granules de stress), des lignées de l'écotype Col-0 ont été utilisées exprimant des protéines de fusion pPABP2-tRFP-PABP2 [161] et pDCP1-YFP-DCP1 [162] généreusement données par Rémy Merret (LGDP, CNRS, Perpignan).

3. *Arabidopsis thaliana* pour l'étude de DExH1

Les lignées d'insertion *dexh1^{KO}* (SALK_062060, SALK_071845, SALK_015779, SALK_045272, SALK72746, SALK_061049, SALK_062061 et SALK_061564), écotype Col-0, proviennent du centre *Arabidopsis Biological Resource Center*.

4. *Arabidopsis thaliana* écotype Landsberg

Des lignées cellulaires issues de l'écotype Landsberg (Ler0) sont cultivées en milieu liquide pour la préparation de protoplastes.

5. *Nicotiana benthamiana*

Des plantes de *Nicotiana benthamiana* ont été utilisées pour la visualisation de constructions fluorescentes par agro-infiltration. Les plantes sont cultivées en serre avec 16 h de photopériode.

II. Souches bactériennes

1. *Escherichia coli*

Pour l'expression de plasmides, la souche TOP10 (ThermoFisher Scientific) a été utilisée : F-*mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80*lacZ* Δ M15 Δ *lacX74* *recA1* *araD139* Δ (*ara-leu*)7697 *galU galK rpsL* (Str^R) *endA1 nupG*.

2. *Agrobacterium tumefaciens*

La souche GV3101::pMP90 a été utilisée pour les transformations d'Arabidopsis par *floral dip*. Elle est résistante à la rifampicine et à la gentamycine et contient le plasmide Ti (*Tumor induced*) pMP90 désarmé permettant l'expression des gènes de virulence nécessaires au transfert de l'ADN-T dans le génome de la plante hôte.

III. Vecteurs

1. pKI1.1R

Le plasmide pKI1.1R (Addgene, plasmide #85808) a été conçu par Tsutsui & Higashiyama [163] pour la transformation d'Arabidopsis par CRISPR/Cas9. Il dérive du vecteur pFAST-R01. Il contient les séquences d'ARN guide sous le promoteur AtU6.26 et le gène codant pour la SpCas9 originaire de *Streptococcus pyogenes* sous le promoteur AtRPS5A. Deux sites de restriction *AarI* permettent l'insertion des séquences des ARN guides. Ce plasmide contient le gène de résistance à la spectinomycine (Sm^R) permettant de sélectionner les bactéries transformées. Il a aussi l'intérêt de posséder comme marqueur de sélection le gène *OLE1* fusionné au TagRFP sous contrôle du promoteur endogène. *OLE1* est une protéine retrouvée dans les corps lipidiques et accumulée dans les graines d'Arabidopsis. Fusionnée à la RFP, elle permet la sélection, par fluorescence, des graines ayant intégrées l'ADN-T. Ce plasmide contient un deuxième marqueur de sélection le gène de résistance à l'hygromycine (Hyg^R).

2. pLITMUS 28i

Le vecteur pLITMUS 28i (New England BioLabs Inc.) est conçu pour le clonage et la transcription *in vitro*. Sa cassette de clonage est encadrée par deux promoteurs de l'ARN polymérase du phage T7. La sélection des bactéries transformées par ce vecteur contenant l'insert peut se faire par criblage « blanc/bleu » grâce au gène codant la β -galactosidase (*lacZ α*) et par résistance à l'ampicilline grâce au gène codant pour la β -lactamase (*bla*). Ce vecteur a été utilisé pour sa taille relativement petite (2,8 kb) en tant qu'entraîneur pour la transfection de petits ARN non-codants dans des protoplastes d'*A. thaliana*.

3. pDONR207

Le plasmide pDONR207 (Invitrogen) est un vecteur donneur utilisé pour le système de clonage Gateway. Il possède les sites *attP* permettant le premier événement de recombinaison (réaction BP) essentiels pour obtenir un clone d'entrée. Il contient le gène *ccdB*, codant un inhibiteur de la gyrase bactérienne, permettant de sélectionner les bactéries ayant intégré le vecteur recombiné. Ce plasmide contient le gène de résistance à la gentamycine (Gm^R) permettant de sélectionner les bactéries transformées.

4. pB7FWG2

Le vecteur pB7FWG2 [164] est utilisé comme vecteur de destination du système de clonage Gateway. Il possède les sites *attR* permettant le deuxième événement de recombinaison (réaction LR). Entre les sites *attR1* et *attR2* est situé le gène *ccdB*, déclenchant la mort cellulaire des bactéries ayant intégré le vecteur non recombiné. Le gène d'intérêt est inséré entre les sites *attR*. Il est placé sous le contrôle du promoteur p35S. A l'extrémité *attR2* se trouve la protéine fluorescente eGFP qui permet l'expression de la protéine d'intérêt fusionnée à la eGFP en C-terminal. Ce vecteur contient le gène de résistance à la spectinomycine (Sp^R) permettant de sélectionner les bactéries transformées.

5. pK7WGF2

Le vecteur pK7WGF2 [164] est semblable au pB7FWG2. Il est aussi utilisé comme vecteur de destination du système de clonage Gateway. Le gène d'intérêt est aussi placé sous le contrôle du promoteur p35S mais à l'extrémité *attR1* se trouve la eGFP qui permet l'expression de la protéine d'intérêt fusionnée à la eGFP en N-terminal. Ce vecteur contient le gène de résistance à la spectinomycine (Sp^R) permettant de sélectionner les bactéries transformées.

6. pALiCE

Le vecteur pALiCE[®] (LenioBio GmbH) est utilisé pour la production de protéines *in vitro*. Il est utilisé avec un kit commercial qui contient du lysat de cellules de tabac (système végétal) qui permet de transcrire la séquence d'intérêt et de la traduire *in vitro*. Le vecteur pALiCE a été modifié par la Plateforme de production de protéine de l'IBMP (P3P) pour utiliser les techniques de clonage Goldengate. Le vecteur est ainsi conçu pour insérer la séquence d'intérêt par restriction-ligation avec l'enzyme *Bsal*. Entre les sites *Bsal* se trouve le gène *ccdB* qui induit la mort cellulaire bactérienne si la séquence d'intérêt n'a pas été insérée dans le vecteur. La séquence d'intérêt est sous le contrôle du promoteur T7. Le vecteur contient le gène de résistance à l'ampicilline (Amp^R) permettant de sélectionner les bactéries transformées.

IV. Numéros d'accessions

Toutes les séquences peuvent être retrouvées dans les banques de données (*Arabidopsis Genome Initiative* ou *GenBank/EMBL*) avec les numéros d'accessions suivants : RNS1, At2g02990 ; RNS2, At2g39780 ; RNS3, At1g26820 ; RNS4, At1g14210 ; DExH1, At2g35920 ; ACT2, At3g18780 ; GAPDH, At3g26650.

V. Oligonucléotides

Les oligonucléotides ont été synthétisés par Integrated DNA Technology.

VI. Anticorps

Les anticorps dirigés contre les RNS (@RNS1 25 kDa, @RNS2 29 kDa, @RNS3 25 kDa et @RNS4 28 kDa) ont été produits auparavant par la société Covalab à partir des protéines recombinantes produites en système bactérien et purifiées au laboratoire. Ce sont des anticorps polyclonaux de lapin purifiés sur colonne Protein A HiTrap HP (GE) par la plateforme P3P [113].

Les anticorps polyclonaux de lapins anti-eEF1 α (@eEF1 α , 50 kDa) sont dirigés contre le facteur d'élongation de la traduction eEF1 α du maïs (*Zea mays*). Ils ont été fournis par le Dr. Brenda Hunter de l'Université d'Arizona.

VII. Outils informatiques

1. CRISPOR

L'outil informatique, disponible en ligne, CRISPOR version 4.97 a été utilisé pour concevoir les séquences des ARN guides pour la transformation par le système CRISPR/Cas9 (<http://crispor.tefor.net/>; [165]). La séquence d'ADNg du gène *RNS2* d'*A. thaliana* écotype Ws a servi de matrice pour prédire les séquences guides et les *off-target* ont été prédits en se basant sur le génome d'*A. thaliana* (TAIR10).

2. MacVector

Le logiciel MacVector[®] a été utilisé pour l'analyse de séquences d'ADN, notamment pour des alignements multiples lors de comparaison des séquences de plantes transformées par le système CRISPR/Cas9.

3. ImageJ

Le logiciel libre Image J (NIH) a été utilisé pour l'analyse d'images de microscopie et la réalisation de figures. Il a aussi servi à la visualisation et à la quantification de signaux obtenus après exposition de membrane contre une plaque de PhosphorImager. Les plugins et fonctions suivants ont été utilisés :

- Find Maxima : cette fonction d'ImageJ a permis de détecter les foci sur des images de microscopie. Les pixels ayant une forte intensité et de manière ponctuelle sont détectés. Un seuil de détection est fixé par l'utilisateur.

- 2CH_Spots_Tool.ijm : cette macro a été fournie par Jérôme Mutterer (IBMP). Elle permet de détecter les foci avec la fonction Find Maxima et ensuite de quantifier l'intensité des pixels dans 2 canaux à l'endroit où les foci ont été détectés (grâce aux coordonnées x et y des foci).
- NeuronJ [166] : ce plugin a permis de mesurer la longueur de racines de plantes à partir de photos en format JPEG ou TIF. Brièvement, les photos sont converties au format 8-bit-grey. Le début et la fin de la racine à mesurer est indiqué par l'utilisateur, le logiciel va indiquer un tracé correspondant à la racine. La longueur de la racine est ensuite calculée par le logiciel en se basant sur la calibration de l'image fournie au préalable par l'utilisateur.
- FigureJ [167] : ce plugin a permis de faire les figures d'images de microscopie. Il permet de lire les images aux formats LSM et LIF et garde les métadonnées associées.

4. Alignements de séquence

D'une manière générale, les alignements de séquence ont été réalisés avec les outils BLAST développés et mis à disposition par le NCBI (*National Center of Biotechnology Information*) :

- Nucleotide BLAST : alignement de séquences nucléiques
- Protein BLAST : alignement de séquences protéiques
- Smart BLAST : recherche des séquences protéiques similaires ou proche de la séquence d'intérêt, dans tous les organismes modèles.
- Multiple Alignment : permet d'aligner plusieurs séquences simultanément en se basant sur les domaines/séquences en commun.

METHODES

La composition des solutions, tampons et milieux indiqués en gras dans le texte est détaillée en Annexe II. L'eau utilisée pour l'ensemble des expérimentations est de l'eau ultrapure filtrée par des distributeurs Milli-Q®.

I. Méthodes relatives aux organismes

1. Méthodes relatives au matériel bactérien

i. Préparation de bactéries compétentes d'Escherichia coli

Les bactéries issues d'un glycérol stock sont étalées sur une boîte de Pétri contenant du **milieu LB-agar** additionné de streptomycine (25 µg/mL). Après une nuit à 37°C, une colonie isolée est utilisée pour réaliser une pré-culture de 3 mL de **milieu LB** additionnée de streptomycine (25 µg/mL) qui est ensuite incubée une nuit à 37°C. La pré-culture est repiquée dans 500 mL de milieu LB additionné de streptomycine (25 µg/mL) et 10 mM de MgCl₂ puis incubée plusieurs heures à 18°C avec agitation jusqu'à obtenir un DO_{600nm} de 0,4. La culture est stoppée en mettant l'erlen sur glace puis les bactéries sont culotées en les centrifugeant 10 min à 2700 g à 4°C. Le culot de bactéries est repris dans 80 mL de **tampon TB** froid et incubé 10 min à 4°C. Après une centrifugation de 10 min à 4500 g à 4°C, le culot est finalement repris dans 20 mL de tampon TB + DMSO 7% (p/v) et incubé 10 min sur glace. Les bactéries sont ensuite aliquotées par 50 µL, congelées à l'azote liquide et conservées à -80°C.

ii. Transformation d'E. coli

Les bactéries compétentes TOP10 sont transformées par choc thermique. Pour cela, environ 50 ng de plasmide ou le produit de ligation sont mélangés à 50 µL de bactéries décongelées extemporanément. Après incubation 10 à 30 min sur glace, les bactéries subissent un choc thermique à 42°C durant 30 à 40 sec puis sont à nouveau placées sur glace durant 2 min. Après avoir ajouté 250 µL de milieu LB aux bactéries, elles sont incubées 1h à 37°C. Le tout est finalement étalé sur des boîtes de Pétri contenant 25 mL de milieu LB-agar additionné des antibiotiques (Annexe III) pour permettre la sélection des bactéries transformées. Les boîtes sont finalement incubées à 37°C pendant au minimum 15 h.

iii. Préparation d'Agrobacterium tumefaciens compétentes

Les agrobactéries issues d'un glycérol stock sont étalées sur une boîte de Pétri contenant du milieu LB-agar additionné de rifampicine (50 µg/mL) et gentamycine (15 µg/mL). Après deux jours à 28°C, une colonie isolée est utilisée pour réaliser une pré-culture de 10 mL de milieu LB additionné de rifampicine (50 µg/mL) et gentamycine (15 µg/mL) qui est ensuite incubée une nuit à 28°C. Cinq mL de pré-culture sont repiqués dans 200 mL de milieu LB puis incubée plusieurs heures à 28°C avec agitation jusqu'à obtenir un DO_{600nm} comprise entre 0,5 et 1.

La culture est stoppée en mettant l'erlen sur glace pendant 5 min puis les bactéries sont culotées en les centrifugeant 10 min à 2700 g à 4°C. Le culot de bactéries est repris dans 4 mL de CaCl₂ 20 mM. Les bactéries sont ensuite aliquotées par 100 µL, congelées à l'azote liquide et conservées à -80°C.

iv. Transformation d'Agrobacterium tumefaciens

Les bactéries compétentes GV3101 sont transformées par choc thermique. Environ 800 ng de plasmide sont ajoutés à 100 µL de bactéries décongelées extemporanément. Les bactéries subissent ensuite un choc thermique à 37°C pendant 5 min puis 900 µL de milieu LB sont ajoutés. Elles sont ensuite incubées 2 h à 28°C puis étalées sur une boîte de Pétri contenant 25 mL de milieu LB-agar additionné des antibiotiques pour la sélection (Annexe III). Après 2 à 3 jours d'incubation à 28°C, les colonies de bactéries transformées sont observables.

2. Méthodes relatives au matériel végétal

i. Stérilisation de graines d'Arabidopsis

Environ 25 mg de graines à stériliser sont incubées dans 1 mL de **solution de stérilisation** durant 15 min à température ambiante sur un agitateur rotatif. Ensuite, en conditions stériles, la totalité de la solution est éliminée puis les graines sont incubées dans 1 mL d'éthanol 100 % durant 2 min sur un agitateur rotatif. Cette étape est répétée une seconde fois puis les graines sont séchées en laissant le tube ouvert sous hotte stérile pendant au minimum 1 h.

ii. Culture d'Arabidopsis

Culture in vitro en milieu solide

Les graines préalablement stérilisées sont semées sur des boîtes de Pétri rectangulaires (120 mm x 120 mm) contenant 50 mL de milieu gélosé (**milieu gélosé MS 222, milieu gélosé +Pi ou milieu gélosé -Pi**). Les boîtes sont scellées avec deux couches de ruban adhésif poreux (Micropore, 3M) et placées à l'obscurité à 4°C durant 2 à 5 jours pour vernalisation. Elles sont ensuite disposées horizontalement en chambre de culture à jours longs (16 h jour, 21°C – 8 h nuit, 17°C).

Culture in vitro en milieu liquide

Les graines préalablement stérilisées sont semées dans des erlens de 250 mL contenant 80 mL de milieu liquide (**milieu liquide MS 231, milieu liquide +Pi ou milieu liquide -Pi**). Les erlens sont placés à l'obscurité à 4°C durant 2 à 5 jours pour vernalisation. Ils sont ensuite disposés sur un agitateur orbital, à environ 130 rpm, en chambre de culture à lumière continue à 24°C. Les plantules sont récoltées après 7 jours en chambre de culture.

Culture en terre

Les graines sont, au préalable, placées à 4°C durant 2 à 5 jours pour vernalisation. Elles sont ensuite semées et cultivées en jours courts (12 h jour, 23°C – 12 h nuit, 23°C) durant 3 à 4 semaines puis, après repiquage, elles sont cultivées en jours longs (16 h jour, 23°C – 8 h nuit, 23°C).

Culture cellulaire

Les cultures cellulaires sont repiquées tous les 7 jours. Environ 5 mL d'une culture de 7 jours sont repiqués dans 100 mL de nouveau **milieu de culture cellulaire**. La culture est mise en chambre de culture à lumière continue à 24°C en agitation orbitale à 130 rpm.

iii. Transformation stable d'*Arabidopsis* par floral dip*Préparation des plantes*

Les plantes à transformer sont cultivées en terre, en jours longs. La première hampe florale de chaque plante est coupée à sa base pour favoriser la croissance d'hampes secondaires. Les plantes sont transformées une semaine après lorsqu'un maximum de boutons floraux sont encore fermés.

Préparation des agrobactéries

Une pré-culture de 5 mL de milieu LB additionné des antibiotiques permettant la sélection des agrobactéries transformées est réalisée à partir de colonies cultivées sur boîte. Après 48 h de croissance à 28°C sous agitation, 200 mL de nouveau milieu LB additionné des antibiotiques sontensemencés avec 200 µL de la pré-culture. Après 24 h de croissance à 28°C sous agitation, la culture est centrifugée 15 min à 5000 g. Le culot bactérien est repris dans 50 mL de **tampon d'agro-transformation**. La DO_{600nm} de la suspension est mesurée puis elle est diluée pour être amenée à la valeur de 1. La suspension est laissée au minimum 1 h à température ambiante.

Floral dip

Les hampes florales des plantes à transformer sont immergées dans la suspension bactérienne pendant 1 min. Elles sont ensuite légèrement égouttées, tuteurées et placées dans une mini-serre à l'obscurité pendant 48 h. Les plantes sont finalement mises à chambre de culture (jours longs) jusqu'à maturité des graines.

iv. Sélection de graines fluorescentes

Les graines issues de transformation avec le plasmide pK11.1R sont sélectionnées par fluorescence sous un stéréomicroscope (AxioZoom, Zeiss) avec le filtre mRFP. Les graines émettant de la fluorescence (tRFP) sont semées sur terre directement après avoir été prélevé une par une à l'aide d'un cure-dent humide.

v. Agro-infiltration de feuille de *N. benthamiana*

Préparation des agrobactéries

Une pré-culture de 1 mL de milieu LB additionné des antibiotiques permettant la sélection des agrobactéries transformées est réalisée à partir de glycérol stock. Après environ 7 h de croissance à 28°C sous agitation, 20 mL de nouveau milieu LB additionné des antibiotiques sontensemencés avec 500 µL de la pré-culture. Après 50 h de croissance à 28°C sous agitation, la culture est centrifugée 10 min à 3200g. Le culot bactérien est repris dans 5 mL de MgCl₂ 10 mM. La DO_{600nm} de la suspension est mesurée puis elle est diluée avec du MgCl₂ 10 mM pour être amenée à la valeur de 0,5.

Infiltration de feuilles avec les bactéries

Une blessure est réalisée à l'aide d'une aiguille sur des jeunes feuilles de *N. benthamiana* (agés de 3 semaines environ). La suspension de bactéries est injectée dans les feuilles à l'aide d'une seringue au niveau des blessures. Les plantes de *N. benthamiana* sont ensuite remises en culture pendant 2 jours pour permettre l'expression de la protéine d'intérêt.

vi. Préparation et transfection de protoplastes d'*Arabidopsis*

La préparation et la transfection de protoplastes sont décrits dans l'article méthodologique pour lequel j'ai participé à la réalisation des expériences et à la rédaction et dont je suis second auteur, « Transfection of small non-coding RNAs into *Arabidopsis thaliana* protoplasts » paru dans un volume spécial du journal *Methods in Molecular Biology* [168](Annexe IV). Les détails des oligonucléotides utilisés pour la transfection sont donnés en Annexe V.

Transfection des protoplastes avec des plasmides

Pour localiser DExH1, les protoplastes sont transfectés avec les vecteurs Gateway exprimant DExH1 fusionnée à la eGFP. Pour chaque condition, 100 µL de protoplastes (soit environ 300 000 protoplastes) sont transfectés avec 15 µg de vecteur, 15 µg de plasmide entraîneur (pLitmus) et 100 µL de tampon de transfection. La transfection a lieu durant 15 min. Les protoplastes transfectés sont ensuite lavés comme expliqué dans l'article (Annexe IV) et repris dans 100 µL de tampon W1. Ils sont transférés dans une plaque 24 puits qui est scellée avec du scotch puis mise à l'obscurité pendant une nuit, à 25°C, le temps que la protéine d'intérêt soit exprimée.

Transfection de protoplastes avec des oligoribonucléotides fluorescents

Pour localiser les tRFs, les protoplastes sont transfectés avec des oligoribonucléotides possédant un fluorochrome. Pour chaque condition, 100 µL de protoplastes (soit environ 3.10⁶ protoplastes) sont transfectés avec 1,5 µL d'oligoribonucléotide fluorescent à 100 µM (Annexe V), 22,5 µg de plasmide entraîneur (pLitmus) et 100 µL de tampon de transfection. La transfection a lieu durant 30 min, en agitation sur une roue à une vitesse de 10 rpm. Les protoplastes transfectés sont ensuite lavés comme expliqué dans l'article (Annexe IV). Ils sont directement utilisés pour observer la localisation des tRF transfectés au microscope confocal.

Transfection de protoplastes avec des oligoribonucléotides biotinylés

Pour identifier les partenaires protéiques des tRFs par spectrométrie de masse après chromatographie, environ 3.10^6 protoplastes (dans 1 mL) sont transfectés avec 15 μL d'oligonucléotide biotinylé à 100 μM (Annexe V), 225 μg de plasmide entraîneur (pLitmus) et 950 μL de tampon de transfection. La transfection a lieu durant 30 min, en agitation sur une roue à une vitesse de 10 rpm. Les protoplastes transfectés sont lavés comme expliqué dans l'article (Annexe IV). Ils sont ensuite utilisés pour la chromatographie d'affinité à la streptavidine (voir IV.4.i, page 94).

II. Méthodes relatives à l'ADN

1. Extraction d'ADN génomique d'Arabidopsis

i. Extraction par broyage

L'ADN génomique (ADNg) de plantes d'Arabidopsis est extrait selon le protocole suivant. Un tube à vis Precellys® (Ozyme) est rempli de 100 mg de billes de verre de 0,75-1mm de diamètre et de 400 μL de **tampon d'extraction d'ADNg**. Un morceau de feuille caulinaire ou de feuille de rosette (environ 1 cm^2) est prélevé sur la plante et placé dans le tube. Les feuilles sont broyées avec l'appareil Precellys Evolution deux fois de suite, interrompues par une pause de 5 sec, à 5500 rpm durant 30 sec. Après centrifugation du broyat 5 min à 16000 g, 200 μL de surnageant sont transvasés dans un nouveau tube. Au surnageant est ajouté 150 μL d'isopropanol 100 % puis le tout est incubé 5 min à température ambiante. Après centrifugation 5 min à 16000 g, le surnageant est éliminé et le culot est lavé dans 900 μL d'éthanol 70 %. Le culot est ensuite séché puis repris dans 50 μL de **tampon T**. Un traitement à la RNase A (concentration finale 40 $\mu\text{g}/\text{mL}$) est réalisé durant 1 h à température ambiante. L'ADN est ensuite précipité avec 2 μL de NaCl à 5 M et 125 μL d'éthanol 100 % durant au minimum 1 h à -20°C . Après centrifugation 30 min à 16000 g à 4°C , le surnageant est éliminé et le culot d'ADN est finalement repris dans 30 μL de tampon T.

ii. Extraction rapide par chauffage

Une extraction d'ADN de feuille peut aussi être réalisée en plaque PCR 96 puits (plaque avec puit de 200 μL , compatible avec un thermocycler), adaptée du protocole précédent, sans broyage.

Les tubes de la plaque sont préalablement remplis avec 100 μL de Tampon d'extraction d'ADNg supplémenté de 0,5 % de SDS. Un morceau de feuille est prélevé et placé dans chaque puit. Après avoir scellé la plaque, celle-ci est chauffée à 95°C durant 10 min puis placée sur glace. Après avoir transféré 90 μL de surnageant dans une nouvelle plaque, 100 μL d'isopropanol 100 % sont ajoutés puis le tout est incubé 20 min à température ambiante. Après centrifugation 20 min à 3200 g à 4°C , le surnageant est éliminé et les culots est lavé dans 150 μL d'éthanol 75 %. Après une deuxième centrifugation, le culot est séché puis finalement repris dans 60 μL de tampon T.

2. Clonages

i. Clonage pour la technique CRISPR/Cas9

Digestion du plasmide pKI1.1R

Le plasmide pKI1.1R doit être au préalable digéré pour pouvoir y insérer les séquences des ARN guides. La digestion a lieu dans un volume final de 20 µL comprenant 1 µg de plasmide pKI1.1R, 0,5 µmol d'oligonucléotide 50X (ThermoFisher), 2 µL de tampon AarI 10X (ThermoFisher) et 2 µL de AarI (2U/µL, ThermoFischer). Cette enzyme cible des sites spécifiques et a une activité optimale lorsqu'il y a au moins deux de ces sites présents. Comme le plasmide pKI1.1R ne possède qu'un site de reconnaissance AarI, il est recommandé d'ajouter des oligonucléotides mimant ces sites spécifiques pour améliorer le rendement de digestion par l'enzyme. Le tube est ensuite incubé 4 h à 37°C. Le plasmide digéré est purifié avec le kit PCR Clean-up (Macherey & Nagel) en suivant le protocole du fournisseur puis élué avec durant 1 min à 70°C avec 15 µL d'eau chauffée à 70°C. Le plasmide est récupéré après une centrifugation, 1 min à 11000 g. Un dosage est réalisé avec le Nanodrop pour estimer la quantité de plasmide.

Hybridation des ARN guides

Avant d'être insérés dans le plasmide pKI1.1R, les séquences d'ADN simple brin correspondant aux ARN guides (Annexe VI) ont besoin d'être hybridées pour former des doubles brins d'ADN. Pour ce faire, dans un tube sont mélangés 0,5 nmol d'oligonucléotide sens, 0,5 nmol d'oligonucléotide antisens et 1,1 µL de **tampon de ligation 10X** (Promega). Le tube est chauffé au bain-marie 10 min à 70°C puis refroidi lentement à température ambiante pour que les deux oligonucléotides s'hybrident. Une dilution au $1/1800$ dans de l'eau est ensuite réalisée.

Ligation des ARN guides au plasmide pKI1.1R

La ligation a lieu dans un volume final de 5 µL comprenant 1 U de T4 DNA Ligase (Promega), 2,5 µL de tampon de ligation 2X (Promega), 0,20 µL de la dilution d'oligonucléotides hybridés et le plasmide pKI1.1R digéré en respectant le rapport molaire insert / plasmide égal à 1. La ligation s'effectue à 4°C pendant environ 15 h. Le mix de réaction est directement utilisé pour transformer des bactéries compétentes TOP10. Elles permettent de produire un stock de plasmide purifié (voir II.6, page 88).

ii. Clonage Gateway

Amplification de la séquence d'intérêt par PCR à haute-fidélité

Les conditions des différentes PCR selon l'ADN à amplifier et les oligonucléotides utilisés sont détaillés en Annexe VII.

La première étape consiste à récupérer la séquence à insérer dans le vecteur Gateway. La séquence peut être disponible dans un plasmide déjà existant ou être synthétisée à partir d'ARN rétrotranscrit en ADNc (voir III.2, page 89). La séquence d'ADN ou ADNc d'intérêt doit être amplifiée avec une polymérase à haute-fidélité qui ne créera pas d'erreurs aléatoires dans la séquence. La PCR peut être réalisée en utilisant directement les amorces qui ajouteront les sites *attB1* et *attB2* à la séquence d'intérêt. Pour cela, une réaction de PCR est réalisée dans un volume de 60 μ L composé de 12 μ L de tampon 5X (ThermoFisher), 1,2 μ L d'un mix de chaque désoxynucléotide (dATP/dTTP/dGTP/dCTP) à 10 mM chacun, 30 pmol de chacun des deux oligonucléotides sens et antisens, entre 50 et 100 ng de cDNA et 0,6 μ L d'ADN polymérase Phusion (à 2 U/ μ L, ThermoFisher). La réaction de PCR commence par une étape de dénaturation 30 sec à 98°C puis de 30 cycles d'amplification se décomposant en 3 étapes. Une première étape de dénaturation 10 sec à 98°C, une seconde étape d'hybridation 20 sec à une température adaptée au T_m des amorces (voir Annexe VII) et une troisième étape d'élongation 30 sec (30 sec/kb, selon la taille du fragment à amplifier) à 72°C. La réaction se termine par une étape d'élongation 5 min à 72°C. Il peut être parfois nécessaire de réaliser une nouvelle PCR dans les mêmes conditions mais à partir du produit de PCR précédent dans le cas de gènes peu exprimés et donc en faible quantité. Les produits de PCR sont purifiés avec le kit PCR Clean-up (Macherey & Nagel) en suivant le protocole du fournisseur.

Réaction BP

La réaction BP permet d'insérer le gène d'intérêt encadré des sites *attB1* et *attB2* dans un vecteur donneur ayant les sites *attP1* et *attP2* (par exemple le pDONR207). Il en résulte un vecteur d'entrée contenant le gène d'intérêt encadré par les sites *attL1* et *attL2*. Dans un tube, pour un volume final de 14 μ L, sont mélangés 200 ng de produit PCR, 300 ng de pDONR207 et 2 μ L de BP clonase (Invitrogen). La réaction a lieu durant 1h à 25°C puis est arrêtée par ajout de 1 μ L de protéinase K (à 2 μ g/ μ L, Invitrogen) et incubation 10 min à 37°C. Le mix de réaction est directement utilisé pour transformer des bactéries compétentes TOP10. Elles permettent de produire un stock de plasmide d'entrée. Les plasmides sont purifiés (voir II.6, page 88) puis vérifiés par séquençage et utilisés pour la réaction LR.

Réaction LR

La réaction LR permet d'insérer le gène d'intérêt contenu dans le vecteur d'entrée et encadré par les sites *attL1* et *attL2* dans le vecteur de destination contenant les sites *attR1* et *attR2*. Dans un tube, pour volume final de 10 μ L, sont mélangés 300 ng de vecteur de destination (pK7WGF2 ou pB7FWG2), entre 100 et 200 ng de pDONR207 issu de la réaction BP et 2 μ L de LR clonase (Invitrogen). La réaction a lieu durant 2 jours à 20°C puis est arrêtée par ajout de 1 μ L de protéinase K (2 μ g/ μ L, Invitrogen) et incubation 10 min à 37°C. Le mix de réaction est directement utilisé pour transformer des bactéries compétentes TOP10. Elles permettent de produire un stock de vecteur qui sera purifié (voir II.6, page 88) puis vérifié par séquençage.

iii. Clonage ALiCE

Amplification de la séquence d'intérêt par PCR à haute-fidélité

Le clonage ALiCE est basé sur la technique Goldengate. La séquence du gène d'intérêt est insérée dans le vecteur pALiCE par restriction-ligation avec l'enzyme *Bsal*. Pour cela, la séquence est amplifiée par PCR avec des amorces qui vont ajouter les sites *Bsal* à ses extrémités. Aucun site *Bsal* ne doit se trouver à l'intérieur de la séquence. Les sites internes sont mutés par PCR avec des amorces spécifiques. Le détail des amorces et des conditions de PCR est détaillé en Annexe VII.

Les séquences d'intérêt, sont amplifiées par PCR haute-fidélité à partir des vecteurs recombinants pK7WGF2 et pB7FWG2 (Nter-GFP-DExH1, DExH1-GFP-Cter et GFP). La réaction de PCR est réalisée dans un volume de 50 μL composé de 4 μL de tampon 5X High Fidelity (ThermoFisher), 0,4 μL d'un mix de chaque désoxynucléotide (dATP/dTTP/dGTP/dCTP) à 10 mM chacun, 25 pmol de chacun des deux oligonucléotides sens et antisens, 20 ng de vecteur pK7WGF2 ou pB7FWG2 et 0,5 μL d'ADN polymérase Phusion (à 2 U/ μL , ThermoFisher). La réaction de PCR commence par une étape de dénaturation 30 sec à 98°C puis de 30 cycles d'amplification se décomposant en 3 étapes. Une première étape de dénaturation 10 sec à 98°C, une seconde étape d'hybridation 20 sec à une température adaptée au T_m des amorces (voir Annexe VII) et une troisième étape d'élongation de 30 sec/kb, selon la taille du fragment à amplifier, à 72°C. La réaction se termine par une étape d'élongation 5 min à 72°C. Les produits de PCR sont purifiés avec le kit PCR Clean-up (Macherey & Nagel) en suivant le protocole du fournisseur.

*Restriction-ligation par *Bsal**

Les fragments de PCR et le vecteur pALiCE sont digérés par *Bsal* puis ligués entre eux. La digestion-ligation se fait simultanément dans un volume final de 20 μL contenant 10 μL de tampon T4 ligase 10X (Promega), 1 μL de T4 ligase (3 U/ μL , Promega), 1 μL de dATP 10 mM, 1 μL de *Bsal* (10 U/ μL , ThermoFisher), les fragments de PCR à insérer et 75 ng de vecteur pALiCE en respectant un ratio 2 : 1 (insert : vecteur). Durant la première étape de la réaction, le mélange est chauffé 3 min à 37°C puis 4 min à 16°C. Cette première étape est répétée 26 fois. Le mélange est ensuite chauffé 5 min à 50°C puis 5 min à 80°C. Le mix de réaction est directement utilisé pour transformer des bactéries compétentes TOP10. Elles permettent de produire un stock de vecteur. Le vecteur recombinant est vérifié par PCR directement à partir des colonies bactériennes ou après purification du plasmide (voir II.3, page 87) qui sera vérifié par séquençage.

3. Amplification d'ADN par PCR

Les conditions des différentes PCR selon l'ADN à amplifier et les oligonucléotides utilisés sont détaillés en Annexe VII.

En règle générale, une réaction de PCR dans un volume de 25 μL est composée de 5 μL de tampon GoTaq 5X (Promega), 2,5 μL de MgCl_2 à 25 mM, 0,5 μL d'un mix de chaque désoxynucléotide (dATP/dTTP/dGTP/dCTP) à 10 mM chacun, 10 pmol de chacun des deux oligonucléotides sens et antisens, 0,125 μL d'ADN polymérase GoTaq® (Promega) et de l'ADN matrice (issu d'extraction d'ADNg de plante ou de colonie bactérienne). La réaction de PCR commence par une étape de dénaturation 5 min à 95°C, puis de 30 à 35 cycles d'amplification se décomposant en 3 étapes. Une première étape de dénaturation 30 sec à 95°C, une seconde étape d'hybridation 30 sec à une température environ égale au T_m des amorces – 5°C et une troisième étape d'élongation 30 sec à 1 min (1 kb/min, selon la taille du fragment à amplifier) à 72°C. La réaction se termine par une étape d'élongation 5 min à 72°C.

4. Reverse Transcription - PCR quantitative (RT-qPCR)

La qPCR est réalisée avec de l'ADNc préparé à partir d'ARN total traité à la DNase et rétrotranscrit (voir III.2, page 89). L'ADNc est généralement dilué au $1/30^{\text{e}}$ voire $1/50^{\text{e}}$. Le mix de qPCR est réalisé dans un volume final de 10 μL composé de 4 μL d'ADNc (ou de Standard), 1 μL d'un couple d'amorce à 2,5 μM chacune et 5 μL de SYBR 2X Taykon qPCR Mastermix Blue (Eurogentec). Après une dénaturation de 7 min à 95°C, 45 cycles d'amplification sont réalisés de 15 sec à 95°C, suivi de 15 sec à 60°C et 15 sec à 72°C. La PCR est réalisée avec l'appareil LightCycler480 (Roche) et analysée à l'aide du logiciel correspondant. Le logiciel analyse les courbes d'amplification pour chaque échantillon et détermine les cycles seuils (Ct) par la méthode des dérivés secondes. Les efficacités d'amplification de chaque couple d'amorces ont été déterminées avec le logiciel LinReg. L'abondance relative en ARN dans les échantillons a été analysée par la méthode des delta-delta Ct. Le détail des amorces utilisées est donné en Annexe VIII.

5. Électrophorèse sur gel d'agarose

Les produits d'amplification de PCR sont déposés sur un gel d'agarose de pourcentage adapté selon la taille du produit et de **tampon TAE 0,5X**. Les produits sont soit d'abord chargés avant d'être déposés ($1/6^{\text{ième}}$ de volume de **tampon de charge ADN 6X**), soit ils sont directement déposés (tampon de charge déjà présent dans le mix de réaction PCR). La migration s'effectue sous une tension constante de 100 V dans du tampon TAE 0,5X. L'ADN est ensuite visualisé par fluorescence sous UV grâce au bromure d'éthidium (BET) incorporé dans le gel.

6. Purification de plasmides

Pour une petite préparation, une culture contenant 3 mL milieu LB additionné des antibiotiques adéquats pour la sélection estensemencée à partir d'une colonie de bactéries cultivées sur boîte après transformation. Elle est incubée environ 15 h à 37°C sous agitation puis elle est utilisée pour réaliser l'extraction d'ADN plasmidique avec le kit Nucleospin Pasmid Quick Pure™ (Macherey-Nagel), en suivant les recommandations du fournisseur.

Pour une préparation plus importante, une pré-culture de 5 mL de milieu LB additionné des antibiotiques adéquats estensemencée à partir d'une colonie de bactéries. La pré-culture est placée à 37°C durant environ 8h sous agitation. Une culture est ensuite préparée avec 100 mL de milieu LB additionné des antibiotiques en l'inoculant avec 100 µL de pré-culture. Elle est incubée à 37°C durant 15h sous agitation. Cette culture est ensuite utilisée pour extraire l'ADN plasmidique avec le kit NucleoBond® Xtra Midi (Macherey-Nagel), en suivant les recommandations du fournisseur.

7. Séquençage d'ADN

Le séquençage d'ADN est basé sur la technique Sanger et est réalisé par la plateforme d'analyse de l'expression génique de l'IBMP. Les amorces utilisées pour le séquençage sont détaillées en Annexe IX.

i. Séquençage à partir de produits de PCR

Pour le séquençage de régions d'ADNg amplifiées par PCR, 10 pmol d'amorce oligonucléotidique (même amorce que l'une des deux utilisées pour la réaction de PCR) sont ajoutées à 4 µL de produit de PCR, le tout dans un volume final de 5 µL.

ii. Séquençage d'ADN plasmidique

L'ADN plasmidique est séquencé après avoir été exprimé en système bactérien et purifié. Environ 500 à 800 ng d'ADN plasmidique sont ajoutés à 5 pmol d'amorce oligonucléotidique, le tout dans un volume final de 10 µL.

8. Marquage de sondes oligonucléotidiques

Les oligonucléotides (Annexe X) sont marqués radioactivement en 5'. Pour ce faire, 10 µM d'oligonucléotide à marquer, 1 µL de **tampon PNK 10X** (Fermentas), 1 µL de T4 polynucléotide kinase (PNK) à 10 U/µL (Fermentas) et 2,5 µL de [$\gamma^{32}\text{P}$] ATP (10 µCi/µL, activité spécifique 3 µCi/pmol) sont mélangés dans un volume final de 10 µL. Après marquage durant 30 min à 37°C, 50 µL d'eau sont ajoutés au mélange. La radioactivité non incorporée est éliminée sur colonne de chromatographie d'exclusion de type Sephadex G50 (en seringue de 1 mL, séchée juste au préalable par centrifugation 45 sec à 400g), en centrifugeant la sonde déposée sur la colonne 45 sec à 400 g. Finalement, 500 µL de PerfectHyb™ Plus (Sigma) sont ajoutés à la sonde radiomarquée.

III. Méthodes relatives à l'ARN

1. Extraction des ARN de feuilles d'*A. thaliana*

i. Extraction des ARN totaux

Les ARN totaux sont extraits à partir de feuilles de rosette d'*A. thaliana*. Les feuilles prélevées sont broyées très finement dans l'azote liquide à l'aide d'un mortier et d'un pilon. Pour 200 mg de poudre, 1,5 ml de TRI Reagent® (Molecular Research Center, Inc.) sont ajoutés et l'ensemble est vortexé durant 10 min. Après centrifugation 10 min à 12000 g à 4°C, le surnageant est prélevé et 300 µL de chloroforme y sont ajoutés. L'ensemble est à nouveau vortexé durant 10 min. Après centrifugation 15 min à 12000 g à 4°C, 900 µL de phase aqueuse sont récupérés et 450 µL d'isopropanol 100 % sont ajoutés. L'ensemble est homogénéisé puis incubé 10 min à température ambiante. Après centrifugation 10 min à 12000 g à 4°C, le surnageant est éliminé et le culot contenant les ARN totaux est lavé à l'éthanol 75 %. Après avoir été séché à l'air libre, le culot d'acides nucléiques est finalement dissout dans 20 µL d'eau.

ii. Extraction des petits ARN

Une fois les ARN totaux dissouts dans l'eau, 0,7 V de **tampon de précipitation 1X** sont ajoutés. Le tube est placé 1 h minimum à 4°C. Après centrifugation 30 min à 16000g à 4°C, le surnageant est prélevé (il contient les petits ARN, le culot contient les ARN supérieurs une taille de 200 nt). Les petits ARN du surnageant sont précipités en ajoutant 0,1 V d'acétate de sodium 1M (pH 4,5), 2,5 V d'éthanol 100 % et 0,5 µL de glycogène à 20 mg/mL (ThermoFisher Scientific) et l'ensemble est placé 1h minimum à -20°C. Après centrifugation 30 min à 16000g à 4°C, le culot est repris dans un volume adéquat d'eau.

2. Préparation d'ADNc

i. Traitement DNase

Avant la rétrotranscription, l'élimination de toute trace d'ADN dans l'échantillon d'ARN est essentielle. Le traitement DNase est réalisé dans un volume de 100 µL final composé de 15 µg d'ARN à traiter, 10 µL de tampon DNase RQ1 10X (Promega) et 2,5 µL de DNase RQ1 (1 U/µL, Promega), durant 30 min à 37°C. La réaction est arrêtée par l'ajout de 10 µL de solution STOP RQ1 (Promega) suivi d'une incubation de 10 min à 65°C. L'ARN est ensuite mis à précipiter avec 0,1 V d'acétate de sodium 1M pH 4,5 et de 2,5 V d'éthanol 100 % pendant au moins 3 h à -20°C. Après centrifugation 30 min, à 16000 g à 4°C, le culot d'ARN est repris dans 20 µL d'eau.

ii. Rétrotranscription

Dans un tube sont mélangés 2,5 µg d'ARN traités à la DNase, 1 µL de Random Hexamers Primers à 500 ng/µL (Promega), 1 µL d'oligo dT 100 µM, 1 µL d'un mix de chaque désoxynucléotide (dATP/dTTP/dGTP/dCTP) à 10 mM chacun, le tout dans un volume final de 13 µL. L'ensemble est chauffé 5 min à 65°C puis, sur glace, sont ajoutés 4 µL de tampon SuperScript IV 5X (Invitrogen), 1 µL de DTT 100 mM, 1 µL de RNase OUT (40 U/µL, Invitrogen), 1 µL de SuperScript IV reverse transcriptase (200 U/µL, Invitrogen). Le mix réactionnel est chauffé 10 min à 23°C puis 10 min à 50°C et enfin 10 min à 80°C. L'ADNc obtenu peut être utilisé pour des clonages ou directement pour une RT-qPCR.

3. Electrophorèse sur gel de polyacrylamide dénaturant

Les petits ARN sont fractionnés par électrophorèse sur gel de polyacrylamide en conditions dénaturantes. Un volume de **tampon de charge ARN 2X** est d'abord ajouté aux échantillons d'ARN avant d'être déposés sur un **gel de polyacrylamide 15 %**. La migration se fait à un voltage constant de 150 V dans du **tampon TBE 1X**. Pour visualiser les ARN sous UV après migration, le gel est incubé 10 min dans une solution de BET à 0,5 µg/L.

4. Northern blot

i. Transfert des ARN sur membrane

Les ARN séparés sur gel de polyacrylamide sont transférés, par électrophorèse, sur une membrane de nylon Hybond-N+ (Amersham Pharmacia Biotech). Le gel est positionné sur la membrane et le tout est entouré de deux couches de papier Whatman et d'une éponge de part et d'autre. Le transfert s'effectue dans du **tampon TBE 0,5X** durant 1h15 à 300 mA à 4°C. Après transfert, la membrane est posée sur deux papiers Whatman imbibés de **tampon SSC 2X** durant 5 min puis les ARN sont fixés par cross-linking aux UVs dans un Stratalinker® (2 x 120 mJ/cm²).

ii. Hybridation avec une sonde radiomarquée

La membrane est placée dans un tube d'hybridation puis imbibée avec au minimum 4,5 mL de PerfectHyb™ Plus (Sigma). La sonde oligonucléotidique radiomarquée (cf § II.8) est ensuite ajoutée dans le tube. L'hybridation s'effectue durant une nuit (environ 15 h) à 42°C. La membrane est ensuite lavée deux fois dans son tube avec du tampon SSC 2X pendant 10 min à 42°C. Un dernier lavage est réalisé avec du tampon SSC 2X, SDS 0,1 % (p/v) durant 30 min à 42 °C. La membrane est finalement séchée sur du papier Whatman puis est exposée contre une plaque de PhosphorImager (Fujifilm) ou contre un film autoradiographique.

5. Préparation des banques de petits ARNs pour sRNA-seq

Les banques de petits ARNs sont réalisées à partir d'ARN totaux extraits de feuilles de rosette. Ils sont quantifiés avec le kit Quant-iT Broad Range RNA Assay (Invitrogen) au fluoromètre Qubit (ThermoFisher) et leur qualité et intégrité sont vérifiées sur une puce Plant RNA Nano avec le Bioanalyzer (Agilent).

i. Traitement des ARN à la phosphatase et à la polynucléotide kinase (PNK)

Pour les traitements, le protocole « *Directional mRNA seq sample preparation* » (Illumina) a été suivi avec quelques modifications. L'étape de fragmentation des ARN n'a pas été réalisée. Pour chaque échantillon, les traitements ont été réalisés en triplicats techniques. Pour le traitement à la phosphatase, dans un tube sont mélangés 1 µg d'ARN dans 16 µL d'eau, 2 µL de tampon phosphatase 10X (NEB), 1 µL d'Antartic phosphatase (5 U/µL, NEB) et 1 µL de RNase OUT (40 U/µL, Invitrogen). L'ensemble est incubé 30 min à 37°C puis 5 min à 65°C. Pour le traitement à la PNK, dans ce même tube sont rajoutés 17 µL d'eau, 5 µL de tampon PNK 10X (NEB), 5 µL d'ATP 10 mM, 1 µL de RNase OUT (40 U/µL, Invitrogen) et 2 µL de PNK (10 U/µL, NEB). L'ensemble est incubé 1 h à 37°C. Les ARN sont ensuite mis à précipiter en ajoutant au tube 130 µL d'eau, 18 µL d'acétate de sodium 3 M, 4 µL de glycogène (à 5 mg/µL) et 600 µL d'éthanol 100 % et laissés une nuit à -80°C. Après centrifugation 30 min à 16000 g à 4°C et un lavage avec 200 µL d'éthanol 70 %, le culot d'ARN est repris dans 5 µL d'eau et les trois réplicats techniques sont rassemblés.

ii. Ligation des adaptateurs pour le séquençage

Le kit NEBNext Multiplex Small RNA library Prep set for Illumina a été utilisé pour préparer les banques. Ainsi, 5 µL d'ARN traité ont été utilisés pour la ligation des adaptateurs en 5' et 3'. Dans un volume final de 7 µL, 5 µL d'ARN ont été mélangés à 1 µL d'adaptateur 3' (NEB). Après une incubation de 2 min à 70 °C, 10 µL de tampon de ligation 3' 2X (NEB) et 3 µL d'enzyme de ligation 3' (NEB) ont été ajoutés. Après 1 h à 25°C, 4,5 µL d'eau et 1 µL de SR RT Primer (NEB) sont ajoutés et l'ensemble est incubé 5 min à 75°C, puis 15 min à 37°C et enfin 15 min à 25°C. Ensuite, le deuxième adaptateur est ligué en 5' en ajoutant au mix précédent 1 µL d'adaptateur 5' préalablement préparé (NEB), 1 µL de tampon de ligation 5' 10X (NEB) et 2,5 µL d'enzyme de ligation 5' (NEB). La ligation a lieu durant 1 h à 25°C.

iii. Retrotranscription et amplification par PCR

Les ARN possédant les adaptateurs en 5' et 3' sont ensuite rétrotranscrits. Pour cela, 30 µL d'ARN avec adaptateur sont mélangés à 8 µL de tampon de réaction pour la synthèse du premier brin (NEB), 1 µL d'inhibiteur de RNases (NEB) et 1 µL de reverse transcriptase ProtoScript II (NEB). La retrotranscription a lieu durant 1 h à 50°C. Ensuite sont ajoutés 50 µL du mix LongAmp Taq 2X Master Mix (NEB), 2,5 µL de primers (NEB), 2,5 µL d'amorces pour Index (NEB) et 5 µL d'eau. La réaction de PCR commence par une étape de dénaturation 30 sec à 95°C, puis de 15 cycles d'amplification se décomposant en 3 étapes.

Une première étape de dénaturation 15 sec à 94°C, une seconde étape d'hybridation 30 sec à 62°C et une troisième étape d'élongation 15 sec à 70°C. La réaction se termine par une étape d'élongation 5 min à 70°C. Les produits de PCR sont purifiés avec le kit PCR Clean-up (Macherey & Nagel) en suivant le protocole du fournisseur et vérifiés sur une puce DNA High Sensitivity avec le Bioanalyzer (Agilent). Finalement, après quantification, 4 cycles de PCR supplémentaire ont été réalisés pour obtenir suffisamment de matériel.

iv. Découpe sur gel et purification des fragments

L'ADNc est ensuite fractionné par électrophorèse sur gel de polyacrylamide 6%, TBE (Novex, Thermofisher). Après coloration du gel au SYBR Gold (Invitrogen), les fragments correspondants aux tailles entre 120 et 190 pb sont découpés et répartis en plusieurs fractions, puis ils sont purifiés avec le kit QIAquick Gel Extraction (Qiagen). L'ADN est ensuite précipité avec 1/10^e du volume d'acétate de sodium 1M, 0,5 µL de GlycoBlue (Thermofisher) et 2,5 volumes d'éthanol 100 % durant 3 h à -20°C. Après centrifugation 30 min à 16000 g à 4°C et un lavage avec de l'éthanol 70 %, le culot d'ADN est repris dans 10 µL d'eau. La quantité et qualité de l'ADN sont vérifiés sur une puce DNA High Sensitivity avec le Bioanalyzer (Agilent). Les fractions conservées pour le séquençage sont mélangées ensemble avec 4 µL de chaque fraction.

v. Séquençage Illumina MiSeq

Le séquençage a été réalisé par Sandrine Koechler de la plateforme Analyse de l'Expression Génique (AEG) de l'IBMP. Les échantillons ont été normalisés à 4 nM puis dénaturés avec du NaOH à 1 nM pour être finalement dilué à une concentration de 10 pM. Le contrôle PhiX (Illumina) est ajouté pour être à 0,5 % (v/v échantillon). Les kits utilisés pour le séquençage sont le kit MiSeq V3 Reagent kit 150 cycles PE box 1 (Illumina) et le kit MiSeq V3 Reagent kit box 2. Le séquençage a été réalisé selon les paramètres suivants : Run SE 1X61 cycles, index bases read 6.

6. Détermination de la présence de structures type G-quadruplex

Pour déterminer la présence de structure G-quadruplex, une solution de N-Methyl Mesoporphyrin IX (NMM, LiveChem) est utilisée. Lorsqu'il est excité à 399 nm, ce composé va émettre une fluorescence à 610 nm uniquement lorsqu'il s'est intercalé dans des structures G-quadruplex.

i. Électrophorèse sur gel de polyacrylamide non-dénaturant

Les ARN sont d'abord mis dans une solution contenant des sels pour favoriser et stabiliser les structures G-quadruplex. Ici, les ARN utilisés sont des oligoribonucléotides biotinylés, les mêmes que ceux utilisés pour la transfection de protoplastes (Annexe V). Un mélange composé de 1 µL d'oligonucléotide 100 µM, 5 µL de KCl 0,2 M et de 4 µL d'eau est incubé 10 min à 95°C puis laissé refroidir 1 h 30 min à température ambiante (sur la paillasse).

Ensuite, 4 μL de NMM 250 μM sont mélangés au mélange précédent et laissé 30 min à température ambiante, à l'abri de la lumière. Les échantillons sont alourdis pour être déposés sur gel avec 7 μL de glycérol 50 %. Les échantillons sont ensuite déposés sur un **gel de polyacrylamide 15 % non-dénaturant** préparé le jour même. La migration a lieu durant 1 h à 80 V constants puis à 110 V constants durant encore 1 h. Le NMM contenu dans les échantillons d'ARN est visualisé directement après migration à la longueur d'onde de 530 nm avec l'appareil Fusion FX (Vilber). Les bandes fluorescentes correspondent à l'ARN formant des G-quadruplex uniquement. Pour ensuite visualiser l'ensemble des ARN sous UV, le gel est incubé 10 min dans une solution de BET à 0,5 $\mu\text{g/L}$.

ii. Mesure de la fluorescence – NMM

L'émission de fluorescence du NMM peut aussi être quantifié directement et simultanément pour plusieurs oligonucléotides. Une solution à base de NMM est réalisée. Pour cela, sont mélangés 2 μL de **tampon sodium phosphate 1 M**, 10 μL de KCl 2 M, 0,8 μL de MgCl_2 1 M, 2 μL de **NMM 500 μM** et 185,2 μL d'eau. Dans un plaque 96 puits, 27 μL de la solution précédente sont déposés dans chacun des puits et 3 μL de chaque oligoribonucléotide 100 μM sont déposés par puit. L'intensité de fluorescence émise par le NMM est mesurée à 590 nm, pour une excitation à 355 nm, avec le lecteur de plaque FluoStar (Omega). Des puits sans oligonucléotides servent pour évaluer le bruit de fond.

IV. Méthodes relatives aux protéines

1. Extraction de protéines

L'extraction des protéines est faite à partir de la phase organique obtenue lors de l'extraction des ARN totaux au TRI-Reagent® (voir III.1.i). Après centrifugation 10 min à 12000g à 4°C, le surnageant est prélevé pour en extraire les ARN. Puis, 0,3 V d'éthanol 100 % est ajouté à la phase organique restante. L'ensemble est vortexé puis incubé 15 min à température ambiante. Après centrifugation 5 min à 16000 g à 4°C, le surnageant est prélevé puis mélangé à 3 V d'acétone 100 %. Après incubation 15 min à -20°C, le tube est centrifugé 2 min à 1500 g à 4 °C. Le surnageant est éliminé et le culot de protéines est lavé avec 1 mL d'acétone 80 %. Le culot est finalement repris dans 50 μL de **tampon d'extraction de protéines**.

La concentration en protéine de chaque échantillon est estimée en mesurant l'absorbance à 595 nm par la méthode de Bradford (Biorad® Protein Assay).

2. Électrophorèse sur gel de polyacrylamide dénaturant

Les protéines sont séparées par électrophorèse en conditions dénaturantes sur un gel de polyacrylamide 15 % comprenant un **gel de concentration** et un **gel de séparation**. La polymérisation du gel s'effectue par addition de persulfate d'ammonium 0,1 % (p/v) et de TEMED 0,01 % (v/v).

Un volume de **Laemmli 5X** est ajouté aux extraits protéiques qui sont ensuite chauffés 5 min à 95°C avant d'être déposé sur gel. La migration se fait à un ampérage constant de 30 mA dans du **tampon SDS-PAGE**.

3. Western blot

i. Transfert des protéines sur membrane

Les protéines séparées sur gel de polyacrylamide dénaturant sont transférées, par électrophorèse, sur une membrane Immobilon® -P 0,45 µm (Millipore) préalablement activée dans de l'éthanol 100 %. Le gel est positionné sur la membrane et le tout est entouré de deux couches de papier Whatman et d'une éponge de part et d'autre. Le transfert s'effectue dans du **tampon de transfert** durant 1h à 80 V.

ii. Coloration des protéines

Les protéines transférées sont visualisées par incubation sous agitation de la membrane dans une **solution de coloration**. Lorsque la membrane devient bleue, plusieurs bains successifs de **solution de décoloration** sont réalisés pour faire apparaître les protéines. Après avoir scanné la membrane, celle-ci est entièrement décolorée dans un bain de méthanol 100 %.

iii. Immunodétection

Pour saturer les sites de fixation aspécifiques, la membrane est incubée au moins 1h à température ambiante (ou une nuit à 4°C) sous agitation dans du **tampon bloquant**. Elle est ensuite incubée 1h à température ambiante ou durant la nuit à 4°C avec l'anticorps primaire spécifique de la protéine à révéler dilué du tampon bloquant (dilutions données dans l'Annexe XI). Après 3 lavages de 10 min dans du tampon bloquant, la membrane est incubée 30 min avec l'anticorps secondaire dans du tampon bloquant (dilué au 1/1000^{ième}). La membrane est lavée 10 min dans du tampon bloquant suivi de deux lavages de 10 min chacun dans du **tampon de lavage membrane**. Les protéines d'intérêt sont finalement révélées à l'aide du kit ECL Clarity ou Clarity Max (Bio-Rad) et visualisées avec l'appareil Fusion FX (Vilber) ou par exposition contre un film autoradiographique.

4. Purification de partenaires protéiques de tRFs

i. Lyse des protoplastes transfectés et chromatographie d'affinité à la streptavidine

Après avoir éliminé le surnageant du dernier lavage (voir I.2.vi), 500 µL de **tampon de lyse** sont ajoutés aux protoplastes transfectés. Le tube est mis en agitation sur un vortex à faible vitesse durant 10 min à 4°C. Pendant ce temps, les colonnes pour la chromatographie sont préparées. Ainsi, 20 µL de résine Streptavidin Sephrose™ (GE Healthcare) sont placés dans des colonnes (Macherey & Nagel, *Receiver Column* avec filtre de 20 µm) puis 500 µL de **tampon de lavage 1X** sont ajoutés. Après avoir centrifugé les colonnes 15 sec à 11000 g, le liquide collecté dans des tubes collecteurs est éliminé. Un deuxième lavage est ensuite réalisé.

Après centrifugation des tubes contenant les protoplastes, 30 sec à 100 g à 4°C, 500 µL de surnageant sont transférés sur les colonnes de streptavidine. Après avoir mis les colonnes en agitation 20 min sur roue à 5 rpm à 4°C, elles sont centrifugées 30 sec à 100 g à 4°C et le liquide collecté est éliminé. Les colonnes sont ensuite lavées avec 500 µL de **tampon SSC 0,1X** puis centrifugées 15 s à 100 g à 4°C et le liquide est éliminé. Au total, cinq lavages comme décrit ci-avant sont réalisées.

ii. Identification des protéines par spectrométrie de masse (nanoLC-MS/MS)

Dans chaque colonne, 50 µL de **tampon d'éluion** sont ajoutés pour reprendre la résine de streptavidine et la transférer dans un nouveau tube. Cette étape est répétée une deuxième fois. Les 100 µL de tampon d'éluion collectés sont chauffés 5 min à 95°C. Après centrifugation 15 sec à 100 g, le surnageant est transféré dans un nouveau tube et congelé dans de l'azote liquide. Il est conservé à -80°C avant d'être envoyé à la plateforme protéomique de Strasbourg - Esplanade.

iii. Analyse des protéines identifiées par spectrométrie de masse

Après une digestion trypsique, les peptides obtenus sont séparés par une nano-chromatographie liquide (nanoLC) avant d'être injectés dans le spectromètre (MS/MS). Les analyses de spectrométrie de masse ont été réalisées à la plateforme de protéomique de Strasbourg-Esplanade. Après digestion à la trypsine, les protéines sont analysées par nanoLC-MS/MS sur un spectromètre QExactive+ (ThermoFisher). Les données obtenues sont ensuite cherchées sur la base de données TAIR10. Les peptides sont identifiés avec l'algorithme Mascot et les résultats sont importés dans le logiciel Proline. L'identification des protéines a été validée avec un FDR < 1 % (False discovery rate). Les protéines ont été quantifiées sur la base du nombre de spectres MS2 qui leur sont attribués après l'étape d'identification. Les données ont ensuite été traitées avec le package R IPInquiry4 mis à disposition par Hélène Zuber sur GitHub (hzuber67/IPInquiry4, [169]). Ce package permet d'identifier les protéines significativement enrichies dans les échantillons, de récupérer les annotations correspondantes aux protéines identifiées, et d'exporter les tableaux de résultats. L'analyse statistique différentielle est basée sur le modèle Genewise Negative Binomial Generalized Linear Model développé par le package edgeR [170].

ANNEXES

Annexe I. Nomenclature des tRFs utilisés

Megel et al. [56]	tDRnamer	ARNt isodécodateur	Nombre d'ARNt isodécodateur	Anticodon	Position	Taille (nt)
tRF-5D Ala	tDR-1:20-Ala-AGC-1-M2	Ala	2	AGC	1 à 20	20
tRF-5D Arg	tDR-1:20-Arg-CCT-2	Arg	1	CCT	1 à 20	19
tRF-5D His	tDR-1:20-His-GTG	<i>ARNt d'origine chloroplastique, non pris en charge par tDRnamer</i>				
tRF-5A Asp	tDR-1:33-Asp-GTC-1-M3	Asp	3	GTC	1 à 33	33
tRF-5D Gly	tDR-1:19-Gly-TCC-1	Gly	1	TCC	1 à 19	18

Annexe II. Composition des solutions, tampons et milieux

Méthodes relatives aux organismes :

Milieu LB : extrait de levure 5 g/L ; bactotryptone 10 g/L ; NaCl 5 g/L ; pH 7,0.

Milieu LB-agar : milieu LB ; agar 15 g/L.

Tampon TB : PIPES 10 mM ; MnCl₂ 55 mM ; CaCl₂ 15 mM ; KCl 250 mM.

Solution de stérilisation : Javel 4 % ; Tween-20 4 % ou Tween-20 0,1 % ; éthanol 70%.

Milieu gélosé MS 222 : Murashige & Skoog 222 (MS) 2,21 g/L ; sucrose 10 g/L ; MES 0,5 g/L ; agar 0,68 % (p/v). Le pH est ajusté à 5,7 avec du KOH 1N.

Milieu gélosé +Pi : MSP01 (CaissonLabs) 2,16 g/L ; sucrose 10 g/L ; MES 0,5 g/L ; agar 0,68 % (p/v). Le pH est ajusté à 5,7 avec du KOH 1N.

Milieu gélosé -Pi : MSP11 (CaissonLabs) 2,08 g/L ; sucrose 10 g/L ; MES 0,5 g/L ; agar 0,68 % (p/v). Le pH est ajusté à 5,7 avec du KOH 1N.

Milieu liquide MS 231 : Murashige & Skoog 231 (MS) 2,21 g/L ; sucrose 25 g/L ; MES 0,5 g/L. Le pH est ajusté à 5,8 avec du KOH 1N.

Milieu liquide +Pi : MSP01 (CaissonLabs) 2,16 g/L ; sucrose 10 g/L ; MES 0,5 g/L. Le pH est ajusté à 5,7 avec du KOH 1N.

Milieu liquide -Pi : MSP11 (CaissonLabs) 2,08 g/L ; sucrose 10 g/L ; MES 0,5 g/L. Le pH est ajusté à 5,7 avec du KOH 1N.

Milieu de culture cellulaire : MS 256 4,41 g/L ; sucrose 30 g/L ; acide α -naphtaleneacetic (NAA) 5,4 μ M ; kinétine 0,23 μ M. Le pH est ajusté à 5,6 avec du KOH 1N.

Tampon d'agro-transformation : MS 222 2,2 g/L ; sucrose 5 g/L ; acétoseringone 200 μ M ; Silwett 77 0,05 % (v/v).

Méthodes relatives à l'ADN :

Tampon d'extraction d'ADNg : Tris-HCl 200 mM, pH 7,5 ; NaCl 250 mM ; EDTA 25 mM, pH 8,0.

Tampon T : Tris-HCl 50 mM, pH 8,5.

Tampon de ligation 10X : Tris-HCl 300 mM, pH 7,8 ; $MgCl_2$ 100 mM ; DTT 100 mM ; ATP 10 mM.

TAE 0,5X : Tris-Acétate 20 mM, pH 8 ; EDTA 0,5 mM.

Tampon de charge ADN 6X (ThermoFisher Scientific) : Tris-HCl 10 mM, pH 7,6 ; bleu de bromophenol 0,03 % (p/v) ; glycérol 60 % (v/v) ; EDTA 60 mM.

Tampon PNK 10X : Tris-HCl 500 mM, pH 7,6 ; $MgCl_2$ 100 mM ; DTT 50 mM ; spermidine 1 mM.

Méthodes relatives à l'ARN :

Tampon de précipitation 1X : LiCl 5 M ; Tris-HCl 50 mM, pH 7,5.

Tampon de charge ARN 2X : formamide 95 % (v/v) ; EDTA 20 mM, pH 8,0 ; bleu de bromophénol 0,05 % (p/v) ; xylène cyanol 0,05 % (p/v).

Gel de polyacrylamide 15 % : acrylamide/bisacrylamide (19/1) 15 % (v/v) ; urée 7M ; TBE 1X ; persulfate d'ammonium 0,07 % (p/v) ; TEMED 0,035 % (p/v).

Tampon TBE 1X : Tris 90 mM ; acide borique 90 mM ; EDTA 2 mM, pH 8,0.

Tampon TBE 0,5X : Tris 45 mM ; acide borique 45 mM ; EDTA 1 mM, pH 8,0.

Tampon SCC 2X : NaCl 0,3 M ; citrate de sodium 0,2 M, pH 4,8.

Gel de polyacrylamide 15 % non-dénaturant : acrylamide/bisacrylamide (19/1) 15 % (v/v) ; TBE 1X ; persulfate d'ammonium 0,07 % (p/v) ; TEMED 0,035 % (p/v).

Tampon sodium phosphate 1 M : NaH_2PO_4 0,735 M ; Na_2HPO_4 0,265 M ; pH 6,4.

NMM 500 μ M : 0,25 mg/mL N-Methyl Mesoporphyrin IX.

Méthodes relatives aux protéines :

Tampon d'extraction de protéines : urée 8 M ; Tris-HCl 50 mM, pH 6,8 ; EDTA 11 mM ; β -mercaptoéthanol 5 %.

Gel de concentration : acrylamide 5 % ; Tris-HCl 100 mM, pH 6,8 ; SDS 0,1 % (p/v).

Gel de séparation : acrylamide 15 % ; Tris-HCl 100 mM, pH 8,8 ; SDS 0,1 % (p/v).

Laemmli 5X : Tris-HCl 250 mM, pH 6,8 ; glycerol 37,5 % (v/v) ; SDS 10 % (p/v) ; β -mercaptoéthanol 10 % (v/v) ; Bleu de Bromophénol 0,125 % (p/v).

Tampon SDS-PAGE : Tris 25 mM ; glycine 250 mM ; SDS 0,1 % (p/v).

Tampon de transfert : Tris 50 mM ; Borate 50 mM.

Solution de coloration : Bleu de Coomassie R250 0,1 % (p/v) ; méthanol 50 % (v/v) ; acide acétique 7 % (v/v).

Solution de décoloration : méthanol 50 % (v/v) ; acide acétique 7 % (v/v).

Tampon bloquant : Tris-HCl 25 mM, pH 7,6 ; NaCl 150 mM ; Tween-20 0,2 % (v/v) ; lait écrémé 5 % (p/v).

Tampon de lavage membrane : Tris-HCl 25 mM, pH 7,6 ; NaCl 150 mM ; Tween-20 0,2 % (v/v).

Tampon de lyse : Tris-HCl 50 mM, pH 7,5 ; NaCl 150 mM ; $MgCl_2$ 2 mM ; NP40 1 %, cocktail inhibiteur de protéases (cOmplete™, Roche).

Tampon de lavage 1X : Tris-HCl 5 mM, pH 7,5 ; NaCl 1 M ; EDTA 1 mM, pH 8,2.

Tampon SSC 0,1X : NaCl 15 mM ; citrate de sodium 1,5 mM, pH 7,0.

Tampon d'élution (Miltenyi Biotec) : Tris-HCl 50 mM, pH 6,8 ; DTT 50 mM ; SDS 1% ; EDTA 1 mM ; bleu de bromophenol 0,005% (p/v) ; glycérol 10 % (v/v).

Annexe III. Antibiotiques utilisés pour la sélection des bactéries

Bactérie	Souche	Transformée avec	Antibiotiques	Concentration finale
<i>E. coli</i>	TOP10	pLitmus 28i	Ampicilline	100 µg/mL
<i>E. coli</i>	TOP10	pKI1.1R	Spectinomycine	100 µg/mL
<i>E. coli</i>	TOP10	pDONR207	Gentamycine	25 µg/mL
<i>E. coli</i>	TOP10	pDONR207	Chloremphenicol	30 µg/mL
<i>E. coli</i>	TOP10	pB7FWG2 ou pK7WGF2	Spectinomycine	100 µg/mL
<i>E. coli</i>	TOP10	pALiCE	Ampicilline	100 µg/mL
<i>A. tumefaciens</i>	GV3101::pMP90	pKI1.1R	Gentamycine	15 µg/mL
			Rifampicine	50 µg/mL
			Spectinomycine	100 µg/mL
<i>A. tumefaciens</i>	GV3101::pMP90	pB7FWG2 ou pK7WGF2	Gentamycine	15 µg/mL
			Rifampicine	50 µg/mL
			Spectinomycine	100 µg/mL

Annexe IV. Article “Transfection of small noncoding RNAs into Arabidopsis thaliana protoplasts”



Transfection of Small Noncoding RNAs into *Arabidopsis thaliana* Protoplasts

Stéphanie Lalande, Marjorie Chery, Elodie Ubrig, Guillaume Hummel, Julie Kubina, Angèle Geldreich, and Laurence Drouard

Abstract

Polyethylene glycol transfection of plant protoplasts represents an efficient method to incorporate foreign DNA and study transient gene expression. Here, we describe an optimized protocol to deliver small noncoding RNAs into *Arabidopsis thaliana* protoplasts. An example of application is provided by demonstrating the incorporation of a 20 nt long small noncoding RNA deriving from the 5' extremity of an *A. thaliana* cytosolic alanine tRNA into freshly isolated protoplasts.

Key words Transfection, Small noncoding RNAs, tRNA-derived fragments, Protoplasts, Northern blot

1 Introduction

Among the various classes of small noncoding RNAs (sncRNAs), tRNA-derived fragments (tRFs) are now recognized as important regulators of numerous biological processes [1–4]. In plants, tRFs represent a large and dynamic repertoire of small RNA fragments [5] but their molecular functions are still poorly understood [2]. Among the different ways to decipher the molecular processes in which tRFs are involved, the use of *in vivo* approaches is rather limited. Indeed, the production of specific tRFs in transgenic plants remains difficult to achieve. Facing this bottleneck, cell transfection represents an interesting option. This technique, mainly developed in the animal field to study the biology of miRNAs or siRNAs, is based on the use of lipofectamine, a reagent that facilitates efficient delivery of small RNA molecules into cells (see for instance [6, 7]). In plants, to deliver nucleic acids, researchers have chosen to work with plant protoplasts (i.e., cells without their cell wall). Nevertheless, chemicals such as lipofectamine appeared toxic and inefficient [8] and thus are rarely utilized. Rather, three other types of

methods have been developed to transfect macromolecules: electroporation (e.g., [9]), microinjection (e.g., [10]), and polyethylene glycol (PEG)-mediated transfection (e.g., [11], the latter method being the most widely used). Such protocols are often employed to deliver DNA into protoplasts for transient gene expression studies, but the incorporation of RNAs into plant protoplasts has also been developed. For instance, viral RNAs, siRNAs, and tRNAs were electroporated into protoplasts [12, 13] and a PEG method allowed viral RNA to be delivered [14]. Here, we have adapted and optimized an efficient DNA transfection method described in details by Yoo et al. [11] to incorporate sncRNAs into *Arabidopsis thaliana* protoplasts. As an example of application, we describe here the transfection of an *A. thaliana* 20 nt long tRF alanine (tRF Ala20) [5] into protoplasts isolated either from an *A. thaliana* cell culture or from young seedlings. To validate the PEG transfection protocol, incorporation of tRF Ala20 was verified by northern blot analysis of the small RNA population extracted from RNase-treated transfected protoplasts. Therefore, we propose this efficient RNA transfection method as a new tool to study the biology of sncRNAs. For example, tagged (e.g., fluorophores, biotin) sncRNAs could be applied for in vivo live imaging or interactome studies.

2 Materials

2.1 Plant Material

1. Culture of individualized cells obtained from the Landsberg *A. thaliana* ecotype.
2. *A. thaliana* Columbia (Col-0) ecotype seeds.

2.2 Nucleic Acids

1. 100 μM Synthetic oligoribonucleotide of interest dissolved in diethyl pyrocarbonate (DEPC)-treated water at a 100 μM concentration.
2. Any plasmid of small size (i.e., 3–4 kb).
3. 100 μM Oligodeoxynucleotide with a sequence complementary to the sequence of the small RNA of interest.

2.3 Solutions

All solutions should be prepared with RNase-free materials and reagents, and with RNase-free, sterile, and ultrapure water. Solutions are autoclaved for 15 min at 120 °C (except otherwise stated). Procedures for preparing classical solutions can be found in Sambrook and Russell [15].

1. Cell culture medium: 4.41 g/L MS256, 30 g/L sucrose, 5.4 μM α -naphthaleneacetic acid (NAA), 0.23 μM kinetin, pH 5.6. Weigh 4.41 g of MS256 and 30 g of sucrose. Transfer to a glass beaker. Add about 500 mL of water. Mix well and adjust pH to 5.6 with 1 N KOH. Add 100 μL of 10 mg/mL

NAA and 50 μL of 1 mg/mL kinetin. Transfer to a cylinder and adjust volume to 1 L with water. Autoclave and store at room temperature.

2. Seed sterilization solution: 0.1% (v/v) Tween-20, 70% (v/v) ethanol. Under a sterile hood, filter solution with a 0.20 μm filter in a sterile bottle. Store at room temperature.
3. Seedling culture medium: 2.21 g/L MS222, 0.5 g/L 2-(*N*-morpholino)ethanesulfonic acid (MES), 10 g/L sucrose, 0.68% agar, pH 5.7. Weigh 2.21 g of MS222, 0.5 g of MES monohydrate, and 10 g of sucrose. Transfer to a glass beaker. Add about 500 mL of water. Mix well and adjust pH to 5.7 with 1 N KOH. Transfer to a cylinder and adjust volume to 1 L with water. Transfer to a bottle and add 6.8 g of agar. Autoclave and store at room temperature.
4. 2 M KCl: Store at room temperature.
5. 1 M MgCl_2 : Store at room temperature.
6. 1 M CaCl_2 : Store at room temperature.
7. 5 M NaCl: Store at room temperature.
8. 0.5 M Mannitol, pH 5.7: Weigh 227.75 g of D-mannitol and 1.75 g of MES monohydrate. Transfer to a glass beaker. Add about 1.5 L of water. Mix and adjust pH to 5.7 with 10 N KOH. Transfer to a cylinder and adjust volume to 2.5 L with water. Autoclave and store at 4 °C.
9. 0.5 M MES, pH 5.7: Weigh 53.3 g of MES monohydrate and transfer to a glass beaker. Add about 400 mL of water. Mix and adjust pH to 5.7 with 10 N KOH. Transfer to a glass cylinder and adjust volume to 500 mL with water. Autoclave and store at room temperature.
10. MM buffer: 0.5 M Mannitol, 4 mM MES, pH 5.7. Pipette 8 mL of 0.5 M MES, pH 5.7, and transfer to a 1 L glass cylinder. Add 0.5 M mannitol, pH 5.7, to a final volume of 1 L and mix well. Autoclave and store at 4 °C.
11. MMG buffer: 0.5 M Mannitol, 4 mM MES, 15 mM MgCl_2 , pH 5.7. Pipette 7.5 mL of 1 M MgCl_2 and transfer to a 1 L glass cylinder. Add MM buffer to a final volume of 500 mL and mix well. Autoclave and store at 4 °C.
12. WI buffer: 0.5 M Mannitol, 4 mM MES, 20 mM KCl, pH 5.7. Pipette 5 mL of 2 M KCl and transfer to a 1 L glass cylinder. Add MM buffer to a final volume of 500 mL and mix well. Autoclave and store at 4 °C.
13. W5 buffer: 2 mM MES, 154 mM NaCl, 125 mM CaCl_2 , 5 mM KCl, pH 5.7. Mix in a glass cylinder 62.5 mL of 1 M CaCl_2 , 15.4 mL of 5 M NaCl, 2 mL of 0.5 M MES pH 5.7, and 1.25 mL of 2 M KCl. Add water to a final volume of 500 mL and mix well. Autoclave and store at 4 °C.

14. TriReagent™ (Sigma-Aldrich) or TRIzol™ Reagent (Ambion): Store at 4 °C.
15. Chloroform: Store at room temperature.
16. Isopropanol: Store at room temperature.
17. 100% Ethanol: Store at room temperature.
18. “Cold” ethanol: 50 mL 100% Ethanol cooled to –20 °C .
19. 95% Ethanol: Mix 47.5 mL of 100% ethanol with 2.5 mL of water. Mix well and store at –20 °C.
20. Diethyl pyrocarbonate (DEPC)-treated water: DEPC is toxic; prepare the solution under a chemical hood. In a bottle, add 0.001% volume of DEPC to 100 mL water. DEPC has limited solubility in water as evidenced by the appearance of globules; thus mix well until globules disappear. Incubate for 16 h, and autoclave. Make aliquots of 2–10 mL, and store at –20 °C.
21. 1 M Tris–HCl, pH 7.5: Weigh 12.11 g of Tris and transfer to a glass beaker. Add about 70 mL of water, mix, and adjust pH to 7.5 with 1 N HCl (*see Note 1*). Transfer to a glass cylinder and adjust volume to 100 mL with water. Autoclave and store at room temperature.
22. 5 M LiCl solution: 5 M LiCl, 50 mM Tris–HCl, pH 7.5. Weigh 21.2 g of LiCl and transfer to a glass beaker. Add 5 mL of 1 M Tris–HCl, pH 7.5, and water to a volume of 90 mL. Mix thoroughly until LiCl is completely dissolved. Transfer to a glass cylinder and adjust volume to 100 mL with water. Autoclave and store at room temperature.
23. 1 M Sodium acetate, pH 4.5: Weigh 8.2 g of sodium acetate and transfer to a glass beaker. Add water to a volume of 70 mL and mix until sodium acetate is completely dissolved. Adjust pH to 4.5 with 99% pure acetic acid. Transfer to a glass cylinder and adjust volume with water to 100 mL. Autoclave and store at room temperature.
24. 0.5 M EDTA, pH 8.0: Weigh 29.2 g of ethylenediaminetetraacetic acid (EDTA) and transfer to a glass beaker. Add 150 mL of water. Mix well and add NaOH pellets until EDTA is completely dissolved (*see Note 2*). Adjust pH to 8.0 with 1 N NaOH. Transfer to a glass cylinder and adjust volume to 200 mL with water. Autoclave and store at room temperature.
25. 2× Loading buffer: 95% (v/v) Formamide, 20 mM EDTA, pH 8.0, 0.05% (w/v), bromophenol blue, 0.05% (w/v) xylene cyanol. Weigh 0.025 g of bromophenol blue and 0.025 g of xylene cyanol and transfer to a 50 mL tube. Add 2 mL of 0.5 M EDTA, pH 8.0, and 0.5 mL of water. Mix by vortexing and add 47.5 mL of formamide. Mix well and store at room temperature.

26. 10× TBE: 900 mM Tris, 20 mM EDTA, 900 mM boric acid. Weigh 109 g of Tris and 55.6 g of boric acid and transfer to a glass beaker. Add 40 mL of 0.5 M EDTA, pH 8.0, and water to a volume of about 900 mL, and mix. Transfer to a glass cylinder and adjust volume to 1 L with water. Autoclave and store at room temperature (*see Note 3*).
27. 1× TBE: 90 mM Tris, 2 mM EDTA, 90 mM boric acid. Mix 100 mL of 10× TBE with 900 mL of water in a glass cylinder. Mix well and store at room temperature.
28. 0.5× TBE: 45 mM Tris, 1 mM EDTA, 45 mM boric acid. Mix 50 mL of 10× TBE with 950 mL of water in a glass cylinder. Mix well and store at room temperature.
29. 15% Denaturing acrylamide gel: 15% Acrylamide/bis-acrylamide 19/1, 7 M urea, 1× TBE. Weigh 84 g of urea and transfer to a glass beaker. Add 75 mL of 40% acrylamide/bis-acrylamide (19:1) and 20 mL of 10× TBE. Add water to a volume of 190 mL and mix thoroughly until the urea is completely dissolved. Transfer to a cylinder and adjust volume to 200 mL with water. Mix well and store at room temperature in a bottle wrapped in aluminum foil to avoid light exposure.
30. 10% (w/v) Ammonium persulfate (APS): Store at 4 °C.
31. *N,N,N,N'*-tetramethylethylenediamine (TEMED): Store at 4 °C.
32. 1% Ethidium bromide: Store at 4 °C.
33. G-50 resin: Pour about 1 mL of Sephadex G-50 resin powder in a 14 mL round-bottom tube. Add about 10 mL of water and mix well. Autoclave and store at 4 °C.
34. PerfectHyb™ Plus hybridization buffer (Sigma-Aldrich).
35. 20× Saline sodium citrate (SSC): 3 M NaCl, 300 mM trisodium citrate, pH 7.0. Weigh 175.3 g of NaCl and 88.2 g of sodium citrate dihydrate and transfer to a glass beaker. Add about 500 mL of water, and mix. Adjust pH to 7.0 with 99% pure acetic acid. Transfer to a glass cylinder and adjust volume to 1 L with water. Autoclave and store at room temperature.
36. 2× SSC: 0.3 M NaCl, 30 mM trisodium citrate, pH 7.0. Mix 50 mL of 20× SSC with 450 mL of water in a glass cylinder. Mix well and store at room temperature.
37. 10% (w/v) Sodium dodecyl sulfate (SDS) (*see Note 4*): Autoclave and store at room temperature.
38. 2× SSC, 0.1% SDS: Transfer 50 mL of 20× SSC into a 500 mL cylinder. Add 400 mL of water and then 5 mL of 10% SDS. Adjust volume to 500 mL by gently adding water to avoid foaming. Mix and store at room temperature (*see Note 5*).

**2.4 Enzymes
and Other Reagents**

1. Cellulase R10 (Yakult Pharmaceutical Industry Co., Ltd).
2. Macerozyme R10 (Yakult Pharmaceutical Industry Co., Ltd).
3. 10% Acetylated BSA.
4. PEG-4000.
5. RNase A/T1 mix: 2 mg/mL RNase A, 5000 units/mL RNase T1.
6. 20 mg/mL Glycogen.
7. 10 units/ μ L Polynucleotide kinase.
8. 10 \times Polynucleotide kinase buffer.
9. 10 μ Ci/ μ L (2500 Ci/mmol) Radiolabeled [γ ³²P]-adenosine triphosphate (ATP).

**2.5 Equipment
and Miscellaneous
Material**

1. Erlenmeyer of 250 mL.
2. Water bath.
3. Centrifuge allowing deceleration without brake, with rotors compatible for 2 mL, 14 mL, and 50 mL tubes.
4. 2 mL Round-bottom tubes.
5. 14 mL Round-bottom tubes.
6. 50 mL Tubes.
7. Wide-mouth pipettes (tips of at least a 2 mm diameter).
8. Micropipette tips cut at their extremity to enlarge the opening (2 mm diameter).
9. Incubator.
10. Petri dishes (100 mm diameter; 20 mm height).
11. Square Petri dish (120 mm \times 120 mm).
12. Scalpel blade.
13. 12-Well plate (3.65 cm²/well) for cell culture.
14. Micropore gas-permeable tape.
15. Orbital shaker.
16. Tube rotator.
17. Vacuum chamber.
18. Miracloth.
19. Inverted bright-field microscope.
20. Fuchs-Rosenthal cell.
21. Mini PROTEAN™ TetraHandcast System (Biorad).
22. Mini PROTEAN™ TetraCell System (Biorad).
23. 15-Well comb.
24. Power supply Power Pac basic (Biorad).
25. UV transilluminator.

26. Trans-Blot™ Plus cell (Biorad).
27. Hybond™ N+ nylon membrane.
28. Whatman 3 MM paper.
29. UV stratalinker.
30. 1 mL Syringe.
31. Sephadex G-50.
32. Fiberglass.
33. Hybridization tubes.
34. Hybridization oven.
35. Phosphorimager (with plate) or X-ray developing unit (and X-ray film).

3 Methods

The workflow includes the isolation of protoplasts from either cell cultures or seedlings of *A. thaliana*, transfection of tRFs, extraction of total sncRNA from transfected protoplasts, and a northern blot experiment to confirm the uptake of tRFs into the cells.

Protoplasts must be manipulated very carefully to avoid lysis. Always use wide-mouth pipettes with a large hole. Tips for micro-pipettes must be cut at their extremity to expand the opening (*see Note 6*). Never vortex the protoplasts and always centrifuge them without brake (*see Note 7*). When removing the supernatant after a centrifugation, always let a minimal volume of culture medium to cover protoplasts with liquid and avoid excessive changes of osmotic pressure. Do not directly pour the solution into the tube with protoplasts but rather always allow the solution to flow along the tube wall (*see Note 8*).

3.1 *A. thaliana* Cell Culture

3.1.1 Cell Culture Maintenance

1. Under a sterile hood, pour 100 mL of cell culture medium into a 250 mL Erlenmeyer.
2. Gently stir to homogenize the distribution of cells in a 7-day-old *A. thaliana* cell culture. Delicately transfer 4 mL of the culture into the new medium using a pipette.
3. Incubate for 7 days at 24 °C under constant light on an orbital shaker at 130 rpm.
4. Repeat **steps 1–3** weekly to maintain the cell culture.

3.1.2 Cell Culture for Protoplast Isolation

1. To 100 mL of cell culture medium, add 4 mL of a 7-day-old *A. thaliana* cell culture as described above.
2. Incubate for 10–11 days at 24 °C under constant light condition on an orbital shaker at 130 rpm.

3.2 A. thaliana**Seedlings**

All steps requiring opening of a tube need to be performed under a sterile hood.

**3.2.1 Arabidopsis Seed
Sterilization**

1. Prepare around 50 μL of *A. thaliana* Col-0 seeds in a 1.5 mL microtube (*see Note 9*).
2. Add 1 mL of seed sterilization solution. Mix to resuspend the seeds.
3. Gently mix for 15 min on a tube rotator.
4. Centrifuge at $3000 \times g$ for 10 s to pellet the seeds.
5. Carefully remove the supernatant.
6. Add 1 mL of 100% ethanol and mix to resuspend the seeds.
7. Gently mix on a tube rotator for 2 min.
8. Centrifuge at $3000 \times g$ for 10 s to pellet the seeds.
9. Remove as much supernatant as possible.
10. Repeat **steps 6–9** one more time.
11. Gently hit the tube with a finger to spread seeds, open the tube, and allow the seeds to dry for 1 h.

**3.2.2 Growing
Arabidopsis Seedlings**

1. Heat seedling growth medium in a microwave to melt it. Under a sterile hood, pour 50 mL of the medium into a square Petri dish (120 mm \times 120 mm) and let the dish open until the medium is solidified.
2. Scatter the sterilized seeds onto the solidified medium.
3. Seal the dish with micropore gas-permeable tape.
4. Allow the seeds to vernalize by keeping the dish at 4 °C for 2–5 days.
5. Let seedlings grow for 7 days under long day conditions (8-h dark at 17 °C, 16-h light at 21 °C).

**3.3 Isolation
of Protoplasts**

Digestion buffer must be prepared just before use. The following protocol is given for 50 mL digestion buffer.

**3.3.1 Preparation
of Digestion Buffer**

1. Preheat a water bath to 55 °C.
2. Weigh 0.75 g of cellulase and 0.2 g of macerozyme and transfer to a 50 mL tube.
3. Add 40 mL of 0.5 M mannitol, pH 5.7, 2.5 mL of 0.4 M MES, and 500 μL of 2 M KCl. Vortex and heat at 55 °C until enzymes are completely dissolved.
4. Add 500 μL of 1 M CaCl_2 and 500 μL of acetylated BSA. Adjust volume to 50 mL with water and gently mix.

**3.3.2 Removal of Cell
Walls from Cultured Cells**

25 mL of digestion buffer per 10 mL of cell culture is required.

1. Transfer 10 mL of a 10–11-day-old cell culture (Fig. 1a) into a 50 mL tube using a pipette.
2. Centrifuge the cells for 1 min at $100 \times g$ without brake. Remove most of the supernatant (keep a small volume of medium on the top) by pipetting and gently tapping the tubes to resuspend cells.
3. Wash cells by adding 30 mL of 0.5 M mannitol, pH 5.7. Centrifuge for 1 min at $100 \times g$ without brake. As above, discard supernatant and tap the tubes to resuspend cells.
4. Add 25 mL of digestion buffer and pour the mixture into a Petri dish. Cover with aluminum foil and incubate for 4 h at 26°C at 50 rpm (orbital shaker).

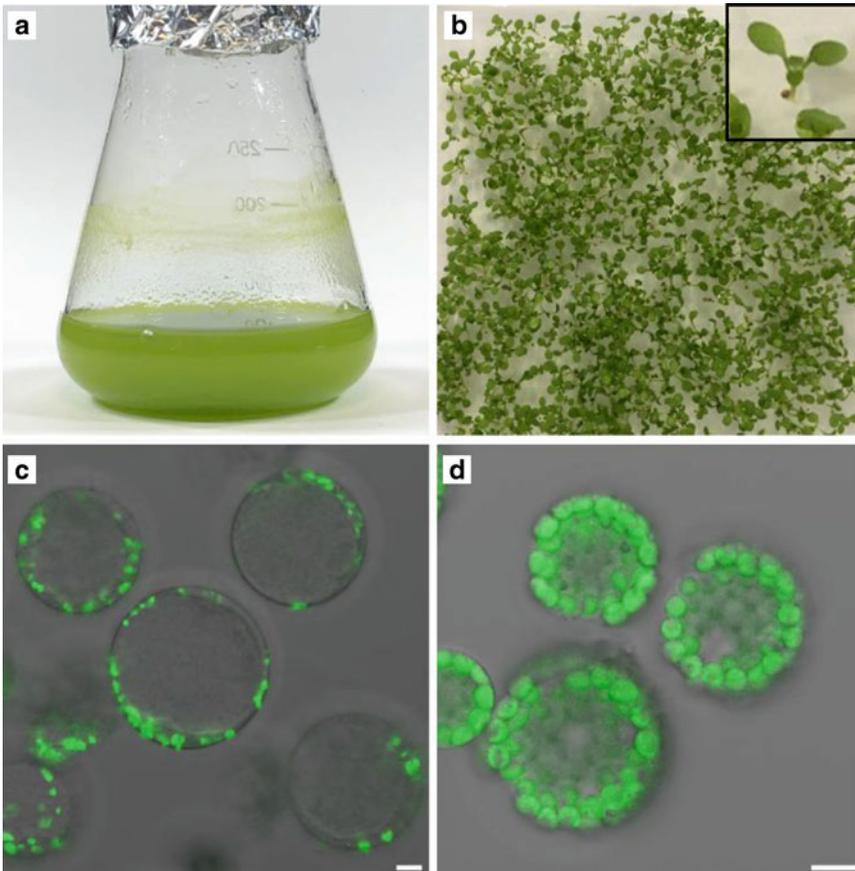


Fig. 1 Protoplast preparation from *A. thaliana* cell culture and seedlings. (a) 11-Day-old cell culture of *A. thaliana* ecotype Landsberg. (b) 7-Day-old seedlings of the *A. thaliana* Col-0 ecotype. Enlargement of a seedling is shown in the top-right edge. (c) Visualization by confocal microscopy of isolated protoplasts prepared from cell culture. (d) Visualization by confocal microscopy of isolated protoplasts from seedlings. Chlorophyll autofluorescence (green) reveals chloroplasts and their integrity (Zeiss LSM780 confocal microscope, 633 nm laser). Scale bars: 10 μm

3.3.3 Removal of Cell Walls from Seedlings

25 mL of digestion buffer per 200 mg of young seedling leaves is required.

1. Place 2 mL of digestion buffer into a Petri dish.
2. From a 7-day-old culture of seedlings (Fig. 1b), cut seedlings with scissors to remove the roots. Weigh 200 mg of the remaining young leaves and transfer them to the Petri dish.
3. Carefully lacerate leaves with a scalpel blade and cover them with 23 mL of digestion buffer. Lacerated leaves must never be in contact with air and must not dry.
4. Incubate in a vacuum chamber at 0.01 bar until bubbles appear. Wait for 1 min to allow buffer infiltration into tissues, and slowly break vacuum.
5. Cover Petri dish with aluminum foil, and incubate at 26 °C overnight at 50 rpm (orbital shaker).

3.3.4 Purification of Protoplasts

1. After cell wall digestion (using cells from either cell culture or seedlings), place a funnel with a single layer of Miracloth on an Erlenmeyer. Wet the Miracloth with 10 mL of MM buffer.
2. Add 10 mL of MM buffer to the protoplast digestion mixture (*see Note 8*). Break cell aggregates by pipetting 2–3 mL of the mixture with a 10 mL pipette and by creating bubbles by pipetting air. Pour the cell culture back into Petri dish, and repeat this operation several times, to allow the release of the majority of protoplasts. Disappearance of aggregates can be followed by observing a drop with an inverted microscope.
3. Filter protoplasts through the Miracloth layer and transfer 10 mL aliquots of the filtrate into 14 mL round-bottom tubes. This step removes the undigested cell walls and tissues.
4. Centrifuge for 3 min at $100 \times g$ without brake. Remove supernatant and resuspend the protoplasts by gentle swirling.
5. Add 5 mL of W5 buffer per tube. Gather the content of two tubes into one. Centrifuge for 1 min at $100 \times g$ without brake.
6. Remove supernatant and softly resuspend protoplasts.
7. Repeat **steps 5 and 6** until all protoplasts are gathered into two tubes.
8. Gently resuspend protoplasts by inversion of the tubes and add 10 mL of MMG buffer per tube. Centrifuge for 1 min at $100 \times g$ without brake. Remove supernatant and softly resuspend protoplasts.
9. Add 2 mL of MMG, gather tubes, and incubate for at least 15 min on ice.
10. Add 10 mL of MMG. Centrifuge for 1 min at $100 \times g$. Remove supernatant and softly resuspend protoplasts in 1 or 2 mL of MMG.

3.3.5 Determination of Protoplast Concentration

1. Clean a Fuchs-Rosenthal cell with 70% ethanol and fix the cover glass with water.
2. Put a drop of protoplast suspension between the slide and the cover glass.
3. Count the number of protoplasts under an inverted microscope, and determine the concentration (*see Note 10*). Protoplasts prepared either from cell culture or from seedlings are shown in Fig. 1.

3.4 Protoplast Transfection

1. Prepare transfection buffer (*see Note 11*). Weigh 3 g of PEG-4000 and transfer to a 50 mL tube. Add 1.5 mL of water, 4 mL of 0.5 M mannitol, pH 5.7, and 1 mL of 1 M CaCl₂. Vortex until PEG-4000 is completely dissolved.
2. Prepare 2 mL round-bottom microtubes each with at least 200,000 protoplasts at a final concentration of two million protoplasts/mL of MMG.
3. Add 1 μ L of 100 μ M small RNA and 15 μ g of plasmid carrier DNA to each tube.
4. Gently mix by slowly inverting the tubes. Do not pipette.
5. Add 1 volume of transfection buffer and homogenize the protoplast suspension by inverting the tubes 2–3 times. Do not pipette.
6. Incubate for 15–30 min at room temperature.

3.5 Washing of Transfected Protoplasts

1. Add 2 volumes of W5 and homogenize the protoplast suspension by gently inverting the tubes.
2. Centrifuge for 1 min at $100 \times g$ without brake. Remove supernatant and gently resuspend protoplasts.
3. Repeat **steps 1** and **2** three times.
4. Add 2 volumes of WI and homogenize the protoplast suspension by gently inverting the tubes.
5. Transfer the suspension to a 12-well plate and close with micropore tape. Wrap the plate in aluminum foil and incubate at 25 °C for the time required for preparing your experiment (e.g., 20 min is sufficient for northern blot analysis) (*see Note 12*).

3.6 RNase Treatment

RNase treatment allows demonstrating that the small RNA is not just bound to the surface of protoplasts but incorporated into *A. thaliana* protoplasts.

1. Delicately pour 100 μ L of the transfected protoplasts into a 2 mL round-bottom tube. Add 1 μ L of the RNase A/T1 mix. Incubate for 10 min at room temperature.

2. Add 1 mL of WI buffer. Centrifuge for 1 min at $100 \times g$ without brake. Remove supernatant by pipetting and gently resuspend the protoplasts.
3. Repeat **step 2**. After the last centrifugation, add 100 μL of WI to the protoplast pellet.

3.7 Small RNA Extraction

Due to the toxicity of TriReagent™, **steps 1–5** must be performed under a chemical hood.

1. Add 300 μL of TriReagent™ to the protoplast pellet. Vortex for 10 min.
2. Centrifuge for 10 min at $12,000 \times g$, at 4 °C. Transfer aqueous phase to a new tube.
3. Add 120 μL of chloroform and vortex for 10 min.
4. Centrifuge for 15 min at $12,000 \times g$, at 4 °C. Transfer aqueous phase to a new tube.
5. Add 240 μL of isopropanol. Delicately homogenize the solution by inverting the tube. Incubate for 5 min at room temperature, and then for 10 min at 4 °C.
6. Centrifuge for 15 min at $12000 \times g$, at 4 °C. Remove supernatant.
7. Wash pellet by adding 175 μL of 95% ethanol.
8. Centrifuge for 15 min at $12,000 \times g$, at 4 °C. Carefully remove supernatant and let the pellet dry at room temperature for 10 min to ensure that no ethanol remains in the tube.
9. Resuspend the pellet in 200 μL of water. Add 133 μL of LiCl solution. Incubate for 16 h at 4 °C (*see Note 13*).
10. Centrifuge for 30 min at $16,000 \times g$, at 4 °C. Transfer supernatant to a new tube.
11. Add 30 μL of 1 M sodium acetate, pH 4.5, 0.5 μL of glycogen, and 2.5 volumes of cold ethanol. Precipitate the nucleic acids for at least 1 h at -20 °C (*see Note 14*).
12. Centrifuge for 30 min at $16,000 \times g$, at 4 °C. Remove supernatant, let the pellet dry, and then dissolve the nucleic acids in 5 μL of DEPC-treated water.

3.8 RNA Fractionation by Polyacrylamide Gel Electrophoresis

1. Assemble plates in the casting frame and fix the frame on the casting stand.
2. Mix 7 mL of 15% acrylamide, 49 μL of APS, and 2.45 μL of TEMED. Quickly pour the solution into the space between the plates. Insert the 15-well comb and wait until the polymerization of the gel is completed.
3. To the samples (5 μL) add 5 μL of $2\times$ loading buffer.

4. Load samples onto the gel. Run the electrophoresis at 150 V in $1\times$ TBE until the bromophenol blue band reaches the bottom of the gel.
5. Prepare a container with water and add ethidium bromide to a final concentration of $0.5\ \mu\text{g}/\text{L}$.
6. Separate the gel plates with a spatula. Carefully take off the gel and transfer it to the ethidium bromide solution. Incubate for 5–10 min at room temperature with gentle agitation. Image the RNA profile with a UV transilluminator.

3.9 Electrophoretic Blotting Procedure

1. Cut four Whatman sheets to the size of the plates, and a Hybond N+ membrane square to a size slightly larger than the gel.
2. Soak a sponge from the electroblot apparatus with $0.5\times$ TBE and place it in the opened electroblotting cassette. Wet two sheets of Whatman paper with $0.5\times$ TBE and place them onto the sponge. Soak the Hybond N+ membrane with $0.5\times$ TBE and place it onto the top of the Whatman sheets. Rinse the gel in $0.5\times$ TBE and transfer the gel onto the nylon membrane. Remove any bubbles (*see Note 15*). Soak two more sheets of Whatman paper and place them on top of the membrane. Wet the second sponge with $0.5\times$ TBE and place it on top. Clamp the cassette closed and place the assembled system into the chamber. Pay attention to the cassette orientation in the chamber to allow transfer of RNAs from the gel onto the membrane. Fill the chamber with $0.5\times$ TBE.
3. Carry out the transfer of the RNAs from the gel to the membrane at 300 mA for 80 min at $4\ ^\circ\text{C}$.
4. Place two sheets of Whatman paper in a container and soak them with $2\times$ SSC. Discard excessive buffer. Disassemble the electroblotting system, remove the nylon membrane, and place it onto the sheets of Whatman paper with the surface carrying the RNA facing up. Incubate for 10 min at room temperature. Place the container into a Stratalinker to cross-link the RNA to the membrane by applying two UV exposures of $120,000\ \mu\text{J}/\text{cm}^2$.

3.10 Labeling of the Oligodeoxynucleotide Probe

All steps requiring manipulation of radioactivity must be carried out in an appropriate room. The safety rules of your institution must be strictly followed.

1. Preheat a water bath at $37\ ^\circ\text{C}$.
2. Mix $4.5\ \mu\text{L}$ of water, $1\ \mu\text{L}$ of $100\ \mu\text{M}$ oligodeoxynucleotide complementary to the small RNA of interest, $1\ \mu\text{L}$ of $10\times$ T4 polynucleotide kinase (PNK) buffer, $1\ \mu\text{L}$ of PNK, and $2.5\ \mu\text{L}$ of $[\gamma^{32}\text{P}]\text{-ATP}$. Incubate for 30 min at $37\ ^\circ\text{C}$.

3. Remove unincorporated [$\gamma^{32}\text{P}$]-ATP with a Sephadex G-50 spin column. To prepare a G-50 column, add 2 mm of fiber-glass into the bottom of a 1 mL syringe. Place the syringe into a 14 mL tube. Load swelling watered G-50 resin into the syringe and allow the water to drop out (but do not let the resin dry). Add resin until the syringe is filled. Centrifuge for 45 s at $2000 \times g$. Place the syringe in a new 14 mL tube.
4. Add 50 μL of water to the PNK reaction mixture and load it onto the top of the G-50 column. Centrifuge for 45 s at $2000 \times g$ to elute the radioactive probe. Discard the syringe in the radioactive waste disposal container.

3.11 Northern Hybridization

1. Wet the membrane with PerfectHyb™ Plus solution and place it into a cylindrical glass hybridization tube. Make sure that the membrane is well attached to the tube wall.
2. Add 4 mL of PerfectHyb™ Plus hybridization solution into the tube.
3. Add 940 μL of PerfectHyb™ Plus to the radiolabeled probe, and transfer it into the hybridization tube.
4. Place the tube into the hybridization oven, and incubate with gentle rotation overnight at 42 °C.
5. Discard the probe into the radioactive waste disposal container or keep it at -20 °C if you wish to reuse the probe another time (the half-time of ^{32}P is 14 days).
6. Pour 10 mL of $2 \times \text{SSC}$ into the tube and wash the membrane for 10 min at 42 °C with gentle rotation.
7. Discard the $2 \times \text{SSC}$ buffer into the waste container dedicated to radioactivity. Repeat **step 6**.
8. Discard $2 \times \text{SSC}$ as above. Add 10 mL of $2 \times \text{SSC}$ and 0.1% SDS into the tube and wash the membrane for 30 min at 42 °C with gentle rotation.
9. Discard the washing solution. Place the membrane on a sheet of Whatman paper to remove excess liquid. Then place the membrane between two sheets of plastic wrap.
10. Expose the radioactive signal of membrane to a phosphorimager screen or to X-ray film. Image the screen using a phosphor-imager or develop the X-ray film (*see Note 16*).
11. As an illustration, the autoradiography of a northern blot showing the successful transfection of sncRNA into *A. thaliana* protoplasts is presented Fig. 2. In this example, an oligoribonucleotide (5'-ACCATCTGAGCTACATCCCC-3') corresponding to a previously identified tRF Ala of 20 nt (tRF Ala20) derived from the 5' extremity of an *A. thaliana* cytosolic alanine tRNA [5] has been incorporated by the protoplasts. The oligoribonucleotide transfection was further

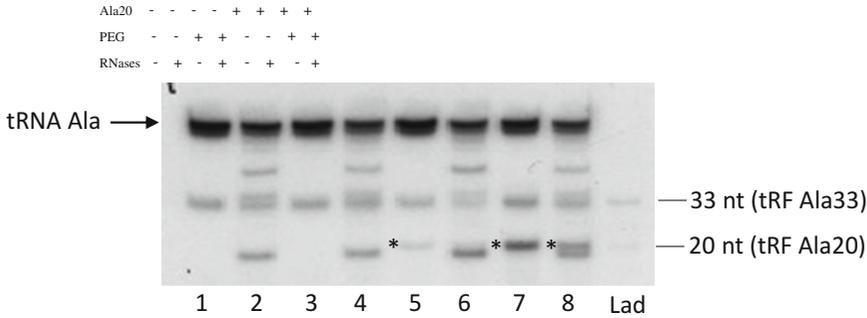


Fig. 2 Analysis by northern blot of the transfection of tRF Ala20 into *A. thaliana* protoplasts. Autoradiography of the northern blot experiment performed with RNA samples extracted from *A. thaliana* protoplasts transfected with an oligoribonucleotide (tRF Ala20) corresponding to the cytosolic *A. thaliana* tRNA alanine (UGC) of 20 nt in size. The same oligoribonucleotide was used as a ladder (Lad), together with an oligoribonucleotide corresponding to tRF Ala33 (33 nt of the 5' extremity of tRNA^{Ala} (UGC)). Protoplasts prepared from an *A. thaliana* cell culture were transfected in the absence (–) or in the presence (+) of tRF Ala20 as well as in the absence (–) or in the presence (+) of PEG. Transfected protoplasts were treated (+) or not treated (–) with a mix of RNase A and RNase T1 (RNases). Extracted RNAs were probed with a ³²P-radiolabeled oligodeoxynucleotide probe complementary to the 5' extremity of tRNA^{Ala} (UGC), (5'-GGGGATGTAGCTCA GATGGT-3'). In all samples, signals corresponding to the endogenous alanine tRNA and the endogenous tRF Ala33 are seen. RNA samples from protoplasts treated with RNases exhibit additional bands, which is likely due to protoplast lysis and accessibility of alanine tRNA to nucleases (lanes 2, 4, 6, and 8). Lane 5 with RNA samples of protoplasts treated for transfection in the absence of PEG shows a band corresponding to tRF Ala20. This signal is not present upon treatment of the protoplasts with RNases (lane 6). By contrast, this signal is higher when protoplasts were treated for transfection in the presence of PEG (lane 7) and this signal is still present when such protoplasts were post-treated with RNases (lane 8), thus demonstrating the successful transfection and internalization of tRF Ala20 into *A. thaliana* protoplasts

analyzed by a northern blot experiment as described above. Control experiments were performed. In the absence of the Ala20 oligoribonucleotide (control with water), or in the absence of PEG, no signal corresponding to the tRF was observed. After RNase treatment, a signal corresponding to the tRF Ala20 was still visible in the transfected protoplasts, thus demonstrating successful incorporation into protoplasts. This oligoribonucleotide transfection method could be widely used to identify the interactome of a biotinylated tRF or the subcellular localization of a fluorescently tagged tRF.

4 Notes

1. The addition of HCl increases the temperature of the solution and distorts pH measurement. Always wait for temperature cooling before measuring pH.

2. EDTA can only be dissolved at a basic pH. Add NaOH pellets and wait for a few minutes. Then add additional pellets one by one until the EDTA powder is dissolved.
3. Boric acid precipitates after long storage (around 1 month). Thus, preparing a large volume of $10\times$ TBE is not advised.
4. Heating is required to dissolve SDS.
5. SDS can precipitate if the room temperature decreases below $20\text{ }^{\circ}\text{C}$. Heat the solution to dissolve SDS before use, if needed.
6. Pipetting with normal tips exposes protoplasts to an excessive osmotic pressure change and can smash them. Better use pipettes with wide-mouth tips, and cut micropipette tips with scissors to enlarge the opening hole. Avoid pipetting and rather pour protoplast solutions to the new containers if possible.
7. Applying braking combined with centrifuge force can smash the cells.
8. Add the buffer slowly along the wall of the tube or the Petri dish to avoid excess of pressure and mechanical force, which could smash the protoplasts.
9. Quality of the seeds is important. Better to use freshly harvested seeds (less than 1 year old). If you use old seeds, sterilization is less efficient. In the case of old seeds or mutant seeds, we have used 80% ethanol and 4% TweenTM 20 (rather than 70% ethanol, 0.1% TweenTM 20) as sterilization solution. Please note that times must be respected to avoid damage to the seeds.
10. Intact protoplasts are perfectly round and well isolated (*see* Fig. 1). Their size can vary, but both large and small ones can be counted if they are not damaged. The number of protoplasts prepared from 10 mL of cell culture is usually around 2×10^6 whereas 200 mg of seedlings will give about 1×10^6 protoplasts.
11. The transfection buffer must be freshly prepared on the day of transfection.
12. Protoplasts can be kept in the 12-well plate for a maximum of 30 h. The longer the protoplasts are kept in the plate, the more of them will die.
13. LiCl causes the precipitation of large RNAs (i.e., >200 nt in size). After centrifugation, the supernatant contains tRNAs and small noncoding RNAs (including tRFs).
14. One hour is usually sufficient for precipitation if ethanol kept at $-20\text{ }^{\circ}\text{C}$ is used. Otherwise, precipitation can also be performed overnight.
15. To avoid bubbles, the membrane should be wet before placing the gel on it. Bubbles can be chased by smoothly pressing a pipette on the gel.

16. To increase signal with X-ray exposure, use a reflective screen placed on the film and incubate the cassette at -80°C . Under these conditions, a strong signal can be obtained after 6 h of exposure. At room temperature, an exposure time of 24–48 h is recommended.

References

1. Shen Y, Yu X, Zhu L, Li T, Yan Z, Guo J (2018) Transfer RNA-derived fragments and tRNA halves: biogenesis, biological functions and their roles in diseases. *J Mol Med* 96:1167–1176
2. Megel C, Morelle G, Lalande S, Duchêne A-M, Small I, Maréchal-Drouard L (2015) Surveillance and cleavage of eukaryotic tRNAs. *Int J Mol Sci* 16:1873–1893
3. Schimmel P (2018) RNA processing and modifications: the emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat Rev Mol Cell Biol* 19:45–58
4. Cristodero M, Polacek N (2017) The multifaceted regulatory potential of tRNA-derived fragments. *Noncoding RNA Investig* 1:7
5. Cognat V, Morelle G, Megel C, Lalande S, Molinier J, Vincent T, Small I, Duchêne A-M, Maréchal-Drouard L (2017) The nuclear and organellar tRNA-derived RNA fragment population in *Arabidopsis thaliana* is highly dynamic. *Nucleic Acids Res* 45:3460–3472
6. Haas G, Cetin S, Messmer M, Chane-Woon-Ming B, Terenzi O, Chicher J, Kuhn L, Hammann P, Pfeffer S (2016) Identification of factors involved in target RNA-directed microRNA degradation. *Nucleic Acids Res* 44:2873–2887
7. Han Z, Ge X, Tan J, Chen F, Gao H, Lei P, Zhang J (2015) Establishment of lipofection protocol for efficient miR-21 transfection into cortical neurons in vitro. *DNA Cell Biol* 34:703–709
8. Spörlein B, Koop H-U (1991) Lipofectin: direct gene transfer to higher plants using cationic liposomes. *Theor Appl Genet* 83:1–5
9. Miao Y, Jiang L (2007) Transient expression of fluorescent fusion proteins in protoplasts of suspension cultured cells. *Nat Protoc* 2:2348–2353
10. Masani MYA, Noll GA, Parveez GKA, Sambanthamurthi R, Prüfer D (2014) Efficient transformation of oil palm protoplasts by PEG-mediated transfection and DNA microinjection. *PLoS One* 9:e96831
11. Yoo S-D, Cho Y-H, Sheen J (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat Protoc* 2:1565–1572
12. Qi Y, Zhong X, Itaya A, Ding B (2004) Dissecting RNA silencing in protoplasts uncovers novel effects of viral suppressors on the silencing pathway at the cellular level. *Nucleic Acids Res* 32:e179
13. Wintz H, Dietrich A (1996) Electroporation of small RNAs into plant protoplasts: mitochondrial uptake of transfer RNAs. *Biochem Biophys Res Commun* 223:204–210
14. McCormack JC, Simon AE (2006) Callus cultures of Arabidopsis. *Curr Protoc Microbiol* 00:16D.1.1–16D.1.9
15. Sambrook J, Russell DW (2001) In: Harbor CS (ed) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Annexe V. Oligonucléotides ARN utilisés pour la transfection de protoplastes d'*A. thaliana*

Référence	Description	Séquence 5' → 3'	Etiquette en -3'
ARS1 ARS3	Correspond aux 20 premiers nucléotides sens de l'extrémité 5' de l'ARNt ^{Ala} (AGC) d' <i>A. thaliana</i> .	GGGGAUGUAGCUCAGAUGGmU	Sp9 - Alexa 647 Sp9 - Biotine
ARS2 ARS4	Correspond à la séquence complémentaire, mais non réverse, de l'ARNt ^{Ala} (AGC) d' <i>A. thaliana</i> (position G3 à G22). Considéré comme une séquence aléatoire.	CCUACAUCGAGUCUACCAUmC	Sp9 - Alexa 647 Sp9 - Biotine
ARS5 ARS6	Correspond aux 19 premiers nucléotides sens de l'extrémité 5' de l'ARNt ^{Arg} (CCT) nucléaire d' <i>A. thaliana</i> .	GCGUCUGUAGCUCAGUGGmA	Sp9 - Alexa 647 Sp9 - Biotine
ARS7 ARS8	Correspond aux 20 premiers nucléotides sens de l'extrémité 5' de l'ARNt ^{His} (GTG) plastidial d' <i>A. thaliana</i> .	GGCGGAUGUAGCCAAGUGGmA	Sp9 - Alexa 647 Sp9 - Biotine
ARS9 ARS10	Correspond aux 33 premiers nucléotides sens de l'extrémité 5' de l'ARNt ^{Asp} (GTC) nucléaire d' <i>A. thaliana</i> .	GUCGUUGUAGUAUAGUGGUAA GUAUUCCCCGCCmU	Sp9 - Alexa 647 Sp9 - Biotine
ARS11 ARS12	Correspond aux 19 premiers nucléotides sens de l'extrémité 5' de l'ARNt ^{Gly} (TCC) nucléaire d' <i>A. thaliana</i> .	GCGUCUGUAGUCCAACGGmU	Sp9 - Alexa 647 Sp9 - Biotine
ARS13 ARS14	Correspond aux 20 premiers nucléotides sens de l'extrémité 5' de l'ARNt ^{Ala} (AGC) nucléaire d' <i>A. thaliana</i> , sauf G2 remplacé par C et G4 par U.	GCGUAUGUAGCUCAGAUGGmU	Sp9 - Alexa 647 Sp9 - Biotine
ARS15 ARS16	Correspond aux 19 premiers nucléotides sens de l'extrémité 5' de l'ARNt ^{Gly} (TCC) nucléaire d' <i>A. thaliana</i> , sauf C2 et U4 remplacés par G.	GGGGCUGUAGUCCAACGGmU	Sp9 - Alexa 647 Sp9 - Biotine
ARS17 ARS18	Correspond aux 21 nucléotides sens de la séquence du promoteur c-Myc.	GGGGAGGGUGGGGAGGGUGGG mG	Sp9 - Alexa 647 Sp9 - Biotine

Annexe VI. Séquences ADN correspondantes aux ARN guides utilisés pour les transformations d'*A. thaliana* par CRISPR/Cas9

Référence	Description	Orientation	Séquence 5' → 3'
1993	ARNg 1 ciblant la séquence codant RNS2	Sens	ATTGCTAATCATGTAGATGGGTTA
1994	ARNg 1 ciblant la séquence codant RNS2	Antisens	AAACTAACCCATCTACATGATTAG
1995	ARNg 2 ciblant la séquence codant RNS2	Sens	ATTGTTATGGCCTGACTATAACGA
1996	ARNg 2 ciblant la séquence codant RNS2	Antisens	AAACTCGTTATAGTCAGGCCATAA
1997	ARNg 3 ciblant la séquence codant RNS2	Sens	ATTGGTGAAAAGATAATCACCTCG
1998	ARNg 3 ciblant la séquence codant RNS2	Antisens	AAACCGAGGTGATTATCTTTTTCAC

Annexe VII. Amorces ADN utilisées pour l'amplification d'ADN par PCR

Référence	Description	Orientation	Séquence 5' → 3'	Température d'hybridation	Taille de l'amplicon
<i>PCR à partir de colonie – TOP10 transformées avec pKI1.1R</i>					
U6 fw	S'hybride à la région du promoteur U6 du plasmide pKI1.1R	Sens	GCCTTCCGCACAATACATCATT		
1994, 1996 ou 1998	ARNg ciblant la séquence codant RNS2 insérée dans le plasmide pKI1.1R	Antisens	AAACTAACCCATCTACATGATTAG AAACTCGTTATAGTCAGGCCATAA AAACCGAGGTGATTATCTTTTCAC	58°C	443 pb
<i>PCR à partir d'ADNg d'A. thaliana – vérification de la présence/absence du T-DNA par amplification de la Cas9</i>					
2093	S'hybride à la séquence Cas9 du plasmide pKI1.1R	Sens	GAAAGAAACTGGTGGACAGCAC	58°C	402 pb
2094	S'hybride à la séquence Cas9 du plasmide pKI1.1R	Antisens	CGTCGTAGGTGTCTTGCTCAG		
<i>PCR à partir d'ADNg d'A. thaliana – contrôle positif ADNg</i>					
AM 975	S'hybride sur le gène FMT (At3g52140)	Sens	CGATGCCCAACCAAGCAAGG	58°C	486 pb
AM 976	S'hybride sur le gène FMT (At3g52140)	Antisens	GGAGCATCAAGGAGGAAGTCC		
<i>PCR à partir d'ADNg d'A. thaliana – amplification de RNS2 au niveau de la région de l'ARNg 1 et de l'ARNg 2 pour séquençage</i>					
2095	S'hybride environ 250 pb en amont de l'ARNg 1 et de l'ARNg 2	Sens	CATTAGATGTGGGAACAAAACAGAG	52°C	556 pb
2096	S'hybride environ 250 pb en aval de l'ARNg 1 et de l'ARNg 2	Antisens	CAAAATGACCCTTTCCCACC		
<i>PCR à partir d'ADNg d'A. thaliana – amplification de RNS2 au niveau de la région de l'ARNg 3 pour séquençage</i>					
2097	S'hybride environ 250 pb en amont de l'ARNg 3	Sens	GATGGTTCGTGGCCTTCATG	52°C	540 pb
2098	S'hybride environ 250 pb en amont de l'ARNg 1	Antisens	GGTTGTGAGATTAAGCATGGAAC		
<i>PCR à partir d'ADNg d'A. thaliana – géotypage lignées insertion T-DNA (FLAG) dans RNS1</i>					
1717	S'hybride sur le gène RNS1	Sens	GGAGCTCTAACCAAAGCC	45°C	500 pb
1716	S'hybride sur le T-DNA (des lignées FLAG)	Antisens	TACAAATTGCCTTTTCTTATCGAC		
1717	S'hybride sur le gène RNS1	Sens	GGAGCTCTAACCAAAGCC	45°C	400 pb
1718	S'hybride sur le gène RNS1	Antisens	CAATCTCCATGAACAAAAGATC		

Référence	Description	Orientation	Séquence 5' → 3'	Température d'hybridation	Taille de l'amplicon
<i>PCR à partir d'ADNg d'A. thaliana – génotypage lignées insertion T-DNA (FLAG) dans RNS3</i>					
1715	S'hybride sur le gène RNS3	Sens	TCTTCATTTTTATTCTAGCGTTAC	45 °C	500 pb
1716	S'hybride sur le T-DNA (des lignées FLAG)	Antisens	TACAAATTGCCTTTTCTTATCGAC		
1715	S'hybride sur le gène RNS3	Sens	TCTTCATTTTTATTCTAGCGTTAC	45°C	500 pb
1714	S'hybride sur le gène RNS3	Antisens	GCACGTACCGTGTTTCTC		
<i>PCR à partir d'ADNg d'A. thaliana – génotypage lignées insertion T-DNA (SALK) dans RNS4</i>					
1719	S'hybride sur le T-DNA (des lignées SALK)	Sens	ATTTTGCCGATTTTCGGAAC	52°C	600pb
2181	S'hybride sur le gène RNS4	Antisens	GACGCAGCGTAATCATTAAAGG		
1489	S'hybride sur le gène RNS4	Sens	ATGAGAGGAATCATCATCGTAAGC	55°C	600 pb
2183	S'hybride sur le gène RNS4	Antisens	ACACAGGTGCCATGCTTATTC		
<i>PCR à partir d'ADNg d'A. thaliana – génotypage lignées insertion T-DNA (FLAG) dans RNS4</i>					
1716	S'hybride sur le T-DNA (des lignées FLAG)	Sens	TACAAATTGCCTTTTCTTATCGAC	52°C	600 pb
2183	S'hybride sur le gène RNS4	Antisens	ACACAGGTGCCATGCTTATTC		
2184	S'hybride sur le gène RNS4	Sens	AAACCAAATTCATGGGGATTC	52°C	1100 pb
2183	S'hybride sur le gène RNS4	Antisens	ACACAGGTGCCATGCTTATTC		
<i>PCR à partir d'ADNg d'A. thaliana – génotypage lignées insertion T-DNA (SALK) dans DExH1</i>					
1719	S'hybride sur le T-DNA (des lignées SALK)	Sens	ATTTTGCCGATTTTCGGAAC	60°C	700 pb
2100	S'hybride sur le gène DExH1	Antisens	TTCTCTTGAAGCATGGTCTGG		
2105	S'hybride sur le gène DExH1	Sens	TGGTAAGACAAAATGGGCTACG	60°C	1000 pb
2100	S'hybride sur le gène DExH1	Antisens	TTCTCTTGAAGCATGGTCTGG		
<i>PCR haute-fidélité à partir d'ADNc d'A. thaliana – amplification de DExH1 (At2g35920)</i>					
2148	S'hybride à partir du codon ATG	Sens	ATGCCTCCTCACGGTCCTAATTC	65,5°C	2988 pb
2149	S'hybride à partir du codon STOP	Antisens	CTAATGTTTTTGGCTACGTAACAACCTCC		

Référence	Description	Orientation	Séquence 5' → 3'	Température d'hybridation	Taille de l'amplicon
<i>PCR haute-fidélité à partir d'ADNc d'A. thaliana – amplification de DExH1 (At2g35920) + ajout des sites attB pour clonage Gateway</i>					
2150	S'hybride à partir du codon ATG + ajoute site <i>attB1</i>	Sens	GGGGACAAGTTTGTACAAAAAAGCAGGCT TCATGCCTCCTCACGGTCCTAATTC	65,5°C	3049 pb
2151	S'hybride à partir du codon STOP + ajout site <i>attB2</i>	Antisens	GGGGACCACTTTGTACAAGAAAGCTGGGT CCTAATGTTTTTGGCTACGTAACTCC		
2152	S'hybride à partir du codon ATG + ajoute site <i>attB1</i>	Sens	GGGGACAAGTTTGTACAAAAAAGCAGGCT TCATGCCTCCTCACGGTCCTAATTC	65,5°C	3046 pb
2153	S'hybride à la fin du gènes + supprime STOP + ajout site <i>attB2</i>	Antisens	GGGGACCACTTTGTACAAGAAAGCTGGGT CATGTTTTTGGCTACGTAACTCC		
<i>PCR haute-fidélité à partir des vecteurs recombinants pK7WGF2 et pB7FWG2 – amplification DExH1 et GFP + ajout des sites Bsal pour clonage ALiCE</i>					
2173	Amplification de la première partie de la GFP+DExH1 à partir du vecteur recombinant pK7WGF2	Sens	GGTCTCACCATATGGTGAGCAAGGGCGAG	64°C	1527 pb
2174		Antisens	GGTCTCCGGTTTCTCCTGAACTACCAAT		
2175	Amplification de la deuxième partie de DExH1 à partir du vecteur recombinant pK7WGF2	Sens	GGTCTCAAACCGGGTGTGGTAAAACAACA CAGCTTC	72°C	2252 pb
2176		Antisens	GGTCTCCAAGCCTAATGTTTTTGGCTACG TAACAACCTCCACCACAGCTGATAC		
2177	Amplification de la première partie de DExH1 à partir du vecteur recombinant pB7FWG2	Sens	GGTCTCACCATATGCCTCCTCACGGTCCT	64°C	777 pb
2174		Antisens	GGTCTCCGGTTTCTCCTGAACTACCAAT		
2175	Amplification de la deuxième partie de DExH1 à partir du vecteur recombinant pB7FWG2	Sens	GGTCTCAAACCGGGTGTGGTAAAACAACA CAGCTTC	72°C	2244 pb
2178		Antisens	GGTCTCCATGTTTTTGGCTACGTAAAC TCCACCACAGCTGATAC		
2179	Amplification de la troisième partie de DExH1+GFP à partir du vecteur recombinant pB7FWG2	Sens	GGTCTCAACATGACCCAGCTTTCTTGT	62°C	774 pb
2180		Antisens	GGTCTCCAAGCTTACTTGTACAGCTCGTC CA		
2173	Amplification de la GFP à partir du vecteur recombinant pB7FWG2	Sens	GGTCTCACCATATGGTGAGCAAGGGCGAG	62°C	742 pb
2180		Antisens	GGTCTCCAAGCTTACTTGTACAGCTCGTC CA		

Référence	Description	Orientation	Séquence 5' → 3'	Température d'hybridation	Taille de l'ampicon
<i>PCR sur colonies transformées pALiCE – amplification de l'insert dans pALiCE</i>					
2185	S'hybride à la région du promoteur T7 du plasmide pALiCE	Sens	TAATACGACTCACTATAGG	52°C	Entre 880 et 3898 pb selon l'insert
2186	S'hybride à la région 3'UTR du plasmide pALiCE	Antisens	ATGCATCTTGACTACCTC		

Annexe VIII. Amorces ADN utilisées pour l'amplification d'ADN par RT-qPCR

Référence	Description	Orientation	Séquence 5' → 3'
2085	Amplification du gène de référence ACT2	Sens	CTTGCACCAAGCAGCATGAA
2086		Antisens	CCGATCCAGACACTGTACTTC
2087	Amplification du gène de référence GAPDH	Sens	TTGGTGACAACAGGTCAAGCA
2088		Antisens	AAACTTGTGCTCAATGCAAT
1694	Amplification du gène RNS4	Sens	TTCAATTTCTTCTACTGGGTCAACT
1695		Antisens	GGCAACATCCTTTTTGGCTA
2112	Amplification du gène DExH1	Sens	GAACAAGATTGAGCAATGGAAA
2113		Antisens	GATCTCCACGGCTAAAATTACG
2114	Amplification du gène DExH1	Sens	TCGTCTGACTCGGGGAATTA
2115		Antisens	CAGATCATCTTTCTTGGATTCTGA
2116	Amplification du gène DExH1	Sens	GCAAGTTGGATCTATAGGCTCATT
2117		Antisens	TCAATAGCATTTCTCAACAGCAAG
2118	Amplification du gène DExH1	Sens	TATGAGGGATATCGGGATGC
2119		Antisens	CCTTAGAGTTACAGGGGAAAAGAAA
2120	Amplification du gène DExH1	Sens	TGGATCTGTCAACGCAAGAG
2121		Antisens	CGTCTTACCTTTTCGCTGT
2122	Amplification du gène DExH1	Sens	GAGGGGAAAGGAGTGGTCTC
2123		Antisens	TGCTTAATACCTAATGTTTTTGGCTA

Annexe IX. Amorces ADN utilisées pour le séquençage

Référence	Description	Orientation	Séquence 5' → 3'
U6 fw	Permet de vérifier la ligation des ARNg dans le plasmide pK1.1R	Sens	GCCTTCCGCACAATACATCATT
2095	Cible RNS2, au niveau de la région de l'ARNg 1 et de l'ARNg 2	Sens	CATTAGATGTGGGAACAAAACAGAG
2096	Cible RNS2, au niveau de la région de l'ARNg 1 et de l'ARNg 2	Antisens	CAAAATGACCCTTTCCCACC
2097	Cible RNS2, au niveau de la région de l'ARNg 3	Sens	GATGGTTCGTGGCCTTCATG
2098	Cible RNS2, au niveau de la région de l'ARNg 3	Antisens	GGTTGTGAGATTAAGCATGGAAC
2148	Cible At2g35920, position 1nt (ATG)	Sens	ATGCCTCCTCACGGTCCTAATTC
2153	Cible At2g35920, position 473 nt	Antisens	CACCAGAGGACCCTGATTCTTGTG
2154	Cible At2g35920, position 553 nt	Sens	AGACCTGATTCTGCTTCCAAATTG
2155	Cible At2g35920, position 1185 nt	Antisens	TGCATAGTGGGAGAGTTCCCAAAG
2114	Cible At2g35920, position 1285 nt	Sens	TCGTCTGACTCGGGGAATTA
2156	Cible At2g35920, position 1754 nt	Antisens	GCTAGTTTCTTTGCTTTACCAC
2116	Cible At2g35920, position 1986 nt	Sens	GCAAGTTGGATCTATAGGCTCATT
2119	Cible At2g35920, position 2395 nt	Antisens	CCTTAGAGTTACAGGGGAAAAGAAA

2120	Cible At2g35920, position 2637 nt	Sens	TGGATCTGTCAACGCAAGAG
2149	Cible At2g35920, position 2988 nt (STOP)	Antisens	CTAATGTTTTTGGCTACGTAACAACCTCC
AM 478	Cible autour de l'insert dans le pDONR207	Sens	TCGCGTTAACGCTAGCATGGATCTC
AM 479	Cible autour de l'insert dans le pDONR207	Antisens	GTAACATCAGAGATTTTGAGACAC
1570	Cible p35S (contenu dans les plasmides pK7WGF2 et pB7FWG2)	Sens	AGTGGATTGATGTGATATCT
1571	Cible eGFP (contenue dans les plasmides pK7WGF2 et pB7FWG2)	Antisens	GTCGCCGTCCAGCTCGACCA
AM 873	Cible T35S (contenu dans les plasmides pK7WGF2 et pB7FWG2)	Antisens	CTGGTGATTTTTGCGGACTC
2185	Cible le promoteur T7 (contenu dans le plasmide pALiCE)	Sens	TAATACGACTCACTATAGG
2186	Cible la région 3'UTR (contenu dans le plasmide pALiCE)	Antisens	ATGCATCTTGACTACCTC

Annexe X. Sondes ADN utilisées pour la détection d'ARN par northern blot

Référence	Description	Séquence
1407	Réverse - complémentaire aux 20 premiers nucléotides de l'extrémité 5' de l'ARN ^{tAla} (AGC) d' <i>A. thaliana</i>	ACCATCTGAGCTACATCCCC
1992	Réverse - complémentaire à la séquence des oligonucléotides ARS2/ARS4	GATGGTAGACTCGATGTAGG
1703	Réverse - complémentaire aux 18 premiers nucléotides de l'extrémité 5' de l'ARN ^{tGly} (TCC) d' <i>A. thaliana</i>	CCGTTGGACTACAGACGC
1409	Réverse - complémentaire aux 26 premiers nucléotides de l'extrémité 5' de l'ARN ^{tAsp} (GTC) d' <i>A. thaliana</i>	AATCCTTACCACTATACTACAACGAC

Annexe XI. Anticorps utilisés pour les western blots

Nom	Description	Type	Dilution utilisée
RNS1	Anticorps dirigé contre RNS1 d' <i>A. thaliana</i>	Polyclonal (Lapin)	1/500 ^e
RNS2	Anticorps dirigé contre RNS2 d' <i>A. thaliana</i>	Polyclonal (Lapin)	1/500 ^e
RNS3	Anticorps dirigé contre RNS3 d' <i>A. thaliana</i>	Polyclonal (Lapin)	1/500 ^e
RNS4	Anticorps dirigé contre RNS4 d' <i>A. thaliana</i>	Polyclonal (Lapin)	1/500 ^e
eEF1 α	Anticorps dirigé contre eEF1 α de <i>Z. mays</i>	Polyclonal (Lapin)	1/5000 ^e
GFP	Anticorps dirigé contre la GFP	Polyclonal (Lapin)	1/10000 ^e

RÉSULTATS

CHAPITRE I : les RNS à l'origine des tRFs chez *A. thaliana*

I. Mise en contexte

Les ribonucléases (couramment appelées RNases) sont des enzymes capables de catalyser la dégradation de molécules d'ARN. Présentes dans tous les domaines de la vie, elles ont des rôles clés dans de nombreux processus comme la maturation de certains ARN ou encore la dégradation des ARN endommagés ou n'ayant plus aucune utilité pour la cellule. Les RNases sont divisées en deux groupes, basés sur leur mode d'action.

Les exoribonucléases dégradent les ARN nucléotide par nucléotide en partant de leur extrémité 5' ou 3'. Elles sont classées en différentes familles selon leurs propriétés catalytiques et leurs séquences. On retrouve chez les eucaryotes, par exemple, le groupe des PNPases et Xrn1 qui participent, entre autres, à la dégradation des ARNm de 3' en 5' et de 5' en 3', respectivement [171,172].

Les endoribonucléases clivent quant à elle les ARN à l'intérieur de leur séquence, libérant des fragments de tailles variés. Certaines endoribonucléases ont une spécificité de substrat alors que d'autres clivent les ARN indifféremment. Elles sont capables de cliver soit de l'ARN simple brin, soit de l'ADN structuré double brin. De façon non exhaustive, par exemple, les RNases P clivent notamment l'extrémité 5' (simple brin) des précurseurs d'ARNt. Les RNases A et T1 sont par exemple couramment utilisés au laboratoire. Les RNases A clivent les ARN en 3' après les résidus Cytosine ou Uracile alors que les RNases T1 clivent après les résidus Guanines [23,24,172–174].

Les endoribonucléases sont à l'origine des tRFs. Selon les organismes, ces RNases appartiennent à différentes familles comme par exemple l'Angiogenine, une RNase 5, qui génère des tRFs chez l'humain. D'après la littérature, chez les plantes, deux types d'endoribonucléases peuvent cliver les ARNt pour former des tRFs : les DCL et les RNS (voir le paragraphe II.2.iii de l'Introduction de ce manuscrit). Les RNS sont l'acteur majeur dans la biogenèse des tRFs chez *A. thaliana*. Elles appartiennent à la famille des RNases T2, plus largement décrites ci-après [170].

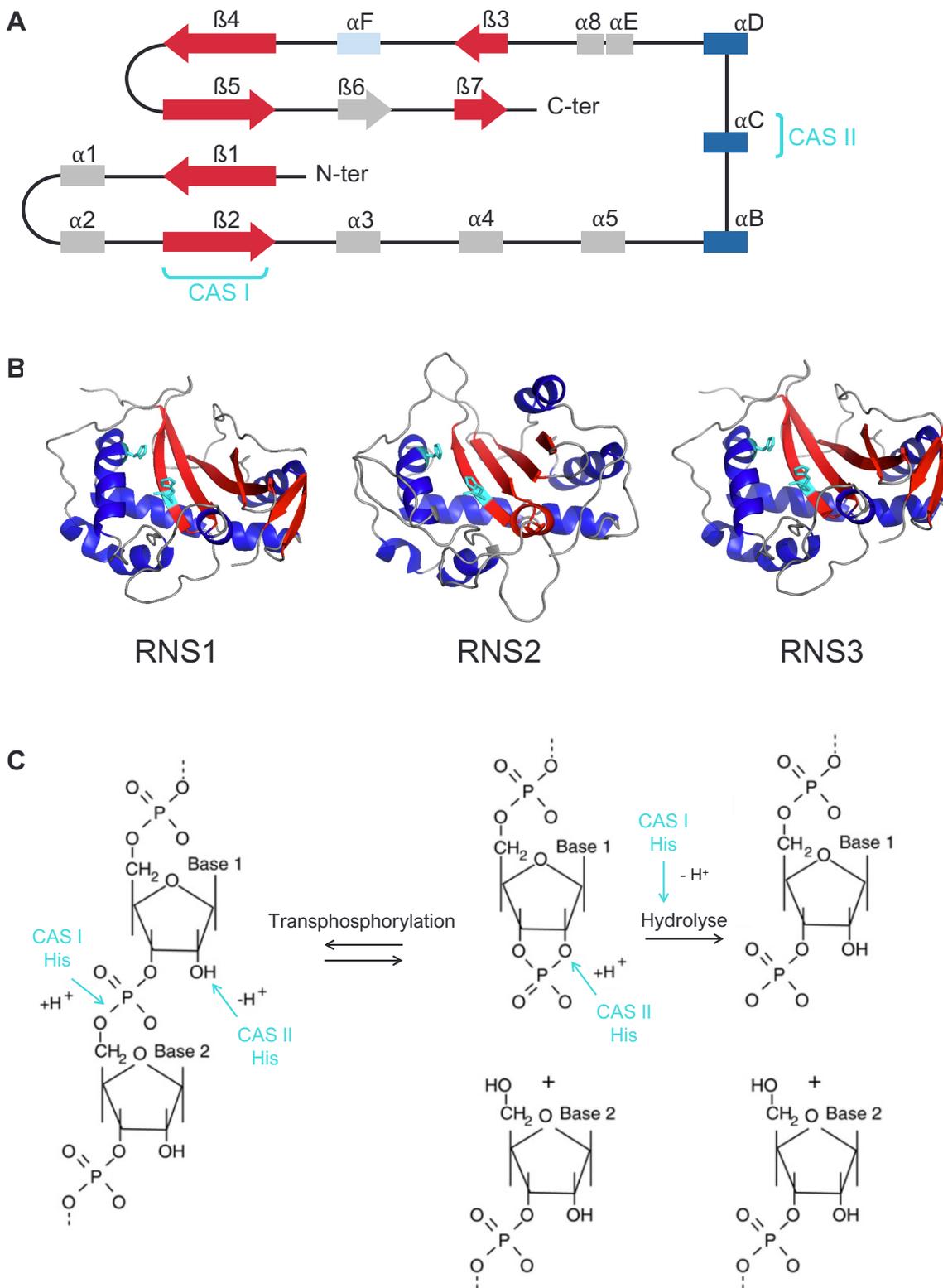


Figure 13 : Structure des RNases T2. A : Diagramme topologique des RNases T2 montrant les structures secondaires (d'après [175]). Les éléments conservés chez tous les membres de cette famille sont représentés en couleurs : les feuilletts β conservés sont représentés par des flèches rouge et les hélices α par des rectangles bleus foncés. Les éléments partiellement conservés et non conservés sont représentés en bleu clair et gris, respectivement. Les sites CAS I et CAS II sont indiqués en cyan. B : Structure tridimensionnelle de RNS1, RNS2 et RNS3 d'*A. thaliana* modélisées avec Phyre2 et visualisées avec le logiciel Pymol [113]. Les hélices α sont représentées en bleu foncé, les feuilletts β en rouge et les deux résidus His des sites CAS I et CAS II en cyan. C : Mécanisme catalytique des RNases T2 (adapté de [177]). Les deux étapes, transphosphorylation et hydrolyse, sont montrées. L'implication des résidus Histidines des sites CAS I et CAS II lors des deux étapes est indiqué en cyan.

1. Description des RNases T2

Les RNases T2 sont présentes chez les eucaryotes mais aussi chez certaines bactéries. Elles ont une masse moléculaire variant entre 19 et 97 kDa mais la majorité des enzymes ont une masse comprise entre 20 et 40 kDa. En plus de la variabilité de leurs séquences primaires, l'hétérogénéité de leur masse moléculaire a été attribuée à la présence de modifications post-traductionnelles telles que des N-glycosylations et O-glycosylations [175,176].

i. Structure

Bien que la séquence protéique des RNases T2 soit très peu conservée entre les procaryotes et les eucaryotes (seulement 35 % d'identité ont été observés entre bactéries et plantes), leur structure tridimensionnelle reste cependant semblable [175]. Les analyses cristallographiques montrent que les RNases T2 possèdent toutes des résidus hydrophobes à des positions similaires, créant des noyaux hydrophobes. Leur structure conservée est aussi caractérisée par quatre feuillets β antiparallèles ($\beta 1$, $\beta 2$, $\beta 4$ et $\beta 5$), deux plus petits feuillets β antiparallèles ($\beta 3$ et $\beta 7$) et trois hélices α (αB , αC et αD) (Figure 13 A et B). La structure de la protéine est stabilisée en conformation active par des ponts C-C entre les résidus cystéines dont deux qui sont conservés dans toutes les RNases T2 [175].

ii. Activité enzymatique

Les RNases T2 sont des endoribonucléases qui produisent des oligo- ou monoribonucléotides avec une extrémité 3'-phosphate (3'-P) *via* un intermédiaire 2',3' cyclic phosphate (2',3'-cP). D'une manière générale, les RNases T2 n'ont pas de préférences de substrat. Elles sont capables de cliver n'importe quelle base de l'ARN [175,177].

Les RNases T2 possèdent deux régions conservées qui correspondent aux deux sites actifs appelés CAS I et CAS II. Leur séquence consensus est F/WTL/IHGLWP pour CAS I et FWXHEWXKHGTC pour CAS II. Plusieurs études ont montré que les résidus histidines (en gras sur les séquences consensus) sont essentiels pour l'activité enzymatique. Ces résidus His sont d'ailleurs retrouvés dans les parties conservées des RNases T2 à savoir dans le feuillet $\beta 2$ et l'hélice αC [175,177].

Un mécanisme catalytique en deux étapes a été proposé (Figure 13 C)[178,179]. D'abord une première étape de transphosphorylation entre les deux résidus His des sites CAS I et CAS II ayant une extrémité 2',3'-cP en intermédiaire. Puis une deuxième étape d'hydrolyse libérant une extrémité 3'-P et 5'-OH (hydroxyle). Les différentes études montrent que la majorité des RNases T2 ont une activité enzymatique optimale lorsque le pH est acide (entre 4 et 5,5) [175,180]. C'est le cas pour les RNases de champignons, de plantes et d'animaux [176,181,182]. Il existe cependant des RNases T2 qui ont une activité à pH neutre comme Rny1 de levure ou RNS2 d'Arabidopsis [175,183], voire même basique comme celles de bactérie par exemple [184].

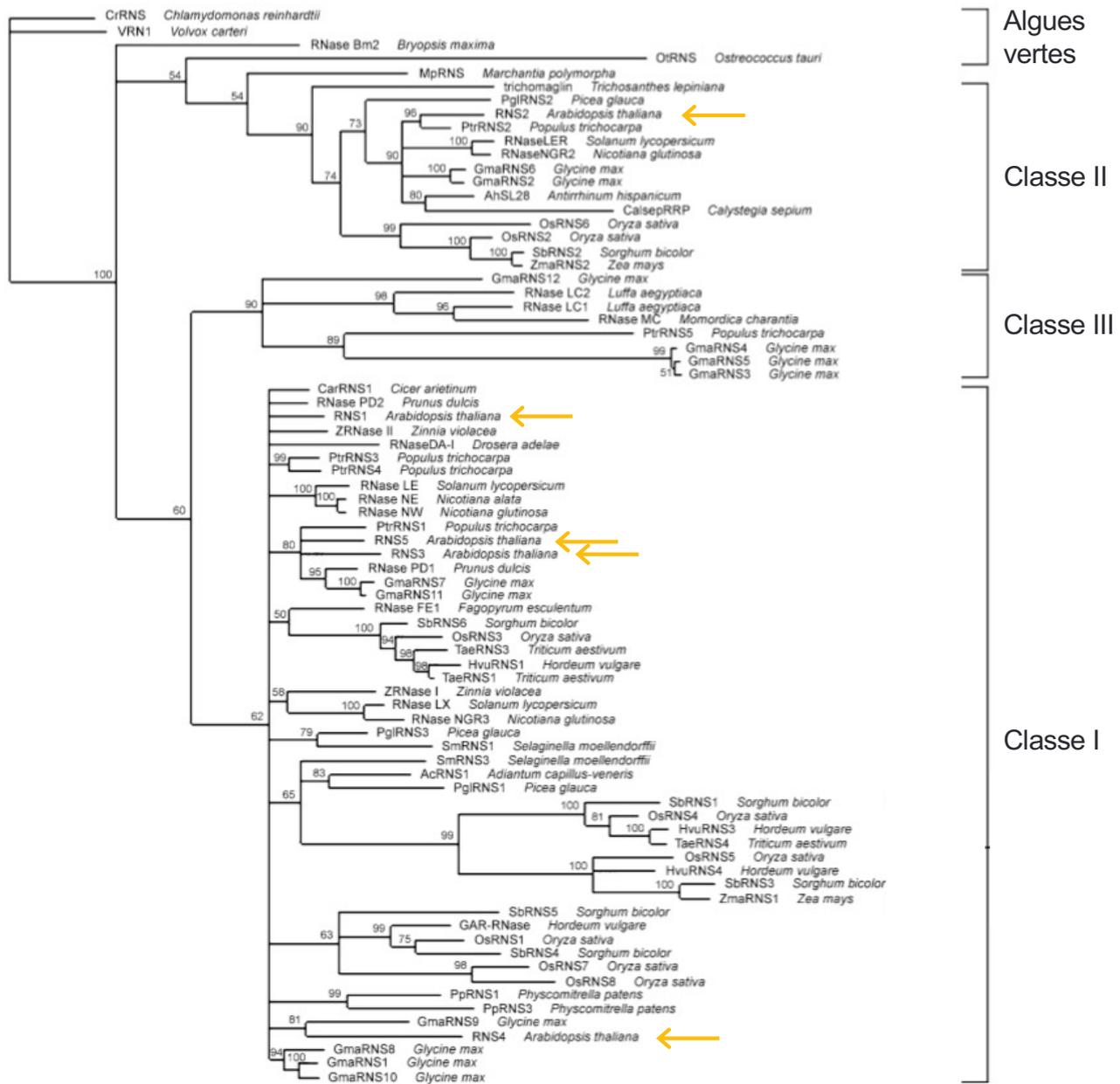


Figure 14 : Arbre phylogénétique des RNases T2 de plantes. Adapté de [185]. Cet arbre a été estimé par la méthode Bayésienne en utilisant uniquement les régions conservées des RNases T2. Les valeurs de Bootstrap > 50% sont indiquées sur les branches. L'arbre est enraciné sur le groupe des algues vertes. Les RNases T2 de classe I, classe II et classe III sont indiquées ainsi que des algues vertes sont indiquées par des accolades. Les RNS d'*A. thaliana* (RNS1, RNS2, RNS3, RNS4 et RNS5) sont indiquées sur l'arbre par des flèches oranges.

iii. *Évolution classe I/II*

Des RNases T2 sont retrouvées chez de nombreux organismes procaryotiques et eucaryotiques mais pas chez les archées.

Chez les plantes, les RNases T2 sont codés par plusieurs gènes, comparé à d'autres organismes comme chez la plupart des espèces animales où un seul gène n'est présent. Chez les plantes, les RNases T2 sont ordonnées en trois classes. Les gènes de la classe I ont connu de nombreux événements de duplications conduisant à une grande diversification selon les espèces (Figure 14). Les gènes de cette classe ont la particularité d'être exprimés de manière spécifique selon les tissus et sont souvent régulés par des stress. Les gènes de la classe III partagent le même ancêtre commun que ceux de la classe I. Ce sont des gènes codants des S-RNases (*Self-incompatibility*) et potentiellement leurs précurseurs sont impliqués dans le processus d'auto-incompatibilité gamétophytique. Les enzymes de la classe II semblent quant à elles avoir conservé des caractéristiques plus ancestrales. Ces gènes sont très conservés chez les plantes et généralement un gène de cette classe est retrouvé par génome. Contrairement aux gènes de la classe I, ceux de la classe II sont constitutivement exprimés [175,185].

2. Les RNS d'*A. thaliana* : des RNases T2 à l'origine de la biogenèse des tRFs

Le génome d'*A. thaliana* code pour 5 RNases T2 : RNS1, RNS2, RNS3, RNS4 et RNS5 (Figure 15). Ces RNases possèdent les deux sites catalytiques CAS I et CAS II conservés chez les RNases T2 [113,186]. RNS2, qui diverge phylogénétiquement des autres, est la seule RNase constitutive de classe II alors que RNS1, RNS3, RNS4 et RNS5 appartiennent à la classe I [187]. Aucune RNase de classe III n'a été retrouvée dans le génome d'*A. thaliana*.

L'ensemble des données sur les RNS, détaillées dans les paragraphes suivants, sont résumées dans la Figure 16.

i. *Profils d'expression des RNS*

Concernant leur expression, les RNS présentent des profils différents. Aucune expression de RNS5 n'a été détecté et RNS4 semble exprimée uniquement dans les racines, à un niveau relativement faible comparé aux RNS1, RNS2 et RNS3 [113]. C'est pourquoi la plupart des études n'ont été réalisées qu'uniquement sur ces trois dernières RNases.

L'étude récente de Megel et al. [113] montre que RNS1 et RNS3 sont exprimées principalement dans les inflorescences. Elles sont notamment plus fortement exprimées dans les graines (au niveau de la chalaze et de l'endosperme) des siliques sénescents. RNS3 est aussi exprimée dans les racines et les tiges. RNS2 présente un profil d'expression différent. Elle est fortement exprimée dans les feuilles sénescents et dans les fleurs. Elle est aussi retrouvée dans les siliques mais durant les premiers stades de formation des siliques contrairement à RNS1 et RNS3 qui sont, elles, exprimées dans les siliques plus âgées.

RNS1	1	<u>MKILLASLCLISLLVILPSVFSASSSSSEDFDFYFVQWPGSYCDTQKKCCYP</u>	53
RNS2	1	<u>MASRLCLLLLVACIAGAFAGDVIELNRSQREFDYFALS LQWPGTYCRGTRHCCSK</u>	55
RNS3	1	<u>MK - - FFIFILALQQLYVQS - - FAQDFDFYFVLQWPGAYCDSRHSCCYP</u>	45
RNS4	1	<u>MRGI IIVSLLILQSLVSSSQTEPDFNFYVWNYWPGAICDSQKGCCPP</u>	49
RNS5	1	<u>MGAKGCVN - - VLLKLLVFQGLFVSR - - PQEDDFYFVLQWPGAYCDTSRACCP</u>	51
Signal peptidique			
RNS1	54	- - - - NSGKPAADFGIHLWPNYKDGTYPSNCDASKPFDSSTISDLLTSMKK - - - S	101
RNS2	56	NACCRGSDAPTQFTIHGLWPDYNDGSWPSCCYR - SDFKEKEI STLMDGLEK - - - Y	106
RNS3	46	- - - - QTGKPAADFGIHLWPNYKTGGWPQNCNPDSRFDDL RVSDLMSDLQR - - - E	93
RNS4	50	- - - - TKGNTASDFI IHGLWPQFNNGTWP AFC DQTNLFDISKISDLVCQMEKKWTE	100
RNS5	52	- - - - TSGKPAADFGIHLWPNYNGGSWP SNCDPDSQFDRSQISDLVSSLKK - - - N	99
CAS I			
RNS1	102	WPTLACP - - - - SGSGEAFWEHEWEKHGTCSESVI - DQHEYFQTALNLKQKTNLL	150
RNS2	107	WPSLSCGSPSSCNGGKGSFWGHEWEKHGTCSSPVFHDEYNYFL TTLNLYLKHNVT	161
RNS3	94	WPTLSCP - - - - SNDGMKFWTHEWEKHGTC AESEL - DQHDYFEAGLKLKQKANLL	142
RNS4	101	WGVWACP - - - - SNE - TNLWEHEWNKHGTCVQSIF - DQHSYFRTNLKFHKVHLL	148
RNS5	100	WPTLSCP - - - - SNEG FNFWEHEWEKHGTCSE SVM - DQHEYFENALKLQKANLL	148
CAS II			
RNS1	151	GALTKAG - INPDGKSYSLESIRDSIKESIGFTPWVECNRDGSGNSQLYQVYLCVD	204
RNS2	162	DVLYQAGYVASNSEKYPLGGIVTAIQNAFHITPEVVCCKRDAID - - - EIRICFYKD	213
RNS3	143	HALTNAG - IKPDDK FYEMKDIENTIKQVVGFAPIECNKDSSHNSQLYQIYLCVD	196
RNS4	149	NILIQKG - IKPNDGFYSLDEIKNAIKCAIGFAPGIECNEDVKGNKQLFQIYICLD	202
RNS5	149	QILKNSG - INPDDGFYNLDKITNAIKDGIGFTPGIECNKDPERNAQLHQIYICVD	202
RNS1	205	RSGSGLIECP - VFPHGKCGAEIEFP SF	230
RNS2	214	FKPRDCVGSQDLTSRKSCP KYVSLPEYTPLDGEAMVLKMP TEREAL	259
RNS3	197	TSASKFINCP - VMPHGRCDSRVQFPKF	222
RNS4	203	NYAKEFVECP - YVPDKSCASKIKF PSLPERDSLNESLSVMSVSTT	247
RNS5	203	TSGTEFIECP - VLPRGSCPSQIQFSKF	228

Figure 15 : Alignement des séquences protéiques des RNS d'*A. thaliana*. L'alignement des séquences a été réalisé avec le programme d'alignement multiple ClustalW et visualisé avec le logiciel MacVector. Les séquences du signal peptidique prédites par le serveur SignalP sont soulignées en orange. Les séquences des sites catalytiques CAS I et CAS II sont encadrés en cyan. Les résidus Histidines importants pour l'activité catalytique sont indiqués par des flèches de couleur cyan.

Ces profils d'expression pour les ces trois RNS sont en concordance avec la littérature plus ancienne [186,187].

En plus de leur expression lors du développement de la plante, les RNS d'*A. thaliana* présentent aussi des profils d'expression différents en cas de stress. Il a été montré par plusieurs équipes indépendantes que RNS1, et RNS2 dans une moindre mesure, s'expriment et s'accumulent fortement lors d'une carence en phosphate organique (Pi) [113,188]. Les premières études suggéraient que RNS1 était impliquée dans le recyclage du phosphate contenu dans les molécules d'ARN pour pouvoir le remobiliser en cas de carence, ce qui expliquait pourquoi elle était fortement induite [188]. Chez des lignées où l'expression de RNS1 ou RNS2 est inhibée, il a même été observé une plus forte accumulation d'anthocyanes lors de carence en Pi. Il a été suggéré que puisque les RNS n'étaient plus présentes et ne permettaient plus la remobilisation du phosphate, la réponse au stress était alors encore plus importante, expliquant la forte accumulation d'anthocyanes [189]. Depuis, les résultats de Megel et al. [113] montrent que l'induction de RNS1 en cas de carence en Pi provoque une plus forte activité de clivage de certains ARNt. Ceci conduit à une accumulation de tRFs en particulier, notamment les tRF-5D et -5A issus de l'ARNt^{Gly} (UCC) et le tRF-5A issu de l'ARNt^{Asp} (GUC). Ces résultats sont d'ailleurs en concordance avec une précédente étude réalisée par Hsieh et al. [190].

ii. Localisation subcellulaire des RNS

Les analyses *in silico* des séquences de RNS1, RNS2 et RNS3 montrent la présence de domaines N-terminaux de sécrétion *via* un transit par le réticulum endoplasmique (RE) et l'appareil de Golgi. RNS2 est la seule à avoir une extension en C-terminal d'adressage à la vacuole [187,188]. Par la suite, Bariola et al. [189] ont montré par l'analyse des composants des compartiments cellulaires de protoplastes que RNS1 a une localisation extracellulaire alors que RNS2 reste intracellulaire. La localisation de RNS2 à la fois dans le RE et dans la vacuole a d'ailleurs été confirmée dans des protoplastes issus de feuilles adultes d'*A. thaliana* [183]. Il a aussi été montré que des protéines RNS2 tronquée pour la partie C-terminale, mais toujours fonctionnelles, ne s'accumulaient plus autant dans la vacuole. Elles sont retrouvées sécrétées après transit par le RE, tout comme RNS1 et RNS3 [191].

Finalement, ces résultats de localisation subcellulaire des RNS a été confirmée par mon équipe à l'aide de constructions où les RNS fusionnées à la GFP en C-terminal ont été exprimés transitoirement dans des feuilles de *N. benthamiana*. La localisation au RE a pu être confirmée pour RNS1 et RNS3. Pour RNS2, la localisation à la fois au RE et dans la vacuole a aussi pu être observée [192].

	Expression (conditions normales)										Expression (stress)	Localisation subcellulaire		Génère des tRFs	
	Ra	Hyp	Ti	Feu j.	Feu s.	Fle	Sil v.	Sil j.	Sil s.	Gra	-Pi	RE	Vacuole	Courts	Longs
RNS1	-	-	-	-	-	+	-	++	-	++	++	✓		✓	✓
RNS2	-	-	-	+	++	++	++	+	-	-	+	✓	✓	✓	✓
RNS3	+	+	+	-	-	+	+	++	+	++	-	✓		✓	✓
RNS4	++	++	-	-	-	-	-	-	-	-	-	na	na	na	na

Figure 16 : Tableau récapitulatif des données sur les RNS d'*A. thaliana*. Les données sont issues de la littérature [112, 113, 186-188, 190, 192] et de résultats personnels de l'équipe. Ra : racine, Hyp : hypocotyle, Ti : tige, Feu j : feuille jeune : Feu s : feuille sénescence, Fle : fleur, Sil v : silique verte, Si j : silique jaunissante, Sil s : silique sénescence, Gra : graine, -Pi : carence en phosphate inorganique, RE : réticulum endoplasmique, - : pas exprimée, + : exprimée, ++ : fortement exprimée, na : donnée non disponible.

Ce signal d'adressage en C-terminal des RNases T2 de classe II est conservé chez les plantes (gymnospermes et angiospermes) et les enzymes de cette classe possèdent une localisation vacuolaire et au RE. Les RNases de classe I sont retrouvées sécrétées dans l'apoplaste sauf pour quelques exceptions qui possèdent une extension en C-terminal et se rapprocheraient des enzymes de classe II [191].

iii. Implication des RNS dans la biogenèse des tRFs

Comme évoqué dans l'introduction (II.2.iii), Martinez et al. [111] proposent que des tRFs d'*Arabidopsis* puissent être générés par des endonucléases de type Dicer-like (DCL). Toutefois, les travaux de l'équipe [113] ainsi que ceux d'Alves et al. [112] démontrent que chez *A. thaliana*, ce sont des RNases T2 qui sont majoritairement à l'origine de la biogenèse des tRFs. En effet, par l'utilisation de mutants *Knock-Out (KO)* pour les RNS, il a pu être observé que RNS1, RNS2 et RNS3, les trois principales RNS fortement exprimées, sont capables de générer des tRFs courts et longs [113]. Dans leurs conditions d'expériences, les DCL ne sont pas impliquées dans la biogenèse des tRFs étudiés.

Ces résultats ont été obtenus avec l'utilisation de lignées simple *rns^{KO}* (*rns1^{KO}*, *rns2^{KO}*, *rns3^{KO}*) et d'une lignée double *rns1rns3^{KO}*. Or, à l'heure actuelle, il n'existe pas de mutant double *rns1rns2^{KO}* et *rns2rns3^{KO}* ni même de mutant triple *KO* pour les trois principales RNS (*rns1rns2rns3^{KO}*). De plus les lignées *rns1^{KO}* et *rns3^{KO}* sont sur fond génétique Ws alors que la lignée *rns2^{KO}* est sur fond génétique Col-0.

Pour aller plus loin dans la caractérisation des RNS et leur rôle dans la biogenèse des tRFs, un des objectifs de ma thèse était de réaliser ces mutants manquants. Les résultats associés à ce travail sont présentés dans les paragraphes suivants.

II. Obtention de lignées *rns^{KO}*

1. Stratégie employée : CRISPR/Cas9

Les techniques d'édition du génome permettent de modifier des séquences ciblées du génomes contrairement aux technologies plus traditionnelles de mutations aléatoires. Pour éditer le génome, il y a besoin de nucléases qui reconnaissent et clivent des régions spécifiques. Ces nucléases peuvent être des méganucléases, des nucléases à doigt de zinc (ZFN), des nucléases *transcription Activator-Like Effector-based Nucleases* (TALEN) ou encore les nucléases *Clustered Regularly Interspaced Short Palindromic Repeats associated nucleases* (CRISPR/Cas) [193]. C'est la méthode basée sur cette dernière nucléase qui a été employée pour créer les mutants *rns^{KO}*.

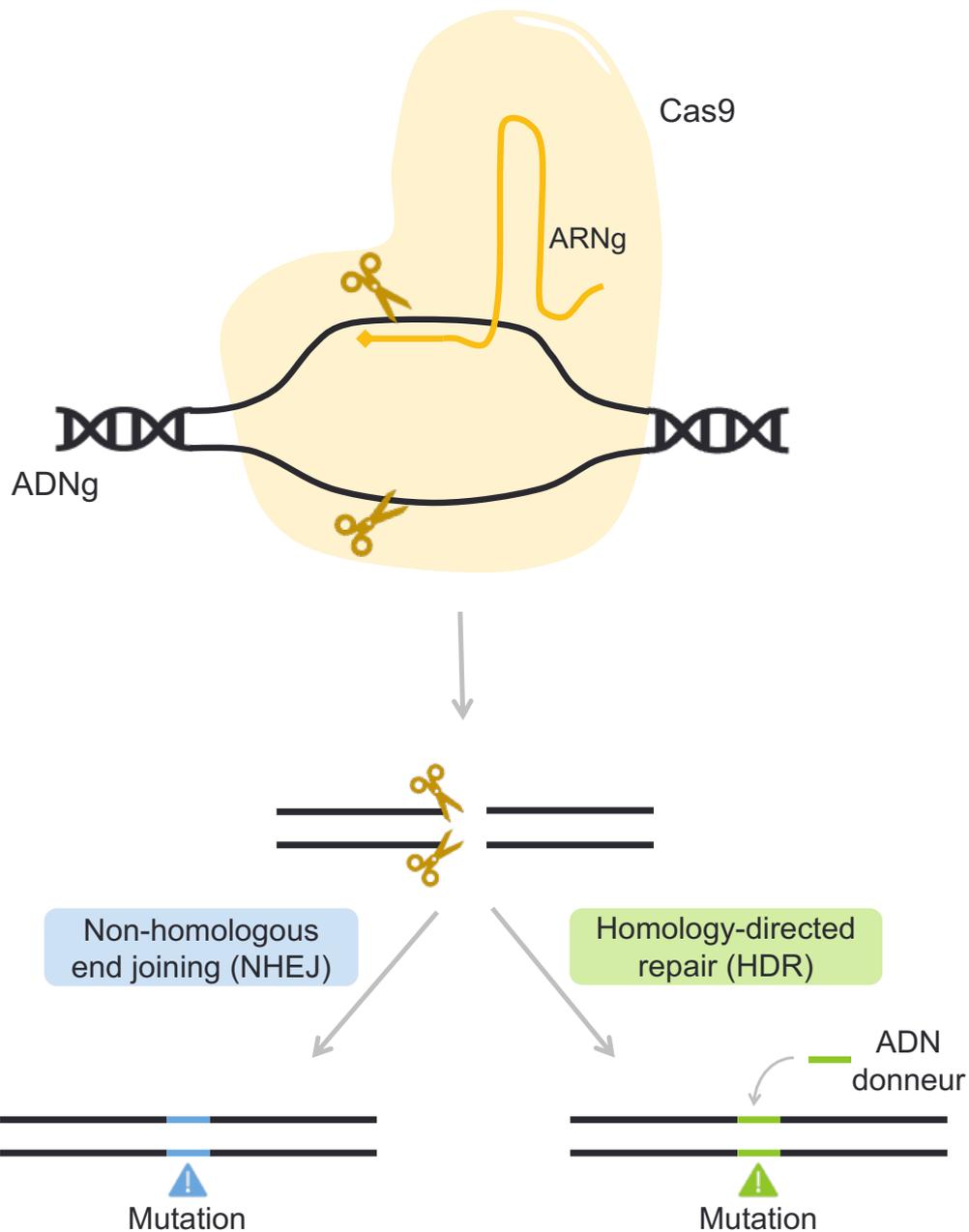


Figure 17 : Schéma récapitulatif d'édition du génome par la technologie CRISPR/Cas9. Formation d'un complexe composé de la Cas9 et de l'ARN guide (en jaune). L'ARN guide est représenté par une ligne jaune avec à son extrémité le motif PAM de reconnaissance (losange jaune). Une fois l'ADN génomique clivé par le complexe, il est réparé soit par la voie *Non-Homologous End Joining* (NHEJ) qui peut introduire des mutations par erreur (en bleu) ; soit par la voie *Homology-Directed Repair* (HDR) qui introduit la séquence de l'ADN donneur au moment de la réparation (en vert).

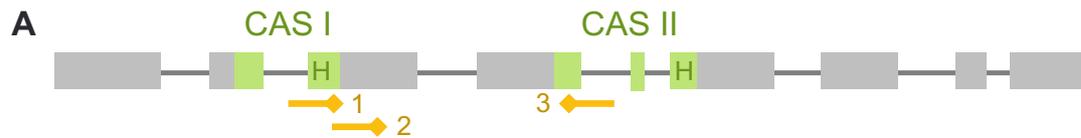
L'édition de génome par CRISPR/Cas a été premièrement compris chez les bactéries et les archées. Ce système existe naturellement chez ces organismes et est en fait un mécanisme de défense contre de l'ADN étranger introduit lors d'infections, majoritairement par des virus. Il a ensuite été détourné pour être utilisé comme outil pour l'édition du génome [194]. Cette découverte a fait l'objet d'un prix Nobel de chimie en 2020, attribué à Emmanuelle Charpentier et Jennifer Doudna.

Les endonucléases Cas possèdent des domaines nucléases qui leur permettent de cliver de l'ADN double brin et de former une cassure (*double-strand break*, DSB). L'ADN endommagé par les Cas est ensuite réparé par la machinerie cellulaire de réparation de l'ADN. Il existe deux types de réparations qui peuvent avoir lieu : la réparation *Non-Homologous End Joining* (NHEJ) et la réparation *Homology-Directed Repair* (HDR). La réparation NHEJ n'est pas toujours parfaite et induit des mutations aléatoires (délétions, insertions, substitution de nucléotides). La réparation HDR quant à elle utilise un fragment d'ADN (simple ou double brin) donneur qui sera inséré à l'endroit de la cassure au moment de la réparation (Figure 17) (pour une revue, voir [195]).

Les endonucléases Cas peuvent cibler des séquences spécifiques du génome. Pour cela deux conditions sont nécessaires. Les Cas ont besoin d'un guide qui va leur indiquer quelle région/séquence du génome cibler. Ce sont des ARN, que l'on appelle ARN guides (ARNg), qui vont jouer ce rôle. Les ARNg sont composés de deux parties. La première partie de la séquence est reconnue par les Cas et permet la prise en charge des ARN par les Cas. La seconde partie de l'ARNg est une séquence complémentaire à la séquence d'ADN cible. C'est cette deuxième partie qui est défini par l'expérimentateur pour cibler la séquence d'ADN qu'il souhaite modifier. Cependant, toutes les régions de l'ADN ne peuvent être ciblées. Il faut que la deuxième condition soit remplie. La deuxième condition nécessaire est la présence d'un motif particulier adjacent à la séquence d'ADN ciblée. Ce motif, appelé PAM (*Protospacer Adjacent Motif*) sert de signal aux Cas pour démarrer leur activité nucléolytique. A l'heure actuelle, il existe différentes nucléases Cas identifiées chez de nombreux organismes qui présentent des spécificités de clivage et reconnaissent différents motifs PAM [195].

i. Conception des ARN guides in silico

Pour l'obtention de mutants *rns^{KO}*, la Cas9 de *Streptococcus pyogenes* a été employée et un seul ARNg pour modifier la séquence d'ADN d'intérêt a été utilisé. Cette SpCas9 nécessite le motif PAM 5'-NGG-3' présent sur l'ADN cible où N peut être n'importe quel nucléotide. Le clivage par la SpCas9 a lieu 3 ou 4 nucléotides en amont (en 5') du motif NNG.



B

ARN guide	Sens	PAM	<i>Off-targets</i> prédits ≤ 12 pb du PAM	<i>Off-targets</i> prédits > 12 pb du PAM
ARNg 1	fw	TGG	0	6
ARNg 2	fw	TGG	0	3
ARNg 3	rv	TGG	0	4

Figure 18 : Conception des ARN guides. A : Représentation schématique de l'ADNg du gène *RNS2* de 5' en 3' (de gauche à droite). Les introns sont représentés par une ligne fine, les exons par des rectangles. Les portions d'exons en vert symbolisent les régions codant les sites catalytiques CAS I et CAS II avec les résidus Histidines (H, en vert) essentiels pour leur activité. Les régions ciblées par les ARN guides sont matérialisées par une ligne orange dont l'extrémité en losange représente le site PAM. B : Tableau récapitulatif des caractéristiques des ARN guides : le sens, le site PAM, les *off-targets* prédits par CRISPOR dans les 12 paires de bases (pb) adjacentes au site PAM et les autres *off-targets* prédits.

Étant donné que les mutants *rns1^{KO}*, *rns3^{KO}* et *rns1rns3^{KO}* étaient déjà présents dans l'équipe, il a été décidé de partir de ces lignées pour obtenir les mutant *KO* manquants. Pour cela, il faut cibler le gène *RNS2* par CRISPR/Cas9. A l'aide de l'outil informatique CRISPOR, j'ai choisi des ARNg qui ciblent au maximum les séquences codants les sites actifs CAS I et CAS II de *RNS2*. Les ARNg 1 et ARNg 2 ciblent le site CAS I. Le premier cible la jonction entre l'intron 2 et l'exon 3 et le deuxième guide cible uniquement l'exon 3. L'ARNg 3 cible lui le site CAS II et plus particulièrement la jonction entre l'exon 4 et l'intron 4 (Figure 18). Ces ARN guides ont été principalement choisi en fonction de la région ciblée sur le gène *RNS2* et pour leur faible prédiction de *off-targets*.

ii. Transformation de plantes d'*A. thaliana* par des agrobactéries avec le plasmide pKI1.1R

Pour éditer le génome par la technologie CRISPR/Cas, il est nécessaire de délivrer l'enzyme Cas et l'ARNg au sein des cellules. Ici, la méthode traditionnelle d'infection par *Agrobacterium tumefaciens*, couramment employée chez les plantes, a été employée pour délivrer ces composants dans les cellules. Le plasmide pKI1.1R [163] spécialement conçu pour l'édition du génome par CRISPR/Cas9 a été utilisé. Ce plasmide permet l'expression de l'ARNg sous le contrôle du promoteur AtU6 et l'expression de la SpCas9 sous le contrôle du promoteur AtRPS5A. Ce plasmide a aussi la particularité d'exprimer le gène OLE1, une protéine impliquée dans l'accumulation de réserves lipidiques dans les graines, fusionnée à la tRFP, sous le contrôle de son propre promoteur. Cette protéine de fusion permet de rendre les graines fluorescentes lorsque l'ADN-T a été inséré dans le génome. Cela reflète la présence du gène de la SpCas9 dans le génome de la plante.

Ce plasmide permet une transformation des plantes par *floral dip*. Les graines T₁ issues de la transformation et ayant intégré l'ADN-T sont fluorescentes. Les plantes générées à partir de ces graines ont donc la SpCas9 exprimée et fonctionnelle. La Cas9 clive au niveau des sites définis par les ARNg et des erreurs lors de réparation de type NHEJ peuvent apparaître. Les plantes de cette génération T₁ génèrent des graines qui, par ségrégation, contiennent encore l'ADN-T ou l'ont perdu. Ceci se reflète par la fluorescence des graines. Les graines T₂ sont alors sélectionnées de façon à ne garder que des grines non fluorescentes. Ces graines portent les mutations éventuelles causées par la Cas9 mais la ségrégation a fait qu'ils n'ont pas hérité des allèles portant l'ADN-T dans leur génome. Cette génération T₂ a potentiellement le gène *RNS2* édité et ne possède plus de Cas9 exprimée (Figure 19).

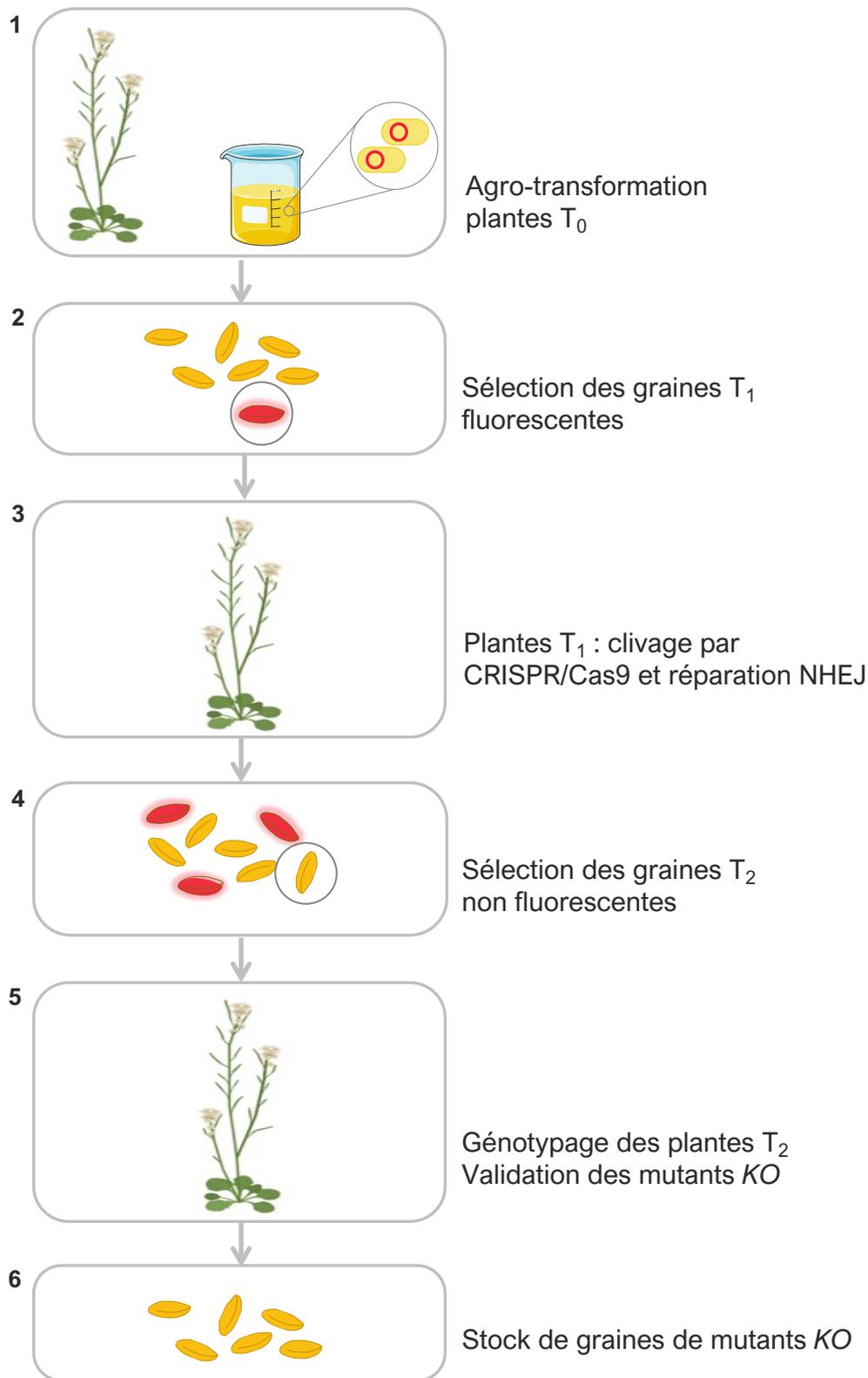


Figure 19 : Schéma des étapes de sélection pour l'obtention de lignées stables modifiées par CRISPR/Cas9. 1 : Les plantes T₀ sont transformées par floral-dip avec des *A. tumefaciens* possédant le plasmide pKI1.1R. 2 : Les graines T₁ issues des plantes T₀ exprimant la RFP sont sélectionnées. 3 : Les plantes T₁ ont la Cas9 exprimée qui va cliver l'ADN ciblé par les ARNg et des erreurs vont apparaître lors de la réparation NHEJ. 4 : Les graines T₂ issues des plantes T₂, n'exprimant pas la RFP sont contre-sélectionnées. 5 : Les plantes T₂ sont génotypées par PCR puis par séquençage. Les mutants KO identifiés sont validés par western blot. 6 : Les graines T₃ issues des plantes T₂ servent de stock de graines de mutants.

Ainsi, les lignées *Ws*, *rns1^{KO}*, *rns3^{KO}* et *rns1rns3^{KO}* ont été transformées pour générer les mutants *rns2^{KO}*, *rns1rns2^{KO}*, *rns2rns3^{KO}* et *rns1rns2rns3^{KO}* respectivement. Pour optimiser les chances d'obtenir des individus mutés, chaque lignée de départ a été transformée avec les 3 ARNg (4 plantes par ARNg). La sélection des générations T₁ et T₂ a été réalisée en suivant les étapes décrites précédemment et dans la Figure 19. En plus de la sélection des graines basée sur leur fluorescence, l'absence de la Cas9 dans le génome des plantes T₂ a été vérifiée par PCR.

2. Résultats des lignées *rns^{KO}* obtenues

Les plantes ayant passé l'ensemble des étapes de sélection ont été séquencées pour le gène *RNS2* au niveau de la zone ciblée par les ARNg.

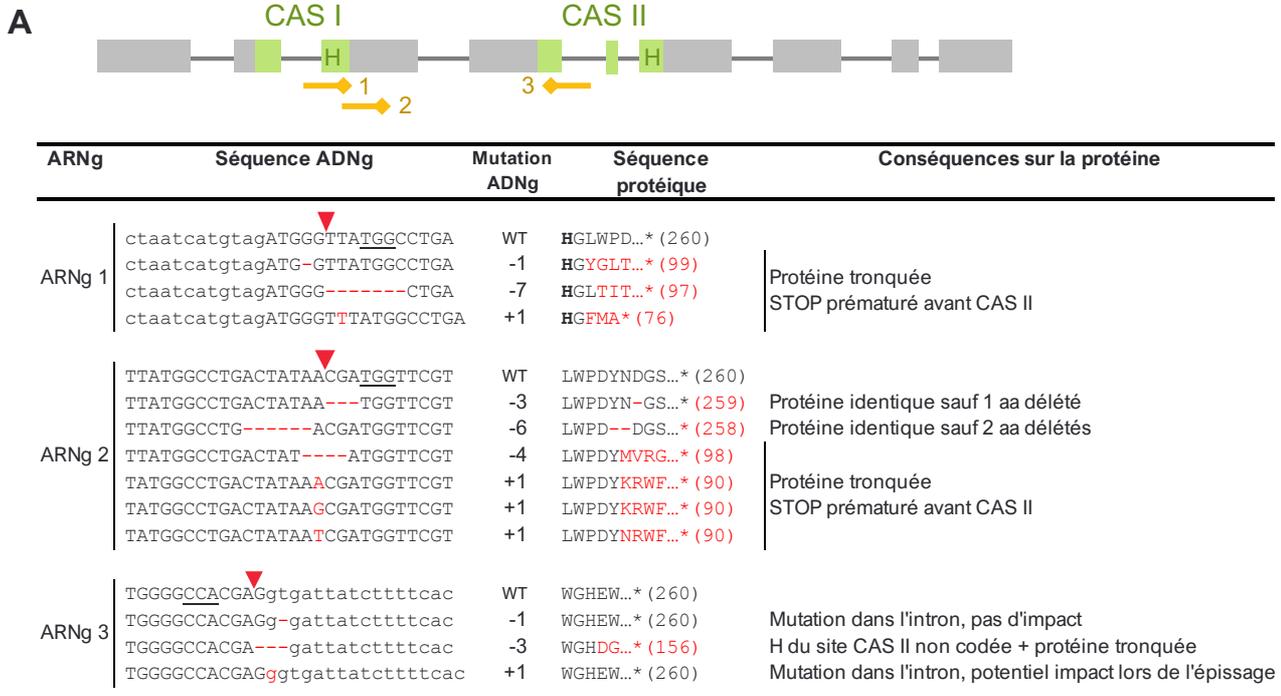
i. Résultats de séquençage des mutants obtenus

Le séquençage de la zone ciblée par les ARNg sur le gène *RNS2* a permis de mettre en évidence des mutations dues à une mauvaise réparation après clivage par la Cas9. Globalement, les 3 ARNg ont été efficaces et ont ciblé les zones prédites. Les différents types de mutation retrouvées dans ces zones sont détaillés dans la Figure 20 A.

L'ARNg 1 a engendré trois mutations différentes (délétion de 1 ou 7 nucléotides et insertion de 1 nucléotide). Ces mutations engendrent un décalage du cadre de lecture lors de la traduction. La séquence en acides aminés est modifiée à partir du site CAS I et un codon STOP prématuré génère une protéine tronquée, sans le deuxième site CAS II.

L'ARNg 2 a engendré des mutations qui conduisent à la traduction de deux types de protéines. Soit la délétion est un multiple de 3 (délétion de 3 ou 6 nucléotides) et ne génère pas de décalage du cadre de lecture. Il manque alors 1 ou 2 acides aminés à la protéine prédite mais le reste de la séquence est inchangé. Étant donné que les acides aminés manquants sont situés dans le site catalytique CAS I, il se peut que la conformation et/ou la stabilité du site catalytique soit modifiée(s) et que la protéine ne soit plus fonctionnelle. Les autres mutations engendrées par l'ARNg 2 (délétion de 4 nucléotides et insertion de 1 nucléotide) modifient le cadre de lecture et un codon STOP apparaît prématurément, formant une protéine tronquée sans le site CAS II.

L'ARNg 3, qui cible quant à lui le site CAS II, a engendré des mutations à la jonction d'un exon et d'un intron. Une délétion de 3 nucléotides, dont 1 dans l'exon, laisse supposer que cela génère un décalage du cadre de lecture lors de la traduction. Si cela s'avère juste, le résidu histidine, essentiel à l'activité du site CAS II n'est plus codé et la protéine sera tronquée (dû à un codon STOP prématuré). Les mutations dans l'intron (délétion ou insertion de 1 nucléotide) n'ont peut-être aucun impact sur la séquence codante. Cependant, étant proches de la jonction exon-intron, il se peut que ces mutations aient un impact lors du mécanisme d'épissage de l'ARNm.



B

Lignée de départ	ARNg	Types de mutation	Lignée mutante stable
Ws	ARNg 1 ARNg 2 ARNg 3	2 2 0	4 <i>rns2^{KO}</i>
<i>rns1^{KO}</i>	ARNg 1 ARNg 2 ARNg 3	0 2 2	4 <i>rns1rns2^{KO}</i>
<i>rns3^{KO}</i>	ARNg 1 ARNg 2 ARNg 3	1 4 2	7 <i>rns2rns3^{KO}</i>
<i>rns1rns3^{KO}</i>	ARNg 1 ARNg 2 ARNg 3	0 1 0	1 <i>rns1rns2rns3^{KO}</i>

Figure 20 : Tableaux récapitulatifs des lignées obtenues par CRISPR/Cas9. A : Tableau montrant les différents types de mutations retrouvées sur le gène *RNS2* au niveau des régions ciblées par les ARN guides dans les plantes T₂. Les lettres en minuscules correspondent aux introns, en majuscule aux exons et celles soulignées dans la séquence WT correspondent au motif PAM. Le site de clivage prédit est indiqué par un triangle rouge. Les mutations sont indiquées en rouge. Les étoiles * représentent les codons STOP et leur position est indiquée entre parenthèses. B : Tableau récapitulatif des lignées transformées et du nombre de lignées stables *rns^{KO}* obtenues.

Chaque lignée de départ (*Ws*, *rns1^{KO}*, *rns3^{KO}* et *rns1rns3^{KO}*) a été transformée avec chaque ARNg (séparément) pour engendrer des mutations différentes sur un même génotype. Ainsi, au final, 4 mutants différents *rns2^{KO}*, 4 mutants *rns1rns2^{KO}*, 7 mutants *rns3rns3^{KO}* et 1 mutant *rns1rns2rns3^{KO}* ont été obtenus (Figure 20 B).

ii. Validation des mutants *rns^{KO}* par western blot

Pour valider au niveau protéique l'absence de RNS2, des western blots à partir d'inflorescences des mutants identifiés ont été réalisés. Un western blot représentatif est montré dans la Figure 21. Les lignées de départ, utilisées pour le floral-dip, ont été utilisées comme contrôle positif. La bande correspondant à RNS2 est retrouvée dans les contrôles mais pas dans les mutants identifiés. Finalement, pour l'ensemble des prédictions de séquences protéiques mutées et/ou tronquées, la protéine RNS2 n'a pas été détectée par western blot. Les protéines devenues aberrantes sont certainement dégradées et ne sont pas présentes dans ces mutants. Étant donné que les anticorps utilisés pour les western blots (@RNS2) sont polyclonaux, il est peu probable que les anticorps ciblent uniquement un seul épitope qui serait absent sur l'extrémité tronquée de la protéine.

Les mutants identifiés par séquençage et vérifiés par western blots ont été validés comme étant *rns2^{KO}*. Ils ont été mis en culture pour obtenir des stocks de graines prêts pour être utilisés comme outil génétique pour la compréhension de la biogenèse des tRFs.

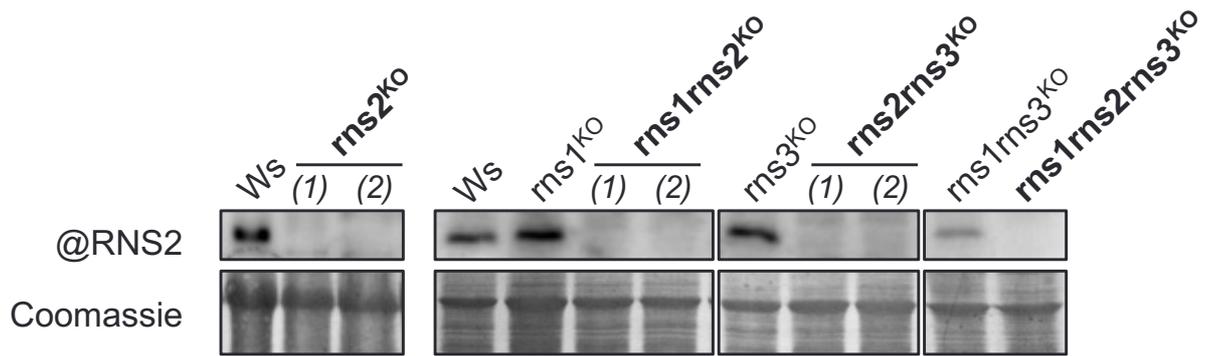


Figure 21 : Validation des lignées obtenues par CIRSPR/Cas9. Western blot avec l'anticorps RNS2 sur des extraits protéiques d'inflorescences. Deux lignées représentatives, (1) et (2), sont montrées pour chaque lignée obtenue. Pour chaque lignée générée, la lignée de départ est utilisée comme contrôle positif. Coloration des membranes au bleu de Coomassie comme contrôle de charge.

III. Caractérisation des mutants *rns*^{KO} – résultats préliminaires

Comme mentionnées dans les paragraphes précédents, plusieurs études ont montré une induction de l'expression de RNS1, et plus faiblement RNS2, lorsque que la plante est soumise à une carence en phosphate [113,188]. Il a aussi pu être montré, par northern blot et par analyse de banques de petits ARN, que certains tRFs sont accumulés lors de telles conditions. On retrouve notamment une accumulation du tRF-5D et tRF-5A issu de l'ARNt^{Gly} (UCC) et du tRF-5A de l'ARNt^{Asp} (GUC) [113,190]. Ce n'est pas le cas pour tous les tRFs. Par exemple, les tRF-5D et -5A issus de l'ARNt^{Ala} ont un profil similaire en présence ou en absence de phosphate. De plus, il a été montré, par l'utilisation du mutant *rns1*^{KO}, que RNS1 était responsable, au moins en partie, de cette forte accumulation de tRFs particuliers lors de carence en phosphate [113].

Le phosphate est un élément essentiel à tout organisme vivant y compris les plantes. Elles ont développé des mécanismes et des voies de régulation pour pouvoir contrôler l'homéostasie et s'adapter à leur environnement en cas de carences par exemple en phosphate inorganique (Pi, forme assimilable). Entre autres, pour augmenter l'assimilation du Pi, les plantes sécrètent au niveau de leurs racines des acides organiques qui vont solubiliser les complexes de phosphore présents dans le sol ; ou encore elles vont augmenter la surface de prospection du sol par l'intermédiaire d'un champignon lors de symbiose mycorhizienne. Pour pallier une telle carence, différents métabolismes vont aussi permettre de recycler le Pi présent dans les molécules au sein de leurs cellules pour le remobiliser [196].

Le système racinaire est un organe clé dans le signalement et l'assimilation du Pi. Il est capable de détecter la quantité de phosphate présent dans le sol et d'assimiler le Pi disponible à proximité. Lors d'une carence en phosphate, la croissance de la racine primaire est stoppée alors que la croissance des racines secondaires et poils racinaires sont favorisées pour permettre à la plante d'avoir une plus grande surface de prospection du sol [196,197].

Étant donné que les RNS sont induites lors de carence en phosphate, elles pourraient avoir un rôle important dans l'adaptation de la plante à un tel stress. Pour initier le travail de caractérisation des mutants *rns*^{KO} générés par la technologie CRISPR/Cas9, la carence en phosphate a été utilisée pour observer leur phénotype et étudier les populations de tRFs.

1. Analyse phénotypique

En conditions normales de croissance, à l'échelle macroscopique, l'ensemble des mutants présentent un développement et une croissance similaires au sauvage Ws. Tous les mutants sont viables et produisent des graines fertiles. Aucun phénotype particulier n'a donc été observé chez les mutants *rns*^{KO}, y compris chez le mutant triple *rns1rns2rns3*^{KO}.

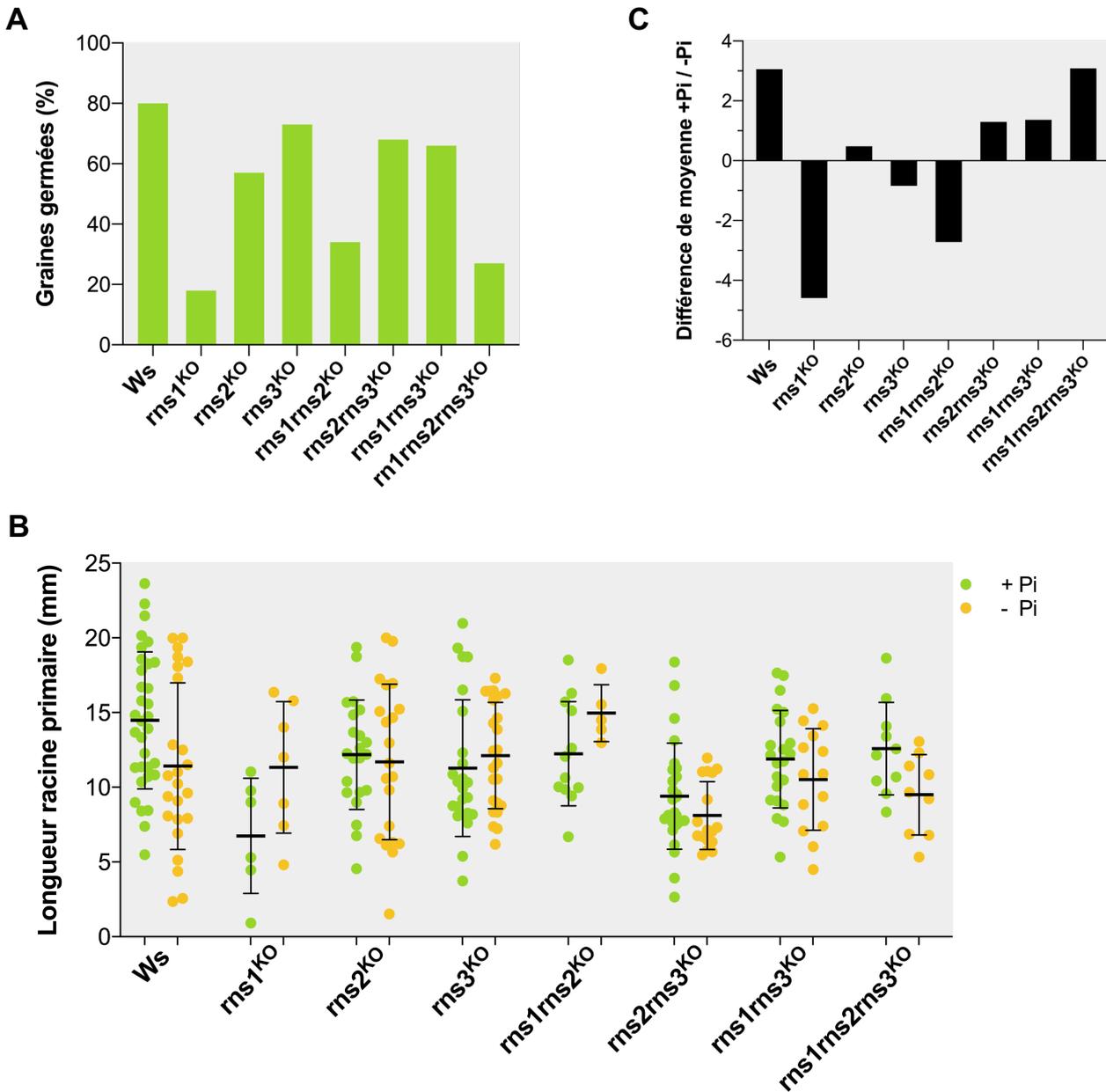


Figure 22 : Analyse phénotypique des mutants rns^{KO} lors d'une carence en phosphate. A : Histogramme représentant le pourcentage de graines germées *in vitro* sur milieu agar contenant du Pi. Après 7 jours de culture les graines ayant germées ont été comptabilisées. Cette expérience n'a été réalisée que sur un seul lot de graines. B : Graphique en nuage de points représentant la longueur de la racine primaire de plantules à 6 jours de culture *in vitro* sur milieu agar en présence de Pi (+ Pi, points verts) ou en absence de Pi (- Pi, points en jaunes). Les barres noires représentent la moyenne et l'écart-type. Pour chaque plantule, la valeur de sa longueur de racine primaire est montrée par un point. Au 6^e jour de culture, la longueur de la racine primaire a été mesurée avec le plugin NeuronJ de ImageJ. Cette expérience a été réalisée à partir d'un seul lot de graine. C : Analyse des différences de longueur de racines primaires observées en B. Chaque barre représente la différence entre la moyenne + Pi et celle en - Pi pour chaque mutant rns^{KO} .

Comme RNS1 et RNS3 sont préférentiellement exprimées dans les graines, en particulier au niveau de l'endosperme, nous avons voulu étudier la germination des lignées *rns*^{KO}. Les graines ont été semées *in vitro* en milieu contenant du Pi. Au bout de 7 jours de culture, les graines germées et non germées ont été comptées. Les résultats sont montrés sous forme d'histogramme dans la Figure 22 A. Le sauvage Ws a un taux de germination des graines proche de 80% en conditions normales. Le taux de germination des mutants *rns2*^{KO}, *rns3*^{KO}, *rns2rns3*^{KO} et *rns1rns3*^{KO} reste proche de celui du sauvage Ws avec 57%, 73%, 68% et 66% respectivement. En revanche, on observe un taux de germination très faible pour *rns1*^{KO}, *rns1rns2*^{KO} et *rns1rns2rns3*^{KO} (moins de 20%, 35% et 30% respectivement). Le point commun de ces lignées est l'absence de RNS1.

À ce stade il est encore trop tôt pour conclure une éventuelle importance de RNS1 dans la germination des graines. L'expérience a été réalisée une seule fois avec un seul lot de graines. Les lots de graines sont assez hétérogènes. Ils proviennent de plantes cultivées simultanément en conditions de cultures contrôlées, cependant nous ne pouvons exclure des variations d'un lot à un autre. La quantité de graines viables produites est dépendante des conditions de cultures (température, arrosage, contamination de champignons, attaques de parasites, ...). Pour s'assurer que le phénomène de germination observé est bien dépendant uniquement de la présence ou absence des RNS, il faudra répéter cette expérience plusieurs fois avec des lots de graines indépendants.

Puisque les RNS, notamment RNS1 et RNS2, sont induites lors de carence en phosphate, nous avons voulu regarder l'impact qu'elles pouvaient avoir dans l'adaptation des plantes à ce stress. Pour cela, le phénotype racinaire a été observé. Les graines ont été semées *in vitro* sur milieu agar contenant du Pi ou non. A partir du 3^e jour de culture, la longueur de la racine primaire de chaque plantule a été mesurée de façon journalière. Le graphique représenté dans la Figure 22 B montre la moyenne des longueurs des racines primaires au 6^e jour après germination. Chez la plante sauvage Ws, des racines primaires plus grandes en présence de Pi (14,5 mm en moyenne) qu'en absence de Pi (11,5 mm) sont observées. Ce résultat est en accord avec la littérature [196–198]. Cette différence de longueur entre + Pi et - Pi est aussi observée chez le mutant *rns1rns2rns3*^{KO} (différence de 3 mm également) et plus faiblement chez les mutants *rns2*^{KO}, *rns2rns3*^{KO} et *rns1rns3*^{KO} (0,5 mm, 1,3 mm et 1,4 mm de différence entre + Pi et - Pi, respectivement). En revanche, de manière surprenante ce phénotype est inversé pour *rns1*^{KO} et *rns1rns2*^{KO} où les racines sont en moyenne plus grandes en - Pi (Figure 22 C). Ce résultat est en l'état difficilement interprétable dans la mesure où ce phénomène est observé chez les mutants dont on dispose de moins de données (dû à un plus faible taux de germination). La répétition de cette expérience avec plusieurs lots de graines indépendant est indispensable pour interpréter correctement ces résultats. De plus, il serait intéressant de mesurer le nombre et la longueur des racines secondaires qui normalement se développent davantage lors de carence en phosphate [197,198].

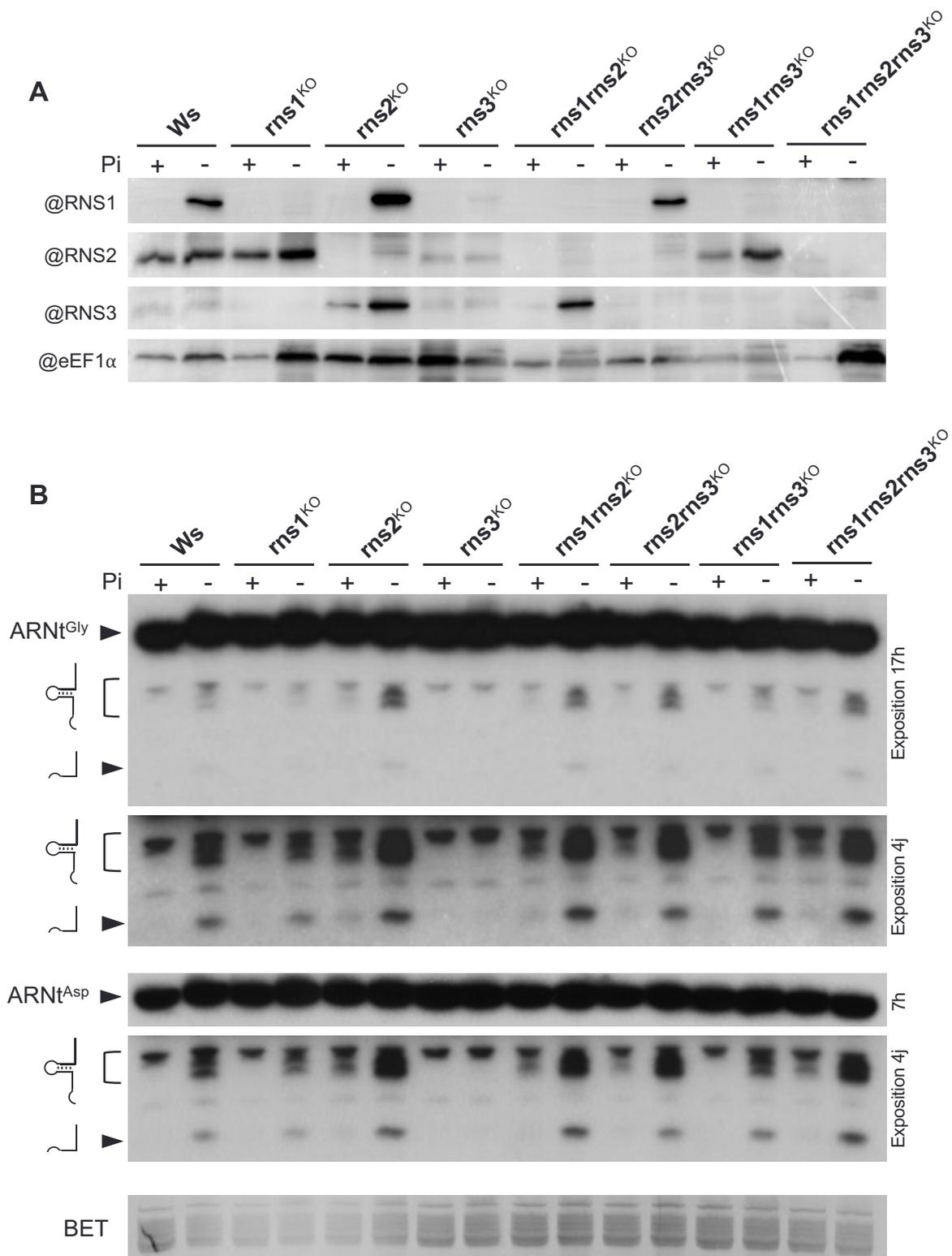


Figure 23 : Expression des RNS et observation des tRFs de plantules d'*A. thaliana rms^{KO}* en condition de carence en phosphate. A : Western blot sur des protéines extraites de plantules (sauvages ou *rms^{KO}*) exposées ou non à une carence en phosphate (+ ou - Pi) durant 7 jours. L'anticorps @eEF1α est utilisé comme contrôle de charge. B : Northern blot sur des petits ARN extraits de plantules (sauvages ou *rms^{KO}*) exposées ou non à une carence en phosphate (+ ou - Pi). Sonde radioactive complémentaire à l'ARNt^{Gly} (UCC) et ARNt^{Asp} (GUC). Différentes expositions sont montrées pour chaque sonde. La coloration du gel au bromure d'éthidium (BET) avant transfert sur membrane est utilisée comme contrôle de charge.

D'un point de vue plus global, aucune tendance cohérente ne se dessine quant à la germination des graines. Si RNS1 et RNS3 étaient impliquées de manière similaire dans les processus de germination, les simples mutants *rns1^{KO}* et *rns3^{KO}* se comporteraient de manière identique. Or, ce n'est pas le cas. On s'attendrait aussi que le mutant double *rns1rns3^{KO}* soit lui aussi fortement impacté et présente un phénotype différent de celui du Ws. Les trois mutants ayant un faible taux de germination ont en commun l'absence de RNS1. Ils présentent un phénotype de germination semblable. Mais si la présence de RNS1 est essentielle pour la germination, on attendrait que le mutant *rns1rns3^{KO}* présente aussi ce phénotype de faible taux de germination. Concernant le phénotype racinaire lors de carence en phosphate, il est aussi difficile d'interpréter l'importance des RNS dans l'adaptation de la plante à ce stress.

2. Observations de la biogenèse des tRFs

A l'échelle moléculaire, l'expression des RNS et les populations de tRFs ont été étudiés dans les mutants *rns^{KO}* en condition de carence en phosphate. Pour cela, des graines ont été semées *in vitro* dans un milieu liquide contenant du phosphate ou non (+ Pi ou - Pi). Après 7 jours de culture, les plantules sont récoltées puis à la fois les ARN et les protéines ont été extraits.

Comme observée dans la littérature, les résultats des western blots montrent une induction de RNS1 et plus faiblement de RNS2 chez le sauvage Ws lors de carence en phosphate (Figure 23 A). RNS2 est constitutivement exprimée dans les plantules Ws, *rns1^{KO}* et *rns1rns3^{KO}*. Une faible induction de RNS2 en -Pi est observée dans ces mutants. Dans le mutant *rns3^{KO}*, son expression est faible mais ne semble pas induite en - Pi. RNS3 est quant à elle induite dans les mutants n'ayant pas RNS2, c'est-à-dire chez *rns2^{KO}* et *rns1rns2^{KO}*. RNS3 semble prendre le relais lorsque RNS2 n'est pas présente. Il est aussi observé que lorsque RNS1 est absente (dans *rns1^{KO}* notamment), RNS3 n'est pas détectée, et inversement. Ce phénomène, qui pourrait être qualifié de compensatoire, a déjà été observé auparavant par plusieurs membres de l'équipe. Ils seront à confirmer et à réaliser en triplicat.

Concernant la production de tRFs lors de carence en phosphate, les northern blots montrent une accumulation des tRF-5D et -5A issus des ARN^{t^{Gly}} (UCC) et ARN^{t^{Asp}} (GUC) chez le sauvage Ws en -Pi, en concordance avec la littérature [113,190] (Figure 23 B). Hormis chez *rns3^{KO}* où le profil de ces tRFs ne change pas en fonction de la teneur en Pi, cette accumulation de tRFs particuliers en absence de phosphate est retrouvée dans l'ensemble des mutants *rns^{KO}*, y compris dans le mutant triple *rns1rns2rns3^{KO}*. Or, chez ce mutant, les western blots de la Figure 23 A ne révèlent pas d'expression ni pour RNS1, RNS2 ou RNS3.

A ce jour les résultats antérieurs de l'équipe et de la littérature suggèrent que les tRFs sont générés par RNS1, RNS2 et RNS3 chez *A. thaliana* [112,113]. Ces résultats soulèvent donc de nouvelles questions : qui est à l'origine de la biogenèse de ces tRFs lorsque RNS1, RNS2 et RNS3 sont absentes ? Est-ce que RNS4 qui a été observée exprimée uniquement dans les racines peut être à l'origine de ces tRFs dans ces conditions ? Est-ce que d'autres endonucléases sont impliquées ?

Étant donné que des plantules entières ont été utilisées pour extraire les ARN et protéines ayant servi aux western et northern blots de la Figure 23, il ne faut pas négliger la part de tissus racinaires dans ces extraits. A ce stade de récolte, les plantules ont un système racinaire quasiment équivalent aux tissus aériens car uniquement les cotylédons sont développés. Il pourrait donc y avoir une activité de RNS4 sous-jacente qui n'est pas prise en compte dans l'interprétation de ces résultats. Un western blot avec l'anticorps dirigé contre RNS4 a été réalisé sur les extraits protéiques utilisés pour la Figure 23 mais la qualité de l'anticorps ne nous permet pas de conclure quant à une expression de RNS4 dans ces échantillons. Pour s'affranchir de l'utilisation de cet anticorps RNS4, il faudrait quantifier l'ARNm de RNS4 par RT-qPCR dans ces extraits. Cette expérience permettrait de déterminer si RNS4 est plus fortement exprimée en l'absence des trois autres RNS et si la carence en phosphate induit son expression, ce qui n'a jamais été observé auparavant.

3. Conclusion

Les différents mutants *rns*^{KO} ont été générés par la technologie CRISPR/cas9 et validés. Ils sont donc maintenant disponibles pour être utilisés comme outil génétique pour approfondir les connaissances sur la biogenèse des tRFs. Les résultats préliminaires ne montrent aucun phénotype particulier de ces mutants, y compris pour *rns1rns2rns3*^{KO} en conditions normales de culture. Le phénotype devra continuer à être étudié en cas de stress (carence en phosphate ou autre). Les expériences doivent être réalisées en triplicat à partir de lots de graines et de lignées indépendantes pour s'affranchir de phénotypes liés à des paramètres indépendants de l'expérience.

Ces premiers résultats soulèvent aussi de nouvelles questions quant à la biogenèse des tRFs. Dans des conditions de carence en phosphate, RNS1, RNS2 et RNS3 ne semblent pas être à l'origine de la biogenèse des tRFs dans le cas du mutant triple *rns1rns2rns3*^{KO}. La seule autre RNS exprimée et non prise en compte dans ces expériences étant RNS4, il faudrait à présent vérifier son implication ou non dans la biogenèse de ces tRFs. Au besoin, il faudra peut-être réaliser le mutant quadruple *rns1rns2rns3rns4*^{KO}. Ceci est tout à fait envisageable puisque des mutants simple *rns4*^{KO} en fond Ws existent dans les banques de mutants d'insertion. Le génotypage de ces simples mutants a déjà été initié pour obtenir des homozygotes (FLAG_327E02 sur fond génétique Ws et SALK_055304 sur fond génétique Col-0). Ils pourront par la suite être croisés au mutant triple *rns1rns2rns3*^{KO} pour obtenir le mutant quadruple KO. Enfin, selon les résultats obtenus, l'implication d'autres endoribonucléases impliquées dans la genèse des tRFs dans ces conditions devra être envisagée. De fait, il a aussi été montré que les Dicer-like sont des endonucléases capables de générer des tRFs chez les plantes [111]. Il serait intéressant d'observer leur expression et leur activité dans le mutant triple *rns1rns2rns3*^{KO}. Finalement, d'autres endoribonucléases non connues à ce jour sont peut-être responsables de la biogenèse de tRFs. Une nouvelle page dans la compréhension et la régulation de cette biogenèse chez *A. thaliana* reste alors encore à écrire.

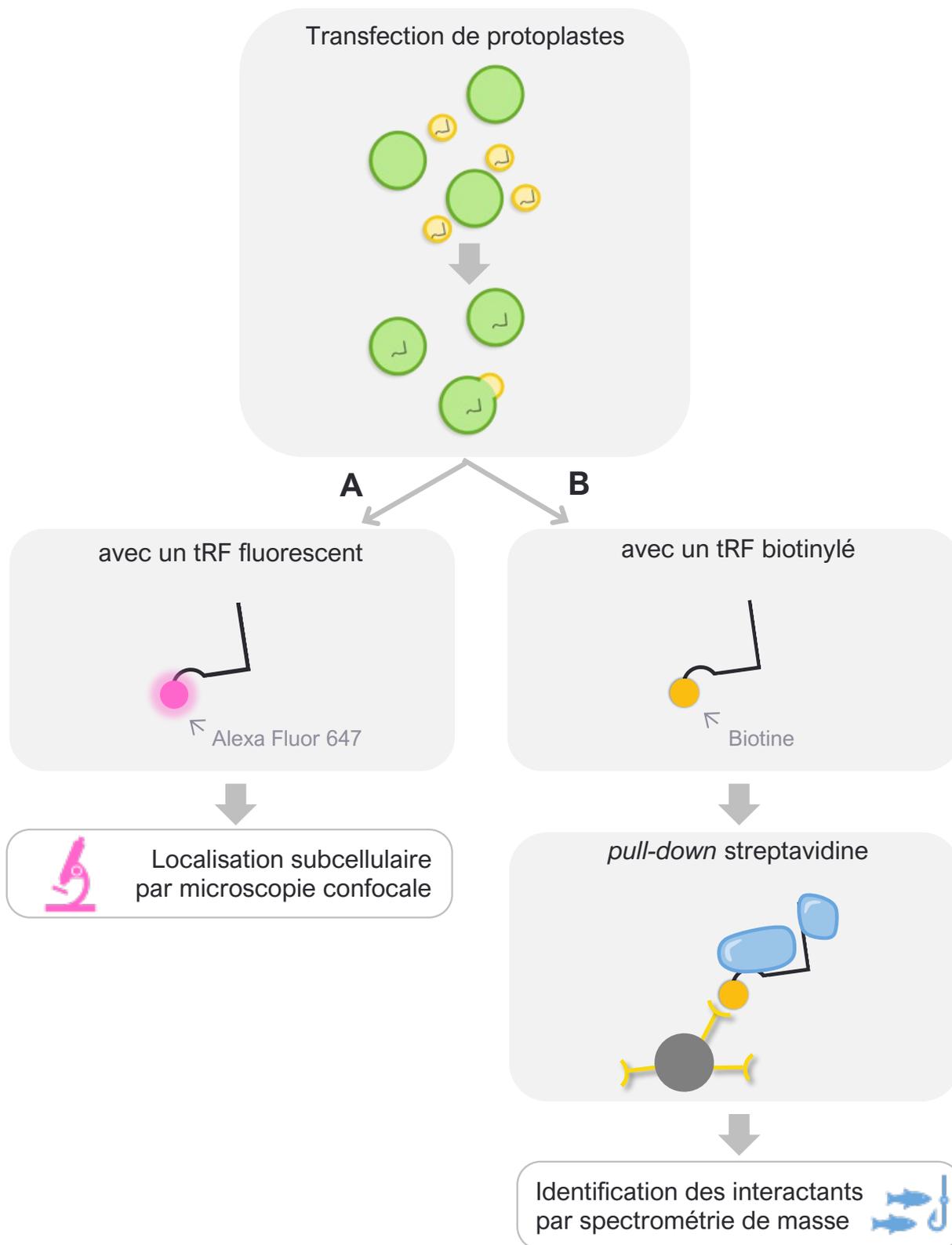


Figure 24 : Schéma de la stratégie expérimentale employée pour l'étude de la fonction du tRF Ala. Pour étudier la localisation subcellulaire du tRF Ala (A), des protoplastes sont transfectés avec le tRF fusionné à l'Alexa Fluor 647. Une fois transfectés, les protoplastes sont directement observés au microscope confocal. Pour identifier ses partenaires protéiques (B), les protoplastes sont transfectés avec le tRF fusionné à une biotine. Après transfection, les protoplastes sont lysés et les interactants du tRFs sont isolés par *pull-down* grâce à la forte affinité de la biotine pour des billes recouvertes de streptavidine. Les interactants isolés sont ensuite identifiés par spectrométrie de masse (LC-MS/MS).

CHAPITRE II : Étude de la fonction biologique du tRF modèle Ala chez *A. thaliana*

I. Mise en contexte

Les fonctions des tRFs sont relativement variées. Ils permettent de réguler l'expression de gènes ciblés en modulant leur transcription ou leur traduction, ceci en réponse à des conditions spécifiques. Concernant leurs fonctions chez les plantes, les données restent encore parcellaires. La plupart des fonctions décrites à l'heure actuelle montrent que les tRFs régulent la traduction de gènes par le mécanisme de *silencing*. Or, il existe de nombreux tRFs qui ne sont pas retrouvés associés à AGO1 et donc qui ne seraient pas impliqués dans ce mécanisme.

L'équipe dans laquelle j'ai réalisé mes travaux de thèse cherche à comprendre quelles pourraient être les fonctions de ces tRFs chez *A. thaliana*. J'ai donc poursuivi le travail initié précédemment sur la détermination des fonctions potentielles du tRF-5D modèle issu de l'ARNt^{Ala}.

1. Pourquoi le tRF-5D Ala comme modèle

Le tRF-5D issu de l'ARNt^{Ala} nucléaire a été retrouvé dans plusieurs banques comme étant le plus abondant chez *A. thaliana* [87,147]. De plus, lorsque ce tRF est généré par les RNS, il n'est pas retrouvé enrichi avec les protéines AGO1 [87]. L'étude de Lalande et al [151] a mis en évidence la capacité de certains tRFs à inhiber la traduction en système *in vitro*, dont le tRF-5D Ala de 20 nt. Cette régulation de la traduction ne se fait pas par complémentarité entre le tRF Ala et le transcrit à réguler. Ces deux études suggèrent que le tRF Ala est capable d'inhiber la traduction mais de manière distincte de celle du mécanisme de *silencing*.

Dans ce contexte, j'ai poursuivi le travail mené dans l'équipe pour tenter de comprendre les mécanismes dans lesquels ce tRF Ala est impliqué et qui conduisent à une régulation de la traduction.

2. Stratégie expérimentale employée pour étudier sa fonction biologique

Pour déterminer la fonction biologique de ce tRF, sa localisation subcellulaire et ses partenaires protéiques ont été étudiés. Ces deux approches complémentaires ont été réalisées à l'aide de la technique de transfection de protoplastes (Lalande et al. [168], voir Annexe IV de la Partie Matériel & Méthodes).

La localisation subcellulaire du tRF Ala a été déterminée par transfection de protoplastes avec un oligoribonucléotide auquel un fluorochrome a été fusionné (Alexa Fluor 647). Ce tRF synthétique est transfecté dans des protoplastes d'*A. thaliana* et sa localisation est directement observée après transfection au microscope confocal (Figure 24 A).

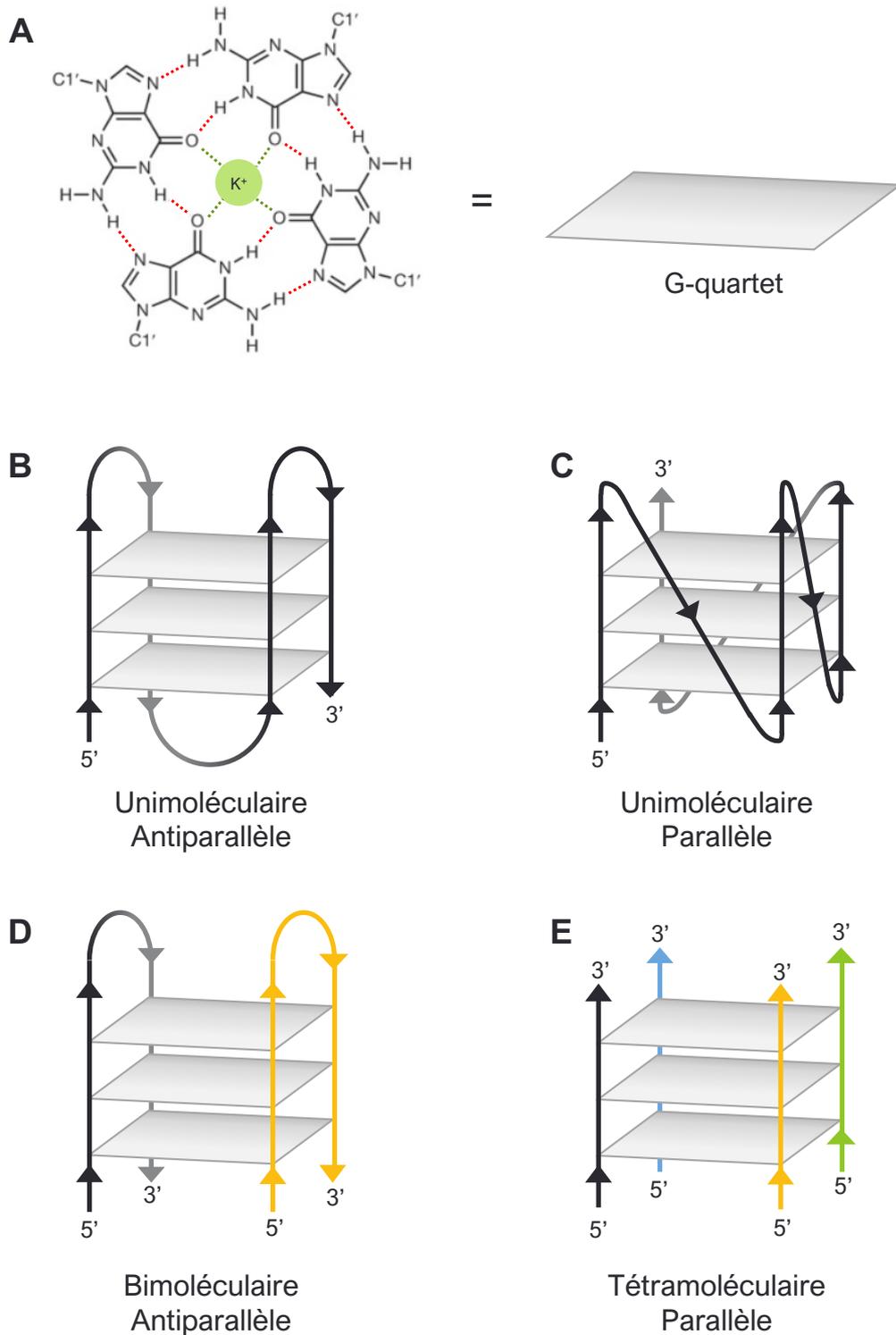


Figure 25 : Structure et types de G-quadruplex ARN. A : Structure chimique d'un G-quartet. Les quatre guanines sont reliées entre elles par des liaisons hydrogènes de Hoogsteen (ligne en pointillés rouges). Le cation, ici K^+ , interagit avec les guanines par liaisons à leur atome d'oxygène (lignes en pointillés verts), adaptée de [203]. B - E : Exemples de types de G-quadruplex à 3 quartets. Les quartets sont représentés par des polygones gris et les molécules d'ARN par des lignes. L'orientation de la molécule d'ARN est montrée par des flèches. B : G-quadruplex unimoléculaire antiparallèle. C : G-quadruplex unimoléculaire parallèle. D : G-quadruplex bimoléculaire antiparallèle (la première molécule d'ARN est représentée par une ligne noire, la seconde par une ligne jaune). E : G-quadruplex tétramoléculaire parallèle (les quatre molécules d'ARN sont représentées par des lignes de différentes couleurs).

Basé sur cette technique, les interactants protéiques du tRF Ala ont été étudiés. Pour se faire, un oligoribonucléotide, fusionné cette fois-ci avec une biotine, a été transfecté dans des protoplastes. Après transfection, les interactants ont été enrichis par *pull-down* grâce à la streptavidine. Les protéines enrichies ont ensuite été identifiées par spectrométrie de masse (Nano LC-MS/MS) (Figure 24 B).

L'ensemble des résultats de ces deux approches sont décrit dans le paragraphe III Résultats de la localisation subcellulaire et des interactants du tRF Ala, ci-après, sous forme d'un article en cours de rédaction dans lequel je suis première auteure.

II. Aide à la compréhension des résultats

1. Les structures G-quadruplex d'ARN

i. Description

Parmi les structures secondaires que peuvent adopter les ARN figurent les G-quadruplex. Ce sont des structures non-canoniques basées sur des appariements entre les bases guanines par des liaisons hydrogènes de type Hoogsteen. Les G-quadruplex sont constitués généralement de trois couches de tétrades de guanine (G-quartets) planes superposées les unes au-dessus des autres (Figure 25). Cette structure est possible pour les ARN possédant une séquence riche en G. A l'origine, les séquences pouvant former des G-quadruplex étaient définies comme ceci : GGGN₁₋₇GGGN₁₋₇GGGN₁₋₇GGG où N peut être n'importe quelle base [199,200]. Plus tard, il a été découvert qu'une succession de seulement 2 guanines au lieu de 3 suffisaient pour former des G-quadruplex, ainsi qu'une séquence variable supérieure à 7 bases [201]. L'ensemble des G-quartets est stabilisé par un cation monovalent interagissant au centre des G-quartets. Cette interaction est plus ou moins forte selon les cations et confère une stabilité plus ou moins grande du G-quadruplex (forte interaction/stabilité > K⁺ > Na⁺ > Li⁺ > faible interaction/stabilité) (Figure 25 A) [202,203]. Il a aussi été observé que plus le G-quadruplex possède des couches de G-quartets, plus sa structure est stable. Différents types de G-quadruplex existent selon le nombre de molécules enfermées dans cette structure, la séquence et le nombre de G impliqués, et le repliement des brins d'ARN. Ainsi, il peut exister des G-quadruplex unimoléculaires, bimoléculaires ou tétramoléculaires. L'orientation des brins d'ARN peut conduire à la formation de G-quadruplex dits parallèles ou antiparallèles (Figure 25 B-E) [202].

Ces structures sont retrouvées à la fois sur l'ADN et sur l'ARN. Les G-quadruplex ARN ont été identifiés dans des régions non traduites de certains ARN (5'-UTR et 3'-UTR) ou encore sur des ARN non codants. Étant prédites en très grand nombre dans le génome humain (correspondant à environ 13 000 régions du transcriptome), ces structures sont considérées comme essentielles et ayant des fonctions importantes. Elles sont impliquées dans des mécanismes variés : la transcription des ARNm, le transport des ARN, leur stockage, leur dégradation ou encore la biogenèse des microARN. Il a été observé que des dysfonctionnements dans la régulation de ces structures provoquent des dérèglements cellulaires [204].

Les nombreuses études émergentes de ces dernières années sur les G-quadruplex ont permis de démontrer que ce sont des *RNA-Binding Proteins* (RBPs) qui sont à l'origine de la régulation de ces structures. Ces protéines ont un rôle soit dans la stabilisation des G-quadruplex, soit dans leur dissociation. Elles permettent de passer d'un ARN non structuré à un état structuré en G-quadruplex, et inversement.

La base de données G4IPDB (<http://bsbe.iiti.ac.in/bsbe/ipdb/index.php>) recense les protéines interagissant avec les G-quadruplex [205]. L'analyse de la composition en acides aminés de ces protéines a permis d'identifier une région riche en résidus arginines et guanines, commune aux protéines interagissant avec les G-quadruplex. Dans la littérature, les motifs riches en RG apparaissent sous le terme de motif RGG ou motif NIQI (*Novel Interesting Quadruplex Interaction motif*) et diffèrent légèrement par leur séquence consensus [206–208]. Dans plusieurs études il a été montré que ce motif RGG/NIQI est nécessaire pour la reconnaissance des G-quadruplex [209–211].

Les hélicases à ARN participent au déroulement et à la dissociation des G-quadruplex. Jusqu'à présent, les hélicases à ARN identifiées pour intervenir dans la régulation de ces structures sont issues des superfamilles SF1 et SF2. Généralement, elles sont capables d'interagir avec les G-quadruplex ARN et ADN. En plus des hélicases, des RBPs participent au maintien de l'ARN sous forme non structuré. Il est aussi suggéré que des RBPs participent au recrutement des hélicases [204].

Chez l'humain, plusieurs hélicases capables d'interagir avec des G-quadruplex ARN ont été identifiées notamment DHX36, DHX9, DDX21, DDX5/17 ; DHX1 et eIF4A [212,213]. L'hélicase DHX36, aussi retrouvée sous les noms RHAU, MLEL1 et G4R1, est bien décrite dans la littérature. Elle est capable de prendre en charge des G-quadruplex ADN ou ARN, avec une meilleure affinité pour les G-quadruplex parallèles. Deux domaines sont impliqués dans la reconnaissance du G-quadruplex : le domaine DSM en N-terminal et le domaine OB en C-terminal. Elle possède en N-terminal, en plus du domaine DSM, le motif RGG [214–216].

ii. Les G-quadruplex chez les plantes

Chez *A. thaliana*, les analyses *in silico* ont permis d'identifier plus de 65000 régions du transcriptome riches en G et prédites comme étant capables de former des G-quadruplex. Cependant, les résultats d'analyses *in vitro* réalisées par Yang et al. [217] montrent que seulement 3000 de ces régions possèdent effectivement une structure G-quadruplex. Il est donc probable qu'une grande partie des prédictions de G-quadruplex n'existent pas *in vivo*. De plus, cette équipe a montré que la présence d'autres structures secondaires sur l'ARN pouvaient impacter la formation de G-quadruplex. Finalement, il est montré que seulement quelques centaines de régions riches en G forment des G-quadruplex *in vivo*. Ces résultats ont aussi été observés chez le riz, suggérant que les G-quadruplex existent *in vivo* chez les monocotylédones et dicotylédones [217–219].

Les fonctions des structures G-quadruplex de l'ARN chez les plantes sont encore peu documentées. Une étude a mis en évidence le rôle de G-quadruplex dans la régulation de la traduction chez *A. thaliana*. La protéine à doigt de zinc JUL induit la formation d'un G-quadruplex sur la région 5'-UTR de l'ARNm de SMXL4/5. Ce dernier joue un rôle clé dans la différenciation du phloème. En favorisant la formation d'un G-quadruplex sur cet ARN, JUL induit indirectement l'inhibition de sa traduction [220]. L'équipe de Kwok et al. [221] a montré *in vivo* qu'un G-quadruplex situé sur la région 5'-UTR de l'ARNm du gène ATR inhibe sa traduction. D'autres résultats basés sur des analyses *in silico*, suggèrent que les G-quadruplex pourraient avoir un rôle lors de stress. Par exemple, lors de sécheresse, la concentration en potassium intracellulaires (K^+) a tendance à augmenter. Étant donné que c'est un cation stabilisateur de G-quadruplex, il est suggéré un lien de cause à effet qui pourrait être à l'origine de la régulation de l'expression de certains gènes particuliers [218]. Cependant, ces résultats ne reposent que sur des analyses bio-informatiques à partir de bases de données et ne sont pas, à ce jour, appuyées par des expérimentations *in vitro* ou *in vivo*.

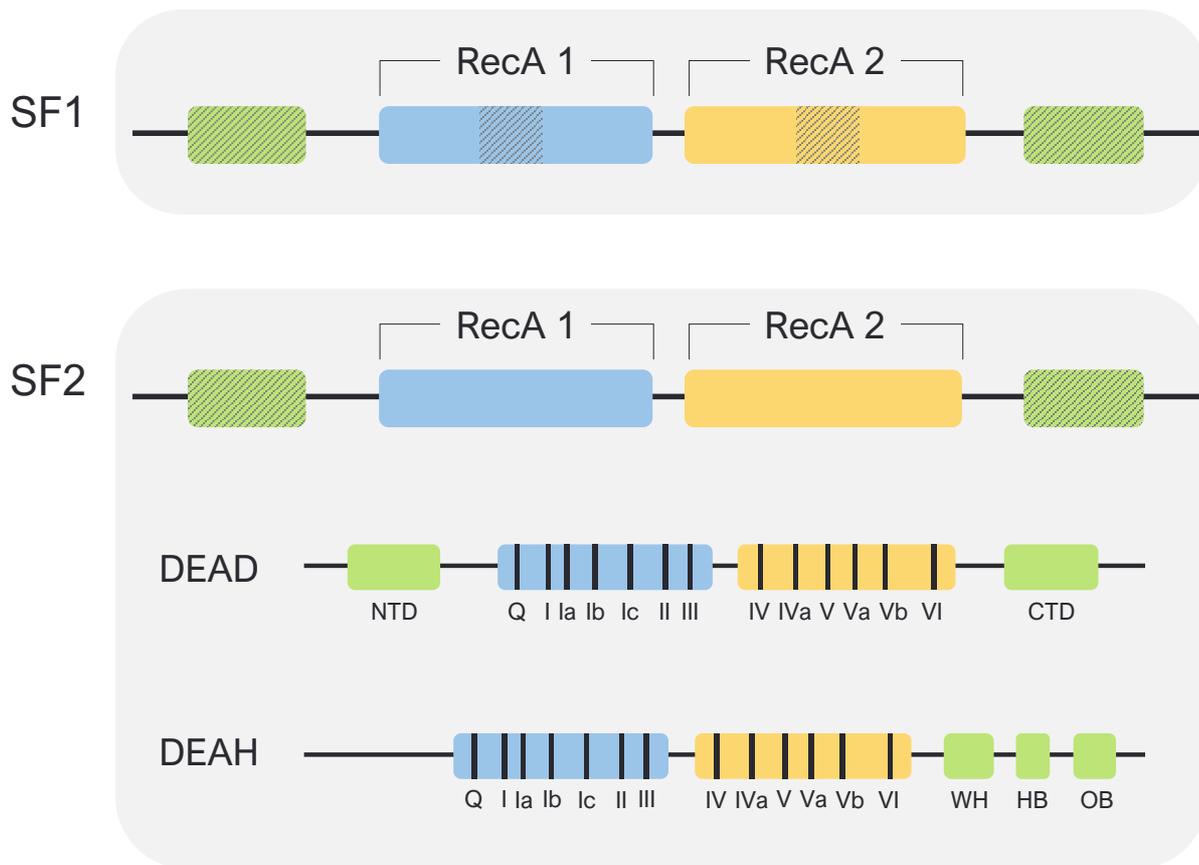


Figure 26 : Organisation des domaines conservés des hélicases des superfamilles SF1 et SF2. Ces deux familles d'hélicases partagent deux domaines RecA 1 (en bleu) et RecA 2 (en jaune) qui constituent le noyau des hélicases (*helicase core*). Elles présentent des domaines en N-terminal et C-terminal (en vert), variables selon les hélicases. Les hélicases SF1 se distinguent des hélicases SF2 par leurs séquences variables en N-terminal, en C-terminal et au sein des domaines RecA 1 et RecA 2 (parties hachurées). Parmi les hélicases SF2, ne sont représentés ici que les familles d'hélicases DEAD et DEAH. Ces deux familles partagent les mêmes domaines RecA 1 et RecA 2 mais se distinguent par leurs extrémités N-terminal et C-terminal. Les domaines des hélicases des superfamilles SF3, SF4, SF5 et SF6 ne sont pas représentés dans cette figure. Les domaines ne sont pas représentés à l'échelle. Abréviations : NTD : domaine N-terminal, CTD : domaine C-terminal, WH : *winged helix*, HB : *helix bundle*, OB : site de liaison aux oligosaccharides.

2. Les hélicases à ARN

i. Généralités

Les hélicases à ARN sont des enzymes capables de dissocier les structures double brins d'ARN, en utilisant de l'énergie sous forme de NTP (le plus souvent d'ATP). Les hélicases à ARN sont présentes chez quasiment tous les organismes allant des procaryotes aux eucaryotes [222].

La séquence des hélicases ARN est relativement bien conservée. Les hélicases sont identifiées par la présence de sept à huit motifs conservés. Ces motifs ont des fonctions particulières et sont impliqués dans l'interaction et l'hydrolyse de l'ATP ou dans la liaison avec le substrat ARN. Les hélicases sont classées selon les motifs qu'elles possèdent. Les hélicases des superfamilles SF1 et SF2 sont considérées actives en monomères ou dimères alors que celles des superfamilles SF3 à SF6 sont actives en complexes hexamériques. Les hélicases des superfamilles SF1 et SF2 possèdent en commun les deux premiers motifs (I et II) de liaison à l'ATP. Elles sont ensuite différenciées selon les six autres motifs qu'elles possèdent. On retrouve ainsi les hélicases de la famille des DEAD-box, DEAH-box et la famille des DExD/H-box (x étant un acide aminé quelconque) basé sur la séquence du motif II (Figure 26) [222–224].

Généralement, les hélicases à ARN interagissent avec les ARN via leur squelette ose-phosphate. De ce fait, le plus souvent les hélicases ne reconnaissent pas une séquence particulière présente sur l'ARN. Certaines hélicases parcourent les ARN jusqu'à reconnaître une structure particulière alors que d'autres reconnaissent directement des structures secondaires locales. La plupart des hélicases dissocient les structures secondaires mais certaines n'ont qu'une activité de reconnaissance/liaison à l'ARN, sans perturber les structures secondaires. Partant de cette observation, les hélicases à ARN peuvent réguler les structures présentes sur les ARN mais aussi les interactions ARN-ARN et ARN-protéines. Le rôle des hélicases dans la régulation de l'expression des gènes a été largement étudié chez l'humain. Les hélicases ARN sont, entre autres, capables de réguler l'épissage des ARNm, leur exportation vers le cytosol, la traduction des ARNm, leur stockage et leur dégradation (pour une revue, voir [225]).

0% 100%
Expression potentielle

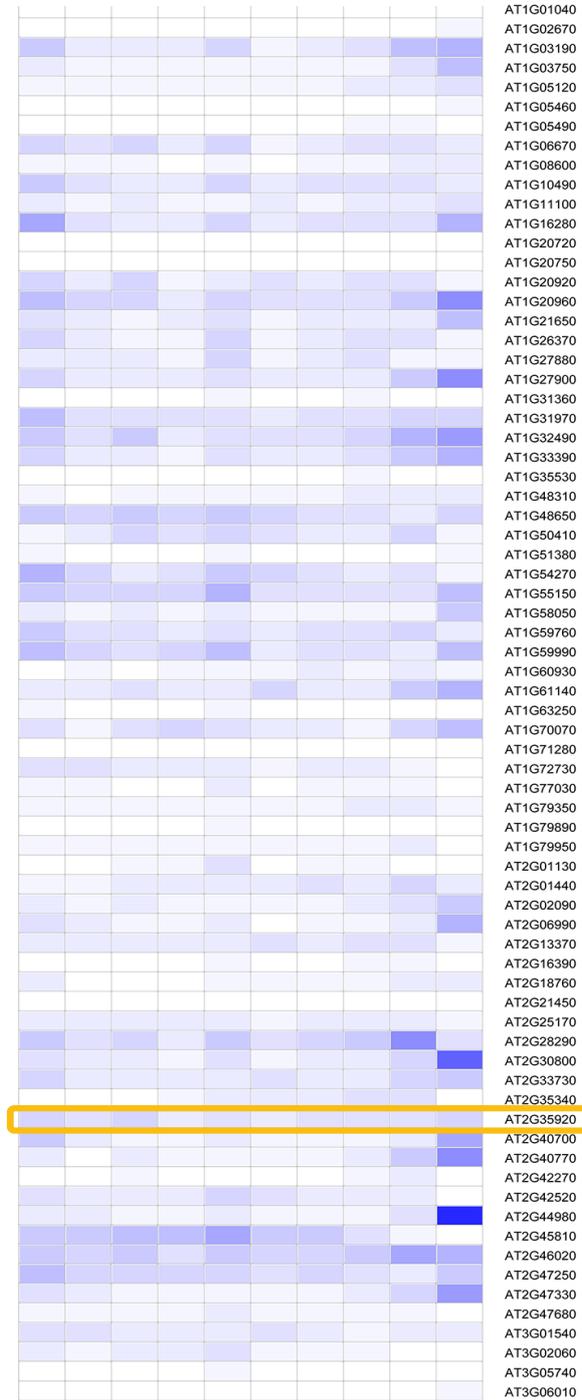


Figure 27 : Profil d'expression des gènes d'hélicases à ARN chez *A. thaliana*. Adaptée de [226]. Le profil d'expression des gènes codant les 144 hélicases à ARN est donné pour les différents stades de développement de la plante *A. thaliana*. Le niveau d'expression est représenté par un dégradé de couleur bleue. Un niveau d'expression faible est représenté pas une couleur bleue claire alors qu'un niveau d'expression élevé est représenté par une couleur bleue foncée. Le profil d'expression de DExH1 (At2g35920) est encadré en orange.

ii. Les hélicases à ARN chez *A. thaliana*

Au total, 161 gènes codant des hélicases à ARN ont été répertoriés dans le génome d'*A. thaliana*. Ils ont été classés en trois sous-familles basés sur leurs motifs conservés : 50 gènes codent pour des hélicases DEAD-box, 40 pour des DEAH-box et 71 pour des DExD/H-box [226]. Les gènes d'hélicases à ARN sont répartis sur l'ensemble des cinq chromosomes que compte le génome d'*Arabidopsis*. L'analyse des banques de données effectuée par Xu et al. [226] montre que 144 gènes sont exprimés au moins à une étape du développement de la plante à des niveaux différents (ce qui représente 89,4% des hélicases à ARN). Parmi eux, plus de la moitié sont exprimés, à niveaux d'expression variables, lors de plusieurs étapes du développement. Hormis six hélicases qui sont fortement exprimées lors d'étapes de sénescence (AT2G44980, AT3G19760, AT3G53110 (LOS4), AT4G18465, AT5G11200 et AT5G51280), le restant des hélicases (soit 138) ne semble pas avoir de profil d'expression spécifique (Figure 27). Parmi toutes ces hélicases, nous retrouvons DExH1. C'est une hélicase de la famille des DExH-box. Cette hélicase ne fait pas partie des hélicases les plus exprimées chez *A. thaliana* (Figure 27).

III. Résultats de la localisation subcellulaire et des interactants du tRF Ala

Les résultats sont présentés dans les pages suivantes sous forme d'un article en cours de rédaction.

Arabidopsis DEA(D/H)-box RNA helicase interacts with the foci-localized tRF Ala

Marjorie Chery¹, Christina Berissou¹, Guillaume Hummel^{1,2}, Thalia Salinas¹, and Laurence Drouard¹

¹Institut de biologie moléculaire des plantes - CNRS, Université de Strasbourg, Strasbourg France.

²Present address: Institute of Biology, University of Hohenheim, Gabenstrasse 30, 70599, Stuttgart, Germany.

¹To whom correspondence should be addressed. Email: Laurence.drouard@cnr.fr

SUMMARY

Like a wide variety of organisms, tRNA-derived small RNA fragments (tRFs, also named tDRs) also exist in plants and likely serve many functions. We previously showed that tRFs are present under normal growth conditions and that tRFs deriving from alanine tRNAs are the most abundant in *Arabidopsis* leaves. We also determined that tRF Ala deriving from the mature tRNA^{Ala} (AGC) can block *in vitro* protein translation, but the molecular mechanism behind this inhibition has yet to be established. Here, we report that the tRF Ala (AGC) is mainly localized within peculiar foci in the cell that are neither P-bodies nor stress granules. We also provide evidence that tRF Ala (AGC) assembles into intramolecular RNA G-quadruplex (rG4) structures and specifically interacts with the DExH1 protein, a DEA(D/H) RNA helicase containing a glycine-arginine domain with RGG/RG/GR/GRR motifs known to be necessary for rG4 interaction. Mutation of the 4Gs at the 5' extremity of tRF Ala abolishes its rG4 structure assembly, its association with the DExH1 protein and foci formation. However, 4Gs mutation does not prevent protein translation inhibition *in vitro*. As a whole, our data suggest that sequestration of tRF Ala into rG4 complexes represents an efficient and rapid means of modulating accessible and functional tRF in the plant cell.

INTRODUCTION

Over the past ten years, numerous studies in a large variety of organisms (bacteria, archaea, and eukaryotes) have revealed the existence of a plethora of small non-coding RNAs derived from precursor or mature tRNAs. Several nomenclatures have been proposed for their naming. Here, we will use the term tRF, but the useful nomenclature proposed by the tDRnamer software (<http://trna.ucsc.edu/tDRnamer/>) will also be given. A growing list of important regulatory functions has been assigned to tRFs, for instance, as translational or transcriptional regulators or as modulators of ribosome biogenesis. The tRFs population varies across cell types and tissues, and numerous tRFs were shown to be upregulated upon stress. Several reviews provide updated and detailed knowledge on tRNA-derived small RNAs (e.g. (1)).

In photosynthetic organisms, the presence of tRFs and their biological importance have also been demonstrated. For example, the crucial role of rhizobial tRFs in regulating nodulation has been recently described (2). The existence of tRFs associated with Argonaute proteins has been shown (3) and there is a report showing that AGO1-associated tRFs can target and cleave transposable elements (4). Hsieh and colleagues provided evidence that some tRFs are upregulated upon phosphate starvation (5). A study in pumpkin shows the presence of tRFs in phloem and the authors have proposed their involvement in protein synthesis inhibition (6). We recently showed that a subset of Arabidopsis tRFs can associate with polyribosomes and modulate translation *in vitro* (7). However, how this is achieved *in vivo* is still unknown. Arabidopsis tRFs deriving from mature tRNAs^{Ala} represent exciting models to address this question. We previously showed that, under normal growth conditions, tRFs Ala of 20nt produced by cleavage of tRNAs^{Ala} (AGC) and tRNA^{Ala} (UGC) in the D region (tRF-5D) are by far the most abundant tRFs in leaves (8). Furthermore, Arabidopsis tRF-5D Ala was shown to efficiently inhibit translation *in vitro*, likely by association with heavy polyribosomes, suggesting a role as a translation modulator for this family of tRFs. Interestingly, tRF-derived tRNA^{Ala} also plays a crucial role in human translation regulation. Indeed, in response to cellular stresses, a subset of human tRFs (named tiRNAs for tRNA-derived stress-induced) is produced by a ribonuclease A, the angiogenin, upon cleavage in the anticodon region of tRNAs (i.e., producing half tRNAs), (9). Among this tRFs subset, the tiRNA^{Ala} was shown to inhibit translation initiation and promote the formation of stress granules (10). Finally, Lyons et al. provided evidence that the stretch of 5 guanine residues at the 5' extremity of human

tiRNAs^{Ala} allows their assembly into intermolecular RNA G-quadruplexes (rG4), rG4s being, in turn, essential to repress translation *in vivo* (11).

By contrast, in *Arabidopsis*, the 20 nt long tRF-5D Ala (AGC) (abbreviated hereafter as tRF Ala, and also named tDR-1:20-Ala-AGC-1-M2 according to tDRnamer; Supplementary Table S1) is present in large amounts in non-stress conditions, thus raising the question of its biological function and its mode of action. Here, taking advantage of an optimized protocol for *Arabidopsis* protoplasts transfection with RNAs, we first show that this tRF Ala is localized in foci that do not correspond to stress granules, consistent with the fact that no stress was applied. We then demonstrate that the stretch of 4Gs at its 5' extremity is necessary to promote assembly in rG4 structures essential for the interaction with DExH1, a DEA(D/H)-box RNA helicase likely involved in rG4 unwinding in *A. thaliana*. The assembly of tRFs into rG4 structures seems to be a conserved process. However, contrary to the situation in humans, we propose that tRF Ala, present in large amounts in *Arabidopsis* cells, is sequestered as rG4 complexes under normal growth conditions. We propose that the RNA helicase DExH1 resolves rG4 complexes and releases tRF Ala to perform its regulatory functions under still unknown conditions.

RESULTS

Arabidopsis tRF Ala can assemble into specific foci *in vivo*

As a first attempt to determine the potential *in vivo* subcellular localization of the 20 nt long tRF Ala, *Arabidopsis* protoplasts were transfected with a fluorescently-labelled synthetic oligoribonucleotide followed by visualization with confocal microscopy (Figure 1a). Two controls were used: transfection with H₂O (*i.e.*, no RNA) and transfection with a randomly generated 20 nt long oligoribonucleotide (named RNA R) with no sequence homology with the *Arabidopsis* genome (Figure 1b). In both controls (no RNA and RNA R), only the chlorophyll fluorescence is visible on the merged images (Figure 1b and Supplementary Figure 1). By contrast, multiple distinctive cytosolic foci were observed when protoplasts were transfected with fluorescently-labelled tRF Ala. As an average, z-stacks analysis shows the presence of circa 200 foci per protoplast (Figures 1b and 1c). We wanted to verify that the absence of visible foci when transfection is performed in the presence of RNA R was not due to an inefficient transfection of this RNA into protoplasts or to its instability. For this, *A. thaliana* protoplasts were co-transfected with equal amounts of both tRF Ala and RNA R, and upon transfection, protoplasts were post-treated or not with a mix of RNases A and T1.

The RNAs were then extracted and analyzed by northern blots with probes specific for each RNA (Figure 1d). In both cases, bands of similar intensities corresponding to either tRF Ala or RNA R were observed in the RNase post-treated protoplasts. Altogether, this shows that RNA R has the same transfection efficiency and stability behavior as tRF Ala, but it cannot assemble into foci.

tRF Ala specific foci are of unknown origin

As the presence of foci was observed for tRF Ala but not for a random RNA fragment, we then wondered if other Arabidopsis tRFs also have the propensity to be localized in foci. To address this question, we synthesized four other fluorescent sequences representative of the Arabidopsis tRFs population that we previously analyzed (Figure 2a) (8). While the very abundant tRF Ala was mainly not retrieved in AGO1 immunoprecipitates, the four others were associated with AGO1 to a different extent. The tRF-5D Arg (19 nt) and the tRF-5D His (21 nt; derived from an organellar tRNA^{His}) are among the most abundant ones associated with AGO1; the tRF-5A Asp is a long tRF of 33 nt and the tRF-5D Gly is upregulated under phosphate deprivation (12). As shown in Figure 2b and 2c and Supplementary Figure 2, a very few foci were observed upon transfection of Arabidopsis protoplasts with tRF Arg, and no foci were visible with tRF Gly, tRF Asp, and tRF His. Although non-exhaustive, these results strongly suggest that the localization in foci remains a specificity of tRF Ala, or at least of a limited set of tRFs.

The size of the tRF Ala containing foci, circa 1 μm , is compatible with several known structures such as mitochondria, P-bodies, or stress granules. To determine whether these foci correspond to one of these structures, we used Mitotracker staining and Arabidopsis lines stably expressing fluorescent fusion proteins representative of P-bodies (DCP1-YFP) or stress granules (tRFP-PAB2) (13,14). Confocal microscopy images (Figure 3a and Supplementary Figure 3) followed by dot plots quantification shows no correlation between the signals of the markers and the tRFs Ala (Mitotracker/YFP/tRFP on the x-axis and Alexa Fluor 467 on the y-axis), (Figure 3b). This likely indicates that the tRF Ala foci are neither mitochondria nor P-bodies or stress granules. It is worthy to note that stress granules were observed under stress conditions (*i.e.*, 1h at 37°C), and no stress granules are visible under normal conditions (Figure 3c), thus strengthening the fact that the recruitment of tRF Ala in specific foci is not induced by stress.

A DEAH-box RNA helicase, DExH1, strongly interacts with tRF Ala

In the absence of identification of the foci containing the tRF Ala, we decided to determine which proteins interact with this tRF *in vivo* using an RNA pull-down assay approach (Figure 1a). For this purpose, upon transfection of Arabidopsis protoplasts with a biotinylated tRF Ala, proteins interacting with it were co-purified on streptavidin beads and identified by LC-MS/MS. Two controls, i.e., H₂O or biotinylated RNA R used instead of tRF Ala during the transfection step, were performed under the same conditions. We observed that the protein encoded by *AT2G35920* was the only significantly enriched in the tRF Ala pull-down assay compared to the two controls (Figures 4a, Supplementary Figure 4, Supplementary Table S3). Similarly, when we transfected protoplasts with biotinylated tRFs corresponding to the four fluorescent tRFs showing no foci localization (tRF Arg, tRF His, tRF Asp, and tRF Gly; Figure 2), as for RNA R, the *AT2G35920.1* protein was not enriched in the pull-down assays (Figure 4b and 4c, Supplementary Table S4). These data suggest that only tRF Ala specifically interacts with the *AT2G35920.1* protein among the tested small RNAs. According to UniProt, this corresponds to a DExH-box ATP-dependent RNA helicase, also named DExH1. Blast search shows that *A. thaliana* DExH1 is the homolog of the human DEA(D/H)-box RNA helicase DHX36 (Supplementary Figure 5) and shares the different domains of a structural core (Figure 4d), (15). Interestingly, the human RNA helicase DHX36 belongs to the family of helicases able to bind and resolve DNA and RNA G-quadruplexes (abbreviated G4 and rG4 respectively; Figure 5a), (16). The N-terminal motif specific to DHX36 and implicated in G4 binding is absent in DExH1 (Supplementary Figure 5). In contrast, we found at the N-terminal extremity, in both DExH1 and DHX36, a glycine-rich region containing a glycine-arginine domain. This domain, named GAR, was shown to be present in many rG4 interactors and to be essential for rG4 interactions (17). The occurrence of RGG/RG/GR/GRR motifs in the GAR domain of DExH1 is high (Figure 4d) and corresponds to a high-confidence rG4 interactor (17). In *A. thaliana*, the DEA(D/H)-box RNA helicase family is large (18) (19) but no other member of the RNA helicase family was enriched in the tRF Ala pull-down assays (Figure 4c). Importantly, sequence alignment of the *A. thaliana* DEA(D/H)-box RNA helicases (18) shows that the GAR domain is only present in the N-terminal extremity of the DExH1 protein (Figure 4e). Thus, we hypothesized that biotinylated tRF Ala pulled down the DExH1 protein because it can form an rG4 structure and thus associate with this RNA helicase.

tRF Ala forms a G-quadruplex structure essential to interact with the DExH1 RNA helicase

Indeed, it has already been demonstrated that a human tRF Ala (AGC) (also named 5'tiRNA^{Ala}) of 30 nt and an Arabidopsis tRF Ala (CGC) of 25 nt possess a stretch of five guanine residues at their 5' extremities that can assemble into rG4s. By contrast, a tRF derived from an Arabidopsis tRNA^{Leu} (At112A) with only 4Gs at its 5' extremity cannot form rG4 structures (Figure 5a) (20). Here, the Arabidopsis tRF Ala (AGC) of 20 nt has only 4 guanines at the 5' extremity, thus raising the question of whether it can assemble into rG4 structures (Figure 5b). We evaluated the capacity of the tRF Ala (AGC) to form an rG4 structure (Figure 5c) by its aptitude to incorporate NMM fluorescent ligand either by fluorescence quantification (Figure 5d) or by native polyacrylamide gel (Figure 5e). As controls, we used the c-Myc RNA sequence that can assemble into an intramolecular G-quadruplex structure and incorporates NMM ligand (positive control, in grey) and the RNA R sequence that has no G stretch (negative control, in yellow). Quantification of the fluorescence shows that the fluorescence intensity of the tRF Ala is higher than the negative control, although inferior to the c-Myc signal, and suggests that it can assemble into an rG4 structure (Figure 5d). We obtained a piece of more direct evidence when we ran the samples on a native gel, and observed a retarded migrating band stained by NMM, characteristic of intermolecular rG4 complexes formed by the assembly of four individual tRF Ala (Figure 5c and 5e) as in (20). Strengthening our data, a mutated version of tRF Ala where GCGU has replaced the 5'-GGGG sequence shows no NMM incorporation neither by fluorescence quantification nor by staining on a native gel. Interestingly, a tRF Gly where GGGG has replaced the 5'-GCGU remains unable to form rG4 structures (Figures 5d and 5e). This result agrees with the absence of rG4 structures observed for At112A by Lyons and colleagues, meaning that the presence of 4 Gs at the 5' extremity of a tRF is essential but not sufficient to promote G-quadruplex structures. Furthermore, protoplasts transfection with the different tRFs shows that DExH1 is pulled-down only by c-Myc RNA and tRF Ala, the only two RNAs able to form rG4 structures (Figure 5f, Supplementary Table S5). However, while, as expected, neither the mutated versions of tRF Ala (5'-GCGU) nor tRF Gly (5'-GGGG) is localized within foci from transfected protoplasts, the c-Myc RNA is also not detected in foci (Figure 5g, Supplementary Figure 7). This observation suggests that either the stretch of 4Gs and the tRF Ala sequence are needed to form rG4 structure and to accumulate in foci.

DExH1 does not promote foci biogenesis

To better understand the role of DExH1 in its interaction with tRF Ala, its subcellular localization was studied. Protoplasts transfected with plasmids expressing DExH1 fused to GFP either at its N-terminal or C-terminal extremity were observed in confocal microscopy. Both constructs show a cytosolic localization of the fusion proteins (Figure 6a, Supplementary Figure 8), and no foci were detected. Although we cannot exclude that the cytosolic background hides the visualization of potential foci, this result suggests that, *in vivo*, the DExH1 protein is not a stable component of the foci. In agreement with this result, when protoplasts are prepared from a *dexh1* T-DNA insertion knock-out Col-0 mutant, upon transfection with the fluorescent tRF Ala, foci are visible, and their number is not significantly different from that obtained with wild-type Col-0 plants (Figures 6b and 6c and Supplementary Figure 9). These data suggest that DExH1 protein does not participate in the tRF Ala foci biogenesis.

DISCUSSION

Small non-coding RNAs derived from tRNAs are now recognized as essential players in gene expression regulation. In plants, how tRFs act at the molecular level remains poorly understood. The data obtained here allow us to propose a model for the tRF Ala mode of action (Figure 7). Under normal growth conditions, tRF Ala present in large amounts in Arabidopsis cells is sequestered as foci, presumably in the form of rG4 structures. In agreement with that, our mutagenesis studies show that the lack of a stretch of 4Gs in tRF Ala prevents the formation of G-quadruplexes and that foci are no longer detectable in transfected protoplasts. However, under conditions that remain to be determined, we hypothesize that rG4 complexes of tRF Ala can dissociate from these foci and become substrates of the DExH1 RNA helicase, which actively resolves rG4 structures, releasing tRF Ala monomers. The free tRF Ala can then play its role as a translation modulator. It could also be implicated in other molecular processes, such as the modulation of gene expression via the RNA silencing pathway. However, we have not found tRF Ala to be strongly associated with Argonaute proteins in our analysis (8). A recent study suggests that in Arabidopsis, a tRF Ala of 19 nt associates with AGO1 and negatively regulates the expression of the Cytochrome P450 71A13 by cleavage of the target mRNA (21).

Overall, we propose that Arabidopsis tRF Ala rG4s are transiently structures that fluctuate dynamically in the plant cells according to tRF Ala demand.

In Arabidopsis cells, tRF Ala is produced in large amounts with potentially high local concentration, thus facilitating aggregation into rG4 complexes and their sequestration in the form of foci. On the one hand, this could allow avoiding a deleterious effect if the “free” tRF is present in too large quantities in the cell and, on the other hand, to be very rapidly operational if released within the cytosol when necessary. Both human and Arabidopsis rG4 structures can be formed from alanine tRNA-derived RNAs, suggesting an evolutionarily conserved process, as already established by Lyons et al. (20). However, their mode of action might be different. In humans, rG4 structures are assembled from human 30 nt long 5'tiRNA^{Ala} produced after a stress-induced cleavage in the anticodon region by angiogenin. They appear essential to repress translation *in vivo* by displacement of eIF4F from capped mRNAs thereby triggering the formation of stress granules (10,22,23) (11). Here, in Arabidopsis, 20 nt long tRFs Ala are highly and constitutively produced, likely by RNases T2 (24), and rG4 complexes can therefore be continuously present. While protein synthesis is repressed by human rG4s, we propose that the monomers of Arabidopsis tRF Ala play this role, whereas tRF Ala rG4s prevent their action. As previously demonstrated, neither a 5' (GCGU) tRF Ala nor a tRF Asn assemble into rG4 complexes, but both of them inhibit translation *in vitro* (7).

Over the last decade, DNA G4s have been studied in detail (e.g.(25) (26), and less is known on RNA rG4s although knowledge in the field is growing rapidly (27). In plants, G4 and rG4 motifs appear widespread, but their potential biological roles remain to be better understood (28) (29). It was recently demonstrated that rG4 structures exist *in vivo* and act as regulators in plants (30). However, rG4s have been mainly described in mRNAs and correspond to intramolecular G-quadruplex. For example, it was shown that a conserved G-quadruplex RNA module present in the 5' UTR of an mRNA is involved in the translational regulation of a gene implicated in phloem establishment (31). The importance of rG4s in small non-coding RNAs is yet to be largely studied in plants, and our data provide the first set of evidence of their potential importance.

The Arabidopsis 20 nt long tRF Ala is short and cannot *a priori* form a hairpin structure, thus is more prone to form stable intermolecular rG4s. However, it is to note that the presence of at least 4 Gs at the 5' extremity is not sufficient to promote G-quadruplex formation. As shown by Lyons et al. (20), an Arabidopsis tRF Leu (with 5Gs) or here a tRF Gly (where an artificial 5' end with 4Gs has been created) cannot form rG4s *in vitro*. Whether the sequence of tRF Ala and/or a peculiar unknown structure is crucial so that rG4 can be formed remains to be elucidated. The *in vivo* stability of the rG4 structure remains unclear. So far, it is admitted that although numerous rG4 structures can be formed in mRNAs *in vitro*, most of them are disassembled *in vivo* (32). Our data suggest that tRF Ala rG4 complexes could be relatively stable within foci. No protein factors other than the RNA helicase DExH1 have been enriched in the pull-down assays, but we cannot exclude that such factors do exist and are present within the foci. Future experiments, including cross-linking ones, will address this question. While the question of rG4 stability *in vivo* is under investigation, so is that of their disassembly *in vivo*. What are the proteins involved in rG4s resolving? Several helicases, including DHX9, DHX21, and DHX36, bind and resolve rG4 structures in humans. In addition, an affinity proteomics approach identified 80 potential interactors of rG4s located in the 5' UTR of mRNAs (17). This high number of rG4-interacting proteins suggests that rG4s do exist *in vivo* and that the unwinding of rG4s is biologically relevant. Many of these proteins contain a GAR domain enabling interaction with the rG4 structure. Human G4-binding proteins with another motif (RGRGRGRGGSGGSGGRGRG) called NIKI for "Novel Interesting Quadruplex Interaction" were also found (33). Plant proteins containing either a GAR or a NIKI motif have been identified and speculated to interact with G4 structures (29). Among them, 7 RNA helicases were identified, including DExH1 (Supplementary Figure 6). While for 6 RNA helicases, the GAR or NIKI motif is localized at their C-terminal extremities, DExH1 is the only RNA helicase with a glycine-arginine rich domain localized at the N-terminal. Here, we provide the first evidence that a plant RNA helicase, DExH1, can interact with the rG4 structure of a small non-coding RNA, i. e. the tRF Ala. If this protein has an affinity for other RNAs *in vivo* remains to be studied, and the biological context in which this RNA helicase acts on tRF Ala rG4 structures to resolve them *in planta*.

EXPERIMENTAL PROCEDURES

Plant material and growth conditions

Arabidopsis cell culture is of Landsberg ecotype. The cell culture is maintained by transferring every week the old culture to a new medium containing MS256, 3% (w/v) sucrose, 5.4 μM α -naphthaleneacetic acid (NAA), 0.23 μM kinetin, pH 5.6. Other plants used in this study are of Columbia (Col-0) ecotype. T-DNA insertion line *dexh1*^{KO} (SALK_062060) was obtained from the Arabidopsis Biological Resource Center. Lines expressing pDCP1-DCP1-YFP (14) and pPAB2-tRFP-PAB2 (13) were provided by R. Merret.

Details of growth conditions used for protoplast preparation are given in (34). For *in vitro* seedlings culture, sterilized seeds are sown on agar plates containing MS255 supplemented with 1% (w/v) sucrose at pH 5.7. After 2 to 4 days of synchronization at 4°C, seedlings are grown 7 days at 22°C under a 16 h light photoperiod. For induction of stress granules, transfected protoplasts were incubated in a water bath 1 h at 37°C.

Protoplasts isolation and transfection

Details about protoplasts isolation and transfection are provided in (34). Briefly, protoplasts were isolated from cell culture or 7-day-old seedlings. After digestion overnight at 26°C with gentle shaking and washes with W5 buffer, protoplasts were counted with a hemocytometer. Protoplasts concentration were adjusted to 6×10^6 cells/mL in MMG buffer. For tRF subcellular localization, 1.5 μL of 100 μM oligoribonucleotide, at least 2,5 μg of carrier DNA plasmid and 100 μL of polyethylene glycol (PEG) transfection solution were added to 100 μL of protoplasts preparation. For tRF pulldown, 15 μL of 100 μM oligoribonucleotide, at least 40 μg of carrier DNA plasmid and 1 mL of PEG transfection solution were added to 1 mL of protoplasts preparation. For both experiments, after gentle mixing, tubes were incubated 30 min at room temperature with moderate rotation. After transfection, protoplasts were washed three times and either resuspended in 100 μL of W1 buffer for microscopy analysis or pelleted and used directly for streptavidin pull-down. For localization of DEXH1-GFP, 60 μg of DEXH1-GFP Gateway vector, 60 μg of carrier DNA plasmid and 400 μL of PEG transfection solution were added to 400 μL of protoplasts preparation. After gentle mixing, samples were left for 15 min at room temperature. After transfection, protoplasts were washed three times with W5 buffer, resuspended in 400 μL of W1 buffer and transferred to a 24 wells plate. Protoplasts were incubated overnight, at 25°C and in the dark, to allow the expression of DEXH1-GFP.

To follow the incorporation and stability of tRF Ala and RNA R, after transfection, small RNAs were extracted from RNase-treated protoplasts and analyzed by northern blot essentially as described in (34).

Subcellular localization in protoplasts

Subcellular localization of fluorescently labelled-tRFs and RFP/YFP fusion proteins were analyzed using a LSM780 confocal microscope (Zeiss) with a 20x magnification lens with the following settings: excitation at 633 nm with collection bandwidth 650-665 nm for Alexa Fluor 647-labelled tRFs; excitation at 561 nm with collection bandwidth 571-588 nm for Mitotracker; excitation at 488 nm with collection bandwidth at 487-569 nm, 532-614 nm, 455-570 nm for YFP, tRFP and GFP respectively; chlorophyll autofluorescence was collected with a bandwidth 523-756 nm. For mitochondria visualization, transfected protoplasts were incubated 15 min with Orange Mitotracker at a final concentration of 20 nM. Foci number per protoplast were determined by local maxima detection method with ImageJ. Co-localization analysis was performed with ImageJ. Foci in each image were determined with a user-supervised local maxima detection method (script from J. Mutterer, <https://bit.ly/mcfoci>). For every detected foci, pixel intensities in channels visualizing Alexa 680 and RFP/YFP fusion proteins were measured. These measures are reported on scatter plot.

Affinity pull-down of biotinylated RNA

After protoplast transfection and washes, totality of transfected protoplasts (i.e. 6×10^6 protoplasts) were pelleted, resuspended in 500 μ L of lysis buffer (Tris-HCl 50 mM, pH 7,5 ; NaCl 150 mM ; MgCl₂ 2 mM ; NP40 1 %, protease inhibitors (cOmplete, EDTA-free Protease Inhibitor Cocktail, Roche)) and gently shaken for 10 min at 4°C. Streptavidin columns prepared by adding 20 μ L of Streptavidin Sepharose High Performance (GE Healthcare) into receiver columns (35 μ m, Macherey & Nagel) were washed twice with 500 μ L of washing buffer (Tris-HCl 5 mM, pH 7.5 ; NaCl 1 M ; EDTA 1 mM, pH 8.2) followed by centrifugation at 11000 g for 15 sec at 4°C. To remove cell debris, lysed protoplasts were centrifuged at 200 g for 30 sec, 4°C. Supernatants loaded on the Streptavidin Sepharose columns were incubated on a rotor wheel for 20 min at 4°C. Columns were washed 5 times with 500 μ L of SSC 0,1x followed each time by a centrifugation step (15 sec at 200 g, 4°C).

The elution step was performed in 100 μ L of elution buffer (Tris-HCl 50 mM, pH 6.8 ; DTT 50 mM ; SDS 1% ; EDTA 1 mM ; bromophenol blue 0.005% (p/v) ; glycerol 10 % (v/v)) by heating for 5 min at 95°C. After centrifugation for 15 sec at 100g, the supernatants containing the eluted proteins were used for mass spectrometry.

Mass spectrometry analysis and data processing

All steps were performed at the "Plateforme Protéomique Strasbourg-Esplanade" (<http://www-ibmc.u-strasbg.fr/proteo/Web/accueil.htm>). Eluted proteins were digested with trypsin (Promega) and analyzed by nano LC-MS/MS on a QEactive+ mass spectrometer (ThermoFisher). Data were searched against the TAIR10 database, peptides were identified with Mascot algorithm (version 2.5, Matrix Science), and data were imported into Proline software (version 1.4). Three biological replicates were performed for each set of experiment. Protein identification was validated using FDR < 1% (False discovery rate). The statistical analysis based on spectral counts was performed using a homemade R package (<https://github.com/hzuber67/IPinquiry4>) as described in (35). Volcano plots display the adjusted p values and fold changes in Y- and X-axis, respectively, and show the enrichment of proteins co-purified with biotinylated tRF Ala as compared to controls.

NMM Fluorescence analysis

The biotinylated oligoribonucleotides used for protoplast transfection were analyzed. For analysis on native RNA gel, oligoribonucleotides were diluted to 10 μ M in 100 mM KCl solution, heated 10 min at 95°C and left to cool to room temperature for 1 hour 30 min. N-Methyl Mesoporphyrin IX (NMM, Livechem Logistics), a molecule that has a strong affinity for R4 structures, was added to obtain a final concentration of 40 ng/ μ L and left 30 min. Glycerol was added to samples before loading on 15% polyacrylamide-TBE gel. NMM was visualized at 530 nm using a fluorescent imaging system (Fusion FX, Vilber). Gel was then post-stained with Ethidium Bromide and visualized using a classical UV transilluminator. For NMM fluorescence measurements, the protocol adapted from (36) , was used. Briefly, 3 μ L of 100 μ M oligoribonucleotide were mixed with 27 μ L of a solution containing 10 mM sodium phosphate buffer pH 6.4, 100 mM KCl, 4 mM MgCl₂ and 2,5 ng/ μ L of NMM.

The fluorescence intensity was measured using FLUOstar Omega (BMG Labtech) plate reader with excitation wavelength of 355 nm and excitation of 590 nm. Empty wells fluorescence intensities values were subtracted from samples fluorescence intensities values. Graph shows means of two repeated measurements

Miscellaneous

The oligoribonucleotide and desoxyribonucleotide sequences used in this study are listed in Supplementary Table S2. They were chemically synthesized and HPLC purified by the company Integrated DNA Technologies IDT.

ACKNOWLEDGEMENTS

This work was supported by the Centre National de la Recherche Scientifique (CNRS) in association with the University of Strasbourg. This work of the Interdisciplinary Thematic Institute IMCBio, as part of the ITI 2021-2028 program of the University of Strasbourg, CNRS and Inserm, was supported by IdEx Unistra (ANR-10-IDEX-0002), STRAT'US (ANR 20-SFRI-0012) and EUR IMCBio (ANR-17-EURE-0023) under the framework of the French Investments for the Future Program. We wish to thank the Gene Expression platform of IBMP for DNA sequencing and the Strasbourg-Esplanade Proteomics facility for the proteomic analysis. We are grateful to Clara Nuninger, Elodie Ubrig and Marie Bernhard for technical supports. R. Merret (LGDP, Perpignan, France) is acknowledged for the gift of *A. thaliana* Col-0 seeds expressing the pPABP2-tRFP-PABP2 and pDCP1-YFP-DCP1 fusion proteins.

AUTHOR CONTRIBUTIONS

MC performed most of the experiments, designed work, analyzed the data, critically participated in the writing of the manuscript and preparing the figures. TS-G, GH participated to some experiments of protoplasts transfection and confocal microscopy visualization and CB participated to experiments dedicated to G-quadruplex analysis. LD supervised the work, design experiments, analyzed the data, wrote the manuscript.

CONFLICT OF INTEREST

No potential conflict of interest was reported by the authors.

DATA AVAILABILITY STATEMENT

All the relevant data can be found within the manuscript and its supporting materials.

SUPPORTING INFORMATION

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

Uncropped blots and gels are provided as source data (Supplementary Figure 11 and 12).

REFERENCES

1. Magee, R. and Rigoutsos, I. (2020) On the expanding roles of tRNA fragments in modulating cell behavior. *Nucleic acids research*, **48**, 9433-9448.
2. Ren, B., Wang, X., Duan, J. and Ma, J. (2019) Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. *Science*, **365**, 919-922.
3. Loss-Morais, G., Waterhouse, P.M. and Margis, R. (2013) Description of plant tRNA-derived RNA fragments (tRFs) associated with argonate and identification of their putative targets. *Biology direct*, **8**, 6.
4. Martinez, G., Choudury, S.G. and Slotkin, R.K. (2017) tRNA-derived small RNAs target transposable element transcripts. *Nucleic acids research*, **45**, 5142-5152.
5. Hsieh, L.C., Lin, S.I., Kuo, H.F. and Chiou, T.J. (2010) Abundance of tRNA-derived small RNAs in phosphate-starved Arabidopsis roots. *Plant signaling & behavior*, **5**.
6. Zhang, S., Sun, L. and Kragler, F. (2009) The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation. *Plant physiology*, **150**, 378-387.
7. Lalande, S., Merret, R., Salinas-Giege, T. and Drouard, L. (2020) Arabidopsis tRNA-derived fragments as potential modulators of translation. *RNA Biol*, 1-12.
8. Cognat, V., Morelle, G., Megel, C., Lalande, S., Molinier, J., Vincent, T., Small, I., Duchene, A.M. and Marechal-Drouard, L. (2017) The nuclear and organellar tRNA-derived RNA fragment population in Arabidopsis thaliana is highly dynamic. *Nucleic acids research*, **45**, 3460-3472.
9. Anderson, P. and Ivanov, P. (2014) tRNA fragments in human health and disease. *FEBS letters*.
10. Yamasaki, S., Ivanov, P., Hu, G.f. and Anderson, P. (2009) Angiogenin cleaves tRNA and promotes stress-induced translational repression. *The Journal of Cell Biology*, **185**, 35-42.

11. Lyons, S.M., Kharel, P., Akiyama, Y., Ojha, S., Dave, D., Tsvetkov, V., Merrick, W., Ivanov, P. and Anderson, P. (2020) eIF4G has intrinsic G-quadruplex binding activity that is required for tiRNA function. *Nucleic acids research*, **48**, 6223-6233.
12. Hsieh, L.C., Lin, S.I., Shih, A.C., Chen, J.W., Lin, W.Y., Tseng, C.Y., Li, W.H. and Chiou, T.J. (2009) Uncovering small RNA-mediated responses to phosphate deficiency in Arabidopsis by deep sequencing. *Plant physiology*, **151**, 2120-2132.
13. Merret, R., Carpentier, M.C., Favory, J.J., Picart, C., Descombin, J., Bousquet-Antonelli, C., Tillard, P., Lejay, L., Deragon, J.M. and Charng, Y.Y. (2017) Heat Shock Protein HSP101 Affects the Release of Ribosomal Protein mRNAs for Recovery after Heat Shock. *Plant physiology*, **174**, 1216-1225.
14. Merret, R., Descombin, J., Juan, Y.T., Favory, J.J., Carpentier, M.C., Chaparro, C., Charng, Y.Y., Deragon, J.M. and Bousquet-Antonelli, C. (2013) XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. *Cell reports*, **5**, 1279-1293.
15. Chen, M.C., Tippana, R., Demeshkina, N.A., Murat, P., Balasubramanian, S., Myong, S. and Ferre-D'Amare, A.R. (2018) Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature*, **558**, 465-469.
16. Vaughn, J.P., Creacy, S.D., Routh, E.D., Joyner-Butt, C., Jenkins, G.S., Pauli, S., Nagamine, Y. and Akman, S.A. (2005) The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates. *J Biol Chem*, **280**, 38117-38120.
17. Herdy, B., Mayer, C., Varshney, D., Marsico, G., Murat, P., Taylor, C., D'Santos, C., Tannahill, D. and Balasubramanian, S. (2018) Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts. *Nucleic acids research*, **46**, 11592-11604.
18. Mingam, A., Toffano-Nioche, C., Brunaud, V., Boudet, N., Kreis, M. and Lecharny, A. (2004) DEAD-box RNA helicases in Arabidopsis thaliana: establishing a link between quantitative expression, gene structure and evolution of a family of genes. *Plant Biotechnol J*, **2**, 401-415.

19. Sulkowska, A., Auber, A., Sikorski, P.J., Silhavy, D.N., Auth, M., Sitkiewicz, E., Jean, V., Merret, R.M., Bousquet-Antonelli, C.C. and Kufel, J. (2020) RNA Helicases from the DEA(D/H)-Box Family Contribute to Plant NMD Efficiency. *Plant Cell Physiol*, **61**, 144-157.
20. Lyons, S.M., Gudanis, D., Coyne, S.M., Gdaniec, Z. and Ivanov, P. (2017) Identification of functional tetramolecular RNA G-quadruplexes derived from transfer RNAs. *Nature communications*, **8**, 1127.
21. Gu, H., Lian, B., Yuan, Y., Kong, C., Li, Y., Liu, C. and Qi, Y. (2022) A 5' tRNA-Ala-derived small RNA regulates anti-fungal defense in plants. *Sci China Life Sci*, **65**, 1-15.
22. Emara, M.M., Ivanov, P., Hickman, T., Dawra, N., Tisdale, S., Kedersha, N., Hu, G.F. and Anderson, P. (2010) Angiogenin-induced tRNA-derived Stress-induced RNAs Promote Stress-induced Stress Granule Assembly. *Journal of Biological Chemistry*, **285**, 10959-10968.
23. Ivanov, P., Emara, Mohamed M., Villen, J., Gygi, Steven P. and Anderson, P. (2011) Angiogenin-Induced tRNA Fragments Inhibit Translation Initiation. *Molecular Cell*, **43**, 613-623.
24. Megel, C., Hummel, G., Lalande, S., Ubrig, E., Cognat, V., Morelle, G., Salinas-Giege, T., Duchene, A.M. and Marechal-Drouard, L. (2019) Plant RNases T2, but not Dicer-like proteins, are major players of tRNA-derived fragments biogenesis. *Nucleic acids research*, **47**, 941-952.
25. Lejault, P., Mitteau, J., Sperti, F.R. and Monchaud, D. (2021) How to untie G-quadruplex knots and why? *Cell Chem Biol*, **28**, 436-455.
26. Spiegel, J., Adhikari, S. and Balasubramanian, S. (2020) The Structure and Function of DNA G-Quadruplexes. *Trends Chem*, **2**, 123-136.
27. Dumas, L., Herviou, P., Dassi, E., Cammas, A. and Millevoi, S. (2021) G-Quadruplexes in RNA Biology: Recent Advances and Future Directions. *Trends Biochem Sci*, **46**, 270-283.
28. Griffin, B.D. and Bass, H.W. (2018) Review: Plant G-quadruplex (G4) motifs in DNA and RNA; abundant, intriguing sequences of unknown function. *Plant science : an international journal of experimental plant biology*, **269**, 143-147.
29. Volnà, A., Bartas, M., Nezval, J., Spunda, V., Pecinka, A. and Cerven, J. (2021) Searching for G-quadruplex-binding proteins in plants: new insights into possible G-quadruplex regulation. *BioTech*, **10**.

30. Yang, X., Cheema, J., Zhang, Y., Deng, H., Duncan, S., Umar, M.I., Zhao, J., Liu, Q., Cao, X., Kwok, C.K. *et al.* (2020) RNA G-quadruplex structures exist and function in vivo in plants. *Genome Biol*, **21**, 226.
31. Cho, H., Cho, H.S., Nam, H., Jo, H., Yoon, J., Park, C., Dang, T.V.T., Kim, E., Jeong, J., Park, S. *et al.* (2018) Translational control of phloem development by RNA G-quadruplex-JULGI determines plant sink strength. *Nat Plants*, **4**, 376-390.
32. Guo, J.U. and Bartel, D.P. (2016) RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science*, **353**.
33. Brazda, V., Cerven, J., Bartas, M., Mikyskova, N., Coufal, J. and Pecinka, P. (2018) The Amino Acid Composition of Quadruplex Binding Proteins Reveals a Shared Motif and Predicts New Potential Quadruplex Interactors. *Molecules*, **23**.
34. Lalande, S., Chery, M., Ubrig, E., Hummel, G., Kubina, J., Geldreich, A. and Drouard, L. (2020) Transfection of Small Noncoding RNAs into Arabidopsis thaliana Protoplasts. *Methods Mol Biol*, **2166**, 413-429.
35. Scheer, H., de Almeida, C., Ferrier, E., Simonnot, Q., Poirier, L., Pflieger, D., Sement, F.M., Koechler, S., Piermaria, C., Krawczyk, P. *et al.* (2021) The TUTase URT1 connects decapping activators and prevents the accumulation of excessively deadenylated mRNAs to avoid siRNA biogenesis. *Nature communications*, **12**, 1298.
36. Ivanov, P., O'Day, E., Emara, M.M., Wagner, G., Lieberman, J. and Anderson, P. (2014) G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. *Proceedings of the National Academy of Sciences of the United States of America*.
37. Mishra, S.K., Jain, N., Shankar, U., Tawani, A., Sharma, T.K. and Kumar, A. (2019) Characterization of highly conserved G-quadruplex motifs as potential drug targets in *Streptococcus pneumoniae*. *Sci Rep*, **9**, 1791.

Figure 1. Subcellular localization of tRF-5D Ala transfected into *A. thaliana* protoplasts. (a) Scheme of the experiments performed upon protoplasts transfection with either fluorescent or biotinylated tRF. (b) Confocal images of protoplasts from *A. thaliana* cell culture transfected with H₂O, fluorescently-labelled RNA R or tRF Ala. Representative images from one biological replicate. Other images are provided in Supplementary Figure 1. White: bright light; blue: chlorophyll autofluorescence; magenta: Alexa Fluor 647 fluorescence; merged: blue and magenta channels; z-stack: z projection of merged planes from the bottom to the top of protoplasts; zoom: zooms of the region insets of the z-stacks are presented. Scale bar: 20 μ m. The random RNA R and tRNA Ala sequences are indicated on the left. (c) Scatter plot of foci number per protoplast. Bars represent the mean and standard deviation from 3 biological replicates. ANOVA followed by Tukey's multiple comparison test was used to determine significance (****: p-value <0,0001; ns: not significant). (d) Northern blot detection of RNAs extracted from *A. thaliana* protoplasts transfected either with H₂O or an equimolar mix of tRF Ala and RNA R, and treated (+) or not (-) with RNases. Hybridizations were performed in the presence of radiolabelled oligonucleotide probes complementary either to the 5' extremity of *A. thaliana* tRNA^{Ala} (AGC) or to the RNA R control.

a

Name	Sequence
RNA R	CCUACAUCGAGUCUACCAUC
tRF Ala	GGGG AUGUAGCUCAGAUGGU
tRF Arg	GCGUCUGUAGCUCAGUGGA
tRF His	GGCGGAUGUAGCCAAGUGGA
tRF Asp	GUCGUUGUAGUAUAGUGGUAAGUAUCCCCGCCU
tRF Gly	GCGUCUGUAGUCCAACGGU

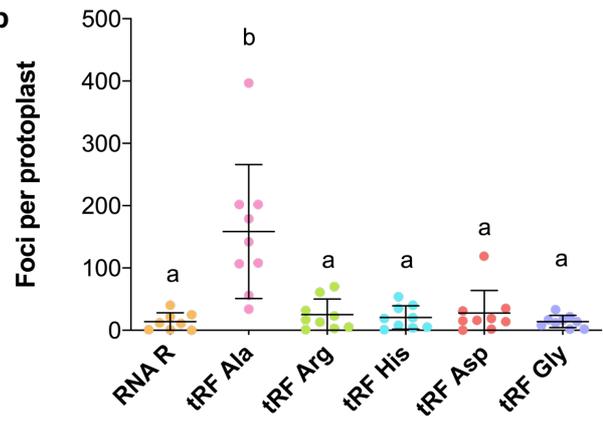
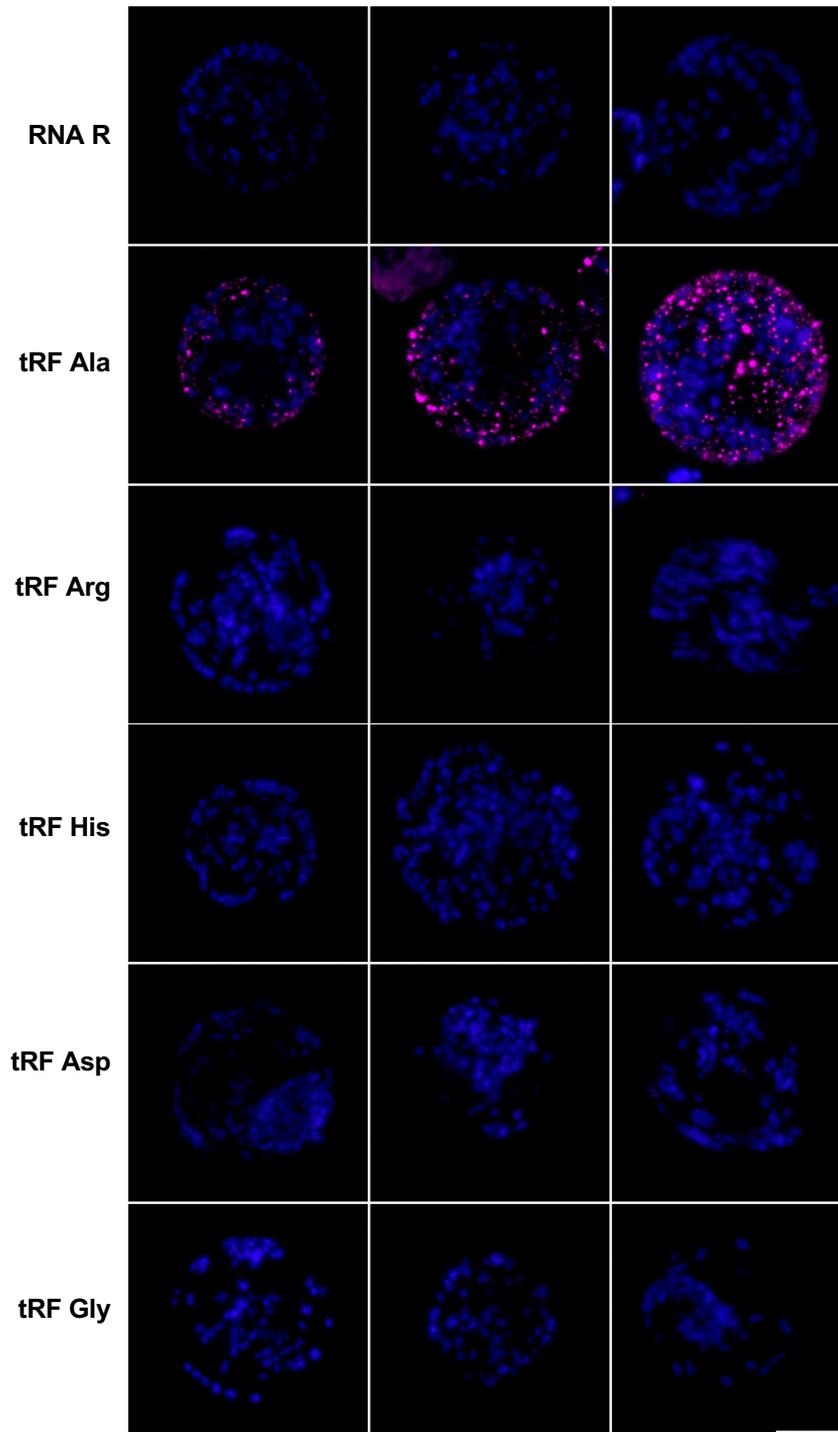
b**c**

Figure 2. Subcellular localization of different tRFs transfected into *A. thaliana* protoplasts. **(a)** Table of tRFs used and their sequences. **(b)** Scatter plot of foci number per protoplast. Bars represent the mean and standard deviation from 3 biological replicates. Significant differences between groups are indicated with letters (a and b) based on ANOVA followed by Tukey's multiple comparison test (p -value $\leq 0,0001$). **(c)** Confocal images (z-projection) of protoplasts from *A. thaliana* cell culture transfected with fluorescently labelled RNA R, tRF Ala, tRF Arg, tRF His, tRF Asp, and tRF Gly. The tRF Asp is a tRF-5A. The others are tRFs-5D. For each tRF, three biological replicates have been done. Only three representative images are shown. Other images are provided in Supplementary Figure 2. Scale bar: 20 μ m. Merged blue (chlorophyll autofluorescence) and magenta (Alexa Fluor 647 fluorescence) channels.

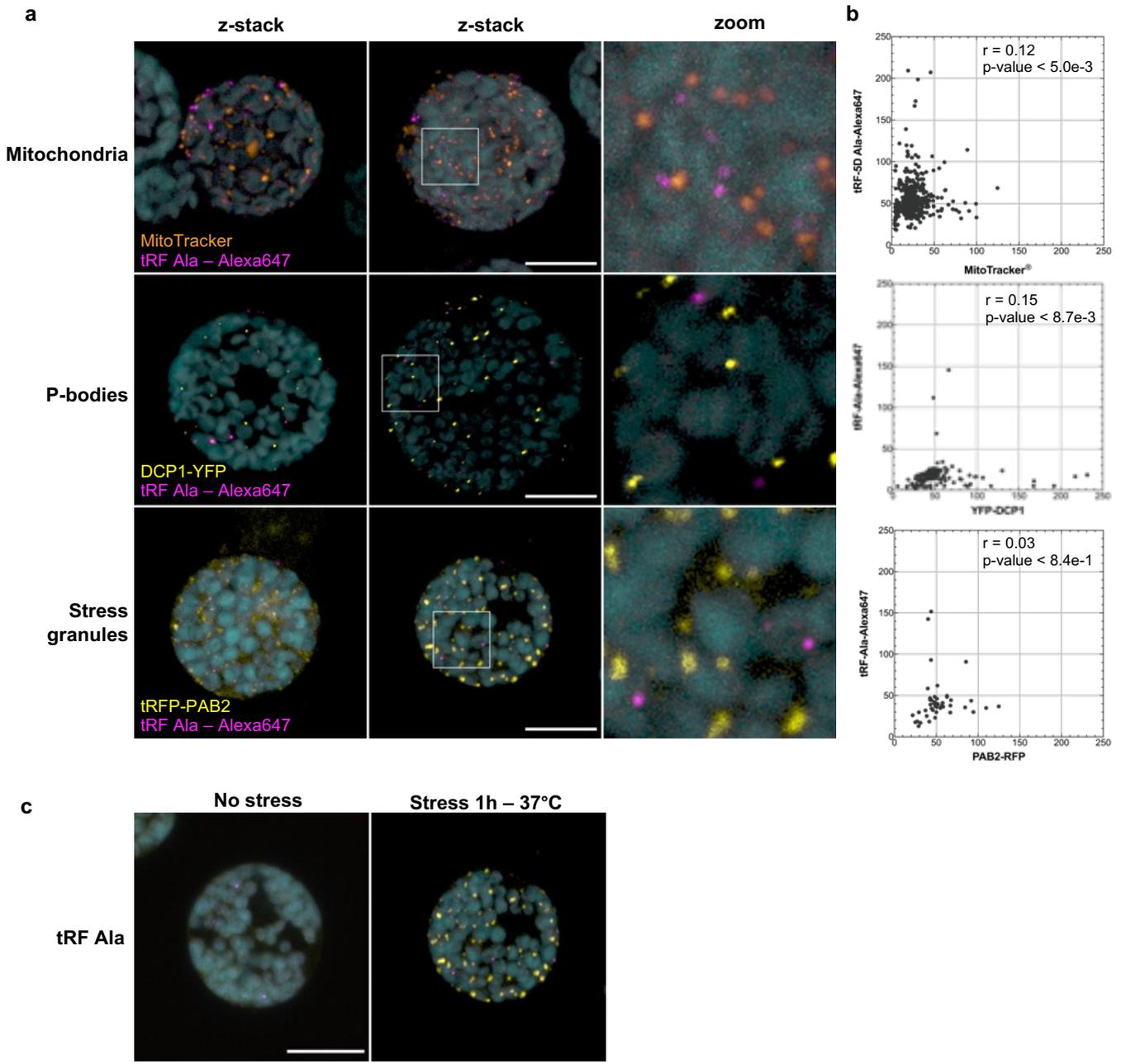
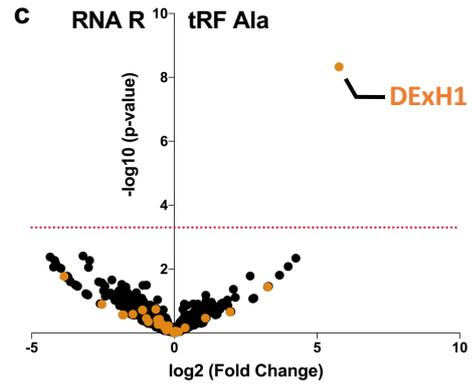


Figure 3. Subcellular localization analysis of tRF Ala foci compared to mitochondria, P-bodies, or stress granules. **(a)** Two representative confocal images (z-projection) of Arabidopsis protoplasts prepared from seedlings of wild-type, DCP1-YFP and tRFP-PAB2 transgenic lines transfected with fluorescently-labelled tRF Ala. Upon transfection, protoplasts from wild-type seedlings were stained with Mitotracker, and protoplasts from tRFP-PAB2 transgenic lines were stressed for 1 hour at 37°C. Scale bar: 20 µm. Cyan: chlorophyll, orange: Mitotracker, magenta: Alexa Fluor 647, and yellow: DCP1-YFP and tRFP-PAB2. Zooms of the region insets on the z-stacks are presented. **(b)** Dot plots quantification from experiments performed in **(a)**. The tRF Ala foci pixel intensities observed in the magenta channel were measured in the second channel (orange for Mitotracker and yellow for DCP1-RFP or tRFP-PAB2). x-axis: Alexa Fluor 647; y-axis: MitoTracker/YFP/tRFP fluorescence. Each plot was obtained from one biological replicate from independent images (Mitochondria plot: n=7; P-bodies plot: n=10; Stress granules plot: n=2). The Pearson's correlation coefficients (r) and p-values are indicated in each plot. **(c)** Arabidopsis protoplasts prepared from seedlings of tRFP-PAB2 transgenic lines transfected with fluorescently labelled tRF Ala. Protoplasts were heat-stressed or not for 1 hour at 37°C upon transfection.

a

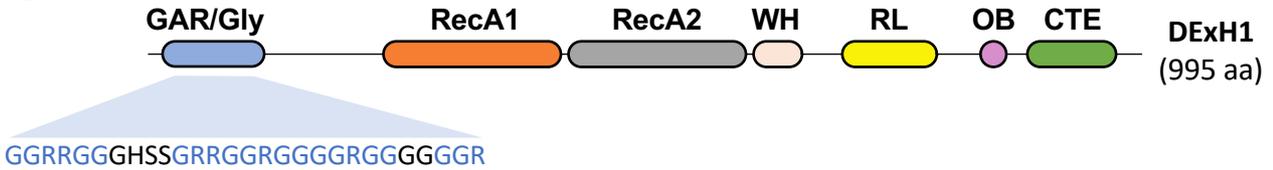
Replicate number	H ₂ O			RNA R			tRF Ala		
	#1	#2	#3	#1	#2	#3	#1	#2	#3
Mean of identified proteins	3354			3419			3450		
AT2G35920 DExH1	0	0	0	0	0	0	31	25	39



b

Replicate number	RNA R			tRF Ala			tRF Arg			tRF His			tRF Asp			tRF Gly		
	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3
Mean of identified proteins	2890			2478			2542			2584			2635			2498		
AT2G35920 DExH1	0	0	1	11	37	15	0	0	1	0	0	2	0	0	0	0	0	0

d



e

DExH1 MPPHGPSQGGRRGGGHSSGRRGGRG-GGRRGGGGGRGEQRWWDVPVWRAERLRQQQAEMEVLDENENWNKIEQWKTGGGEQLIKRNF SRGDQQLSDMALQ
 DHX36 MSYDYHQNWGRDGGPRSSGGYGGGPAGGGHGNRSGGGGGGGGGRRGR-----GRHPGHLKGREIGMWYAKKQGQKN--KEAERQERAVVHMDERREEQIVQL
 RH6 MNNNNNRGRFPPGIGAAGGPDPNFQSRNP
 RH11 MSASWAD----VADSEKAVSQSKPPYVPHLRN
 RH12 MN---TNRGRYPPGVTGTRGAP-----PNP
 RH14 MAATAAASVVRYAPEDHTLPKPWKGLIDDRGTG
 RH46 MAATASA--IRYAPEDPNLPKPWKGLVDSRTGY

Figure 4. tRF Ala interacts with the DEAH-box RNA helicase DExH1. **(a)** Table summarizing the enrichment of DExH1 peptides (highlighted in yellow) upon protoplasts transfection with biotinylated tRF Ala compared to transfection with H₂O or with RNA R. Transfection experiments were followed by pull-down assay on streptavidin beads and LC-MS/MS analysis. The three biological replicates are numbered from #1 to #3, and the number of proteins was identified by Mascot with FDR < 1%. **(b)** Table summarizing the enrichment of DExH1 peptides (highlighted in yellow) upon protoplasts transfection with biotinylated tRF Ala compared to transfection with RNA R, tRF Arg, tRF His, tRF Asp, and tRF Gly. The three biological replicates were performed as in **(a)**. **(c)** Volcano plot of the experiments described in **(b)** showing the DExH1 protein differentially enriched in tRF Ala transfected protoplasts versus RNA R transfected protoplasts. Y- and X-axis display the p-value (-log₁₀ (p-value)) and fold change (log₂ (Fold Change)) respectively. The red dashed line indicates the significant threshold (p-value < 0,0005). Dot color code: RNA helicases in orange, other proteins in black. **(d)** Domain organization of the *A. thaliana* DExH1 RNA helicase (Gene Accession number: At2g35920). The different domains, Gly (Glycine-rich element), RecA1 and RecA2 (RecA-like), WH (winged-helix), RL (ratchet-like), OB (oligosaccharide-binding-fold-like), and CTE (C-terminal extension), were identified by comparison with the *H. sapiens* DEAH/RHA DHX36 sequence (Supplementary Figure 5), (15). The glycine-arginine (GAR) sequence present within the Gly domain is indicated. **(e)** Alignment of the N-terminal sequences of DExH1, DEX36, and of five other *A. thaliana* representative DEA(D/H)-box RNA helicases RH6 (AT2G45810), RH11 (AT3G58510), RH12 (AT3G61240), RH14 (AT3G01540) and RH46 (AT5G14610) (19). The GAR domain of DExH1 is circled in blue.

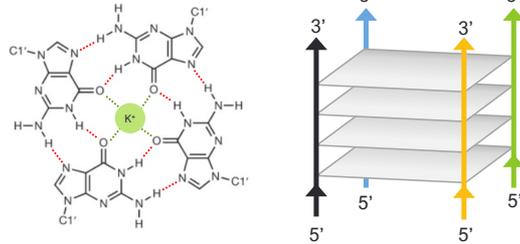
a

Name (organism)	Isodecoder	Sequence	Name in (20)	Ability to fold rG4
tRF Ala (<i>H.s</i>)	AGC/CGC	GGGGG UGUAGCUCAGUGGUAGAGCGCGUGC	5'tiRNA Ala	Yes
tRF Ala (<i>A.t</i>)	CGC	GGGGG UGUAGCUCAUAUUGGUAGAGC	At90	Yes
tRF Leu (<i>A.t</i>)	TAA	GGGG AUAUGGCGAAUUGGUAGACGCCACG	At112A	No

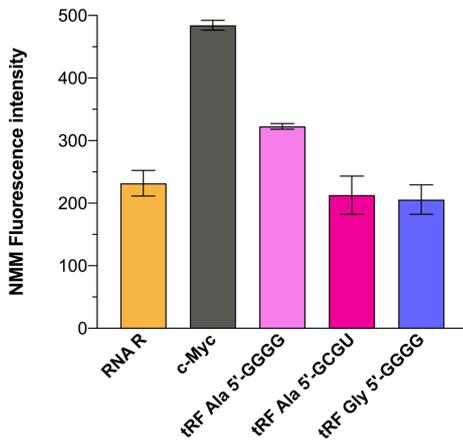
b

Name	Sequence
RNA R	CCUACAUCGAGUCUACCAUC
c-Myc	GGGGAGGGUGGGGAGGGUGGGG
tRF Ala 5'GGGG	GGGG AUGUAGCUCAGAUGGU
tRF Ala 5'GCGU	GCGU AUGUAGCUCAGAUGGU
tRG Gly 5'GGGG	GGGG CUGUAGUCCAACGGU

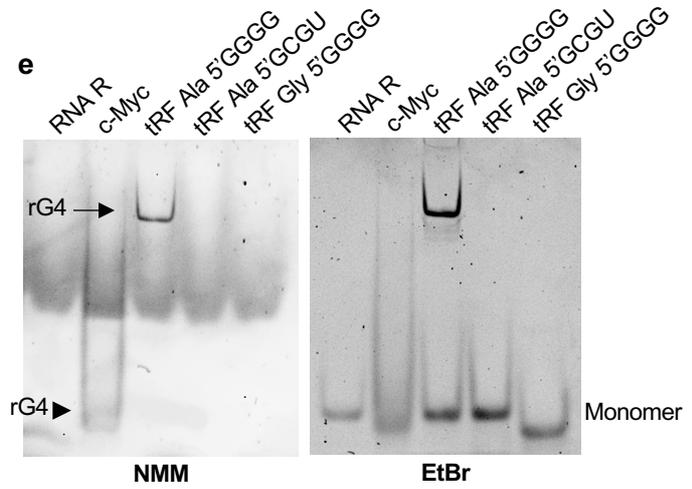
c



d



e



f

	RNA R	c-Myc	tRF Ala 5'-GGGG	tRF Ala 5'-GCGU	tRF Gly 5'-GGGG
Replicate number	#1 #2 #3	#1 #2 #3	#1 #2 #3	#1 #2 #3	#1 #2 #3
Mean of identified proteins	3333	3248	3231	2794	3320
AT2G35920 DExH1	0 0 1	15 19 11	2 13 4	0 0 0	0 0 0

g

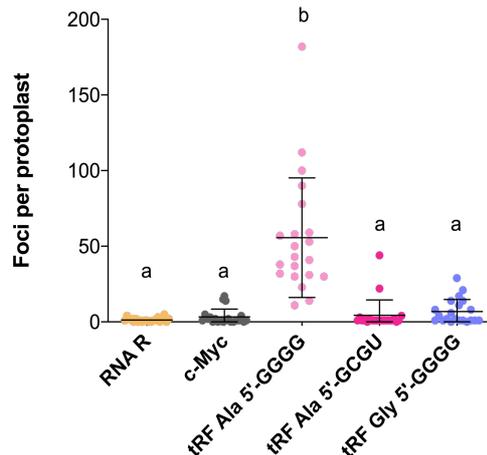


Figure 5. *A. thaliana* tRF Ala can form G-quadruplex presumably required for DExH1 interaction. (a) Table of tRFs known to form rG4. (b) Table of tRFs used for the experiments of this figure. (c) Structure of a G-quartet formed by the Hoogsteen hydrogen-bonded guanines and central cation coordinated to an oxygen atom (adapted from (26)) and intermolecular G-quadruplex parallel topology (adapted from (37)). (d) Histogram showing the NMM fluorescence intensities of the corresponding RNAs. Error bars represent the standard deviation of two independent biological experiments. (e) Analysis of the RNA G-quadruplex status on a native gel. After heat treatment in 100 mM KCl, RNA samples were stained with NMM and visualized at 530 nm following acrylamide gel migration (NMM); the gel was then post-stained with ethidium bromide (EtBr). Intramolecular rG4 formed by c-Myc RNA and intermolecular rG4 formed by tRF Ala are indicated by an arrowhead and an arrow, respectively. (f) Table summarizing the enrichment of DExH1 peptides (highlighted in yellow) upon protoplasts transfection with biotinylated c-Myc RNA and tRF Ala compared to transfection with RNA R, mutated version of tRF Ala (5'-GCGU) or tRF Gly (5'-GGGG). Transfection experiments were followed by a pull-down assay on streptavidin beads and LC-MS/MS analysis. The three biological replicates are numbered from #1 to #3, and the number of proteins was identified by Mascot with FDR < 1%. (g) Scatter plot of foci number per protoplast transfected with the indicated fluorescently-labelled RNAs. Data are from 21 images for each RNA and 3 biological replicates. Significant differences between groups are indicated with letters (a and b) based on ANOVA followed by Tukey's multiple comparison test (p -value ≤ 0.0001).

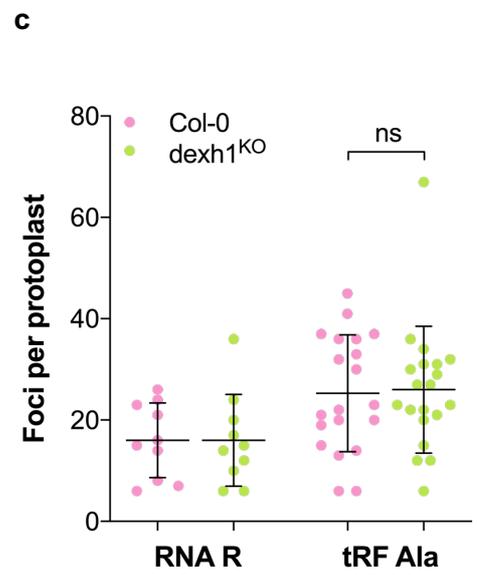
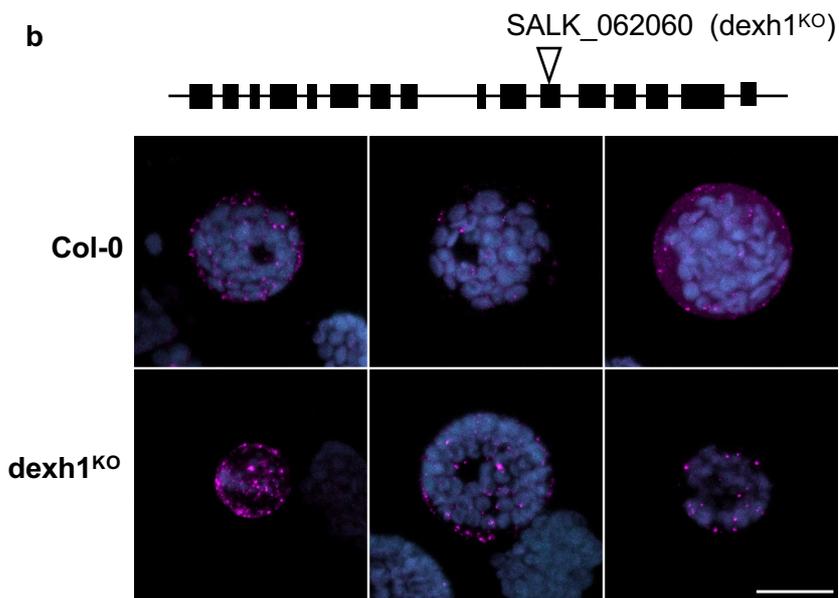
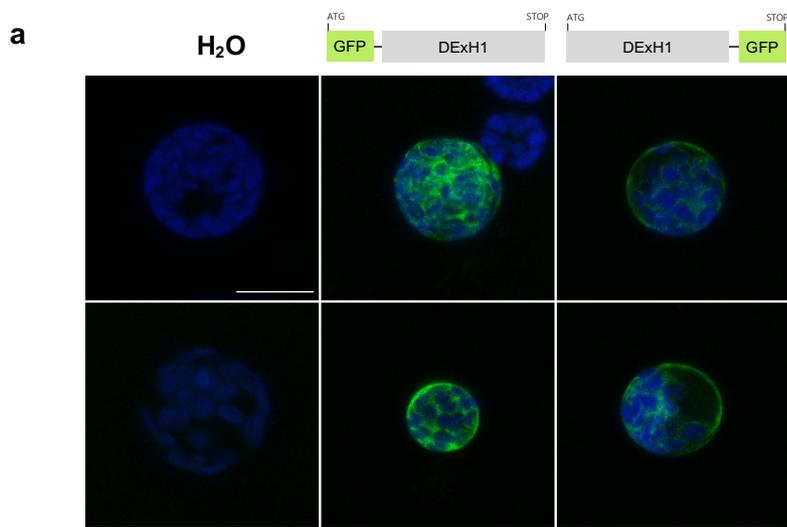
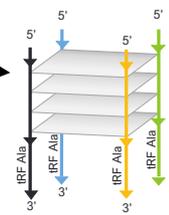


Figure 6. The cytosolic-localized DExH1 protein is not involved in tRF Ala foci biogenesis. **(a)** Confocal images (z-projection) of *A. thaliana* protoplasts transfected with H₂O or plasmid constructs containing the DExH1 sequence fused to the GFP sequence either in the N-terminal or C-terminal. For each condition, two representative images are shown. Other images are provided in Supplementary Figure 8. Scale bar: 20 μm. Merged blue (chlorophyll autofluorescence) and green (GFP fluorescence) channels. **(b)** Confocal images (z-projection) of protoplasts from *A. thaliana* Col-0 wild-type or *dexh1*^{KO} mutant seedlings transfected with fluorescently labelled tRF Ala. For each condition, three representative images are shown. Scale bar: 20 μm. Merged cyan (chlorophyll autofluorescence) and magenta (Alexa Fluor 647 fluorescence) channels. The *A. thaliana* gene (*AT2G35920*) encoding the DExH1 protein is represented above the images. An arrowhead indicates the position of the T-DNA in the insertion line *dexh1*^{KO}. **(c)** Scatter plot of foci number per protoplast transfected either with RNA R or tRF Ala. Bars represent the mean and standard deviation from one biological replicate (10 images for RNA R and 20 images for tRF Ala). ANOVA followed by Tukey's multiple comparison test was used to determine significance (ns: not significant). Data from another biological replicate are shown in supplemental Figure 9.

Conditions
Development
Environment



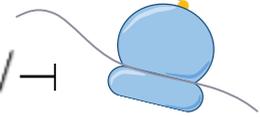
Unfold rG4



Active free
tRF Ala



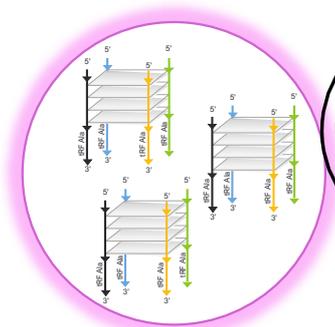
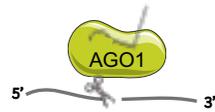
Translation inhibition



Other functions?

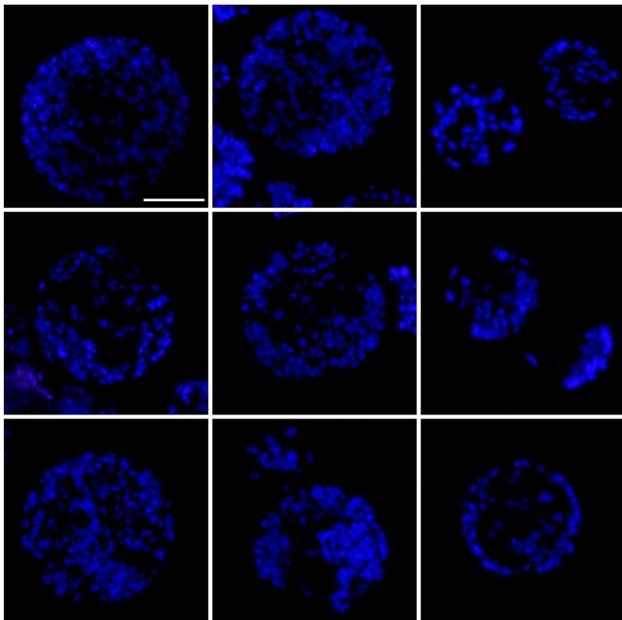
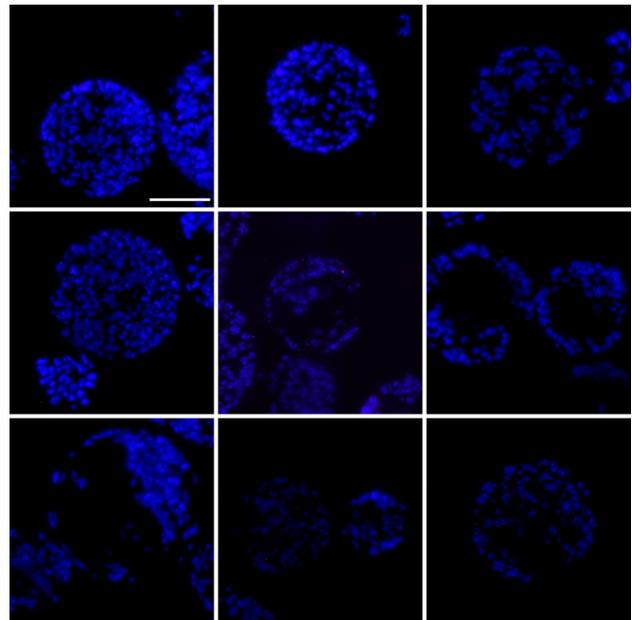
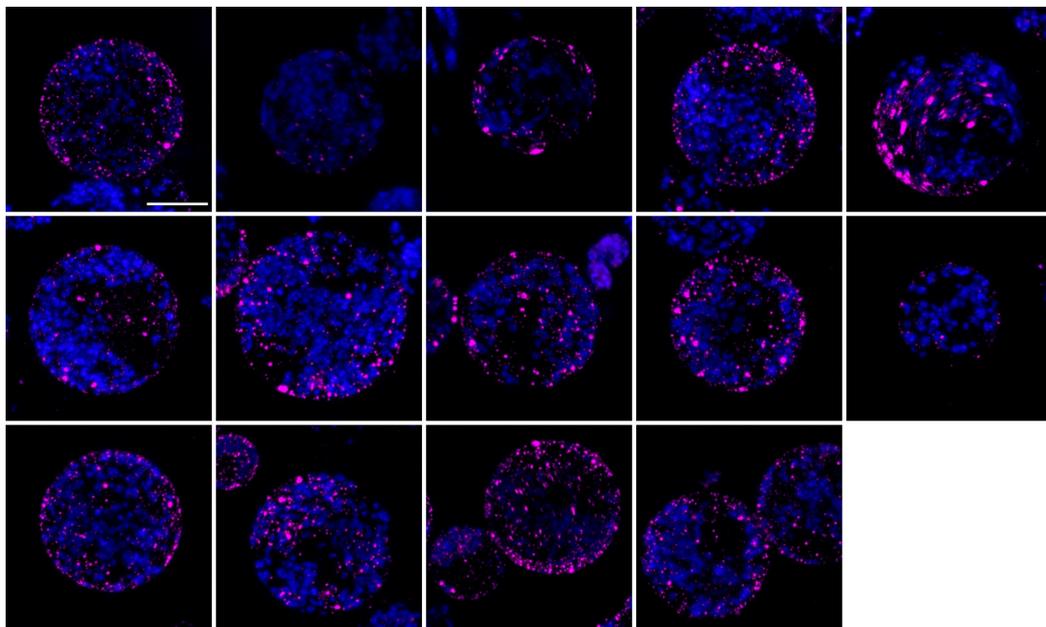


Gene silencing



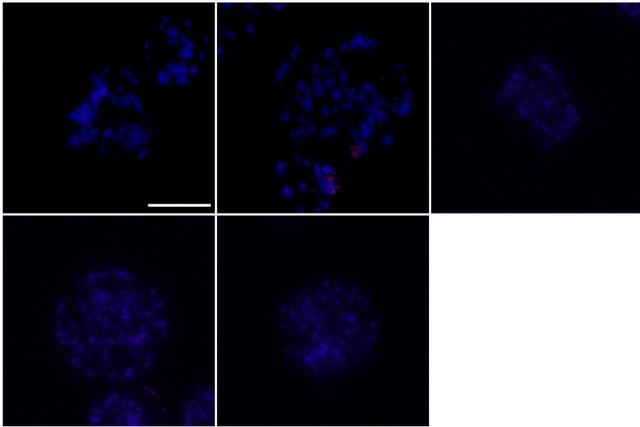
Inactive rG4 tRF Ala in foci

Figure 7. Model of tRF Ala mode of action. Significant amounts of tRF Ala present in plant cells form RNA G-quadruplex (rG4) and are sequestered in foci. Upon unknown conditions, rG4 tRF Ala molecules are released and become accessible to the DExH1 RNA helicase. This RNA helicase binds and unwinds rG4 tRF Ala structures. Free tRF Ala can, in turn, either inhibit global translation, be implicated in specific mRNA degradation (as described in (21)), or play other regulatory, yet unidentified functions.

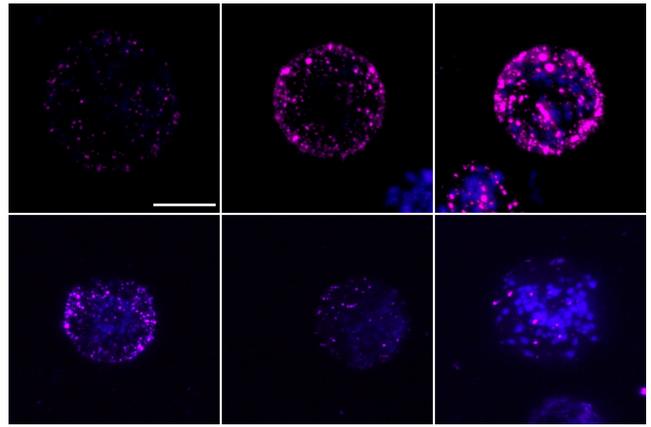
H₂O**RNA R****tRF Ala**

Supplementary Figure 1. Subcellular localization of tRF Ala transfected into *A. thaliana* protoplasts. Confocal images of *A. thaliana* protoplasts transfected with H₂O, RNA R or tRF Ala. Images are merged z-stack images of blue (chlorophyll autofluorescence) and magenta (Alexa Fluor 647 fluorescence; tRFs) channels. Scale bar: 20 μ m.

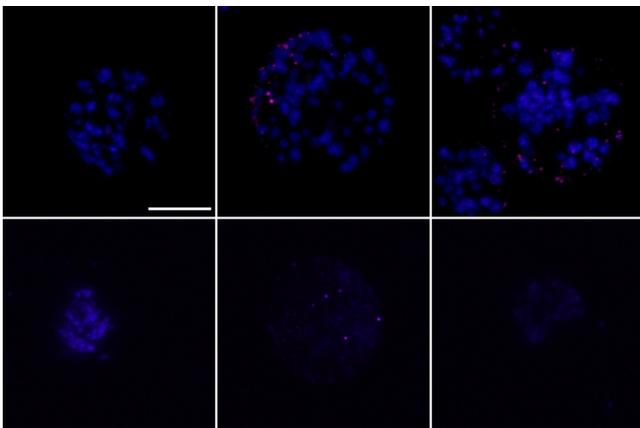
RNA R



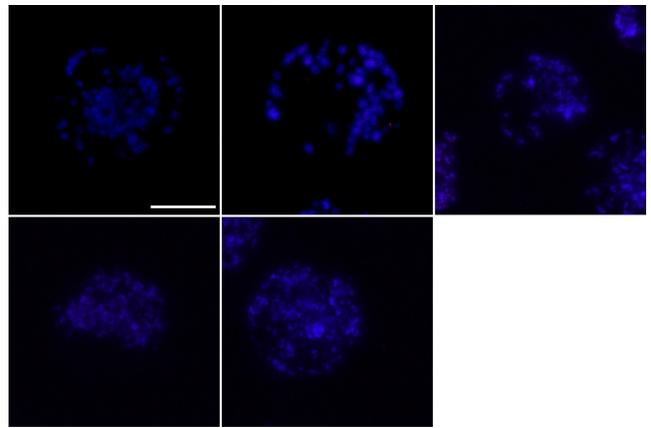
tRF Ala



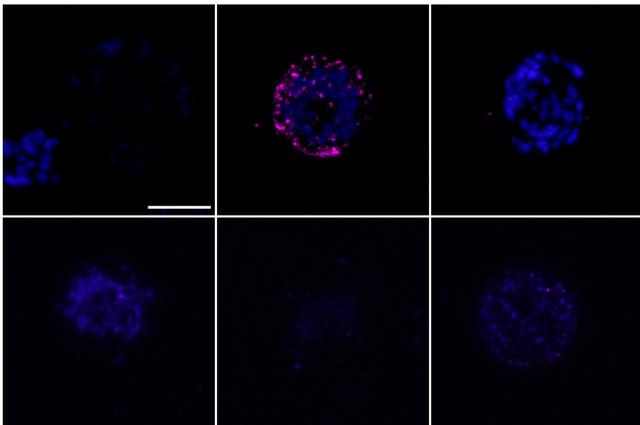
tRF Arg



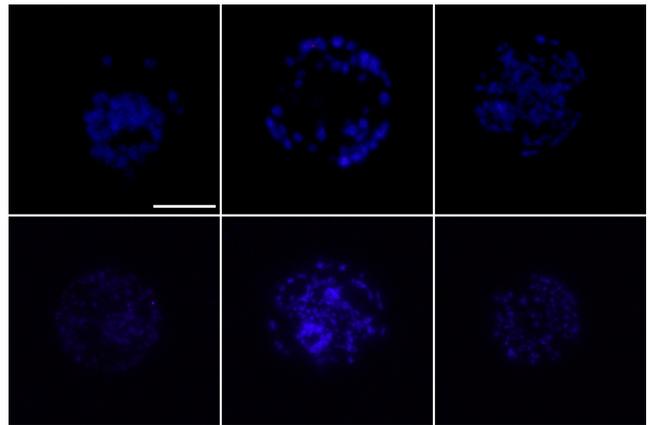
tRF His



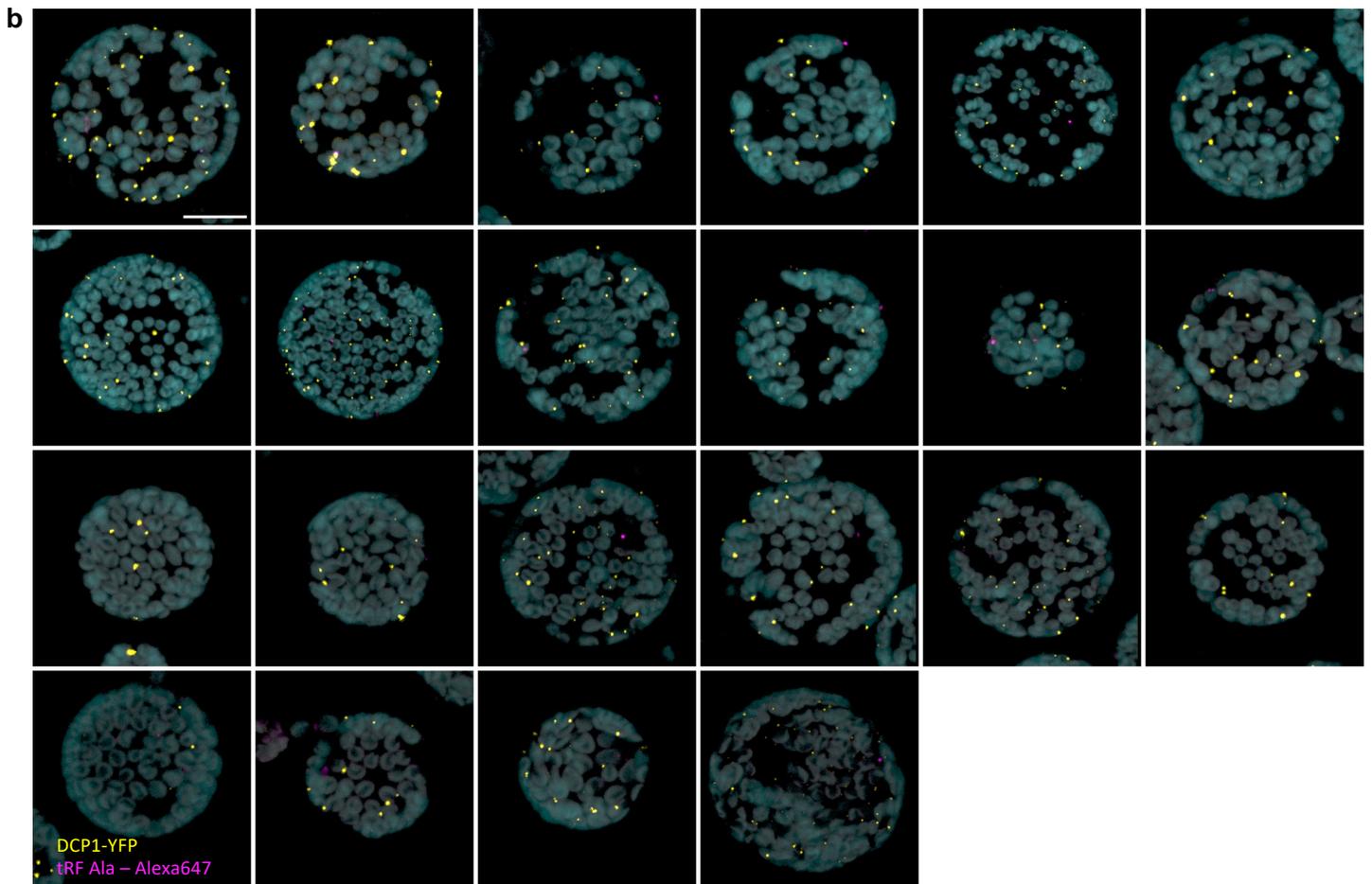
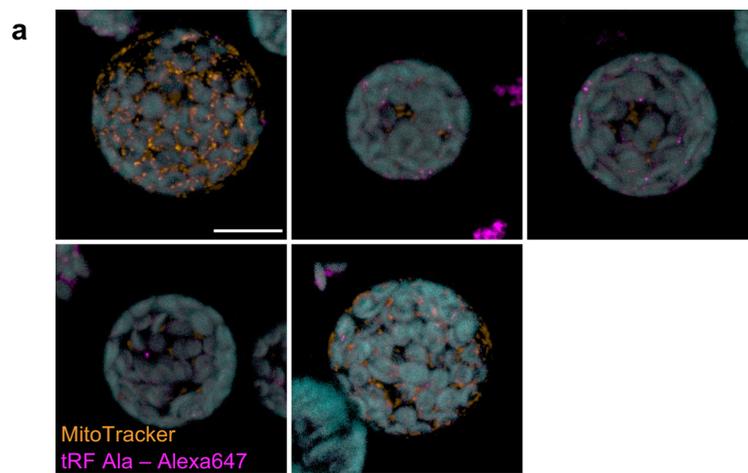
tRF Asp



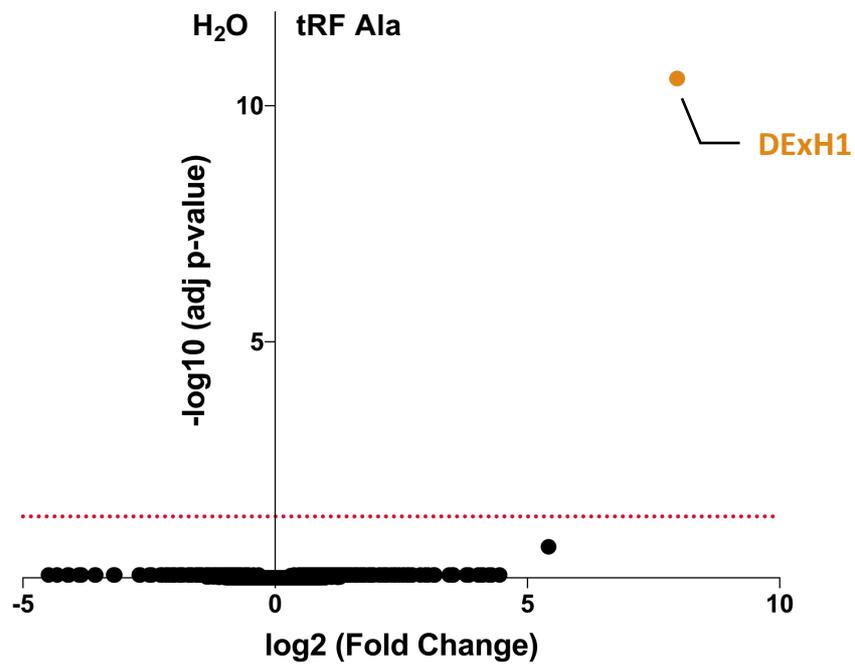
tRF Gly



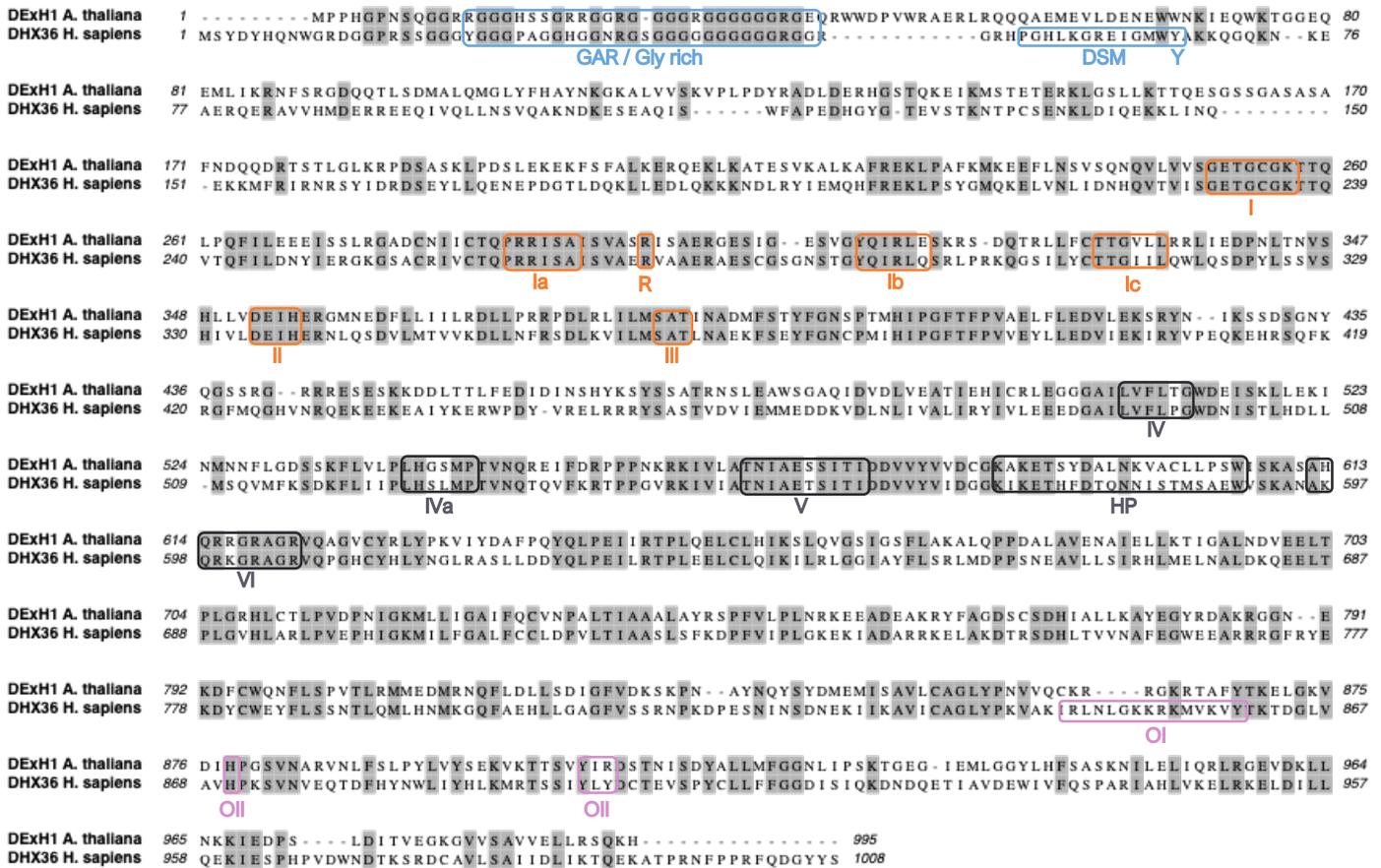
Supplementary Figure 2. Subcellular localization of different tRFs transfected into *A. thaliana* protoplasts. Confocal images (z-projection) of Arabidopsis protoplasts transfected with RNA R, tRF-5D Ala (AGC), tRF-5D Arg (CCT), tRF-5D His (GTG), tRF-5A Asp (GTC), tRF-5D Gly (TCC). Scale bar: 20 μ m. Merged blue (chlorophyll autofluorescence) and magenta (Alexa Fluor 647 fluorescence ; tRFs) channels.



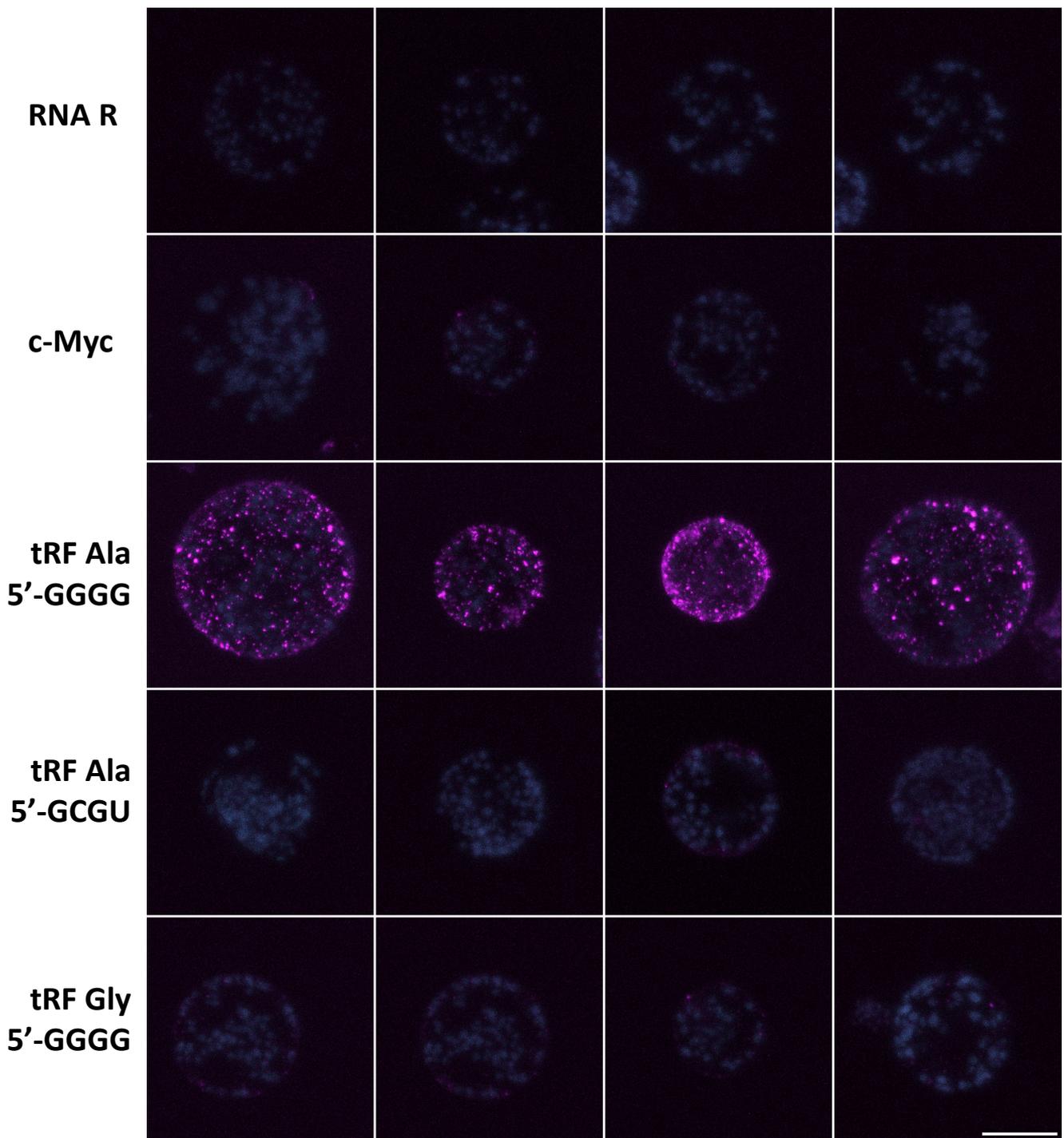
Supplementary Figure 3. Subcellular co-localization analysis between tRF Ala foci and mitochondria, P-bodies or stress granules. Confocal images (z-projection). (a) Arabidopsis protoplasts prepared from seedlings Col-0 transfected with fluorescently labelled tRF Ala and stained with Mitotracker. (b) Arabidopsis protoplasts prepared from seedlings of DCP1-YFP transgenic lines transfected with fluorescently labelled tRF Ala. Scale bar: 20 μ m. Cyan: chlorophyll, orange: mitotracker, magenta: Alexa Fluor 647 and yellow: DCP1-YFP and tRFP-PAB2.



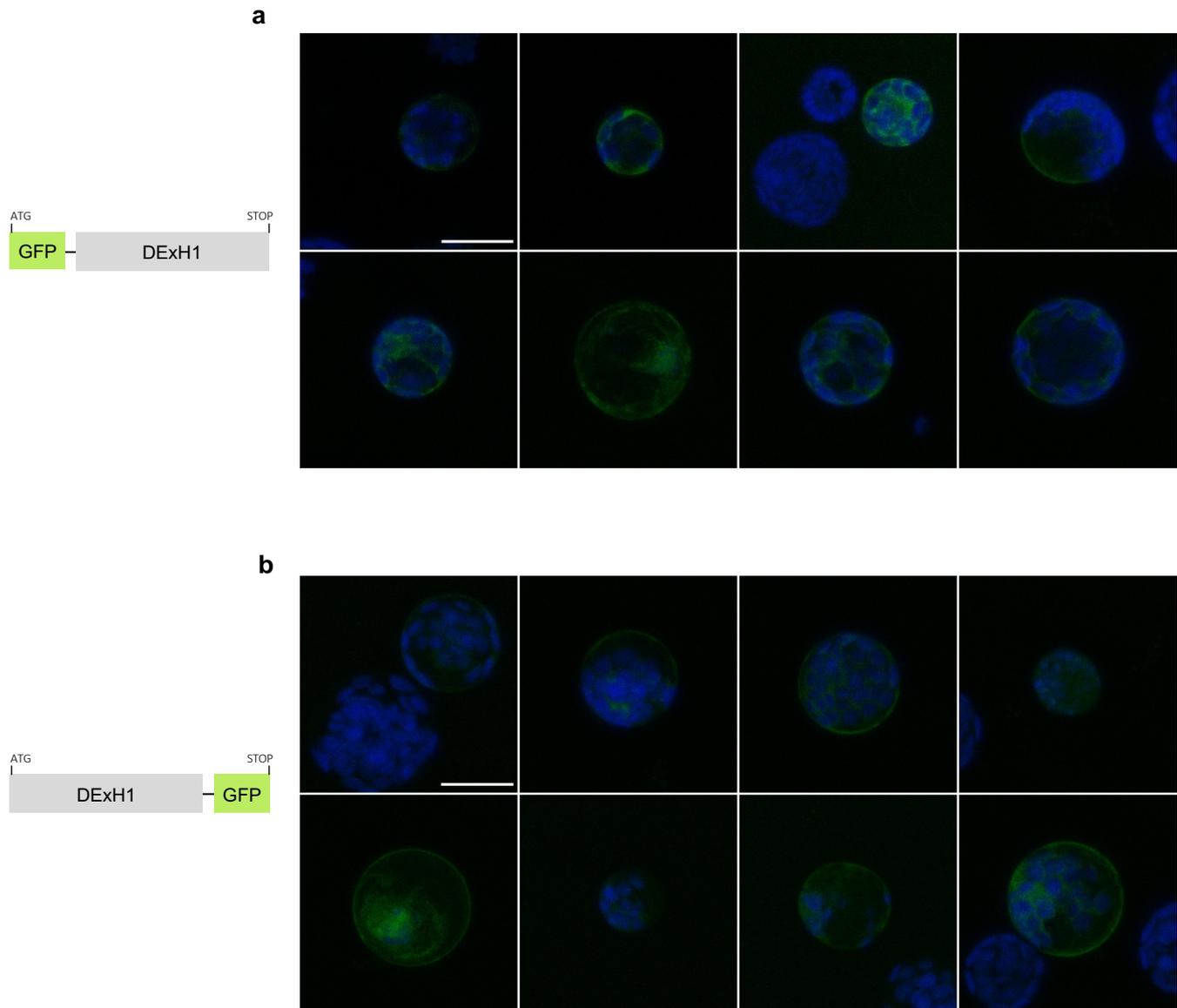
Supplementary Figure 4. Volcano plot of the experiments described in Figure 4a showing the DExH1 protein differentially enriched in tRF Ala transfected protoplasts versus water transfected protoplasts. Y- and X-axis display the adjusted p-value ($-\log_{10}(\text{adj p-value})$) and fold change ($\log_2(\text{Fold Change})$) respectively. The red dashed line indicates the significant threshold (adj p value $<0,05$).



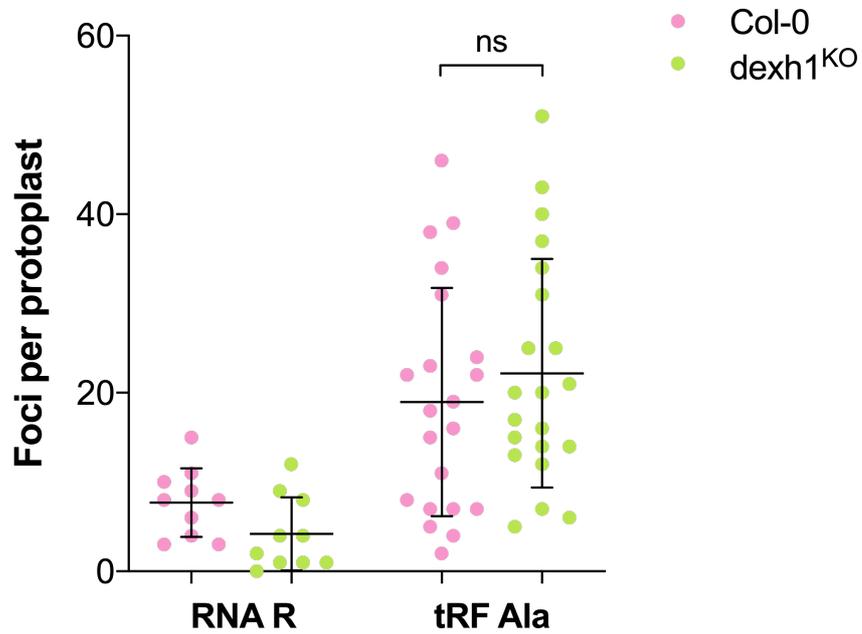
Supplementary Figure 5. Sequences alignment, using ClustalW, between *A. thaliana* DExH1 and *H. sapiens* DHX36. Letters on a grey background shows similarities. Domains presented in Figure 4d are highlighted with colored boxes.



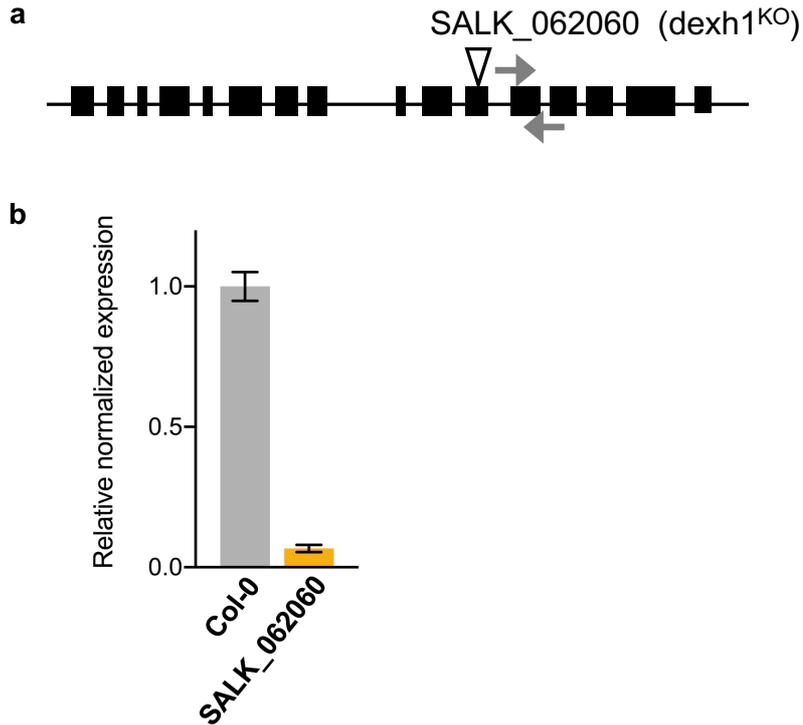
Supplementary Figure 7. Subcellular localization of different tRFs transfected into *A. thaliana* protoplasts. Confocal images (z-projection) of Arabidopsis protoplasts transfected with RNA R, c-Myc RNA, tRF Ala 5'-GGGG, tRF Ala 5'-GCGU, tRF Gly 5'-GGGG. Scale bar: 20 μ m. Merged cyan (chlorophyll autofluorescence) and magenta (Alexa Fluor 647 fluorescence ; tRFs) channels.



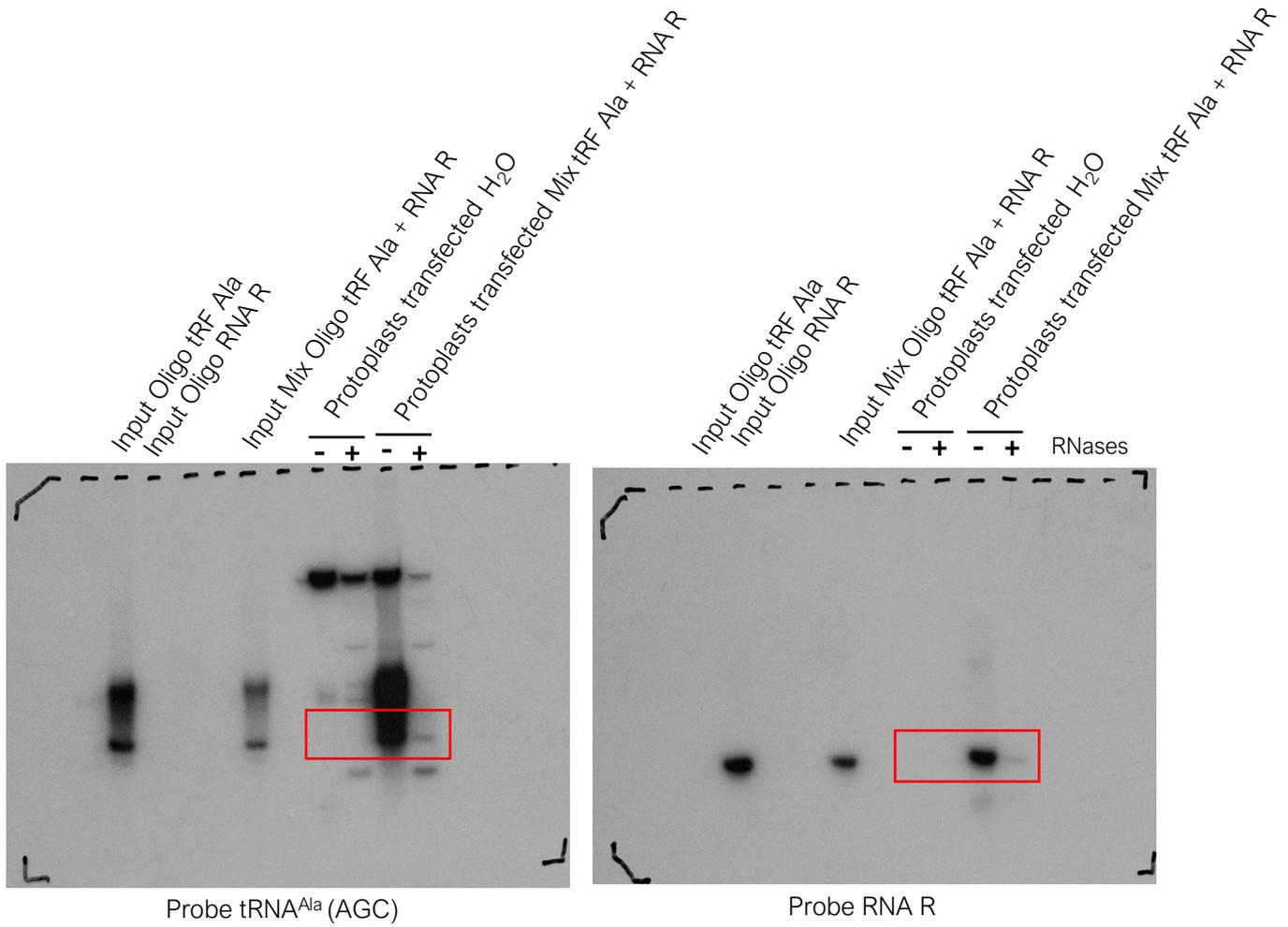
Supplementary Figure 8. Confocal images (z-projection) of protoplasts from *A. thaliana* Col-0 seedlings transfected with DEXH1 fused with GFP (a) in N-terminal or (b) in C-terminal. Scale bar: 20 μ m. Merged blue (chlorophyll autofluorescence) and green (GFP fluorescence) channels.



Supplementary Figure 9. Scatter plot of foci number per protoplast transfected with either RNA R or tRF Ala. Bars represent mean and standard deviation. Data come from one biological replicate, 10 images for RNA R and 21 images for tRF Ala were analyzed. ANOVA followed by Tukey's multiple comparison test was used to determine significance (ns: not significant).

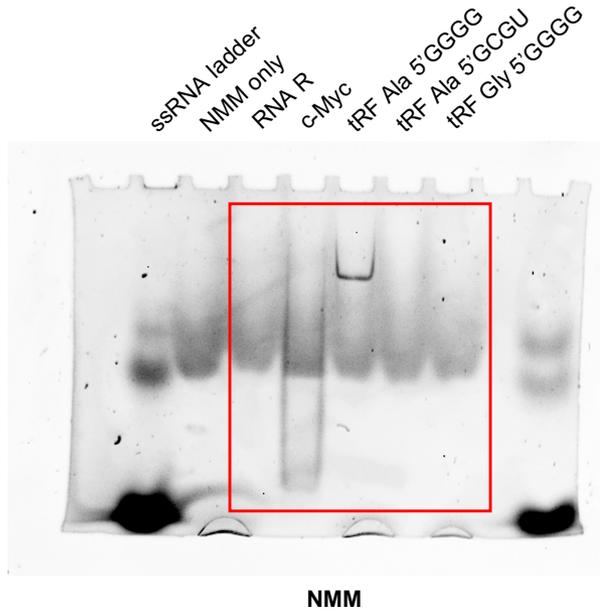


Supplementary Figure 10. Genotyping of the *dexh1* T-DNA insertion line (SALK_062060) and analysis of the expression level of DEXH1 mRNA. (a) The *A. thaliana* gene (AT2G35920) encoding the DEXH1 protein is represented. Position of the T-DNA in the insertion line *dexh1*^{KO} is indicated by an arrowhead. Analysis of the expression level of DEXH1 mRNA by RT-qPCR. Probes used are indicated by grey arrows on the gene represented in (a). Normalization with ACTIN and GAPDH. Relative expression calculated with the $\Delta/\Delta C_t$ method.

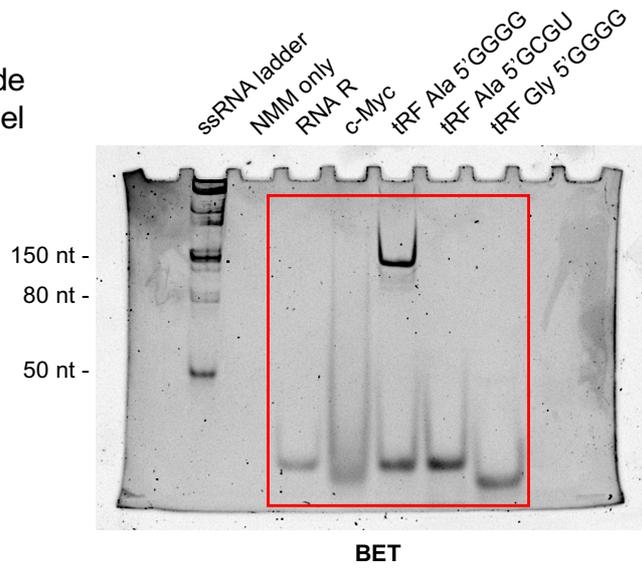


Supplementary Figure 11. Uncropped and unmodified data related to Figure 1.

Native 15% acrylamide
TBE gel



Native 15% acrylamide
TBE gel



Supplementary Figure 12. Uncropped and unmodified data related to Figure 5e.

Table S1: sequences of the tRF analyzed in this study

tRF names	tDRnamer	Sequences
tRF Ala	tDR-1:19-Ala-AGC-1-M2	GGGGAUGUAGCUCAGAUGGU
tRF Arg	tDR-1:20-Arg-CCT-2	GCGUCUGUAGCUCAGUGGA
*ptRF His		GGCGGAUGUAGCCAAGUGGA
tRF Asp	tDR-1:33-Asp-GTC-1-M3	GUCGUUGUAGUAUAGUGGUAAGUAUCCCCGCCU
tRF Gly	tDR-1:20-Gly-TCC-1	GCGUCUGUAGUCCAACGGU

* ptRF His derives from the plastial tRNA His of Arabidopsis. tDRnamer does not include organelar tRFs.

Table S2: Oligoribonucleotides and oligodesoxiribonucleotides used in this work

Name	Sequence (5' to 3')	Experiment
RNA R -fluo	CCUACAUCGAGUCUACCAU-2'Omc-iSp9-Alexa647	Localization by protoplast transfection
c-Myc -fluo	GGGGAGGGUGGGGAGGGUGGG-2'Omg-iSp9-Alexa647	Localization by protoplast transfection
tRF-5D Ala / tRF-5D Ala 5'GGGG -fluo	GGGAUGUAGCUCAGAUGG-2'Omu-iSp9-Alexa647	Localization by protoplast transfection
tRF-5D Ala 5'GCGU -fluo	GCGUAUGUAGCUCAGAUGG-2'Omu-iSp9-Alexa647	Localization by protoplast transfection
tRF-5D Arg -fluo	GCGUCUGUAGCUCAGAUGG-2'Oma-iSp9-Alexa647	Localization by protoplast transfection
tRF-5D His -fluo	GGCGGAUGUAGCCAAGUGG-2'Oma-iSp9-Alexa647	Localization by protoplast transfection
tRF-5D Asp -fluo	GUCGUUGUAGUAUAGUGGUAAGUAUJCCCGCC-2'Omu-iSp9-Alex647	Localization by protoplast transfection
tRF-5D Gly -fluo	GCGUCUGUAGUCCAACGG-2'Omu-iSp9-Alexa647	Localization by protoplast transfection
tRF-5D Gly 5'GGGG -fluo	GGGGCUGUAGUCCAACGG-2'Omu-iSp9-Alexa647	Localization by protoplast transfection
RNA R -bio	CCUACAUCGAGUCUACCAU-2'Omc-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
c-Myc -bio	GGGGAGGGUGGGGAGGGUGGG-2'Omg-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
tRF-5D Ala / tRF-5D Ala 5'GGGG -bio	GGGAUGUAGCUCAGAUGG-2'Omu-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
tRF-5D Ala 5'GCGU -bio	GCGUAUGUAGCUCAGAUGG-2'Omu-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
tRF-5D Arg -bio	GCGUCUGUAGCUCAGAUGG-2'Oma-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
tRF-5D His -bio	GGCGGAUGUAGCCAAGUGG-2'Oma-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
tRF-5D Asp -bio	GUCGUUGUAGUAUAGUGGUAAGUAUJCCCGCC-2'Omu-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
tRF-5D Gly -bio	GCGUCUGUAGUCCAACGG-2'Omu-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
tRF-5D Gly 5'GGGG -bio	GGGGCUGUAGUCCAACGG-2'Omu-iSp9-Biotin	Streptavidin pull down after protoplast transfection + NMM quantification and gel
tRNA Ala probe	ACCATCTGAGCTACATCCCC	Northern blot
RNA R probe	GATGGTAGACTCGATGTAGG	Northern blot
DEXH1-cDNA fw	ATGCCTCCTCACGGTCCTAATTC	PCR cDNA
DEXH1-cDNA rv	CTAATGTTTTGGCTACGTAACAACCTCC	PCR cDNA
DEXH1-GFP-Nter+Cter-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGCCTCCTCACGGTCCTAATTC	Gateway cloning
DEXH1-GFP-Nter-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCCCTAATGTTTTGGCTACGTAACAACCTCC	Gateway cloning
DEXH1-GFP-Cter-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTATGTTTTGGCTACGTAACAACCTCC	Gateway cloning
SALK_062060-RP	TTCTCTTGAAGCATGGTCTGG	PCR Genotyping

Table S3: Raw data of mass spectrometry results shown in Figure 4. Rows are sorted according to the p-value of RNA R vs tRF Ala column.

Identification by Mascot and validation at FDR<1% by Proline (ProFI consortium)	H2O			RNA R			tRF Ala		
	#1	#2	#3	#1	#2	#3	#1	#2	#3
Total number of proteins identified	3210	3401	3447	2943	3885	3518	3245	3622	3377
Total number of spectra identified	29019	30696	30691	25258	37097	30921	28927	34402	29725

spectra = 1 spectrum
 # spectra = 2-5 spectra
 # spectra = 6-10 spectra
 # spectra = 11-30 spectra
 # spectra >= 31 spectra

Accession	Gene name	Description	Basic spectral count / number of spectra by replicate									Statistics with IP-inquiry from H.Zuber (Norm. DEseq)					
			H2O			RNA R			tRF Ala			tRF Ala vs H2O			RNA R vs tRF Ala		
			#1	#2	#3	#1	#2	#3	#1	#2	#3	Log2FC	p-value	adjp	Log2FC	p-value	adjp
AT2G35920.1		RNA helicase family protein							31	25	39	7,97	4,98E-15	2,63E-11	-8,03	1,77E-18	9,33E-15
AT4G36020.1		cold shock domain protein 1, Symbo...	3	14	6	3	9	5	28	17	27	1,60	7,25E-03	8,70E-01	-2,16	9,76E-06	2,58E-02
AT4G32720.1		La protein 1, Symbols: AtLa1, La1			1				6	5	4	3,45	7,71E-04	8,70E-01	-5,40	1,56E-05	2,75E-02
AT3G03710.1	GBSS1	polynucleotide nucleotidyltran...							3	3	10	5,42	8,34E-05	2,20E-01	-5,48	2,47E-05	3,26E-02
AT5G50250.1	MFP2	chloroplast RNA-binding protein 31...			2		1		4	3	6	2,45	1,02E-02	8,70E-01	-3,35	9,09E-04	9,59E-01
AT1G73110.1		P-loop containing nucleoside triph...	3		1				3	2	2	0,70	4,52E-01	1,00E+00	-4,34	2,52E-03	1,00E+00
AT1G32330.1	emb1427	heat shock transcription factor A1...		1	2				3	1	3	1,09	2,67E-01	8,96E-01	-4,34	2,53E-03	1,00E+00
AT1G55150.1		DEA(D/H)-box RNA helicase family p...	1	2					3	2	1	0,87	3,86E-01	1,00E+00	-4,13	4,67E-03	1,00E+00
AT5G23420.1	ABC16	high-mobility group box 6, Symbols...		3			4	2				-3,19	4,52E-02	8,70E-01	4,03	7,62E-03	1,00E+00
AT2G17870.1	CTIMC	cold shock domain protein 3, Symbo...	7	23	12	3	14	8	16	17	21	0,31	5,46E-01	1,00E+00	-1,20	1,01E-02	1,00E+00
AT1G48830.1		Ribosomal protein S7e family protein	1	1					6			1,42	2,51E-01	8,70E-01	-4,17	1,02E-02	1,00E+00
AT4G16660.1	HEMA2	heat shock protein 70 (Hsp 70) fam...		1		1				5	4	2,71	2,15E-02	8,70E-01	-2,81	1,24E-02	1,00E+00
AT5G18900.1	RAP74	2-oxoglutarate (2OG) and Fe(II)-de...		2		3	1	1				-2,68	9,87E-02	8,70E-01	3,82	1,25E-02	1,00E+00
AT1G09330.1	RPS2	unknown protein	3			2	2	1				-3,20	4,48E-02	8,70E-01	3,81	1,27E-02	1,00E+00
AT1G65270.2	GRV2	unknown protein				2	2	1				0,00	1,00E+00	1,00E+00	3,81	1,28E-02	1,00E+00
AT3G22430.1		unknown protein	2	1		1	3	1				-3,20	4,49E-02	8,70E-01	3,79	1,38E-02	1,00E+00
AT3G03940.1	EIF(ISO)4G2	Protein kinase family protein			3		3	2				-3,18	4,57E-02	8,70E-01	3,79	1,41E-02	1,00E+00
AT3G16910.1	ARC5	acyl-activating enzyme 7, Symbols:...	7	8	2	6	3	2	7	15	8	0,74	2,37E-01	8,70E-01	-1,47	1,43E-02	1,00E+00
AT1G22620.1	NDPK3	Phosphoinositide phosphatase famil...	8	3	2	2		1	10	2	3	0,17	8,38E-01	1,00E+00	-2,26	1,44E-02	1,00E+00
AT1G67325.2	G6PD5	Ran BP2/NZF zinc finger-like super...	7	4	1		1	1	7	2	2	-0,17	8,40E-01	1,00E+00	-2,35	1,57E-02	1,00E+00
AT1G06670.1	SEC15B	nuclear DEIH-boxhelicase, Symbols:...	2	2	12	1	2	4	6	6	8	0,29	6,94E-01	1,00E+00	-1,54	1,80E-02	1,00E+00
AT2G37220.1		RNA-binding (RRM/RBD/RNP motifs) f...	1	5	2	1	4	2	8	5	7	1,24	7,59E-02	8,70E-01	-1,55	1,80E-02	1,00E+00
AT3G27700.1	AGO4	zinc finger (CCCH-type) family pro...	4	1	1				4			-0,57	5,85E-01	1,00E+00	-3,59	1,91E-02	1,00E+00
AT1G14900.1	GA2	high mobility group A, Symbols: HMGA	2	4					4			-0,57	6,02E-01	1,00E+00	-3,59	1,91E-02	1,00E+00
AT5G49520.1		WRKY DNA-binding protein 48, Symbo...	1		1				3	1		0,85	4,68E-01	1,00E+00	-3,59	1,93E-02	1,00E+00
AT1G15810.1	ARC6	S15/NS1, RNA-binding protein							2	1	1	3,53	2,53E-02	8,70E-01	-3,58	1,94E-02	1,00E+00
AT3G04340.1	PTAC2	FtsH extracellular protease family...							2	1	1	3,53	2,53E-02	8,70E-01	-3,58	1,94E-02	1,00E+00
AT5G47080.1		casein kinase II beta chain 1, Sy...		4	1				1	1	2	-0,34	7,35E-01	1,00E+00	-3,58	1,95E-02	1,00E+00
AT3G33520.1		actin-related protein 6, Symbols: ...	1		1					1	3	0,83	4,75E-01	1,00E+00	-3,58	1,96E-02	1,00E+00
AT1G17440.1		Transcription initiation factor TF...		1						2	2	1,62	2,07E-01	8,70E-01	-3,57	1,96E-02	1,00E+00
AT2G38670.1	CYP705A19	phosphorylethanolamine cytidyltr...	12	8	3	2	5	4	9	8	9	0,11	8,48E-01	1,00E+00	-1,30	1,96E-02	1,00E+00
AT4G31750.1	EIF3G1	HOPW1-1-interacting 2, Symbols: WIN2								3	1	3,51	2,59E-02	8,70E-01	-3,57	1,97E-02	1,00E+00
AT4G33630.1	DC11	Protein of unknown function (DUF35...								3	1	3,51	2,59E-02	8,70E-01	-3,57	1,97E-02	1,00E+00
AT2G38470.1	RECQL2	WRKY DNA-binding protein 33, Symbo...	2	7	1		1		6	1	1	-0,33	7,15E-01	1,00E+00	-2,70	2,02E-02	1,00E+00
AT5G56760.1	GRF9	serine acetyltransferase 1;1, Symb...				1		3				0,00	1,00E+00	1,00E+00	3,52	2,42E-02	1,00E+00
AT3G07680.1	PFK1	emp24/gp25L/p24 family/GOLD family...				2	1	1				0,00	1,00E+00	1,00E+00	3,52	2,42E-02	1,00E+00

Table S4: Raw data of mass spectrometry results shown in Figure 4. Rows are sorted according to the p-value of RNA R vs tRF Ala column.

Identification by Mascot and validation at FDR<1% by Proline (ProFI consortium)			RNA R			tRF Ala			tRF Arg			tRF His			tRF Asp			tRF Gly			# spectra = 1 spectrum		
#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	# spectra = 2-5 spectra		
Total number of proteins identified			2563	2763	3343	2078	2306	3049	2125	2543	2957	2071	2488	3194	2449	2348	3108	1844	2673	2977	# spectra = 6-10 spectra		
Total number of spectra identified			23218	25699	31543	18720	21520	27208	20786	23868	28443	18222	21991	31307	22055	22018	30611	16136	27271	28416	# spectra = 11-30 spectra		
																					# spectra >= 31 spectra		
Accession	Gene name	Description	Basic spectral count / number of spectra by replicate															Statistics with IP-inquiry from H.Zuber (Norm. DEseq)					
			RNA R			tRF Ala			tRF Arg			tRF His			tRF Asp			tRF Gly			RNA R vs tRF Ala		
			#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	Log2FC	p-value	adjp
AT2G35920.1		RNA helicase family protein	0	0	1	11	37	15	0	0	1	0	0	2	0	0	0	0	0	0	5,77	4,66E-09	2,04E-05
AT1G06890.1	At1g06890	nodulin MtN21 /EamA-like transport...	0	0	0	2	3	0	4	2	0	2	2	0	1	3	0	3	0	0	3,99	8,25E-03	1,00E+00
AT1G15270.1	At1g15270	At1g Translation machinery associated TMA7	0	0	0	3	1	0	2	2	0	1	2	0	1	1	0	1	0	0	3,69	1,55E-02	1,00E+00
AT4G33520.2	PAA1	P-type ATP-ase 1, Symbols: PAA1, HMA6	0	0	0	2	1	0	1	1	0	2	3	0	0	2	0	2	1	0	3,30	3,48E-02	1,00E+00
AT1G32050.1	SCAMP4	SCAMP family protein	0	0	0	4	2	0	2	0	0	2	1	0	1	0	0	3	0	0	4,26	4,55E-03	1,00E+00
AT1G26880.1	RPL34A	Ribosomal protein L34e superfamily...	0	1	1	4	3	1	3	1	3	1	2	2	3	3	4	1	2	4	2,07	4,29E-02	1,00E+00
AT1G54570.1	At1g54570	Esterase/lipase/thioesterase famil...	0	0	0	2	2	0	1	1	0	0	1	0	1	2	0	0	0	0	3,68	1,57E-02	1,00E+00
AT3G56460.1	TSP19_110	At3g GroES-like zinc-binding alcohol de...	0	0	0	1	2	0	2	0	0	1	0	0	0	0	0	2	1	0	3,30	3,51E-02	1,00E+00
AT4G18465.1		RNA helicase family protein	0	0	0	0	2	1	0	0	2	0	0	1	0	0	2	0	1	1	3,28	3,60E-02	1,00E+00
AT2G24200.1	LAP1	Cytosol aminopeptidase family protein	0	0	0	1	1	1	0	1	1	1	1	0	0	0	1	0	0	1	3,28	3,57E-02	1,00E+00
AT5G35910.1	RRP6L2	Polynucleotidyl transferase, ribon...	0	0	0	1	0	2	0	0	0	1	0	1	1	0	1	0	0	0	3,27	3,62E-02	1,00E+00
AT3G06730.1	CITRX	Thioredoxin z, Symbols: TRX P, TRX z	0	0	0	2	0	1	0	0	0	0	0	0	0	0	1	0	0	0	3,29	3,54E-02	1,00E+00
AT1G61670.1	At1g61670	At1g Lung seven transmembrane receptor ...	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	3,30	3,51E-02	1,00E+00
AT4G32720.1	LA1	La protein 1, Symbols: AtLa1, La1	0	0	4	0	7	16	0	0	5	0	0	6	0	0	4	0	0	4	2,66	1,64E-02	1,00E+00
AT3G29970.1	At3g29970	B12D protein	0	0	0	1	1	0	2	1	0	2	2	0	1	2	0	2	1	0	2,78	8,26E-02	1,00E+00
AT3G51010.1	F24M12.50	At3g unknown protein	0	0	0	0	1	0	0	3	1	0	4	0	2	2	1	0	2	0	1,97	2,16E-01	1,00E+00
AT3G08620.1	At3g08620	RNA-binding KH domain-containing p...	0	0	0	0	2	0	4	0	0	2	1	0	1	1	0	4	1	0	2,77	8,31E-02	1,00E+00
AT4G35570.1	HMG5	high mobility group B5, Symbols: H...	0	0	0	1	1	0	1	1	1	1	1	1	1	0	0	1	1	1	2,78	8,26E-02	1,00E+00
AT5G08360.1	At5g08360	F8L1 Protein of unknown function (DUF789)	0	0	0	0	1	0	1	2	0	1	2	0	1	1	0	1	0	0	1,97	2,16E-01	1,00E+00
AT5G23080.1	TGH	SWAP (Suppressor-of-White-Apricot)...	0	0	0	0	0	2	0	0	3	0	0	2	1	1	1	0	0	1	2,75	8,52E-02	1,00E+00
AT3G44850.1	F28D10_40	At3g Protein kinase superfamily protein	0	0	0	0	2	0	0	1	3	0	0	2	0	0	2	0	0	3	2,75	8,52E-02	1,00E+00
AT1G30620.1	MUR4	NAD(P)-binding Rossmann-fold super...	0	0	0	0	1	0	1	1	1	1	1	0	1	0	0	0	0	0	1,97	2,16E-01	1,00E+00
AT4G38240.1	GNT1	alpha-1,3-mannosyl-glycoprotein be...	0	0	0	0	1	1	0	0	1	1	2	0	1	0	1	0	2	0	2,76	8,42E-02	1,00E+00
AT5G02430.1	At5g02430	T22i Transducin/WD40 repeat-like superf...	0	0	0	1	1	0	0	0	1	1	1	1	0	1	1	0	0	1	2,78	8,26E-02	1,00E+00
AT1G20270.1	P4H3	2-oxoglutarate (2OG) and Fe(II)-de...	0	0	0	1	1	0	0	1	1	0	1	0	1	1	1	1	1	1	2,78	8,26E-02	1,00E+00
AT5G13260.1	At5g13260	T31i unknown protein	0	1	0	1	0	2	2	1	2	1	2	2	3	2	2	1	1	1	1,47	2,65E-01	1,00E+00
AT1G18850.1	At1g18850	F6A: unknown protein	0	0	0	0	0	2	0	0	1	0	0	1	0	0	4	0	0	2	2,75	8,52E-02	1,00E+00
AT1G11580.1	PME18	methyltransferase PCR A, Symbols: ATP...	0	0	0	1	1	0	1	0	0	1	0	0	3	0	0	1	0	0	2,78	8,26E-02	1,00E+00
AT1G32360.1	At1g32360	Zinc finger (CCCH-type) family pro...	0	0	1	0	0	1	1	4	2	2	3	1	3	2	0	2	1	0	0,18	9,09E-01	1,00E+00
AT3G06290.1		SAC3/GANP/Nin1/mts3/eIF-3 p25 family	0	0	0	0	0	0	1	1	2	1	0	1	1	0	1	0	0	1	6,41E-16	1,00E+00	1,00E+00
AT5G06140.1	SNX1	sorting nexin 1, Symbols: SNX1, AT...	0	0	0	0	0	0	1	0	0	0	0	3	0	0	3	0	0	0	6,41E-16	1,00E+00	1,00E+00
AT5G10160.1	At5g10160	T31i Thioesterase superfamily protein	0	0	0	0	0	0	0	1	0	0	1	2	0	0	1	0	1	1	6,41E-16	1,00E+00	1,00E+00
AT5G27560.1	At5g27560	F15i Domain of unknown function (DUF1995)	0	0	0	0	0	0	1	0	0	2	1	0	2	0	0	0	1	0	6,41E-16	1,00E+00	1,00E+00
AT5G64440.1	FAAH	fatty acid amide hydrolase, Symbol...	0	0	0	0	0	0	0	0	1	0	0	3	0	0	1	0	0	1	6,41E-16	1,00E+00	1,00E+00
AT4G10640.1	IQD16	IQ-domain IQ-domain 16, Symbols: IQD16	0	0	0	1	1	0	1	0	0	0	2	1	0	0	0	2	0	0	2,78	8,26E-02	1,00E+00
AT1G22870.1	F19G10.17	Protein kinase family protein with...	0	0	0	0	2	0	0	0	0	0	1	0	1	2	0	1	0	0	2,77	8,31E-02	1,00E+00
AT2G35350.1	PLL1	poltergeist like 1, Symbols: PLL1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	0	6,41E-16	1,00E+00	1,00E+00
AT3G14990.1	DJ1A	Class I glutamine amidotransferase...	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0	0	4	2	2,76	8,42E-02	1,00E+00
AT5G24165.1	At5g24165	unknown protein	0	0	0	0	1	0	1	1	0	0	1	0	0	1	0	1	3	0	1,97	2,16E-01	1,00E+00
AT2G19080.1	MTX1	metaxin-related	0	0	0	1	0	0	1	0	1	0	1	1	0	1	0	1	1	1	1,98	2,16E-01	1,00E+00
AT5G33320.1	PPT1	Glucose-6-phosphate/phosphate tran...	0	0	0	1	1	0	1	0	0	1	0	0	0	0	0	1	1	1	2,78	8,26E-02	1,00E+00
AT5G25220.1	KNAT3	KNOTTED1-like homeobox gene 3, Sym...	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	2	1	1,95	2,18E-01	1,00E+00

AT5G61220.1	MAF19.16 MAF1 LYR family of Fe/S cluster biogene...	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	2	2	1,95	2,18E-01	1,00E+00	
AT2G47330.1	RH24 P-loop containing nucleoside triph...	0	0	0	0	0	0	0	0	1	0	0	2	0	0	2	0	1	2	6,41E-16	1,00E+00	1,00E+00
AT3G51820.1	CHLG UbiA prenyltransferase family prot...	0	0	0	0	0	0	0	2	0	1	1	0	0	1	0	0	3	0	6,41E-16	1,00E+00	1,00E+00
AT3G01780.1	AT3G01780.1 ARM repeat superfamily protein, Sy...	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2	0	1	2	6,41E-16	1,00E+00	1,00E+00
ATMG00090.1	RPS3 structural constituent of ribosome...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	6,41E-16	1,00E+00	1,00E+00
AT4G10110.1	At4g10110 At4g RNA-binding (RRM/RBD/RNP motifs) f...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	1	6,41E-16	1,00E+00	1,00E+00
AT3G58640.1	At3g58640 Mitogen activated protein kinase k...	4	0	0	2	1	2	0	3	3	1	3	1	5	3	1	0	0	0	0,45	6,62E-01	1,00E+00
AT3G58730.1	VHA-D vacuolar ATP synthase subunit D (V...	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-3,44	3,22E-02	1,00E+00
AT5G10950.1	At5g10950 T30I Tudor/PWWP/MBT superfamily protein	0	0	5	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	-3,64	2,52E-02	1,00E+00
AT3G02875.1	ILR1 Peptidase M20/M25/M40 family prote...	0	0	4	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	-3,39	3,45E-02	1,00E+00
AT5G13950.3	unknown protein	0	0	4	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	-3,39	3,45E-02	1,00E+00
AT5G56910.1	MHM17.2 MHM Proteinase inhibitor I25, cystatin...	1	1	2	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	-3,43	3,25E-02	1,00E+00
AT3G43190.1	SUS4 sucrose synthase 4, Symbols: SUS4,...	1	2	6	0	0	1	0	1	0	0	0	4	0	0	2	0	1	2	-2,54	2,66E-02	1,00E+00
AT5G04430.2	binding to TOMV RNA 1L (long form)...	1	8	5	0	0	1	0	1	1	1	1	2	1	2	4	0	4	1	-3,20	3,91E-03	1,00E+00
AT3G57410.1	VLN3 villin 3, Symbols: VLN3, ATVLN3	1	0	3	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	-3,41	3,34E-02	1,00E+00
AT5G12150.1	ROPGAP6 Rho GTPase activation protein (Rho...	2	2	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	-3,76	1,66E-02	1,00E+00
AT2G37770.2	AKR4C9 NAD(P)-linked oxidoreductase super...	1	0	3	0	0	0	0	0	1	0	0	0	1	1	1	0	1	1	-3,41	3,34E-02	1,00E+00
AT3G53880.1	AKR4C11 NAD(P)-linked oxidoreductase super...	3	0	5	0	0	0	0	1	1	0	0	1	1	1	5	0	1	3	-4,35	4,20E-03	1,00E+00
AT1G47260.1	GAMMACA2 gamma carbonic anhydrase 2, Symbol...	2	6	0	0	1	0	0	2	1	0	1	0	0	1	1	1	2	0	-2,48	4,80E-02	1,00E+00
AT2G31670.1	UP3 Stress responsive alpha-beta barre...	6	3	7	0	1	3	0	2	3	0	1	5	1	0	2	0	2	2	-1,70	3,59E-02	1,00E+00
AT4G10120.1	Sucrose-phosphate synthase family ...	2	1	3	0	0	0	0	0	2	0	0	3	0	0	0	0	1	1	-3,97	9,47E-03	1,00E+00
AT5G58490.1	MQJ2.6 MQJ2_6 NAD(P)-binding Rossmann-fold super...	3	3	2	0	1	0	1	1	1	0	1	1	0	2	1	0	4	2	-2,43	3,46E-02	1,00E+00
AT2G30860.1	GSTF9 glutathione S-transferase PHI 9, S...	5	3	4	1	0	1	1	1	2	0	2	3	3	4	2	0	5	2	-2,19	2,20E-02	1,00E+00
AT1G69210.1	At1g69210 F4N: Uncharacterised protein family UPF...	0	0	4	0	0	0	0	0	1	0	0	2	0	0	1	0	0	0	-3,39	3,45E-02	1,00E+00
AT3G62310.1	At3g62310 RNA helicase family protein	0	0	6	0	0	0	0	0	1	0	1	4	0	0	2	0	0	1	-3,86	1,70E-02	1,00E+00
AT4G26300.1	EMB1027 Arginyl-tRNA synthetase, class Ic,...	3	0	7	1	0	0	0	0	4	0	0	7	1	0	3	0	0	3	-2,65	2,48E-02	1,00E+00
ATCG01240.1-ATCG rps7-A	ribosomal protein S7, ...	1	3	0	0	0	0	0	1	0	0	1	1	0	0	1	0	3	0	-3,47	3,07E-02	1,00E+00
AT2G45540.2	WD-40 repeat family protein / beig...	3	3	2	0	0	1	0	2	1	0	0	4	1	0	2	0	3	3	-2,44	3,36E-02	1,00E+00
AT2G36390.1	SBE2.1 starch branching enzyme 2.1, Symbo...	13	3	4	0	3	2	1	3	3	2	4	5	8	2	5	1	2	2	-1,81	4,56E-02	1,00E+00
AT1G74560.3	NAP1-related protein 1, Symbols: NRP1	0	1	3	0	0	0	0	0	1	0	0	2	0	0	2	0	1	0	-3,41	3,36E-02	1,00E+00
AT1G74690.1	IQD31 IQ-domain 31, Symbols: IQD31	0	0	4	0	0	0	0	0	1	0	0	3	0	0	3	0	0	4	-3,39	3,45E-02	1,00E+00
AT2G20790.1	AP5M clathrin adaptor complexes medium ...	2	1	1	0	0	0	1	0	0	1	0	1	1	1	1	0	1	1	-3,46	3,13E-02	1,00E+00
AT3G12390.1	At3g12390 Nascent polypeptide-associated com...	3	4	6	0	2	1	1	1	3	1	3	4	1	1	6	1	3	3	-1,76	4,26E-02	1,00E+00
AT5G13850.1	At5g13850 nascent polypeptide-associated com...	3	4	6	0	2	1	1	1	3	1	3	4	1	1	6	1	3	3	-1,76	4,26E-02	1,00E+00
AT5G02050.1	At5g02050 T7H: Mitochondrial glycoprotein family ...	0	1	3	0	0	0	0	0	2	0	0	1	0	0	1	0	0	0	-3,41	3,36E-02	1,00E+00
AT1G62390.1	CLMP1 MADB2 f Octicosapeptide/Phox/Bem1p (PB1) d...	0	0	5	0	0	0	0	0	2	0	0	1	0	0	2	0	0	1	-3,64	2,48E-02	1,00E+00
AT1G20760.1	At1g20760 F2D: Calcium-binding EF hand family pro...	0	4	1	0	0	0	1	1	0	1	1	0	1	0	0	1	1	0	-3,75	1,89E-02	1,00E+00
AT5G46180.1	DELTA-OAT ornithine-delta-aminotransferase, ...	3	4	5	0	1	0	0	2	3	0	2	1	0	0	5	0	4	3	-2,97	5,35E-03	1,00E+00
AT3G45640.1	MPK3 mitogen-activated protein kinase 3...	2	0	3	0	0	0	2	0	1	0	0	2	2	0	2	0	0	1	-3,72	1,83E-02	1,00E+00
AT2G16090.1	ARI2 RING/U-box superfamily protein, Sy...	2	1	2	0	0	0	0	2	0	2	0	0	3	0	0	0	1	0	-3,74	1,68E-02	1,00E+00
AT4G19710.2	AKHSDH2 aspartate kinase-homoserine dehydr...	3	1	3	0	0	0	0	1	3	0	1	1	1	1	3	0	0	2	-4,19	5,33E-03	1,00E+00
AT3G44890.1	RPL9 ribosomal protein L9, Symbols: RPL9	0	6	1	0	0	0	0	3	1	0	1	1	3	2	2	0	4	1	-4,23	8,76E-03	1,00E+00
AT3G57290.1	TIF3E1 eukaryotic translation initiation ...	5	2	16	0	2	3	2	3	8	0	1	8	3	1	12	1	2	7	-1,84	3,02E-02	1,00E+00
AT4G04720.1	CPK21 calcium-dependent protein kinase 2...	1	1	2	0	0	0	0	1	1	0	0	2	0	0	1	0	0	0	-3,43	3,25E-02	1,00E+00
AT1G09430.1	AT1G09430.1 ATP-citrate lyase A-3, Symbols: AC...	1	2	2	0	0	0	0	2	1	0	1	3	0	0	1	0	3	1	-3,74	1,69E-02	1,00E+00
AT5G20080.1	CBR2 FAD/NAD(P)-binding oxidoreductase	0	3	4	0	0	0	0	2	3	0	0	5	0	0	3	0	2	4	-4,16	6,41E-03	1,00E+00
AT2G35790.1	At2g35790 T20I unknown protein	2	8	2	0	1	0	4	3	0	4	2	1	0	4	0	1	7	0	-3,01	8,90E-03	1,00E+00
AT4G31120.1	PMRT15 SHK1 binding protein 1, Symbols: S...	1	0	3	0	0	0	0	0	3	0	0	2	0	0	2	0	0	1	-3,41	3,34E-02	1,00E+00
AT3G15090.1	At3g15090 GroES-like zinc-binding alcohol de...	2	3	1	0	0	0	4	3	0	2	4	2	1	3	2	0	2	0	-4,01	9,12E-03	1,00E+00
AT5G42870.1	PAH2 phosphatidic acid phosphohydrolase...	4	3	3	0	1	1	0	3	3	0	1	4	1	1	3	1	3	0	-1,94	4,82E-02	1,00E+00
AT2G34040.1	At2g34040 T14f Apoptosis inhibitory protein 5 (API5)	0	3	1	0	0	0	1	2	2	1	1	6	0	0	5	0	3	2	-3,45	3,19E-02	1,00E+00
AT1G71270.1	VP552 Vps52 / Sac2 family, Symbols: POK...	0	2	2	0	0	0	1	1	0	1	1	0	0	1	1	2	1	0	-3,43	3,27E-02	1,00E+00
AT1G51580.1	F19C24.19 At1g RNA-binding KH domain-containing p...	2	1	1	0	0	0	0	1	1	0	0	4	0	0	2	0	0	2	-3,46	3,13E-02	1,00E+00
AT1G20630.1	CAT1 catalase 1, Symbols: CAT1	11	10	20	3	6	7	3	11	14	4	12	18	6	6	11	1	10	12	-1,10	3,27E-02	1,00E+00
AT1G80410.2	NAA15 EMB275: tetratricopeptide repeat (TPR)-con...	21	14	25	6	7	13	8	14	19	8	12	23	14	7	21	1	19	20	-0,98	3,16E-02	1,00E+00

AT3G18090.1	NRPD2b	nuclear RNA polymerase D2B, Symbol...	3	1	0	0	1	2	0	0	0	0	0	0	1	0	0	0	0	-0,23	8,36E-01	1,00E+00	
AT1G15730.1	At1g15730 F7H: Cobalamin biosynthesis CobW-like p...		3	5	2	0	1	1	0	1	1	0	1	1	2	1	1	0	7	1	-1,95	5,03E-02	1,00E+00
AT2G18790.1	PHYB	phytochrome B, Symbols: PHYB, HY3,...	6	4	9	1	0	6	0	0	4	1	1	4	4	1	6	1	3	8	-1,20	1,14E-01	1,00E+00
AT5G22640.1	TIC100	MORN (Membrane Occupation and Reco...	1	1	2	1	1	0	0	0	0	0	1	0	0	1	1	0	0	0	-0,66	5,65E-01	1,00E+00
AT5G07180.1	ERL2	ERECTA-like 2, Symbols: ERL2	1	1	2	0	1	1	0	0	0	0	0	2	0	0	1	0	0	0	-0,67	5,55E-01	1,00E+00
AT1G75660.1	XRN3	5'-3' exoribonuclease 3, Symbols: ...	0	0	4	0	1	2	0	0	0	0	0	1	0	0	1	0	0	0	-0,12	9,13E-01	1,00E+00
AT5G58130.1	ROS3 REPRESSO	RNA-binding (RRM/RBD/RNP motifs) f...	0	0	5	0	0	3	0	0	0	0	0	1	0	1	2	0	0	0	-0,40	7,24E-01	1,00E+00
AT3G53710.1	AGD6	ARF-GAP domain 6, Symbols: AGD6	2	2	1	0	0	1	0	0	0	0	1	0	2	1	0	0	0	0	-1,80	1,46E-01	1,00E+00
AT1G16740.1	At1g16740 F19f	Ribosomal protein L20	3	2	1	1	3	0	0	0	0	0	3	0	2	2	1	0	0	0	-0,34	7,26E-01	1,00E+00
AT4G13190.1	PBL24 At4g131f	Protein kinase superfamily protein	2	1	3	0	1	1	0	0	0	0	0	3	0	0	1	0	0	0	-1,21	2,56E-01	1,00E+00
AT5G03650.1	SBE2.2	starch branching enzyme 2.2, Symbo...	6	1	0	0	1	1	0	0	0	0	1	1	2	1	0	0	0	0	-1,56	1,85E-01	1,00E+00
AT5G42480.1	ARC6	Chaperone DnaJ-domain superfamily ...	2	1	1	1	0	1	0	0	0	0	0	1	1	0	2	0	2	2	-0,69	5,47E-01	1,00E+00
AT5G04020.1		calmodulin binding	1	0	3	0	0	3	0	0	0	0	0	3	2	0	2	0	1	2	-0,17	8,81E-01	1,00E+00
AT1G32090.1	At1g32090	early-responsive to dehydration st...	2	1	1	0	0	1	0	0	0	0	1	0	1	1	1	1	1	1	-1,50	2,42E-01	1,00E+00
AT3G49680.1	BCAT3	branched-chain aminotransferase 3,...	0	3	1	0	0	1	0	0	0	0	1	2	0	0	1	0	1	1	-1,50	2,49E-01	1,00E+00
AT1G72550.1	At1g72550	tRNA synthetase beta subunit famil...	1	0	3	0	0	1	0	0	0	0	0	4	0	0	1	0	0	1	-1,46	2,58E-01	1,00E+00
AT5G41880.1		DNA primases;DNA primases, Symbols...	1	0	3	0	0	1	0	0	0	0	0	1	0	0	1	0	1	0	-1,46	2,58E-01	1,00E+00
AT1G79150.1		binding	1	2	2	1	1	3	0	0	0	0	0	2	0	1	2	0	1	2	0,21	8,23E-01	1,00E+00
AT3G19820.1	DIM	cell elongation protein / DWARF1 / ...	0	2	3	0	1	1	0	0	0	0	0	2	0	0	1	0	2	1	-0,95	3,91E-01	1,00E+00
ATCG00380.1	rsp4	chloroplast ribosomal protein S4, ...	1	2	2	0	1	0	0	0	0	0	0	1	1	1	1	0	1	1	-1,77	1,52E-01	1,00E+00
AT3G18480.1	CASP	CCAAT-displacement protein alterna...	0	3	2	0	0	1	0	0	0	1	2	2	0	1	1	0	1	1	-1,78	1,56E-01	1,00E+00
AT2G07360.2	At2g07360 At2g	SH3 domain-containing protein	1	1	3	1	0	1	0	0	0	1	0	2	1	0	0	0	0	2	-0,95	3,88E-01	1,00E+00
AT1G61640.1	At1g61640 T25f	Protein kinase superfamily protein	1	1	3	0	0	2	0	0	0	0	0	2	0	0	1	0	0	2	-0,97	3,83E-01	1,00E+00
AT4G31460.1	At4g31460 F3L1	Ribosomal L28 family	1	3	1	0	0	1	0	0	0	0	0	2	0	0	1	0	1	1	-1,80	1,49E-01	1,00E+00
AT1G15950.1	CCR1	cinnamoyl coa reductase 1, Symbols...	1	1	3	0	0	1	0	0	0	0	0	3	0	0	1	0	0	1	-0,96	3,85E-01	1,00E+00
AT2G17220.1	PIX13 KIN3 At2g	Protein kinase superfamily protein	1	2	3	0	1	2	0	0	0	0	0	3	0	0	2	0	1	1	-0,71	4,79E-01	1,00E+00
AT4G34090.3	F28A23.80 At4g	unknown protein	2	2	2	0	1	1	0	0	0	0	2	0	2	1	1	0	1	0	-1,23	2,50E-01	1,00E+00
AT4G36690.1	U2AF65A	U2 snRNP auxilliary factor, large ...	2	1	3	1	0	2	0	0	0	0	0	2	0	0	2	0	0	1	-0,70	4,80E-01	1,00E+00
AT2G20850.1	SRF1	STRUBBELIG-receptor family 1, Symb...	3	1	2	1	0	0	0	0	0	0	1	1	2	0	0	0	0	1	-2,02	9,41E-02	1,00E+00
AT5G10470.2	KCA1	kinesin like protein for actin bas...	3	1	4	2	3	3	0	0	1	0	1	4	3	2	3	0	1	3	0,23	7,69E-01	1,00E+00
AT2G05940.1	RIPK PBL14 PIXE	Protein kinase superfamily protein	2	2	4	0	2	2	1	0	0	1	0	3	1	0	2	0	1	1	-0,72	4,24E-01	1,00E+00
AT1G48950.1	F27K7.14	C3HC zinc finger-like	1	1	7	1	0	4	0	0	1	1	0	6	4	0	4	0	1	4	-0,54	5,62E-01	1,00E+00
AT1G13220.2	CRWN2	nuclear matrix constituent protein...	2	5	6	2	0	7	0	1	2	2	1	4	2	0	4	1	1	6	-0,31	6,92E-01	1,00E+00
AT1G26630.1	ELF5A-2	Eukaryotic translation initiation ...	23	29	27	31	25	29	34	30	42	28	28	34	34	27	34	32	29	32	0,34	3,49E-01	1,00E+00
AT3G49560.1	T9C5.150 HP30	Mitochondrial import inner membran...	0	3	1	0	2	1	2	1	0	0	0	0	0	0	0	0	3	1	-0,17	8,74E-01	1,00E+00
AT1G43850.1	SEU	SEUSS transcriptional co-regulator...	1	3	0	2	0	0	2	0	0	0	0	0	1	1	0	0	0	0	-0,69	5,62E-01	1,00E+00
AT3G61440.1	CYS1	cysteine synthase C1, Symbols: ATC...	0	4	0	1	1	0	0	1	0	0	0	0	0	1	1	0	5	1	-0,69	5,68E-01	1,00E+00
AT3G09720.1	RH57	P-loop containing nucleoside triph...	0	1	3	0	0	3	0	0	2	0	0	0	0	0	3	0	0	4	-0,16	8,85E-01	1,00E+00
AT3G13120.1	RPS10	Ribosomal protein S10p/S20e family...	2	2	0	0	1	0	1	0	0	0	0	0	2	1	0	1	2	0	-1,51	2,43E-01	1,00E+00
AT2G29460.1	GSTU4	glutathione S-transferase tau 4, S...	1	3	0	0	0	1	0	1	1	0	0	0	0	1	0	0	3	0	-1,53	2,42E-01	1,00E+00
AT4G25770.1	At4g25770 F14f	alpha/beta-Hydrolases superfamily ...	1	2	1	1	0	0	0	0	1	0	0	0	1	1	1	0	1	0	-1,48	2,49E-01	1,00E+00
AT1G60660.1	CB5LP	cytochrome B5-like protein, Symbol...	2	2	0	0	2	0	1	1	0	0	0	0	1	0	0	1	0	0	-0,71	5,45E-01	1,00E+00
AT1G74960.2	KAS2	fatty acid biosynthesis 1, Symbols...	2	1	1	0	0	2	0	0	1	0	0	0	1	0	0	0	0	1	-0,71	5,38E-01	1,00E+00
AT1G04710.1	KAT1	peroxisomal 3-ketoacyl-CoA thiolas...	3	2	0	0	2	1	0	1	0	0	0	0	0	1	0	0	1	0	-0,52	6,28E-01	1,00E+00
AT2G42010.1	PLDBETA1	phospholipase D beta 1, Symbols: P...	0	0	5	0	1	1	0	0	1	0	0	0	0	0	1	0	0	1	-0,89	4,41E-01	1,00E+00
AT5G65960.1	At5g65960 K2A	GTP binding	2	2	3	0	1	2	0	1	2	0	0	1	2	0	1	0	1	2	-0,92	3,39E-01	1,00E+00
AT3G10370.1	SDP6	FAD-dependent oxidoreductase famil...	1	3	4	0	2	1	0	1	1	0	1	0	0	1	2	0	4	1	-1,08	2,56E-01	1,00E+00
AT2G43460.1-AT3G RPL38A		Ribosomal L38e protein...	5	4	1	1	5	4	6	0	0	0	0	1	2	4	6	1	3	1	0,15	8,49E-01	1,00E+00
AT5G51220.1	MWD22.17 MW	ubiquinol-cytochrome C chaperone f...	1	2	1	1	2	0	2	2	0	1	0	0	0	0	0	0	0	0	-0,16	8,82E-01	1,00E+00
AT5G47430.1	At5g47430 MQL	DWNN domain, a CCHC-type zinc finger	0	2	2	1	0	1	0	1	0	0	1	1	0	0	0	2	1	0	-0,66	5,64E-01	1,00E+00
AT1G71310.2	RAD52-1 ODB1 / cobalt ion binding		1	1	2	0	0	2	0	1	0	0	1	1	0	0	0	0	2	1	-0,69	5,51E-01	1,00E+00
AT5G61460.1	SMC6B	P-loop containing nucleoside triph...	1	0	4	0	0	1	0	0	1	0	0	2	0	0	0	0	0	2	-1,74	1,67E-01	1,00E+00
AT4G11220.1	RTNLB2	VRB2-interacting protein 2, Symbo...	1	3	3	0	1	3	0	2	1	1	0	3	0	0	1	0	3	2	-0,55	5,54E-01	1,00E+00
AT3G59820.2		LETM1-like protein	1	3	3	1	2	0	4	2	2	1	2	1	0	1	0	2	1	1	-0,89	3,60E-01	1,00E+00
AT4G33250.1	TIF3K1	eukaryotic translation initiation ...	2	1	4	0	0	1	0	0	1	0	0	2	0	0	1	0	0	2	-2,22	5,93E-02	1,00E+00

Table S5: Raw data of mass spectrometry results shown in Figure 5. Rows are sorted according to the p-value of RNA R vs tRF Ala column and RNA R vs c-Myc.

Identification by Mascot and validation at FDR<1% by Proline (ProFI consortium)	RNA R			c-Myc			tRF Ala 5'GGGG			tRF Ala 5'GCGU			tRF Gly 5'GGGG		
	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3
Total number of proteins identified	3614	2942	3443	3189	3154	3400	3199	2653	3481	2864	2154	3365	3286	3110	3564
Total number of spectra identified	32751	23313	30948	28364	25465	28586	26641	18341	28918	22883	15182	27386	28164	24631	28569

spectra = 1 spectrum
 # spectra = 2-5 spectra
 # spectra = 6-10 spectra
 # spectra = 11-30 spectra
 # spectra >= 31 spectra

Accession	Description	Basic spectral count / number of spectra by replicate															Statistics with IP-inquiry from H.Zuber (Norm. DEseq)						
		RNA R			c-Myc			tRF Ala 5'GGGG			tRF Ala 5'GCGU			tRF Gly 5'GGGG			RNA R vs tRF Ala 5'GGGG			RNA R vs c-Myc			
		#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	#1	#2	#3	Log2FC	p-value	adjp	Log2FC	p-value	adjp	
AT2G35920.1	RNA helicase family protein			1	15	19	11	2	13	4								3,97	7,69E-05	3,47E-01	5,11	7,99E-10	3,65E-06
AT4G39660.1	alanine:glyoxylate aminotransferas...	3		3	9		9	6	1	8	7	1	10	8				1,40	3,99E-02	9,79E-01	1,59	1,13E-01	9,47E-01
AT5G27560.1	Domain of unknown function (DUF1995)				2	2	1	4		2	1	1	1	2	1	2		4,16	7,60E-03	9,79E-01	3,88	1,91E-02	9,47E-01
AT4G37190.1	unknown protein				2	2		2	1	1	1		2	1	1	2		3,61	2,86E-02	9,79E-01	3,58	4,11E-02	9,47E-01
AT1G73460.1	Protein kinase superfamily protein	1		1	4	4	3	6	1	2	2	2	2	3	2	1		2,11	2,73E-02	9,79E-01	2,33	9,05E-03	9,47E-01
AT3G22990.1	ARM repeat superfamily protein, Sy...					2		1	2	1	1	2			1			3,61	2,87E-02	9,79E-01	2,69	1,69E-01	9,47E-01
AT3G16857.2	response regulator 1, Symbols: ARR...					1	1	1	2	1			1			2		3,61	2,87E-02	9,79E-01	2,69	1,61E-01	9,47E-01
AT3G04760.1	Pentatricopeptide repeat (PPR-like...				1		1	1	2	1	1			1		1		3,61	2,87E-02	9,79E-01	2,69	1,60E-01	9,47E-01
AT5G12200.1	pyrimidine 2, Symbols: PYD2						2	1					3			4		3,61	2,98E-02	9,79E-01	2,69	1,68E-01	9,47E-01
AT1G76950.1	Regulator of chromosome condensati...				1		1			1	2	1	1	1		1		1,93	3,26E-01	9,79E-01	2,69	1,60E-01	9,47E-01
AT5G12140.1	cystatin-1, Symbols: ATCYS1, CYS1	1			1	1	1	2	1	3	1	1	4	2	2	1		2,33	5,40E-02	9,79E-01	1,35	3,40E-01	9,47E-01
AT2G16930.2	Ribosomal protein L27 family protein				2	1	3			1			1	1	1	2		1,93	3,26E-01	9,79E-01	4,12	1,00E-02	9,47E-01
AT4G01990.1	Tetratricopeptide repeat (TPR)-lik...				2	1	1			1			1	2		2		1,93	3,26E-01	9,79E-01	3,58	3,68E-02	9,47E-01
AT3G23540.1	alpha/beta-Hydrolases superfamily ...				2		1	1		2	1		2	2	1	2		3,24	6,55E-02	9,79E-01	3,20	7,99E-02	9,47E-01
AT3G13120.1	Ribosomal protein S10p/S20e family...				1	1	1			1			1	2	2	1		1,93	3,26E-01	9,79E-01	3,20	7,42E-02	9,47E-01
AT1G50600.1	scarecrow-like 5, Symbols: SCL5					1	1	1	1	1	1			1	2	2		3,24	6,39E-02	9,79E-01	2,69	1,61E-01	9,47E-01
AT1G62010.1	Mitochondrial transcription termin...				1		1	1		1			1	2	1	1		1,93	3,26E-01	9,79E-01	2,69	1,60E-01	9,47E-01
AT4G03280.1	photosynthetic electron transfer C...			1	4	2	5		1	3	2		3	2	2	2		1,79	1,79E-01	9,79E-01	3,11	2,72E-03	9,47E-01
AT1G23110.1	unknown protein				2	2		1					1	1		1		1,93	3,26E-01	9,79E-01	3,58	4,11E-02	9,47E-01
AT3G20780.1	topoisomerase 6 subunit B, Symbol...		1	1	1	7		1	3		1	3		1	6	1		1,01	3,94E-01	9,79E-01	1,88	1,11E-01	9,47E-01
AT4G10110.1	RNA-binding (RRM/RBD/RNP motifs) f...				1	1	1	1	1			1	1	1	1	1		2,73	1,47E-01	9,79E-01	3,20	7,42E-02	9,47E-01
AT3G48890.1	membrane-associated progesterone b...				1		1	1		1	1		1	1	1	1		2,73	1,47E-01	9,79E-01	2,69	1,60E-01	9,47E-01
AT2G19080.1	metaxin-related				1	1		1		1	1		1	1	1	1		2,73	1,47E-01	9,79E-01	2,69	1,61E-01	9,47E-01
AT5G16620.1	hydroxyproline-rich glycoprotein f...				1		1	2		1			1	1	1	1		3,24	6,55E-02	9,79E-01	2,69	1,60E-01	9,47E-01
AT2G30170.1	Protein phosphatase 2C family protein				1			2	1				2	1		2		3,24	6,55E-02	9,79E-01	1,90	3,38E-01	9,47E-01
AT5G40370.2	Glutaredoxin family protein					1	2		1	1			2	1		2		2,73	1,47E-01	9,79E-01	3,20	8,01E-02	9,47E-01
AT4G32690.1	hemoglobin 3, Symbols: GLB3, ATGLB3				1					1			1	1		2		1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G62730.1	Terpenoid synthases superfamily pr...				1		2	1		1			2	2	1			1,93	3,26E-01	9,79E-01	3,20	7,98E-02	9,47E-01
AT1G10990.2	unknown protein				2		1			2	1		1		1	2		2,73	1,49E-01	9,79E-01	3,20	7,99E-02	9,47E-01
AT3G09810.1	isocitrate dehydrogenase VI, Symbo...						2			1	2			1	2			1,93	3,26E-01	9,79E-01	2,69	1,68E-01	9,47E-01
AT5G17910.1	unknown protein				1	1		1	1		1			1	2			2,73	1,47E-01	9,79E-01	2,69	1,61E-01	9,47E-01
AT5G61770.2	PETER PAN-like protein, Symbols: PPA...					4				1			2		2	1		1,93	3,26E-01	9,79E-01	3,58	6,36E-02	9,47E-01
AT1G03910.2	unknown protein					1				1				2	1			1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G06190.1	Rho termination factor	2	2		3	3	3	3	5	2	3	3	1	4	3	3		1,39	8,58E-02	9,79E-01	1,17	1,48E-01	9,47E-01
AT1G18850.1	unknown protein				3			1	2		1	1			3			3,24	6,55E-02	9,79E-01	3,20	9,20E-02	9,47E-01
AT1G31817.1	Ribosomal L18p/L5e family protein,...			1	2	1	2	1	1	1	1	2	1	3	1	1		1,41	3,12E-01	9,79E-01	2,03	1,11E-01	9,47E-01
AT3G54150.1	S-adenosyl-L-methionine-dependent ...		1	2	3	2	4	3	1	3	1		5	2	2	4		1,26	1,78E-01	9,79E-01	1,54	7,50E-02	9,47E-01
AT1G69740.1	Aldolase superfamily protein, Symb...	1		1	3		2	3		4	2		2	2	1	3		1,76	8,77E-02	9,79E-01	1,24	2,75E-01	9,47E-01
AT5G13570.2	decapping 2, Symbols: DCP2				1		1			1	1		1	1		1		1,93	3,26E-01	9,79E-01	2,69	1,60E-01	9,47E-01

AT3G11770.1	Polynucleotidyl transferase, ribon...			1	1	2			1	1		1	3,24	6,55E-02	9,79E-01	1,90	3,38E-01	9,47E-01
AT3G24030.1	hydroxyethylthiazole kinase family...		2	1	1		1		1	1		1	2,73	1,47E-01	9,79E-01	3,20	8,05E-02	9,47E-01
AT3G04950.1	unknown protein	1	1	1	1		1		1	1		1	2,73	1,47E-01	9,79E-01	3,20	7,42E-02	9,47E-01
AT2G47400.1	CP12 domain-containing protein 1, ...	1	1		1				1	1		1	1,93	3,26E-01	9,79E-01	2,69	1,61E-01	9,47E-01
AT4G32870.1	Polyketide cyclase/dehydrase and l...			1			1		1	1		1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT2G46230.1	PIN domain-like family protein		2			1		1	2	1	1		2,73	1,47E-01	9,79E-01	2,69	1,69E-01	9,47E-01
AT1G26530.1	PIN domain-like family protein		1			1		1	2	1	1		2,73	1,47E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT3G09830.1	Protein kinase superfamily protein				1	1		1	2	1	1		2,73	1,47E-01	9,79E-01	1,90	3,38E-01	9,47E-01
ATCG00710.1	photosystem II reaction center pro...	1	1		1			1		1	1		1,93	3,26E-01	9,79E-01	2,69	1,61E-01	9,47E-01
AT3G25780.1	allene oxide cyclase 3, Symbols: AOC3	1			2			1	1	1	1		1,93	3,26E-01	9,79E-01	3,20	7,98E-02	9,47E-01
AT1G50240.2	Protein kinase family protein with...	1			1			1	1	1	1		1,93	3,26E-01	9,79E-01	2,69	1,60E-01	9,47E-01
AT1G14610.1	valyl-tRNA synthetase / valine-tR...				1			1	1	1	1		1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT5G41940.1	Ypt/Rab-GAP domain of gyp1p superf...	1	1	1		1			1	2			1,93	3,26E-01	9,79E-01	3,20	7,42E-02	9,47E-01
AT3G53120.1	Modifier of rudimentary (Mod(r)) p...			1	2			1	2			2	1,93	3,26E-01	9,79E-01	3,20	8,01E-02	9,47E-01
AT3G25230.2	rotamase FKBP 1, Symbols: ROF1, AT...				1			1		1		2	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT5G55125.1	Ribosomal protein L31	1				1		2	1			2	3,24	6,55E-02	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G47128.1	Granulin repeat cysteine protease...				1			2	1			2	2,73	1,49E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT4G00830.1	RNA-binding (RRM/RBD/RNP motifs) f...			1			3		1			2	3,24	6,95E-02	9,79E-01	1,90	3,38E-01	9,47E-01
AT4G05020.2	NAD(P)H dehydrogenase B2, Symbols:...	1		2	2	2	3	1	2	2		2	2,33	5,38E-02	9,79E-01	2,27	6,14E-02	9,47E-01
AT4G12340.1	copper ion binding	1		2	2	1	1	1	1	2	2	1	1,41	3,12E-01	9,79E-01	2,03	1,12E-01	9,47E-01
AT3G01340.1	Transducin/WD40 repeat-like superf...	1		1	1	1	1	1	2	1		1	1,79	1,76E-01	9,79E-01	1,35	3,39E-01	9,47E-01
AT5G08740.1	NAD(P)H dehydrogenase C1, Symbols:...			1					1	3		1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT3G57330.1	autoinhibited Ca2+-ATPase 11, Symb...	1			1	1		1	1		1	1	2,73	1,47E-01	9,79E-01	2,69	1,60E-01	9,47E-01
AT5G49820.1	Protein of unknown function, DUF64...	1			1			1		1		1	1,93	3,26E-01	9,79E-01	2,69	1,60E-01	9,47E-01
AT5G05060.1	Cystatin/monellin superfamily protein			1			1	1		1		1	2,73	1,47E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT2G02730.1	Protein of unknown function (DUF1664)	1	1			1			1			1	1,93	3,26E-01	9,79E-01	2,69	1,61E-01	9,47E-01
AT5G13630.1	magnesium-chelatase subunit chlH, ...	2				1			1			1	1,93	3,26E-01	9,79E-01	2,69	1,69E-01	9,47E-01
AT2G24090.1	Ribosomal protein L35	1				1			1			1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G26640.1	Amino acid kinase family protein				1			1			1	1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT5G20070.1	nudix hydrolase homolog 19, Symbol...				1			1			1	1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT5G49930.1	zinc knuckle (CCHC-type) family pr...	1						1			2		1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT3G08740.1	elongation factor P (EF-P) family ...	1	1				1	1	1			1	2,73	1,47E-01	9,79E-01	2,69	1,61E-01	9,47E-01
AT4G24550.2	Clathrin adaptor complexes medium ...			1				1	1			1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G50480.1	10-formyltetrahydrofolate syntheta...				1	1		2			1	1	3,24	6,55E-02	9,79E-01	1,90	3,38E-01	9,47E-01
AT4G03520.1	Thioredoxin superfamily protein, S...				2	1		1			1	1	2,73	1,47E-01	9,79E-01	2,69	1,68E-01	9,47E-01
AT3G51840.1	acyl-CoA oxidase 4, Symbols: ACX4,...	1			1			2			1	1	2,73	1,49E-01	9,79E-01	2,69	1,60E-01	9,47E-01
AT3G23280.1	XB3 ortholog 5 in Arabidopsis thal...				1			2			1	1	2,73	1,49E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G16360.1	LEM3 (ligand-effect modulator 3) f...	1			1	1		1			1	1	1,93	3,26E-01	9,79E-01	2,69	1,60E-01	9,47E-01
AT2G06850.1	xyloglucan endotransglucosylase/hy...				2			1			1	1	1,93	3,26E-01	9,79E-01	2,69	1,68E-01	9,47E-01
AT1G06000.1	UDP-Glycosyltransferase superfamil...				2			1			1	1	1,93	3,26E-01	9,79E-01	2,69	1,68E-01	9,47E-01
AT1G79830.4	golgin candidate 5, Symbols: GC5			1				1			1	1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT4G21560.1	vacuolar protein sorting-associate...				1			1			1	1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT3G01390.1	vacuolar membrane ATPase 10, Symbo...				1			1			1	1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G78070.1	Transducin/WD40 repeat-like superf...				1			1			1	1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G01680.1	plant U-box 54, Symbols: ATPUB54, ...				1			1			1	1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT3G47780.1	ABC2 homolog 6, Symbols: ATATH6, ATH6				1			1			1	1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT1G18210.1	Calcium-binding EF-hand family pro...			2	1			1			2		1,93	3,26E-01	9,79E-01	3,20	8,05E-02	9,47E-01
AT3G55605.1	Mitochondrial glycoprotein family ...	1	1			1	1			1		1	2,73	1,47E-01	9,79E-01	2,69	1,61E-01	9,47E-01
AT4G16770.1	2-oxoglutarate (2OG) and Fe(II)-de...	1						1		1		1	1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT3G18140.1	Transducin/WD40 repeat-like superf...			1		2			1			1	2,73	1,49E-01	9,79E-01	1,90	3,38E-01	9,47E-01

AT3G09070.1	Protein of unknown function (DUF740)			1			1		1			1		1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT4G01897.1	unknown protein				1		1					1		1,93	3,26E-01	9,79E-01	1,90	3,38E-01	9,47E-01
AT3G05470.1	Actin-binding FH2 (formin homology...	1	2	2	2	2	2	2	1	1	1	1		2,33	5,35E-02	9,79E-01	2,27	6,18E-02	9,47E-01
AT4G31500.1	cytochrome P450, family 83, subfam...	1	2		1			3	1		1	1		1,41	3,18E-01	9,79E-01	1,35	3,49E-01	9,47E-01
AT3G28480.2	Oxoglutarate/iron-dependent oxygenase	1	2		3		1				2			1,41	3,14E-01	9,79E-01	2,03	1,22E-01	9,47E-01
AT1G25260.1	Ribosomal protein L10 family protein	1			4			2	1	1	1			1,41	3,15E-01	9,79E-01	1,73	2,30E-01	9,47E-01
AT4G34660.1	SH3 domain-containing protein		1	2	2	1	1	2	1		2			1,79	1,76E-01	9,79E-01	2,03	1,12E-01	9,47E-01
AT3G54760.1	dentin sialophosphoprotein-related	1			2				3		1	1	1	1,41	3,18E-01	9,79E-01	0,85	5,95E-01	9,89E-01
AT4G10000.1	Thioredoxin family protein		1	1		1	1	2	1		1	2		1,41	3,14E-01	9,79E-01	0,85	5,88E-01	9,89E-01
AT2G21250.1	NAD(P)-linked oxidoreductase super...		1			2	1	3	1	1	1	2		1,79	1,79E-01	9,79E-01	0,85	5,95E-01	9,89E-01
AT5G22140.1	FAD/NAD(P)-binding oxidoreductase ...		1			1	1	1	2	1	2	2		1,79	1,76E-01	9,79E-01	0,05	9,76E-01	9,89E-01
AT1G60490.1	vacuolar protein sorting 34, Symbo...							5				2		3,91	2,03E-02	9,79E-01			
AT5G48180.1	nitrile specifier protein 5, Symbo...							2				2		2,73	1,49E-01	9,79E-01			
AT5G61060.2	histone deacetylase 5, Symbols: HDA05						1	1	1	1	1	1		2,73	1,47E-01	9,79E-01			
AT5G44730.2	Haloacid dehalogenase-like hydroლა...							1			1	1	1	1,93	3,26E-01	9,79E-01			
AT2G18630.1	Protein of unknown function (DUF677)						1	1	1		1			2,73	1,47E-01	9,79E-01			
AT3G62450.1	unknown protein						1	1	1			1		2,73	1,47E-01	9,79E-01			
AT4G29380.1	protein kinase family protein / WD...							1	1			2		1,93	3,26E-01	9,79E-01			
AT1G10050.1	glycosyl hydrolase family 10 prote...					2					1	1		2,73	1,49E-01	9,79E-01			
AT1G03000.1	peroxin 6, Symbols: PEX6								1		1	1		1,93	3,26E-01	9,79E-01			
AT5G11920.1	6-&1-fructan exohydrolase, Symbols...								1					1,93	3,26E-01	9,79E-01			
AT2G43810.1	Small nuclear ribonucleoprotein fa...							1	1	1				2,73	1,47E-01	9,79E-01			
AT3G51820.1	UbiA prenyltransferase family prot...						1			1				1,93	3,26E-01	9,79E-01			
AT3G60210.1	GroES-like family protein						1		1			2	1	2,73	1,47E-01	9,79E-01			
AT4G15840.1	BTB/POZ domain-containing protein						1	1	1	1				3,24	6,39E-02	9,79E-01			
ATCG00350.1	Photosystem I, PsaA/PsaB protein, ...						1	1			1	1		2,73	1,47E-01	9,79E-01			
AT1G05350.1	NAD(P)-binding Rossmann-fold super...						1		1		1	1		2,73	1,47E-01	9,79E-01			
AT2G21280.1	NAD(P)-binding Rossmann-fold super...						1		1		1			2,73	1,47E-01	9,79E-01			
AT5G58440.1	sorting nexin 2A, Symbols: SNX2a						1				1	1		1,93	3,26E-01	9,79E-01			
AT3G05230.1	Signal peptidase subunit						1				1	1		1,93	3,26E-01	9,79E-01			
AT5G54440.1	CLUB, Symbols: CLUB, AtTRS130						1				1			1,93	3,26E-01	9,79E-01			
AT1G70750.1	Protein of unknown function, DUF593								1		1			1,93	3,26E-01	9,79E-01			
AT3G25530.1	glyoxylate reductase 1, Symbols: G...								1		1			1,93	3,26E-01	9,79E-01			
AT4G23940.1	FtsH extracellular protease family	1			4	4	2	3	2	1	4	1		2,09	1,00E-01	9,79E-01	2,98	5,15E-03	9,47E-01
AT5G64580.1	AAA-type ATPase family protein	1	1		4	4	3	3	2	3	1	2	1	1,31	2,44E-01	9,79E-01	2,33	9,05E-03	9,47E-01
AT3G22740.1	homocysteine S-methyltransferase 3...	1	1		3	3	3	2	1	2	2	2	3	1,31	2,40E-01	9,79E-01	2,05	2,90E-02	9,47E-01
AT2G25870.1	haloacid dehalogenase-like hydroლა...	1			3		1	2		2		1	1	1,79	1,78E-01	9,79E-01	1,73	2,09E-01	9,47E-01
AT5G44070.1	phytochelatin synthase 1 (PCS1), S...			2	1	1	5	1	5		4	1	1	1,55	1,56E-01	9,79E-01	1,70	1,11E-01	9,47E-01
AT1G03400.1	2-oxoglutarate (2OG) and Fe(II)-de...			1			3	1	2		2	1		1,41	3,14E-01	9,79E-01	1,35	3,62E-01	9,53E-01
AT5G13500.1	unknown protein		1		1	1	1	1	1	1	1	1		1,41	3,12E-01	9,79E-01	1,35	3,40E-01	9,47E-01
AT1G06730.1	pfkB-like carbohydrate kinase fami...	2			3	1	1	1	1	2	3	1		1,01	3,91E-01	9,79E-01	1,24	2,75E-01	9,47E-01
AT3G17420.1	glyoxysomal protein kinase 1, Symb...	1	1		2		3	1	1	2	1	4	2	1,01	3,88E-01	9,79E-01	1,24	2,75E-01	9,47E-01
AT2G42210.2	Mitochondrial import inner membran...		2		1	3	1	2	2	1	3	1	1	1,31	2,44E-01	9,79E-01	1,24	2,79E-01	9,47E-01
AT3G45190.1	SlT4 phosphatase-associated family...	1	2		3	2	2	2	1	3	5	1	1	1,05	2,83E-01	9,79E-01	1,20	1,95E-01	9,47E-01
AT4G13200.1	unknown protein	2	1	1	2	2	4	3	4	1	2	1	4	1,07	2,00E-01	9,79E-01	1,01	2,15E-01	9,47E-01
AT2G14720.1	vacuolar sorting receptor 4, Symbo...	1		1	1	1	2	1	2	1	3	1	2	1,01	3,89E-01	9,79E-01	0,95	4,18E-01	9,84E-01
AT4G11220.1	VIRB2-interacting protein 2, Symbo...		2		1	3		2	2	1	2	2	2	1,31	2,44E-01	9,79E-01	0,95	4,45E-01	9,89E-01
AT3G51980.1	ARM repeat superfamily protein			2		4		1		3	1		4	1,01	3,95E-01	9,79E-01	0,95	4,68E-01	9,89E-01
AT2G20010.2	Protein of unknown function (DUF810)		1	1		2		3	1	1	2	1	3	1,31	2,42E-01	9,79E-01	0,95	4,29E-01	9,89E-01
AT3G62910.1	Peptide chain release factor 1, Sy...	1			2			2		1	1		1	1,41	3,14E-01	9,79E-01	0,85	5,95E-01	9,89E-01

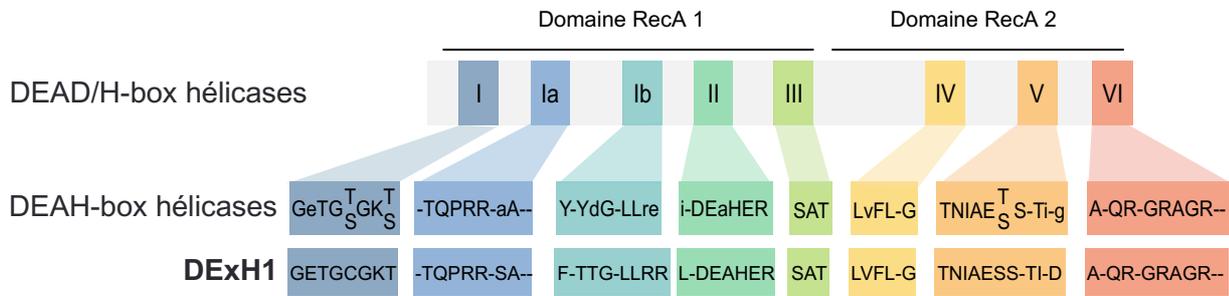


Figure 28 : Motifs des hélicases ARN et de DExH1 de la famille des DEAH-box. Schéma des motifs conservés chez la famille des DEAD/H-box hélicases. Le détail des motifs conservés est donné pour les DEAH-box hélicases et pour DExH1 de *A. thaliana*. Les acides aminés conservés à plus de 80% sont représentés par une lettre capitale (d'après [222]).

IV. Résultats complémentaires à l'article

1. Analyses supplémentaires de DExH1

L'analyse de la séquence protéique de DExH1 montre qu'elle contient les domaines RecA1 et RecA2 aux motifs conservés de la famille des hélicases DExH-box (Figure 28). Les hélicases ayant une forte homologie de séquence avec DExH1 ont en commun ces motifs des domaines Rec1 et RecA2 conservés (Figure 29). Il semblerait donc que ce qui différencie DExH1 des autres hélicases DExH-box soit la région riche en RGG présente en N-ter (voir Figure 4d de l'article présenté dans le paragraphe III ci-dessus). D'autres hélicases présentent une région riche en RGG en C-ter mais uniquement DExH1 possède cette région en N-ter [208].

Concernant sa localisation, d'après les bases de données, DExH1 a une localisation nucléaire. J'ai cependant vérifié sa localisation en expression transitoire dans *N. benthamiana*. Des constructions exprimant DExH1 fusionnée à la GFP, soit en N-terminal soit en C-terminal, ont été agro-infiltrées dans des feuilles de *N. benthamiana*. Après 48h d'expression, les feuilles ont été visualisées au microscope confocal. La fluorescence de la GFP a été observée dans le cytosol des cellules épidermiques. Dans certains cas, la GFP été visible dans des structures rondes, vraisemblablement dans le noyau des cellules (Figure 30). L'expression transitoire de cette même construction dans des protoplastes d'*A. thaliana* a révélé une localisation uniquement cytosolique (voir Figure 6a de l'article présenté dans le paragraphe III ci-dessus). Dans les deux modèles, DExH1 semble avoir une localisation cytosolique. De façon surprenante, il est supposé que DExH1 a été observée dans le noyau des cellules de *N. benthamiana* mais pas dans les protoplastes d'*A. thaliana*. Ces variations peuvent être dues aux plantes modèles utilisées, différentes, dans lesquels sont exprimées la construction. Le noyau des protoplastes n'étant pas forcément visibles dans nos conditions d'observations microscopiques, il serait intéressant de renouveler cette expérience en utilisant un marqueur nucléaire.

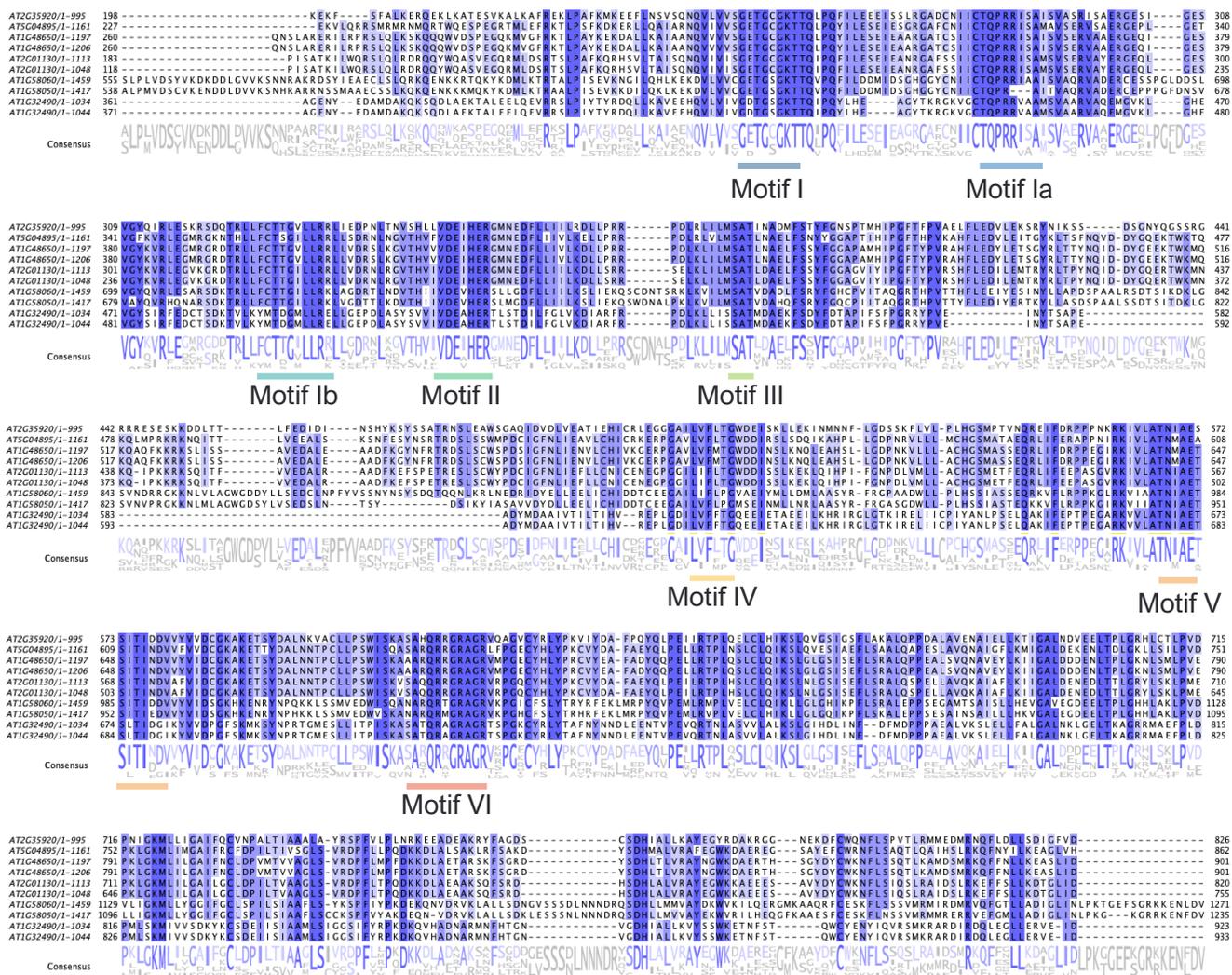


Figure 29 : Alignement de séquences protéiques d'hélicases ayant une forte homologie avec DExH1. La séquence de DExH1 a été utilisée pour rechercher les hélicases proches (NCBI). Les 9 hélicases avec le meilleur score d'alignement ont été utilisées et comparées entre elles avec l'algorithme Muscle et visualisées avec Jalview. Les acides aminés sont colorés par un dégradé de bleu selon leur pourcentage d'apparition (un acide aminé retrouvé dans chaque séquence sera marqué en bleu foncé ; au contraire, un acide aminé retrouvé dans une seule séquence sera blanc).

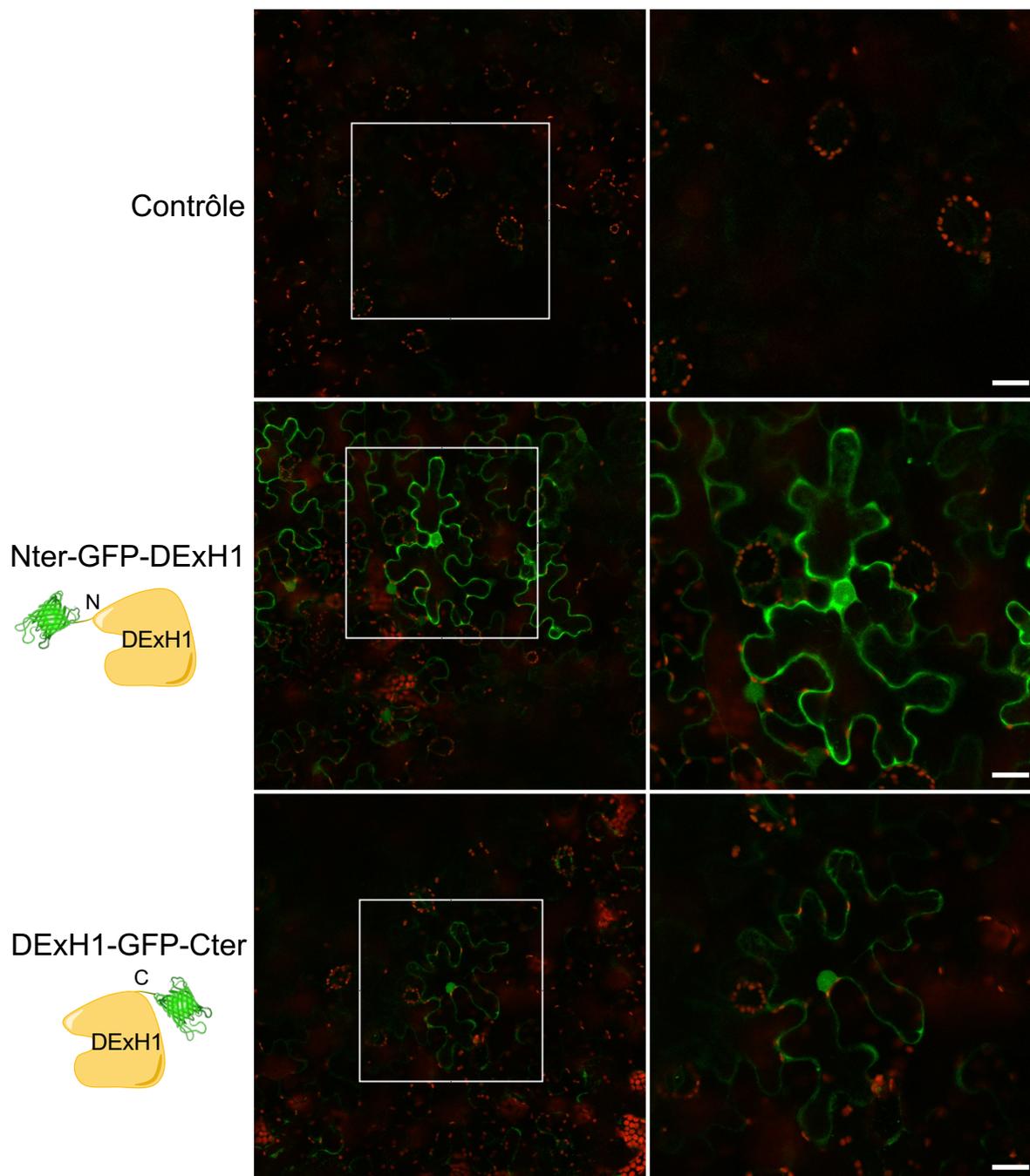


Figure 30 : Localisation subcellulaire de DEXH1 par expression transitoire dans des feuilles des *N. benthamiana*. La protéine DEXH1 fusionnée à la GFP soit en N-ter, soit en C-ter a été exprimée de manière transitoire dans des feuilles de *N. benthamiana* par agro-infiltration. Après 48h d'expression, les feuilles ont été visualisées par microscopie confocale. Une feuille infiltrée uniquement avec le tampon d'agro-infiltration ($MgCl_2$) est utilisée comme contrôle. La chlorophylle est marquée en rouge (autofluorescence), la GFP est marquée en vert. L'image de droite est un zoom de l'image de gauche. L'emplacement du zoom est marqué sur l'image de gauche par un carré blanc. La barre d'échelle en blanc correspond à 20 μm .

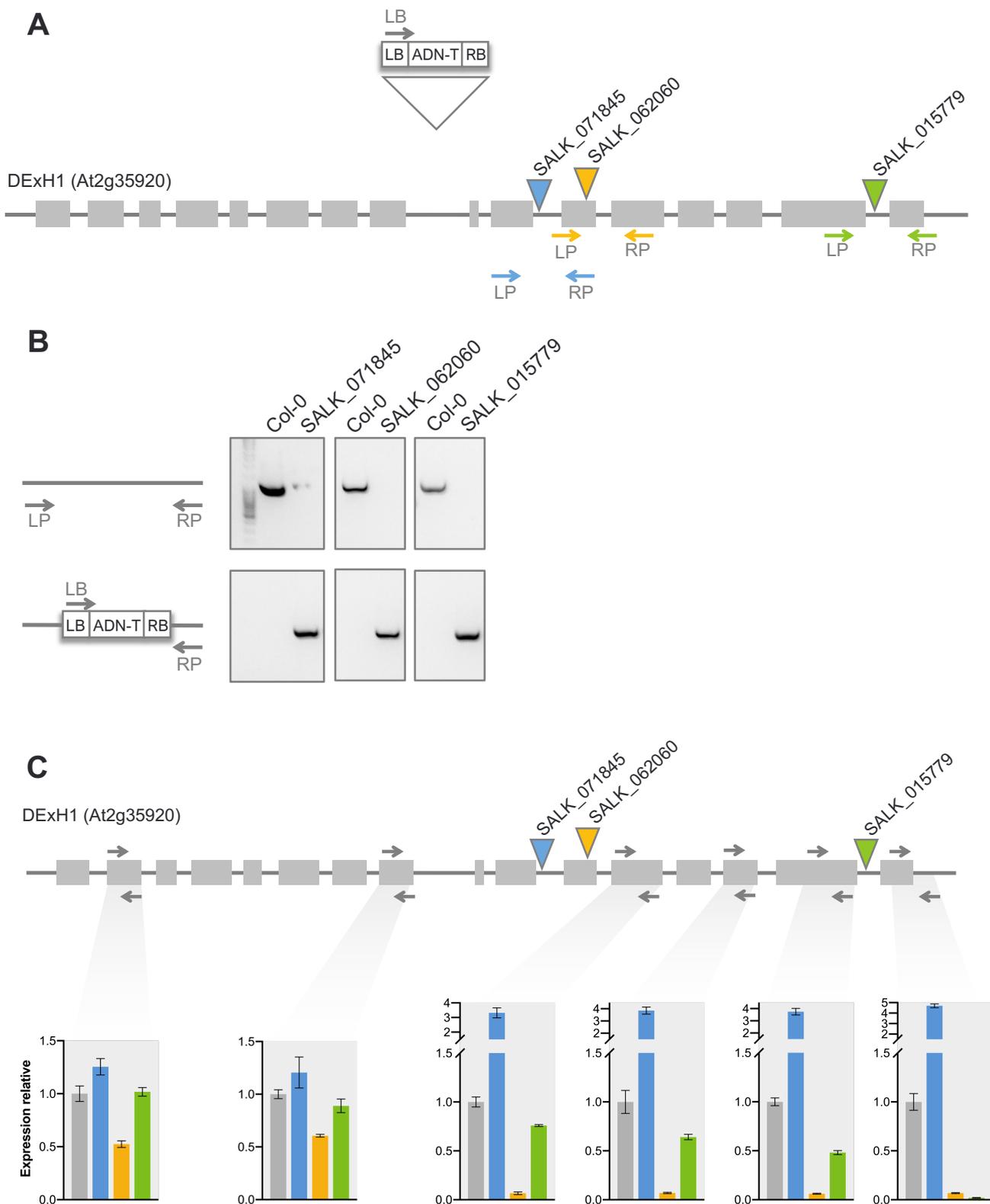


Figure 31 : Génotypage et vérification des mutants d'insertion SALK *dexh1*^{KO}. A : Représentation schématique du gène de DEXH1 avec la localisation des insertions d'ADN-T selon les mutants (triangles de couleurs) et les sites d'hybridation des amorces utilisées pour le génotypage par PCR (flèches de la couleur du mutant correspondant). B : Résultats des PCR réalisés avec les amorces décrites en A. La première PCR avec les amorces LP + RP permet d'amplifier l'ADN dans le cas d'une plante sauvage (Col-0). La deuxième PCR avec les amorces LB + RP permet d'amplifier la jonction entre l'ADN-T et l'ADNg et confirme la présence de l'ADN-T dans les mutants d'insertion SALK. C : Quantification par RT-qPCR de l'expression de l'ARNm de DEXH1 dans les mutants SALK. Les sites d'hybridation des amorces sont indiquées par des flèches sur la représentation de l'ADNg de DEXH1. Pour chaque couple d'amorce, un graphique montre l'expression relative de DEXH1 dans les mutants SALK comparés à Col-0 (histogramme gris).

2. Génotypage et vérification des lignées *dexh1*^{KO}

Pour comprendre la fonction de DEXH1 qui est retrouvée associée au tRF Ala, des lignées *dexh1*^{KO} sont nécessaires. Plusieurs lignées d'insertion sont disponibles dans les banques de mutants d'*A. thaliana*. Au total, 8 lignées ont été commandées (SALK_071845, SALK_015779, SALK_045272, SALK_062060, SALK72746, SALK_061049, SALK_062061 et SALK_061564) et les plantes ont été génotypées.

Le génotypage a été réalisé par PCR. Deux couples d'amorces sont nécessaires pour chaque lignée à génotyper (Figure 31 A). Le premier couple, LP + RP, s'hybride sur l'ADNg sauvage de DEXH1, entourant la zone d'insertion de l'ADN-T prédite par les banques. Une amplification par PCR avec ce couple d'amorce signifie qu'il n'y a pas d'insertion d'ADN-T, et qu'au moins un allèle pour ce gène est sauvage. Le deuxième couple, LB + RP, s'hybride à la fois sur la bordure gauche de l'ADN-T et l'ADNg de DEXH1. Une amplification par PCR avec ce couple montre la présence de l'ADN-T sur au moins l'un des allèles du gène. Si une amplification est observée avec les deux couples d'amorces, ceci signifie que la plante est hétérozygote et possède une version sauvage et une version avec insertion de l'ADN-T.

Le génotypage par PCR a permis d'identifier 3 lignées homozygotes avec insertion d'ADN-T pour le gène de DEXH1 (Figure 31 B). Le séquençage de cette amplification a permis d'identifier avec précision la région où l'ADN-T est inséré. Deux lignées ont une insertion de l'ADN-T dans un intron (SALK_071845 et SALK_015779) et une lignée possède une insertion dans un exon (SALK_062060).

La présence de l'ARNm codant pour DEXH1 a été vérifié par RT-qPCR (une seule expérience, Figure 31 C). En effet, étant donné que l'insertion a lieu dans des introns pour deux des lignées, il se peut qu'elle n'ait pas d'impact au moment de la transcription et de l'épissage de l'ARNm. Plusieurs couples d'amorces ont été utilisés pour couvrir la totalité de l'ARNm. Pour les deux lignées dont l'insertion est dans un intron, les résultats montrent une augmentation ou une diminution mais que au niveau de l'extrémité 3' de l'ARNm pour les lignées SALK_071845 et SALK_015779 respectivement. Dans le cas de la lignée SALK_015779, il est probable que l'insertion dans l'avant-dernier intron produise un transcrit tronqué. Dans ce cas, la présence de la protéine DEXH1 est incertaine. Ces deux lignées n'ont pas été retenues pour la suite des expériences. Concernant la lignée SALK_062060, les résultats montrent une quantité réduite de l'ARNm de DEXH1. Elle devient quasiment nulle après le site d'insertion de l'ADN-T suggérant que si l'ARNm est présent, il l'est en plus petite quantité et n'est dans tous les cas pas complet. S'il est traduit, il est très probable que la protéine DEXH1 ne soit pas complète et fonctionnelle et qu'elle soit dégradée. Ne possédant pas d'anticorps dirigés contre la protéine DEXH1, les lignées d'insertion n'ont pas pu être vérifiées au niveau protéique. La lignée SALK_062060 a finalement été choisie comme lignée *dexh1*^{KO} et utilisée pour les expérimentations.

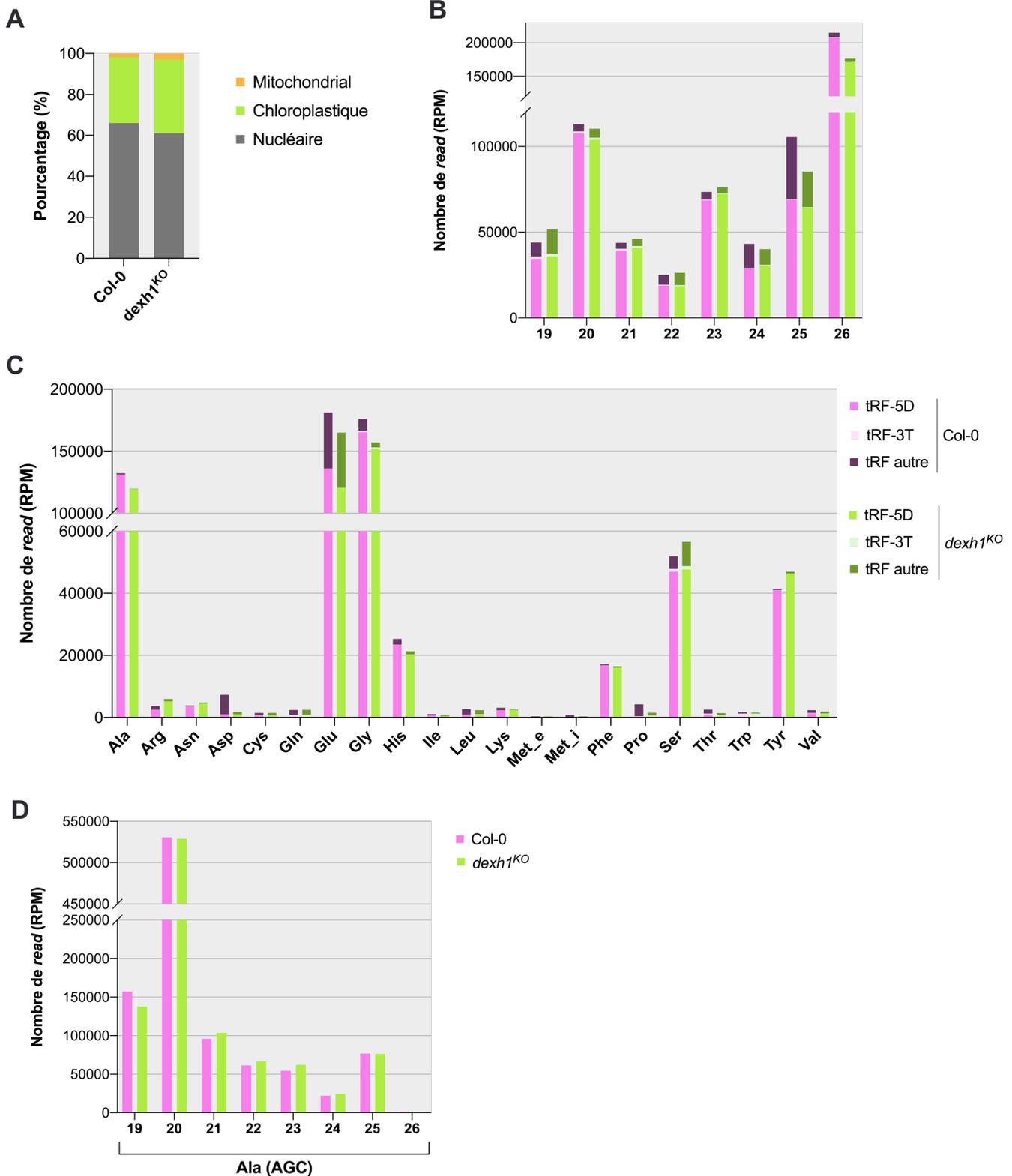


Figure 32 : Comparaison de la distribution des tRFs chez les lignées Col-0 et *dexh1*^{KO} d'*A. thaliana*. Population de tRFs entre 19 et 26 nt issus des banques de petits ARN de feuille de rosette de Col-0 et *dexh1*^{KO} réalisée en un seul réplicat. A : Distribution en pourcentage des tRFs selon l'origine de l'ARNt dont ils sont issus (nucléaire, chloroplastique ou mitochondrial). B : Répartition de la population de tRFs selon leur taille (en nt) et le type (tRF-5D, tRF-3T ou autres tRFs). C : Répartition de la population de tRFs nucléaires en fonction de la famille d'ARNt isoaccepteur dont ils sont originaires et du type de tRF. D : Répartition des tRF-5D issus de l'ARNt^{Ala} (AGC) en fonction de leur taille (en nt).

3. Étude de l'impact de DExH1 sur la biogenèse/stabilité du tRF Ala

L'hélicase DExH1 étant retrouvée en interaction avec le tRF Ala, l'impact qu'elle pourrait avoir sur la biogenèse ou la stabilité du tRF a été étudié. Pour se faire, une banque de petits ARN a été réalisée à partir de feuilles de rosette de plantes Col-0 et de plantes *dexh1^{KO}*, puis séquencée (*Next Generation Sequencing*, Illumina, Mi-seq). Les analyses bio-informatiques ont permis d'identifier les séquences correspondantes aux tRFs. Elles ont été annotées en fonction de la taille des tRFs et de leur ARNt d'origine. Uniquement les tRFs d'une taille comprise entre 19 et 26 nt ont été pris en compte. Les résultats sont résumés sous forme de graphiques dans la Figure 32.

La répartition des tRFs selon leur compartiment d'origine (noyau, chloroplaste ou mitochondrie) semble similaire entre la lignée Col-0 sauvage et la lignée mutante *dexh1^{KO}* (Figure 32 A). La distribution des tRF-5D, tRF-3T et autres selon leur taille a été comparée (Figure 32 B). Globalement, il ne semble pas y avoir de différence notable entre la distribution observée chez les plantes sauvages et *dexh1^{KO}*. Seuls les tRFs de 25 nt semblent plus abondants chez le sauvage Col-0 que chez *dexh1^{KO}*, en particulier lié à une plus grande abondance des tRFs autres dans la lignée sauvage. Aussi, la distribution des tRFs en fonction des ARNt nucléaires dont ils proviennent a été comparée (Figure 32 C). D'une manière générale, très peu de variations sont observées entre les deux lignées. Il est observable que les tRFs autres que -5D et -3T de l'ARNt^{Asp} sont plus abondants dans la lignée sauvage Col-0. Certains tRFs sont légèrement moins abondants dans les plantes sauvages (tRFs Ser et Tyr par exemple) ; d'autres en revanche sont moins abondants chez *dexh1^{KO}* (tRFs Asp, Glu, Gly et His entre autres). Ces différences semblent assez mineures, et ne sont pas considérées comme significatives. Concernant les tRF-5D Ala, allant de 19 à 26 nt, leur abondance, bien que légèrement plus faible dans la lignée KO, ne semble pas significativement différente entre les deux lignées sauvage et *dexh1^{KO}* (Figure 32 C). Finalement, pour le tRF-5D Ala de 20 nt, DExH1 ne semble pas impacter son abondance (Figure 32 D).

La banque d'ARN n'ayant été réalisée qu'en un seul réplicat, il est difficile de conclure sur la significativité des différences observées entre la lignée sauvage et la lignée *dexh1^{KO}*. Ces résultats préliminaires suggèrent toutefois qu'aucune différence importante n'existe, en particulier au niveau des tRF Ala. Pour l'instant, nous n'avons pas jugé prioritaire la réalisation de réplicats.

Concernant la stabilité du tRF Ala, la présence ou absence de l'hélicase DExH1 ne semble pas avoir d'effet majeur. Ces résultats confortent notre hypothèse que DExH1 a un rôle dans la formation/déformation des G-quadruplex qui se forment avec les tRFs Ala mais pas sur l'abondance générale du tRF Ala au sein de la cellule.

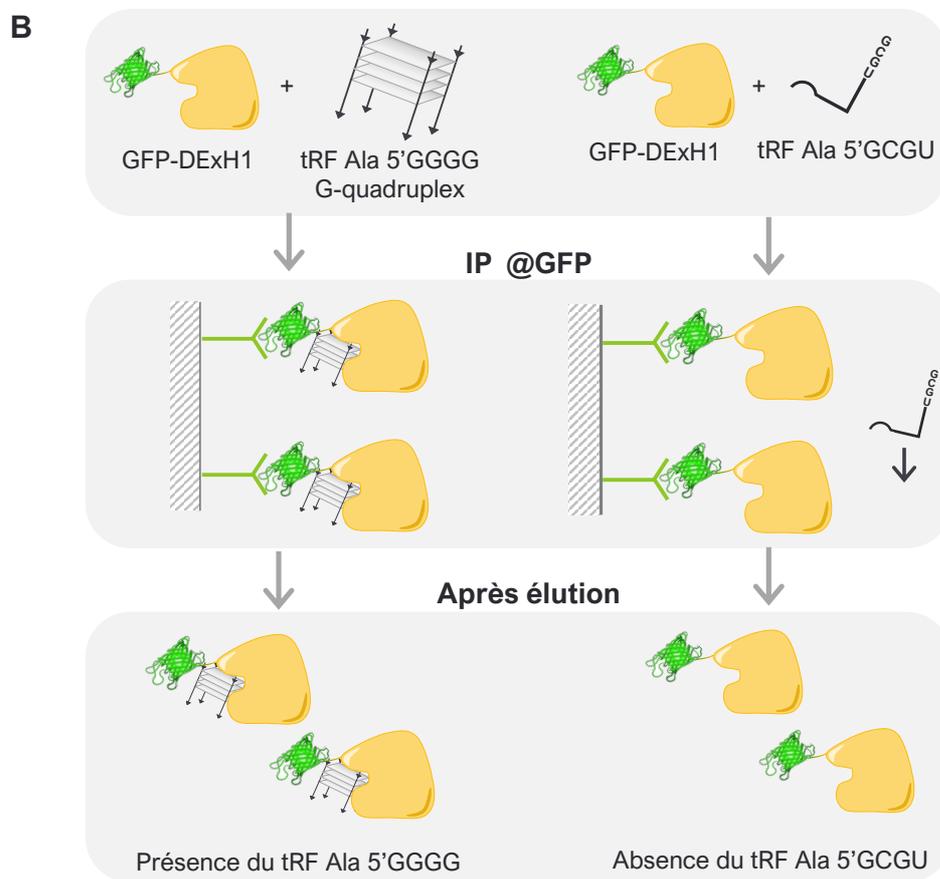
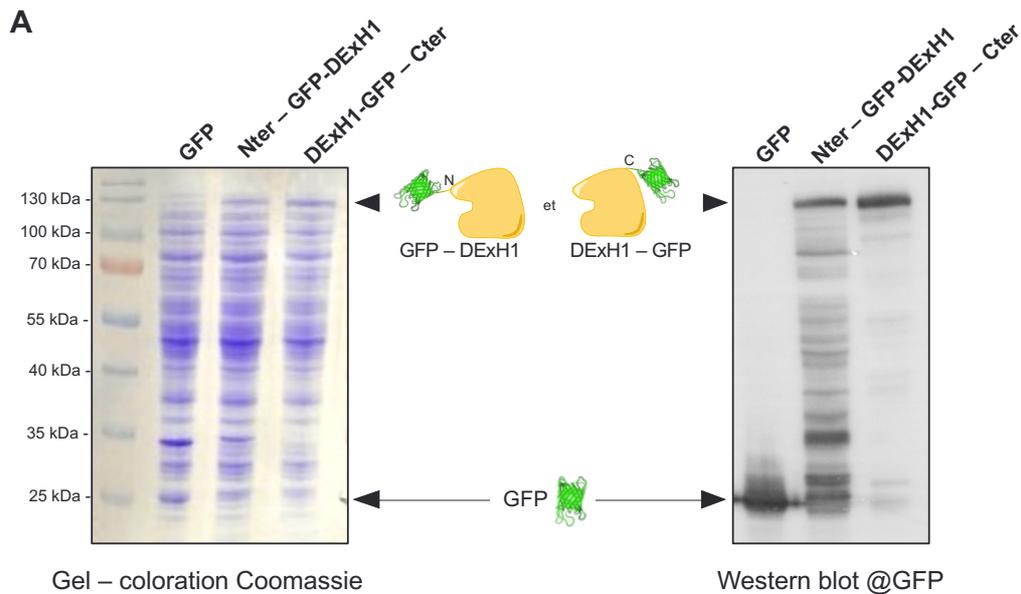


Figure 33 : Production de protéines de fusion DExH1-GFP pour RNA Immunoprecipitation (RIP). A : Les mélanges réactionnels de transcription / traduction du système ALiCE des constructions DExH1-GFP ont été déposés sur gel SDS-PAGE 15%. Le gel a été coloré au bleu de Coomassie (A gauche). Western blot avec un anticorps dirigé contre la GFP (à droite). Les bandes correspondantes aux protéines de fusion sont indiquées par un triangle, celle de la GFP est indiquée par une flèche. B : Schéma théorique du RIP à réaliser. Selon notre modèle, le tRF Ala 5'GGGG est censé interagir avec la protéine de fusion et être présent dans l'éluat (partie gauche du schéma). Le tRF Ala modifié 5'GCGU n'est pas censé interagir avec la protéine de fusion, devrait être éliminé par les lavages et retrouvé absent dans l'éluat (partie droite).

4. Production *in vitro* de protéines de fusion DExH1-GFP

L'interaction entre le tRF Ala et DExH1 a été identifiée en transfectant des protoplastes avec le tRF Ala biotinylé. Après pull-down à la streptavidine, les protéines enrichies ont été identifiées par spectrométrie de masse. L'idée ici est de confirmer cette interaction en utilisant cette fois-ci DExH1 comme appât (*RNA Immunoprecipitation*).

Pour cela, la séquence de DExH1 fusionnée à la GFP en N-terminal et en C-terminal, ainsi la séquence de la GFP (utilisée comme contrôle) ont été clonés dans le vecteur pALiCE® (Lenobio). Basé sur un lysat de cellules de *N. benthamiana*, le système de transcription / traduction ALiCE® permet de produire *in vitro* les protéines d'intérêt. Les résultats préliminaires montrent une forte production des trois constructions souhaitées, à savoir la GFP (contrôle), DExH1 fusionnée à la GFP en N- et C-terminal (Figure 33 A). Toutefois, pour la construction où DExH1 est fusionnée à la GFP en N-terminal, des produits intermédiaires ou de dégradation sont visibles. Une bande correspondant à la GFP seule est aussi apparente. Il est à noter que le codon d'initiation de la traduction AUG est présent au début de la séquence de la GFP et de DExH1, pouvant être probablement à l'origine de protéines de tailles différentes.

Ces extraits protéiques seront ensuite incubés soit avec le tRF Ala (5'-GGGG) soit avec le tRF Ala modifié (5'-GCGU). L'interaction pourra être confirmée par *RNA Immunoprecipitation* grâce à l'utilisation d'anticorps dirigés contre la GFP. La présence des tRFs pourra être évaluée par northern blot ou simplement par migration sur gel en utilisant des tRFs fluorescents (comme les tRFs couplés à l'Alexa 647). Par exemple, le tRF Ala (5'-GGGG) est attendu enrichi avec DExH1-GFP alors que le tRF Ala (5'-GCGU) ne devrait pas être retenu (Figure 33 B).

V. Conclusion

Les études concernant la fonction des tRFs chez les plantes restent encore aujourd'hui parcellaires. Mes travaux de thèse ont consisté à explorer la fonction du tRF-5D Ala chez *A. thaliana*. Étant le plus abondant des petits tRFs chez cet organisme, ce tRF en particulier a été choisi comme modèle. De plus, le tRF Ala n'est pas retrouvé associé à AGO1 ce qui suggère qu'il aurait une fonction indépendante du mécanisme de *silencing*. Il a été montré que ce tRF est capable de réguler la traduction en système *in vitro* [151].

Ainsi, pour comprendre la fonction de ce tRF et les mécanismes moléculaires, sa localisation subcellulaire ainsi que ses partenaires protéiques ont été étudiés par la technique de transfection de protoplastes. Les résultats de localisation obtenus montrent que le tRF Ala s'accumule en *foci* cytoplasmiques. Par comparaison, d'autres tRFs modèles testés, les tRFs His, Gly, Asp et Arg, ne s'accumulent pas en *foci*. Cette observation est donc spécifique du tRF Ala dans ces conditions. La nature de ces *foci* n'a pas pu être déterminée pour le moment. Cependant, les expériences de colocalisation montrent que ces *foci* n'ont pas de lien avec les mitochondries et ne sont ni des *P-bodies* ni des granules de stress.

De plus, ces expériences ont permis de montrer que les *foci* du tRF Ala se forment dans des conditions basales puisqu'ils sont présents avant un stress induisant l'expression des granules de stress.

L'étude de l'interactome du tRF Ala a permis de mettre en évidence une interaction du tRF avec l'hélicase DExH1. Elle est retrouvée enrichie uniquement avec le tRF Ala et pas avec les autres tRFs modèles. C'est donc une interaction spécifique. En l'absence de DExH1, les *foci* du tRF Ala sont toujours observés. DExH1 n'a pas donc pas d'impact sur la biogenèse de ces *foci* ni sur leur maintien au cours du temps. Puisque très peu d'informations sur DExH1 sont disponibles, les données de son homologue chez l'humain, DHX36, ont été analysées pour comprendre d'avantage sa fonction. DExH1 n'est pas l'hélicase la plus abondante chez *A. thaliana* et d'après les analyses *in silico*, elle est exprimée à un faible niveau dans l'ensemble des tissus de la plante sans spécificité particulière [226]. Son expression transitoire dans *N. benthamiana* a montré une localisation cytosolique et nucléaire. DExH1 est une hélicase de la famille des DExH-box et possède les motifs RecA1 et RecA2 conservés de cette famille. Elle possède également un motif RGG en N-terminal. D'autres hélicases chez *A. thaliana* possèdent ce motif RGG [208] mais uniquement DExH1 a été retrouvée enrichie significativement avec le tRF Ala. Les études sur DHX36 et la présence du motif RGG suggèrent que DExH1 interagit avec les structures d'ARN G-quadruplex.

Une étude menée par Lyons et al. [138] a montré la capacité du tRF Ala d'*A. thaliana* de former des G-quadruplex tétramériques. N'étant pas le même isodécoteur que celui utilisé dans mes travaux de thèse, ce résultat a été confirmé avec notre tRF Ala. Ces résultats corroborent avec l'identification de DExH1. L'utilisation de tRFs mutés a permis de montrer que les quatre résidus G à l'extrémité 5' du tRF Ala ainsi que le restant de la séquence sont nécessaires pour former des G-quadruplex et pour interagir avec DExH1. Une première hypothèse où le tRF Ala serait stocké et protégé des dégradations au sein de la cellule sous forme de G-quadruplex a été vérifiée. Les premiers résultats montrent que DExH1 n'a pas d'impact sur l'abondance du tRF Ala ni vraisemblablement sur sa stabilité en conditions normales.

Les résultats de mes travaux de thèse ont donc permis d'avancer dans la compréhension de la fonction du tRF Ala chez les plantes. Il reste cependant encore à déterminer le rôle de DExH1 avec le tRF Ala, les conditions et les facteurs régulant l'activité de DExH1, ainsi que la fonction précise du tRF Ala dans la régulation de la traduction.

DISCUSSION & PERSPECTIVES

La biogenèse des tRFs chez *A. thaliana*

Les données antérieures du laboratoire suggèrent que, chez *A. thaliana*, les tRFs sont générés à partir de clivages d'ARNt réalisés par des RNases T2. Son génome code pour 5 RNases T2 : RNS1-5. Seulement RNS1, RNS2 et RNS3 sont fortement exprimées alors que RNS4 est faiblement exprimée et uniquement dans les racines. Aucune expression de RNS5 dans aucun des tissus de la plante n'a été relevé [113]. De par cette constatation, les études ont été essentiellement consacrées aux RNS1, 2 et 3 et s'appuient sur des lignées *KO* pour ces RNS. Seuls les simples mutants *rns1^{KO}*, *rns2^{KO}*, *rns3^{KO}* et la lignée double *rns1rns3^{KO}* étaient disponibles dans les banques de graines d'Arabidopsis. Mais l'utilisation d'une lignée triple *KO* *rns1rns2rns3^{KO}* devenait nécessaire pour approfondir les connaissances sur les RNS et leur implication dans la biogenèse des tRFs.

Mon travail de thèse a permis d'obtenir les lignées *rns1rns2^{KO}*, *rns2rns3^{KO}* et *rns1rns2rns3^{KO}* par la technologie CRISPR/Cas9. Les résultats préliminaires de caractérisation de ces mutants ne montrent pas de phénotype particulier, y compris pour la lignée *rns1rns2rns3^{KO}*, que ce soit en conditions normales de culture ou en cas de stress, ici une carence en phosphate.

Les analyses phénotypiques devront être poursuivies. Des réplicats devront être réalisés afin de confirmer nos résultats préliminaires. D'autres types de stress pourront être réalisés. Par exemple, l'analyse de banques de tRFs a notamment montré que des tRFs spécifiques sont générés lors de stress froid ou stress UV [87]. Il pourrait alors être intéressant d'observer si l'absence des RNS lors d'un tel stress impacte la production de ces tRFs.

Plusieurs études montrent aujourd'hui chez les plantes que les RNase T2 sont impliquées dans la réponse à des attaques de pathogènes fongiques [175,185,227,228]. Chez la tomate (*S. lycopersicum*), la RNase T2 LE chez la tomate (*S. lycopersicum*) est induite lors d'inoculation avec le champignon pathogène *B. cinerea*. Le mutant *KO* pour la RNase LE présente une plus grande sensibilité après inoculation et le champignon se développe d'avantage [227]. Chez *A. thaliana*, RNS1 est aussi retrouvée impliquée dans la défense virale. En effet, lors d'une infection avec le virus MNV1 (*Murine norovirus*), RNS1 est induite et ralentit l'internalisation du virus dans la plante [229]. Aussi, il a été montré dans plusieurs études que globalement les RNases T2 sont induites lors de blessures mécaniques et donc suggéré qu'elles participent à la défense lors d'un tel stress [175]. L'ensemble de ces études suggèrent que les RNS soient impliquées dans la réponse à des stress mécaniques et aux attaques de pathogènes. Ces types de stress pourront être utilisés pour observer le phénotype du mutant *rns1rns2rns3^{KO}* et leurs impacts sur la biogenèse et les profils des tRFs.

La caractérisation des mutants *rns*^{KO} à l'échelle moléculaire (analyse de l'expression des RNS par western blot et de la biogenèse des tRFs par northern blot) a mis en évidence une production de tRFs dans le mutant *rns1rns2rns3*^{KO}, que ce soit des tRFs longs (tRF-5A et -3A) ou des tRFs courts (tRF-5D ou -3T), lors de carence en phosphate. Dans ce mutant, l'absence de RNS1, RNS2 et RNS3 a été confirmée au niveau protéique par western blot. Pour conforter ses résultats, l'absence des RNS dans les mutants *rns*^{KO} devra être vérifiée par RT-qPCR. De plus, la production de tRFs dans ces mutants, et notamment dans le mutant *rns1rns2rns3*^{KO}, pourra être analysée par une approche plus quantitative, par analyse de banques de tRFs (*RNA-seq*) par exemple. Des banques de différents tissus pourront être analysées, comme des banques de fleurs où les RNS sont fortement exprimées, ainsi que des extraits de plantes soumises à divers stress (carence en phosphate, stress UV, stress hydrique, etc.).

Les résultats obtenus suggèrent qu'il y ait une/des autre(s) RNase(s) capable(s) de générer des tRFs dans ces conditions. Jusqu'à présent les études se focalisaient uniquement sur les trois premières RNS (1 à 3), fortement exprimées chez une plante sauvage. Il est à noter que RNS4 était négligée car peu exprimée dans la lignée sauvage. A présent, dans le mutant *rns1rns2rns3*^{KO} il faut peut-être reconsidérer d'avantage RNS4. Étant donné l'absence des trois autres RNS, il peut être envisagé une compensation par RNS4 pour générer des tRFs. Ceci d'autant plus que RNS4 est exprimée dans les racines et qu'ici les observations de tRFs lors de carence en phosphate sont réalisées sur des plantules jeunes (ayant donc une balance parties aériennes / racines quasiment identique). Pour savoir si RNS4 est impliquée dans la biogenèse de ces tRFs, il faudrait vérifier son niveau d'expression (par RT-qPCR et/ou par western blot). L'obtention et l'utilisation d'une lignée quadruple *KO* *rns1rns2rns3rns4*^{KO} pourrait alors être nécessaire. Si les résultats ne montrent pas de différence d'expression de RNS4 entre la lignée sauvage et la lignée *rns1rns2rns3*^{KO} et qu'il y a toujours présence de tRFs dans le mutant *rns1rns2rns3rns4*^{KO}, il faudra alors envisager que d'autres RNases sont à l'origine de ces tRFs dans ces conditions.

L'équipe de Martinez et al. [111] a montré que la RNase III DCL1 était capable de produire des tRFs courts puisque dans les mutants *dcl1*^{KO}, la quantité de tRFs est réduite. D'un autre côté, les travaux de Megel et al. [113] et de Alves et al. [112] montrent en revanche que la production de tRFs n'est pas impactée dans le mutant *dcl1234*^{KO} et que ce sont plutôt les RNases T2 RNS1, RNS2 et RNS3 qui sont à l'origine des tRFs chez *A. thaliana*. Il semblerait donc y avoir deux types de RNase capable de générer des tRFs chez *A. thaliana*. Il peut alors être supposé que dans le triple mutant *rns1rns2rns3*^{KO} ce soit les DCL qui compensent l'absence des RNS et qui génèrent les tRFs. Dans ce cas, il serait intéressant de comparer le niveau d'expression des DCL entre la lignée triple mutante et la lignée sauvage. Il pourrait aussi être intéressant d'utiliser un mutant *KO* pour les RNS et les DCL. Si dans un tel mutant des tRFs sont toujours observés, il faudra alors explorer la piste d'une autre RNase à l'origine de la biogenèse des tRFs.

Maintenant que la lignée triple *rns1rns2rns3^{KO}* a été obtenue, les premiers résultats montrent une production de tRFs indépendante des RNS1, RNS2 et RNS3 dans ces conditions. Ces résultats soulèvent à présent de nouvelles interrogations concernant l'expression des RNases T2 et l'origine de la biogenèse des tRFs. Comment les RNS participent-elles à l'adaptation de la plante à des stress biotiques et abiotiques ? RNS4 compense-t-elle l'activité de clivage des ARNt des autres RNS en leur absence ? Est-ce que les DCL interviennent pour compenser l'absence des RNS ? Y a-t-il d'autres endoribonucléases impliquées dans la biogenèse des tRFs ? Quelles sont les endoribonucléases impliquées lors de stress comme la carence en phosphate ? Autant de questions soulevées qui, lorsqu'elles seront élucidées, apporteront de nouvelles informations sur le rôle des RNases T2 dans la biogenèse des tRFs chez les plantes.

Un pas vers l'étude fonctionnelle du tRF Ala

Longtemps considérés comme de simples produits de dégradation, les tRFs sont depuis quelques années définis comme des ARN non codants aux fonctions variées et essentielles pour les organismes. Que ce soit chez les procaryotes, les archées ou les eucaryotes, de nombreux résultats montrent leur implication dans divers mécanismes moléculaires. Les tRFs participent à la régulation de l'expression des gènes par *silencing*, à la régulation de la traduction et même à la régulation épigénétique. Au-delà de leur implication à l'échelle moléculaire, ils sont aussi acteurs dans de nombreux processus cellulaires et plus généralement conduisent à des phénotypes particuliers. Chez l'humain, en provoquant ou inhibant la traduction de gènes ciblés ou en régulant des processus cellulaires vitaux, les tRFs participent à l'apparition de maladies et de cancers [230–233]. L'identification de leur rôle précis a permis de faire évoluer les techniques de diagnostic et de mettre au point de nouvelles méthodes thérapeutiques pour ces maladies [232,233].

Les fonctions des tRFs chez les plantes commencent à émerger ces dernières années. Quelques études montrent leur implication dans l'adaptation des plantes à des stress biotiques tels que des attaques de pathogène [89,147,149,153] ou abiotiques comme des carences en nutriments ou des variations climatiques [87,88,113,154]. L'enrichissement des connaissances sur les fonctions des tRFs chez les plantes pourra permettre, à terme, de proposer par exemple des mécanismes de lutte plus ciblés contre les pathogènes ou encore de créer des variétés agricoles plus tolérantes aux conditions climatiques à venir.

C'est dans cette perspective que s'inscrivait une partie de mon travail de thèse qui a consisté à étudier la fonction du tRF-5D Ala de 20 nt chez la plante modèle *A. thaliana*.

L'utilisation de la technique de transfection de protoplastes avec des tRFs a permis de mieux comprendre la localisation et les partenaires protéiques du tRF Ala. Ainsi, il a pu être observé que le tRF Ala est localisé dans des *foci* cytosoliques qui ne correspondent pas à des P-bodies et des granules de stress. La recherche de partenaires protéiques a permis d'identifier une interaction du tRF Ala avec l'hélicase DExH1. Cette hélicase appartient à la famille des DExH-box et présente en plus un motif riche en glycine et en arginine en N-terminal. Il s'avère que cette hélicase interagit avec les ARN ayant une structure G-quadruplex. La présence d'une telle structure secondaire sur le tRF Ala a été confirmée *in vitro*. Ce sont les quatre résidus G à l'extrémité 5' qui interagissent entre eux pour former un G-quadruplex tétramérique. Les résultats obtenus suggèrent que ces résidus G sont essentiels mais pas suffisant pour former cette structure. L'ensemble de la séquence du tRF Ala est nécessaire pour former un G-quadruplex. Ces résultats sont concordants avec l'étude de Lyons et al. [138] montrant la capacité du tRF Ala humain de 30 nt à former des G-quadruplex tétramériques.

Les précédents résultats de l'équipe montrent que le tRF Ala est capable d'inhiber la traduction en système *in vitro* [151]. Cette régulation se fait indépendamment des quatre résidus G en 5'. En revanche, le restant de la séquence du tRF est essentiel pour aboutir à une inhibition de traduction. Cela suggère que les résidus G en 5' ne sont peut-être pas impliqués dans le mécanisme d'inhibition de la traduction mais sont essentiels pour assurer d'autres fonctions. Nous avons montré que ces résidus G sont utilisés pour former des G-quadruplex. Étant une structure secondaire stable, ce pourrait être une forme de stockage du tRF Ala. Ils seraient alors protégés de toute dégradation. Au sein de la cellule, les tRF Ala se trouveraient sous deux formes : libre et assemblés en G-quadruplex. L'abondance du tRF libre peut alors être due à la fois à une plus forte activité de clivage de l'ARNt et aussi à une régulation entre sa forme structurée et sa forme libre. Il est alors supposé que DExH1 intervient dans cette régulation pour déstructurer les G-quadruplex et libérer les tRFs. Cette hypothèse peut être soutenue par les résultats montrant que le tRF Ala est le plus abondant chez *A. thaliana* [87,147]. Contrairement à d'autres tRFs présents en plus faible quantité en conditions normales et dont l'abondance augmente ou diminue en cas de stress [87,88,113,154], le tRF Ala serait en permanence abondant dans la cellule. C'est le ratio forme libre / structuré qui changerait selon les conditions. Puisque les techniques de préparation des banques de petits ARN utilisées dans les études [87,147] ne prennent pas en compte la structure du tRF Ala, il n'est pas possible d'évaluer ce ratio, ni dans quelles conditions il fluctue.

Basées sur nos résultats, d'autres hypothèses peuvent être émises. Des mécanismes moléculaires différents peuvent être à l'origine de ces résultats. L'équipe de Lyons et al. a notamment montré chez l'homme que le tRF-5A Ala possédant 5 résidus G en 5', appelé le motif TOG, est généré par l'Angiogénine en cas de stress [234]. Il est capable de s'assembler en G-quadruplex grâce à ce motif TOG [138]. Ils ont montré que le tRF Ala est capable d'inhiber l'initiation de la traduction et que ce processus conduit à la formation de granules de stress [137,235]. La formation de ces granules se fait cependant de manière indépendante à la formation canonique des granules de stress par phosphorylation de eIF2 α . Plus récemment, ils ont montré que le tRF Ala, sous forme libre ou en G-quadruplex, peut interagir avec le facteur d'initiation eIF4G et plus précisément avec le domaine HEAT1 qu'il contient [236]. Cette interaction perturbe la stabilité du complexe d'initiation eIF4F sur la coiffe des ARNm. L'étape de *scanning* de la 5'UTR des ARNm par le complexe ribosomal 48S est alors inhibé.

Certains de ces résultats présentent des ressemblances avec les résultats que nous avons obtenus : le tRF Ala est capable de former des G-quadruplex, d'inhiber la traduction et conduit à la formation de granules. Cependant, dans notre cas, le tRF n'est pas généré en cas de stress et les *foci* dans lesquels est stocké le tRF Ala ne sont pas des granules de stress. De plus, aucune interaction avec le facteur d'initiation eIF4F n'a été retrouvé dans nos analyses d'interactome du tRF Ala.

Les résultats de notre étude apportent des pistes de réflexion quant à la fonction potentielle du tRF Ala chez les plantes. Il faut cependant noter que nos expériences sont basées sur l'utilisation de tRFs synthétiques qui ne possèdent pas les modifications portées par leurs ARNt d'origine. Bien qu'ici, d'après les prédictions, le tRF-5D Ala ne semble pas porter de modifications sur ses 20 premières bases [237], les modifications présentes sur les ARNt peuvent avoir un impact sur la biogenèse des tRFs en favorisant le clivage des ARNt ou au contraire en l'inhibant [238–240]. Il est à supposer que ces modifications peuvent aussi agir sur la structure et la stabilité du tRF. L'ensemble de ces paramètres peuvent avoir une influence sur la fonction du tRF.

Il est aussi à prendre en considération que la capacité du tRF Ala à former des G-quadruplex n'a été vérifiée que par des expériences *in vitro*. La concentration en KCl utilisée *in vitro* pour stabiliser les G-quadruplex ne reflète pas l'osmolarité *in vivo*. Pour vérifier la présence de tRF Ala en G-quadruplex, il pourrait être envisagé de transférer des protoplastes avec un tRF fluorescent, puis de fixer les protoplastes et ensuite réaliser de l'immunofluorescence avec un anticorps dirigé contre les G-quadruplex. Cet anticorps sera capable de détecter tous les G-quadruplex présents dans la cellule (ADN et ARN), cependant, une colocalisation entre le tRF fluorescent et l'anticorps contre les G-quadruplex devrait être attendue et observable.

De plus, l'inhibition de traduction par le tRF Ala [151] a aussi été observée en système *in vitro* en utilisant un transcrite particulier. Les concentrations en tRF et en transcrits ne reflètent pas une réalité biologique. Pour le moment, aucune expérience *in vivo* n'a permis de montrer un effet global du tRF Ala sur la traduction. Il serait intéressant d'observer le profil d'expression des ARNm en lien avec la quantité de leurs protéines après transfection de protoplastes avec le tRF Ala. Ce résultat pourrait nous indiquer quels sont les ARN les plus affectés dans leur traduction par le tRF Ala. La technique de *Short Tandem Target Mimic* (STTM), initialement développée pour inhiber l'action des miARN, et utilisée par les équipes de Martinez et al. [111] et Gu et al. [147] pour diminuer l'action des tRFs, pourrait être utilisée ici pour étudier la fonction du tRF Ala. Le tRF Ala serait comme capturé et rendu non fonctionnel. Le niveau de traduction des ARNm pourra être analysé dans différents tissus de la plante [147,241].

Finalement, nos résultats reposent sur la technique de transfection de protoplastes, système quelque peu éloigné d'un organisme entier. De plus, les protoplastes utilisés lors de ces expériences sont issus d'un tissu particulier, dans des conditions de cultures bien définies. Bien que ces observations soient un premier pas vers la compréhension de la fonction du tRF Ala, ils devront être confirmés à l'échelle de l'organisme entier. Des variations seront peut-être observées selon le tissu étudié et selon les conditions de culture (en cas de stress notamment).

Les résultats obtenus lors de ma thèse participent à la compréhension de la fonction du tRF Ala mais soulèvent aussi maintenant de nombreuses questions.

Une interaction entre le tRF Ala et DExH1 a été identifiée mais à ce jour les mécanismes de cette interaction ne sont pas encore définis. DExH1 est très peu décrite dans la littérature. Il serait intéressant de la caractériser et d'étudier ses domaines spécifiques d'interaction avec le tRF Ala et les G-quadruplex en général. L'utilisation de constructions où le motif RGG, supposé interagir avec les G-quadruplex, est absent permettrait de confirmer la fonction de ce motif chez DExH1. De plus, les conditions de régulation de DExH1 ne sont pas définies. La technique de transfection de protoplastes, mais cette fois-ci avec une étape de *crosslinking* permettant de fixer les interactions entre protéines, pourrait permettre d'identifier par spectrométrie de masse les partenaires de DExH1 lors de l'interaction avec le tRF Ala. L'identification de ces partenaires pourrait donner à la fois des informations quant à la fonction de DExH1 mais aussi à sa régulation par des facteurs protéiques. Puisque peu de données existent à l'heure actuelle sur DExH1, son profil d'expression pourra aussi être étudié, en particulier lors de stress, en comparant par exemple le mutant *dexh1^{KO}* à la lignée sauvage Col-0.

Finalement, le but sera à terme de connaître la fonction du tRF Ala. Les premières pistes de réflexions vont en faveur d'une régulation de la traduction. Selon notre hypothèse, le tRF Ala libre permettrait de réguler la traduction. Le tRF Ala structuré en G-quadruplex serait une forme de stockage inactif. Les expériences préliminaires de traduction *in vitro* avec le tRF Ala libre ou le tRF inactif en présence de KCl ne sont pas concluantes. Pour s'affranchir de l'utilisation de sels, il pourra être envisagé d'utiliser le tRF Ala avec des modifications, par exemple la 7-deazaguanosine [242], sur les résidus G en 5' qui empêchent la formation de G-quadruplex sans perturber la séquence du tRF Ala. Notre hypothèse, où le tRF Ala libre inhibe la traduction et le tRF Ala structuré, inactif, n'a plus d'effet sur la traduction, pourra alors être confirmée.

En prenant en compte toutes ces données, de nombreuses questions quant à la fonction du tRF Ala restent encore en suspens : dans quel type de foci est stocké le tRF et comment ce stockage est-il régulé ? Où et quand l'interaction entre DExH1 et le tRF Ala a-t-elle lieu ? Dans quelles conditions le tRF Ala régule-t-il la traduction ? Peut-il avoir une autre fonction ? Les réponses à ces questions seront un pas de plus vers l'étude fonctionnelle du tRF Ala chez la plante modèle *A. thaliana*.

Plus globalement, les fonctions des tRFs chez les plantes restent encore à être élucidées. Le tRF Ala a été choisi ici comme modèle parmi la population de tRFs présents chez *A. thaliana* mais d'autres tRFs pourraient faire l'objet d'études plus approfondies. Une longue histoire reste encore à être écrite afin de comprendre et d'identifier les différents rôles des tRFs dans le règne végétal.

BIBLIOGRAPHIE

- [1] M.B. Hoagland, M.L. Stephenson, J.F. Scott, L.I. Hecht, P.C. Zamecnik, A soluble ribonucleic acid intermediate in protein synthesis, *Journal of Biological Chemistry*. 231 (1958) 241–257. [https://doi.org/10.1016/S0021-9258\(19\)77302-5](https://doi.org/10.1016/S0021-9258(19)77302-5).
- [2] J.M. Goodenbour, T. Pan, Diversity of tRNA genes in eukaryotes, *Nucleic Acids Research*. 34 (2006). <https://doi.org/10.1093/nar/gkl725>.
- [3] R.W. Holley, J. Apgar, G.A. Everett, J.T. Madison, M. Marquisee, S.H. Merrill, J.R. Penswick, A. Zamir, Structure of a ribonucleic acid, *Science* (1979). 147 (1965). <https://doi.org/10.1126/science.147.3664.1462>.
- [4] K. Hamashima, M. Tomita, A. Kanai, Expansion of noncanonical V-Arm-containing tRNAs in eukaryotes, *Molecular Biology and Evolution*. 33 (2016). <https://doi.org/10.1093/molbev/msv253>.
- [5] H. Shi, P.B. Moore, The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: A classic structure revisited, *RNA*. 6 (2000) 1091–1105. <https://doi.org/10.1017/S1355838200000364>.
- [6] P.P. Chan, B.Y. Lin, A.J. Mak, T.M. Lowe, TRNAscan-SE 2.0: Improved detection and functional classification of transfer RNA genes, *Nucleic Acids Research*. 49 (2021) 9077–9096. <https://doi.org/10.1093/nar/gkab688>.
- [7] V. Cognat, G. Pawlak, A.M. Duchêne, M. Daujat, A. Gigant, T. Salinas, M. Michaud, B. Gutmann, P. Giegé, A. Gobert, L. Maréchal-Drouard, PlantRNA, a database for tRNAs of photosynthetic eukaryotes, *Nucleic Acids Research*. 41 (2013). <https://doi.org/10.1093/nar/gks935>.
- [8] T. Salinas, A.M. Duchêne, L. Maréchal-Drouard, Recent advances in tRNA mitochondrial import, *Trends in Biochemical Sciences*. 33 (2008). <https://doi.org/10.1016/j.tibs.2008.04.010>.
- [9] A.-M. Duchêne, L. Maréchal-Drouard, The Chloroplast-Derived trnW and trnM-e Genes Are Not Expressed in Arabidopsis Mitochondria, *Biochemical and Biophysical Research Communications*. 285 (2001) 1213–1216. <https://doi.org/10.1006/bbrc.2001.5303>.
- [10] L. Maréchal-Drouard, J.H. Weil, A. Dietrich, Transfer RNAs and transfer RNA genes in plants, *Annual Review of Plant Physiology and Plant Molecular Biology*. 44 (1993). <https://doi.org/10.1146/annurev.pp.44.060193.000305>.
- [11] V. Cognat, G. Pawlak, D. Pflieger, L. Drouard, PlantRNA 2.0: an updated database dedicated to tRNAs of photosynthetic eukaryotes, *BioRxiv*. (2021) 2021.12.21.473619. <https://doi.org/10.1101/2021.12.21.473619>.
- [12] A.-M. Duchêne, C. Pujol, L. Maréchal-Drouard, Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria, *Current Genetics*. 55 (2009) 1–18. <https://doi.org/10.1007/s00294-008-0223-9>.

- [13] M. Michaud, V. Cognat, A.M. Duchêne, L. Maréchal-Drouard, A global picture of tRNA genes in plant genomes, *Plant Journal*. 66 (2011) 80–93. <https://doi.org/10.1111/j.1365-313X.2011.04490.x>.
- [14] G. Galli, H. Hofstetter, M.L. Birnstiel, Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements, *Nature*. 294 (1981). <https://doi.org/10.1038/294626a0>.
- [15] H. Hofstetter, A. Kressmann, M.L. Birnstiel, A split promoter for a eucaryotic tRNA gene, *Cell*. 24 (1981). [https://doi.org/10.1016/0092-8674\(81\)90348-2](https://doi.org/10.1016/0092-8674(81)90348-2).
- [16] S. Sharp, D. DeFranco, T. Dingermann, P. Farrell, D. Söll, Internal control regions for transcription of eukaryotic tRNA genes, *Proc Natl Acad Sci U S A*. 78 (1981). <https://doi.org/10.1073/pnas.78.11.6657>.
- [17] Y. Yukawa, M. Sugita, N. Choise, I. Small, M. Sugiura, The TATA motif, the CAA motif and the poly(T) transcription termination motif are all important for transcription re-initiation on plant tRNA genes, *Plant Journal*. 22 (2000). <https://doi.org/10.1046/j.1365-313X.2000.00752.x>.
- [18] M.D. Berg, C.J. Brandl, Transfer RNAs: diversity in form and function, *RNA Biology*. 18 (2021) 316–339. <https://doi.org/10.1080/15476286.2020.1809197>.
- [19] R. Weinmann, R.G. Roeder, Role of DNA dependent RNA polymerase III in the transcription of the tRNA and 5S RNA genes, *Proc Natl Acad Sci U S A*. 71 (1974). <https://doi.org/10.1073/pnas.71.5.1790>.
- [20] O.S. Gabrielsen, A. Sentenac, RNA polymerase III (C) and its transcription factors, *Trends in Biochemical Sciences*. 16 (1991). [https://doi.org/10.1016/0968-0004\(91\)90166-S](https://doi.org/10.1016/0968-0004(91)90166-S).
- [21] N.H. Blewett, R.J. Maraia, La involvement in tRNA and other RNA processing events including differences among yeast and other eukaryotes, *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. 1861 (2018) 361–372. <https://doi.org/10.1016/j.bbagr.2018.01.013>.
- [22] K. Liere, A. Weihe, T. Börner, The transcription machineries of plant mitochondria and chloroplasts: Composition, function, and regulation, *Journal of Plant Physiology*. 168 (2011). <https://doi.org/10.1016/j.jplph.2011.01.005>.
- [23] B.C. Stark, R. Kole, E.J. Bowman, S. Altman, Ribonuclease P: An enzyme with an essential RNA component, *Proc Natl Acad Sci U S A*. 75 (1978). <https://doi.org/10.1073/pnas.75.8.3717>.
- [24] H.D. Robertson, S. Altman, J.D. Smith, Purification and properties of a specific *Escherichia coli* ribonuclease which cleaves a tyrosine transfer ribonucleic acid precursor., *Journal of Biological Chemistry*. 247 (1972). [https://doi.org/10.1016/s0021-9258\(19\)44963-6](https://doi.org/10.1016/s0021-9258(19)44963-6).
- [25] C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme, *Cell*. 35 (1983). [https://doi.org/10.1016/0092-8674\(83\)90117-4](https://doi.org/10.1016/0092-8674(83)90117-4).

- [26] N.J. Reiter, A. Osterman, A. Torres-Larios, K.K. Swinger, T. Pan, A. Mondragón, Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA, *Nature*. 468 (2010). <https://doi.org/10.1038/nature09516>.
- [27] J. Holzmann, P. Frank, E. Löffler, K.L. Bennett, C. Gerner, W. Rossmann, RNase P without RNA: Identification and Functional Reconstitution of the Human Mitochondrial tRNA Processing Enzyme, *Cell*. 135 (2008). <https://doi.org/10.1016/j.cell.2008.09.013>.
- [28] A. Gobert, B. Gutmann, A. Taschner, M. Göringer, J. Holzmann, R.K. Hartmann, W. Rossmann, P. Giegé, A single Arabidopsis organellar protein has RNase P activity, *Nature Structural and Molecular Biology*. 17 (2010). <https://doi.org/10.1038/nsmb.1812>.
- [29] B. Gutmann, A. Gobert, P. Giegé, PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis, *Genes and Development*. 26 (2012). <https://doi.org/10.1101/gad.189514.112>.
- [30] A. Vogel, O. Schilling, B. Späth, A. Marchfelder, The tRNase Z family of proteins: physiological functions, substrate specificity and structural properties, *Biological Chemistry*. 386 (2005) 1253–1264. <https://doi.org/10.1515/BC.2005.142>.
- [31] G. Canino, E. Bocian, N. Barbezier, M. Echeverría, J. Forner, S. Binder, A. Marchfelder, Arabidopsis encodes four tRNase Z enzymes, *Plant Physiology*. 150 (2009). <https://doi.org/10.1104/pp.109.137950>.
- [32] C. Marck, H. Grosjean, tRNomics: Analysis of tRNA genes from 50 genomes of eukarya, archaea, and bacteria reveals anticodon-sparing strategies and domain-specific features, *RNA*. 8 (2002). <https://doi.org/10.1017/S1355838202022021>.
- [33] S. Vörtler, M. Mörl, tRNA-nucleotidyltransferases: Highly unusual RNA polymerases with vital functions, *FEBS Letters*. 584 (2010). <https://doi.org/10.1016/j.febslet.2009.10.078>.
- [34] S. Schmidt Von Braun, A. Sabetti, P.J. Hanic-Joyce, J. Gu, E. Schleiff, P.B.M. Joyce, Dual targeting of the tRNA nucleotidyltransferase in plants: Not just the signal, *Journal of Experimental Botany*. 58 (2007). <https://doi.org/10.1093/jxb/erm267>.
- [35] P.F. Johnson, J. Abelson, The yeast tRNATyr gene intron is essential for correct modification of its tRNA product, *Nature*. 302 (1983) 681–687. <https://doi.org/10.1038/302681a0>.
- [36] A.K. Hopper, Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*, *Genetics*. 194 (2013). <https://doi.org/10.1534/genetics.112.147470>.
- [37] C.R. Trotta, S. v. Paushkin, M. Patel, H. Li, S.W. Peltz, Cleavage of pre-tRNAs by the splicing endonuclease requires a composite active site, *Nature*. 441 (2006). <https://doi.org/10.1038/nature04741>.
- [38] E.M. Phizicky, R.C. Schwartz, J. Abelson, *Saccharomyces cerevisiae* tRNA ligase. Purification of the protein and isolation of the structural gene, *Journal of Biological Chemistry*. 261 (1986). [https://doi.org/10.1016/s0021-9258\(17\)35882-9](https://doi.org/10.1016/s0021-9258(17)35882-9).
- [39] E.M. de Robertis, P. Black, K. Nishikura, Intranuclear location of the tRNA splicing enzymes, *Cell*. 23 (1981). [https://doi.org/10.1016/0092-8674\(81\)90273-7](https://doi.org/10.1016/0092-8674(81)90273-7).

- [40] T. Yoshihisa, K. Yunoki-Esaki, C. Ohshima, N. Tanaka, T. Endo, Possibility of cytoplasmic pre-tRNA splicing: The yeast tRNA splicing endonuclease mainly localizes on the mitochondria, *Molecular Biology of the Cell*. 14 (2003). <https://doi.org/10.1091/mbc.E02-11-0757>.
- [41] M. Englert, H. Beier, Plant tRNA ligases are multifunctional enzymes that have diverged in sequence and substrate specificity from RNA ligases of other phylogenetic origins, *Nucleic Acids Research*. 33 (2005). <https://doi.org/10.1093/nar/gki174>.
- [42] A.F. de Longevialle, L. Hendrickson, N.L. Taylor, E. Delannoy, C. Lurin, M. Badger, A.H. Millar, I. Small, The pentatricopeptide repeat gene OTP51 with two LAGLIDADG motifs is required for the cis-splicing of plastid ycf3 intron 2 in *Arabidopsis thaliana*, *Plant Journal*. 56 (2008). <https://doi.org/10.1111/j.1365-313X.2008.03581.x>.
- [43] A. Khrouchtchova, R.A. Monde, A. Barkan, A short PPR protein required for the splicing of specific group II introns in angiosperm chloroplasts, *RNA*. 18 (2012). <https://doi.org/10.1261/rna.032623.112>.
- [44] S.D. Nguyen, H. Kang, Comprehensive analysis of chloroplast intron-containing genes and conserved splice sites in dicot and monocot plants, *Science and Technology Development Journal - Natural Sciences*. 1 (2017). <https://doi.org/10.32508/stdjns.v1it1.435>.
- [45] W.A. Cantara, P.F. Crain, J. Rozenski, J.A. McCloskey, K.A. Harris, X. Zhang, F.A.P. Vendeix, D. Fabris, P.F. Agris, The RNA modification database, RNAMDB: 2011 update, *Nucleic Acids Research*. 39 (2011). <https://doi.org/10.1093/nar/gkq1028>.
- [46] P. Boccaletto, F. Stefaniak, A. Ray, A. Cappannini, S. Mukherjee, E. Purta, M. Kurkowska, N. Shirvanizadeh, E. Destefanis, P. Groza, G. Avşar, A. Romitelli, P. Pir, E. Dassi, S.G. Conticello, F. Aguiló, J.M. Bujnicki, MODOMICS: a database of RNA modification pathways. 2021 update, *Nucleic Acids Research*. 50 (2022) D231–D235. <https://doi.org/10.1093/nar/gkab1083>.
- [47] Y. Yoluç, G. Ammann, P. Barraud, M. Jora, P.A. Limbach, Y. Motorin, V. Marchand, C. Tisné, K. Borland, S. Kellner, Instrumental analysis of RNA modifications, *Critical Reviews in Biochemistry and Molecular Biology*. 56 (2021). <https://doi.org/10.1080/10409238.2021.1887807>.
- [48] A. Leger, P.P. Amaral, L. Pandolfini, C. Capitanichik, F. Capraro, V. Miano, V. Migliori, P. Toolan-Kerr, T. Sideri, A.J. Enright, K. Tzelepis, F.J. van Werven, N.M. Luscombe, I. Barbieri, J. Ule, T. Fitzgerald, E. Birney, T. Leonardi, T. Kouzarides, RNA modifications detection by comparative Nanopore direct RNA sequencing, *Nature Communications*. 12 (2021). <https://doi.org/10.1038/s41467-021-27393-3>.
- [49] Y. Motorin, M. Helm, Methods for RNA modification mapping using deep sequencing: Established and new emerging technologies, *Genes (Basel)*. 10 (2019). <https://doi.org/10.3390/genes10010035>.
- [50] E.M. Phizicky, A.K. Hopper, tRNA biology charges to the front, *Genes and Development*. 24 (2010). <https://doi.org/10.1101/gad.1956510>.

- [51] H. Hori, Methylated nucleosides in tRNA and tRNA methyltransferases, *Frontiers in Genetics*. 5 (2014). <https://doi.org/10.3389/fgene.2014.00144>.
- [52] E.M. Phizicky, J.D. Alfonzo, Do all modifications benefit all tRNAs?, *FEBS Letters*. 584 (2010) 265–271. <https://doi.org/10.1016/j.febslet.2009.11.049>.
- [53] P.F. Agris, E.R. Eruysal, A. Narendran, V.Y.P. Väre, S. Vangaveti, S. v. Ranganathan, Celebrating wobble decoding: Half a century and still much is new, *RNA Biology*. 15 (2018). <https://doi.org/10.1080/15476286.2017.1356562>.
- [54] R. Giege, M. Sissler, C. Florentz, Universal rules and idiosyncratic features in tRNA identity, *Nucleic Acids Research*. 26 (1998) 5017–5035. <https://doi.org/10.1093/nar/26.22.5017>.
- [55] C.T.Y. Chan, M. Dyavaiah, M.S. DeMott, K. Taghizadeh, P.C. Dedon, T.J. Begley, A Quantitative Systems Approach Reveals Dynamic Control of tRNA Modifications during Cellular Stress, *PLoS Genetics*. 6 (2010) e1001247. <https://doi.org/10.1371/journal.pgen.1001247>.
- [56] C. Megel, G. Morelle, S. Lalande, A.M. Duchêne, I. Small, L. Maréchal-Drouard, Surveillance and cleavage of eukaryotic tRNAs, *International Journal of Molecular Sciences*. 16 (2015) 1873–1893. <https://doi.org/10.3390/ijms16011873>.
- [57] S. Kadaba, X. Wang, J.T. Anderson, Nuclear RNA surveillance in *Saccharomyces cerevisiae*: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA, *RNA*. 12 (2006). <https://doi.org/10.1261/rna.2305406>.
- [58] J.M. Dewe, J.M. Whipple, I. Chernyakov, L.N. Jaramillo, E.M. Phizicky, The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications, *RNA*. 18 (2012) 1886–1896. <https://doi.org/10.1261/rna.033654.112>.
- [59] L. Kotelawala, E.J. Grayhack, E.M. Phizicky, Identification of yeast tRNA Um44 2'-O-methyltransferase (Trm44) and demonstration of a Trm44 role in sustaining levels of specific tRNA^{Ser} species, *RNA*. 14 (2008). <https://doi.org/10.1261/rna.811008>.
- [60] H.H. Shaheen, A.K. Hopper, Retrograde movement of tRNAs from the cytoplasm to the nucleus in *Saccharomyces cerevisiae*, *Proc Natl Acad Sci U S A*. 102 (2005). <https://doi.org/10.1073/pnas.0503836102>.
- [61] J.E. Wilusz, J.M. Whipple, E.M. Phizicky, P.A. Sharp, tRNAs marked with CCACCA are targeted for degradation, *Science* (1979). 334 (2011). <https://doi.org/10.1126/science.1213671>.
- [62] M. Ibba, D. Soll, Aminoacyl-tRNA synthesis, *Annual Review of Biochemistry*. 69 (2000) 617–650. <https://doi.org/10.1146/annurev.biochem.69.1.617>.
- [63] A.R. Fersht, C. Dingwall, Evidence for the Double-Sieve Editing Mechanism in Protein Synthesis. Steric Exclusion of Isoleucine by Valyl-tRNA Synthetases, *Biochemistry*. 18 (1979). <https://doi.org/10.1021/bi00579a030>.

- [64] V. Rajendran, P. Kalita, H. Shukla, A. Kumar, T. Tripathi, Aminoacyl-tRNA synthetases: Structure, function, and drug discovery, *International Journal of Biological Macromolecules*. 111 (2018). <https://doi.org/10.1016/j.ijbiomac.2017.12.157>.
- [65] H. Himeno, T. Hasegawa, T. Ueda, K. Watanabe, K.I. Miura, M. Shimizu, Role of the extra G-C pair at the end of the acceptor stem of tRNA^{Hb} in aminoacylation, *Nucleic Acids Research*. 17 (1989). <https://doi.org/10.1093/nar/17.19.7855>.
- [66] A.M. Duchêne, A. Giritch, B. Hoffmann, V. Cognat, D. Lancelin, N.M. Peeters, M. Zaepfel, L. Maréchal-Drouard, I.D. Small, Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*, *Proc Natl Acad Sci U S A*. 102 (2005). <https://doi.org/10.1073/pnas.0504682102>.
- [67] J.W.B. Hershey, N. Sonenberg, M.B. Mathews, *Principles of translational control*, Cold Spring Harbor Perspectives in Biology. 11 (2019). <https://doi.org/10.1101/cshperspect.a032607>.
- [68] W.C. Merrick, G.D. Pavitt, Protein synthesis initiation in eukaryotic cells, *Cold Spring Harbor Perspectives in Biology*. 10 (2018). <https://doi.org/10.1101/cshperspect.a033092>.
- [69] T.E. Dever, J.D. Dinman, R. Green, Translation elongation and recoding in eukaryotes, *Cold Spring Harbor Perspectives in Biology*. 10 (2018). <https://doi.org/10.1101/cshperspect.a032649>.
- [70] C.U.T. Hellen, Translation termination and ribosome recycling in eukaryotes, *Cold Spring Harbor Perspectives in Biology*. 10 (2018). <https://doi.org/10.1101/cshperspect.a032656>.
- [71] K.S. Browning, J. Bailey-Serres, Mechanism of Cytoplasmic mRNA Translation, *The Arabidopsis Book*. 13 (2015). <https://doi.org/10.1199/tab.0176>.
- [72] F. Waltz, T.T. Nguyen, M. Arrivé, A. Bochler, J. Chicher, P. Hammann, L. Kuhn, M. Quadrado, H. Mireau, Y. Hashem, P. Giegé, Small is big in *Arabidopsis* mitochondrial ribosome, *Nature Plants*. 5 (2019) 106–117. <https://doi.org/10.1038/s41477-018-0339-y>.
- [73] T. Salinas, C. Remacle, L. Maréchal-Drouard, Mitochondrial Translation in Green Algae and Higher Plants, in: *Translation in Mitochondria and Other Organelles*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013: pp. 181–205. https://doi.org/10.1007/978-3-642-39426-3_8.
- [74] R. Zoschke, R. Bock, Chloroplast translation: Structural and functional organization, operational control, and regulation, *Plant Cell*. 30 (2018) 745–770. <https://doi.org/10.1105/tpc.18.00016>.
- [75] A.G. Hinnebusch, Translational regulation of GCN4 and the general amino acid control of yeast, *Annual Review of Microbiology*. 59 (2005). <https://doi.org/10.1146/annurev.micro.59.031805.133833>.
- [76] I. Avcilar-Kucukgoze, A. Kashina, Hijacking tRNAs From Translation: Regulatory Functions of tRNAs in Mammalian Cell Physiology, *Frontiers in Molecular Biosciences*. 7 (2020). <https://doi.org/10.3389/fmolb.2020.610617>.

- [77] M. Raina, M. Ibba, tRNAs as regulators of biological processes, *Frontiers in Genetics*. 5 (2014). <https://doi.org/10.3389/fgene.2014.00171>.
- [78] A. Mogk, R. Schmidt, B. Bukau, The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies, *Trends in Cell Biology*. 17 (2007). <https://doi.org/10.1016/j.tcb.2007.02.001>.
- [79] C. Schuessele, S.N.W. Hoernstein, S.J. Mueller, M. Rodriguez-Franco, T. Lorenz, D. Lang, G.L. Igloi, R. Reski, Spatio-temporal patterning of arginyl-tRNA protein transferase (ATE) contributes to gametophytic development in a moss, *New Phytologist*. 209 (2016) 1014–1027. <https://doi.org/10.1111/nph.13656>.
- [80] K. Nishimura, Y. Kato, W. Sakamoto, Chloroplast proteases: Updates on proteolysis within and across suborganellar compartments, *Plant Physiology*. 171 (2016). <https://doi.org/10.1104/pp.16.00330>.
- [81] D.J. Gibbs, J. Bacardit, A. Bachmair, M.J. Holdsworth, The eukaryotic N-end rule pathway: Conserved mechanisms and diverse functions, *Trends in Cell Biology*. 24 (2014). <https://doi.org/10.1016/j.tcb.2014.05.001>.
- [82] V. Van, A.T. Smith, ATE1-Mediated Post-Translational Arginylation Is an Essential Regulator of Eukaryotic Cellular Homeostasis, *ACS Chemical Biology*. 15 (2020) 3073–3085. <https://doi.org/10.1021/acscchembio.0c00677>.
- [83] G. Levicán, A. Katz, P. Valenzuela, D. Söll, O. Orellana, A tRNAGlu that uncouples protein and tetrapyrrole biosynthesis, *FEBS Letters*. 579 (2005). <https://doi.org/10.1016/j.febslet.2005.09.100>.
- [84] I. Hwang, H. Sakakibara, Cytokinin biosynthesis and perception, *Physiologia Plantarum*. 126 (2006). <https://doi.org/10.1111/j.1399-3054.2006.00665.x>.
- [85] M. Hrtyan, E. Šliková, J. Hejátko, K. Růžička, RNA processing in auxin and cytokinin pathways, in: *Journal of Experimental Botany*, 2015. <https://doi.org/10.1093/jxb/erv189>.
- [86] F. Borges, R.A. Martienssen, The expanding world of small RNAs in plants, *Nature Reviews Molecular Cell Biology*. 16 (2015) 727–741. <https://doi.org/10.1038/nrm4085>.
- [87] V. Cognat, G. Morelle, C. Megel, S. Lalande, J. Molinier, T. Vincent, I. Small, A.M. Duchene, L. Marechal-Drouard, The nuclear and organellar tRNA-derived RNA fragment population in *Arabidopsis thaliana* is highly dynamic, *Nucleic Acids Research*. 45 (2017) 3460–3472. <https://doi.org/10.1093/nar/gkw1122>.
- [88] S. Zahra, A. Singh, N. Poddar, S. Kumar, Transfer RNA-derived non-coding RNAs (tncRNAs): Hidden regulation of plants' transcriptional regulatory circuits, *Computational and Structural Biotechnology Journal*. 19 (2021) 5278–5291. <https://doi.org/10.1016/j.csbj.2021.09.021>.
- [89] Z. Sun, Y. Hu, Y. Zhou, N. Jiang, S. Hu, L. Li, T. Li, tRNA-derived fragments from wheat are potentially involved in susceptibility to *Fusarium* head blight, *BMC Plant Biology*. 22 (2022). <https://doi.org/10.1186/s12870-021-03393-9>.

- [90] T. Ogawa, K. Tomita, T. Ueda, K. Watanabe, T. Uozumi, H. Masaki, A cytotoxic ribonuclease targeting specific transfer RNA anticodons, *Science* (1979). 283 (1999). <https://doi.org/10.1126/science.283.5410.2097>.
- [91] K. Tomita, T. Ogawa, T. Uozumi, K. Watanabe, H. Masaki, A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops, *Proc Natl Acad Sci U S A*. 97 (2000). <https://doi.org/10.1073/pnas.140213797>.
- [92] I. Morad, D. Chapman-Shimshoni, M. Amitsur, G. Kaufmann, Functional expression and properties of the tRNA(Lys)-specific core anticodon nuclease encoded by *Escherichia coli* prrC, *Journal of Biological Chemistry*. 268 (1993). [https://doi.org/10.1016/s0021-9258\(19\)74188-x](https://doi.org/10.1016/s0021-9258(19)74188-x).
- [93] K. Winther, J.J. Tree, D. Tollervey, K. Gerdes, VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation, *Nucleic Acids Research*. 44 (2016). <https://doi.org/10.1093/nar/gkw781>.
- [94] J.M. Schifano, J.W. Cruz, I.O. Vvedenskaya, R. Edifor, M. Ouyang, R.N. Husson, B.E. Nickels, N.A. Woychik, tRNA is a new target for cleavage by a MazF toxin, *Nucleic Acids Research*. 44 (2016) 1256–1270. <https://doi.org/10.1093/nar/gkv1370>.
- [95] J.M. Lentini, D. Fu, Monitoring the 5-Methoxycarbonylmethyl-2-Thiouridine (mcm5s2U) Modification Utilizing the Gamma-Toxin Endonuclease, in: *Methods in Molecular Biology*, 2021. https://doi.org/10.1007/978-1-0716-1374-0_13.
- [96] J. Lu, B.O. Huang, A. Esberg, M.J.O. Johansson, A.S. Byström, The *Kluyveromyces lactis* γ -toxin targets tRNA anticodons, *RNA*. 11 (2005). <https://doi.org/10.1261/rna.2172105>.
- [97] A.K. Chakravarty, P. Smith, R. Jalan, S. Shuman, Structure, Mechanism, and Specificity of a Eukaryal tRNA Restriction Enzyme Involved in Self-Nonself Discrimination, *Cell Reports*. 7 (2014). <https://doi.org/10.1016/j.celrep.2014.03.034>.
- [98] K.L. Andersen, K. Collins, Several RNase T2 enzymes function in induced tRNA and rRNA turnover in the ciliate *Tetrahymena*, *Molecular Biology of the Cell*. 23 (2012). <https://doi.org/10.1091/mbc.E11-08-0689>.
- [99] D.M. Thompson, R. Parker, The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stress in *Saccharomyces cerevisiae*, *Journal of Cell Biology*. 185 (2009). <https://doi.org/10.1083/jcb.200811119>.
- [100] Y.S. Lee, Y. Shibata, A. Malhotra, A. Dutta, A novel class of small RNAs: tRNA-derived RNA fragments (tRFs), *Genes and Development*. 23 (2009). <https://doi.org/10.1101/gad.1837609>.
- [101] J. Donovan, S. Rath, D. Kolet-Mandrikov, A. Korennykh, Rapid RNase L–driven arrest of protein synthesis in the dsRNA response without degradation of translation machinery, *RNA*. 23 (2017). <https://doi.org/10.1261/rna.062000.117>.
- [102] A.N. Suhasini, R. Sirdeshmukh, Transfer RNA cleavages by onconase reveal unusual cleavage sites, *Journal of Biological Chemistry*. 281 (2006). <https://doi.org/10.1074/jbc.M504488200>.

- [103] J. Sheng, Z. Xu, Three decades of research on angiogenin: A review and perspective, *Acta Biochimica et Biophysica Sinica*. 48 (2016). <https://doi.org/10.1093/abbs/gmv131>.
- [104] Z. Su, C. Kuscu, A. Malik, E. Shibata, A. Dutta, Angiogenin generates specific stress-induced tRNA halves and is not involved in tRF-3-mediated gene silencing, *Journal of Biological Chemistry*. 294 (2019) 16930–16941. <https://doi.org/10.1074/jbc.RA119.009272>.
- [105] S. Li, G.F. Hu, Emerging role of angiogenin in stress response and cell survival under adverse conditions, *Journal of Cellular Physiology*. 227 (2012). <https://doi.org/10.1002/jcp.23051>.
- [106] C. Cole, A. Sobala, C. Lu, S.R. Thatcher, A. Bowman, J.W.S. Brown, P.J. Green, G.J. Barton, G. Hutvagner, Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs, *RNA*. 15 (2009). <https://doi.org/10.1261/rna.1738409>.
- [107] A. Di Fazio, M. Schlackow, S.K. Pong, A. Alagia, M. Gullerova, Dicer dependent tRNA derived small RNAs promote nascent RNA silencing, *Nucleic Acids Research*. 50 (2022) 1734–1752. <https://doi.org/10.1093/nar/gkac022>.
- [108] D. Haussecker, Y. Huang, A. Lau, P. Parameswaran, A.Z. Fire, M.A. Kay, Human tRNA-derived small RNAs in the global regulation of RNA silencing, *RNA*. 16 (2010). <https://doi.org/10.1261/rna.2000810>.
- [109] J.Y. Yang, X.Y. Deng, Y.S. Li, X.C. Ma, J.X. Feng, B. Yu, Y. Chen, Y.L. Luo, X. Wang, M.L. Chen, Z.X. Fang, F.X. Zheng, Y.P. Li, Q. Zhong, T.B. Kang, L.B. Song, R.H. Xu, M.S. Zeng, W. Chen, H. Zhang, W. Xie, S. Gao, Structure of Schlafen13 reveals a new class of tRNA/rRNA-targeting RNase engaged in translational control, *Nature Communications*. 9 (2018). <https://doi.org/10.1038/s41467-018-03544-x>.
- [110] M. Li, E. Kao, D. Malone, X. Gao, J.Y.J. Wang, M. David, DNA damage-induced cell death relies on SLFN11-dependent cleavage of distinct type II tRNAs, *Nature Structural and Molecular Biology*. 25 (2018). <https://doi.org/10.1038/s41594-018-0142-5>.
- [111] G. Martinez, S.G. Choudury, R.K. Slotkin, tRNA-derived small RNAs target transposable element transcripts, *Nucleic Acids Research*. 45 (2017) 5142–5152. <https://doi.org/10.1093/nar/gkx103>.
- [112] C.S. Alves, R. Vicentini, Vitor, F. Pinoti, M. Vincentz, F.T.S. Nogueira, Genome-wide identification and characterization of tRNA-derived RNA fragments in land plants, *Plant Mol Biol*. 93 (2017) 35–48. <https://doi.org/10.1007/s11103-016-0545-9/Published>.
- [113] C. Megel, G. Hummel, S. Lalande, E. Ubrig, V. Cognat, G. Morelle, T. Salinas-Giegé, A.M. Duchêne, L. Maréchal-Drouard, Plant RNases T2, but not Dicer-like proteins, are major players of tRNA-derived fragments biogenesis, *Nucleic Acids Research*. 47 (2019) 941–952. <https://doi.org/10.1093/nar/gky1156>.
- [114] Z. Durdevic, M.B. Mobin, K. Hanna, F. Lyko, M. Schaefer, The RNA methyltransferase dnmt2 is required for efficient dicer-2-dependent siRNA pathway activity in *Drosophila*, *Cell Reports*. 4 (2013). <https://doi.org/10.1016/j.celrep.2013.07.046>.

- [115] S. Blanco, S. Dietmann, J. v Flores, S. Hussain, C. Kutter, P. Humphreys, M. Lukk, P. Lombard, L. Treps, M. Popis, S. Kellner, S.M. Hölter, L. Garrett, W. Wurst, L. Becker, T. Klopstock, H. Fuchs, V. Gailus-Durner, M. Hrabě de Angelis, R.T. Káradóttir, M. Helm, J. Ule, J.G. Gleeson, D.T. Odom, M. Frye, Aberrant methylation of t RNA s links cellular stress to neuro-developmental disorders, *The EMBO Journal*. 33 (2014). <https://doi.org/10.15252/embj.201489282>.
- [116] Y. Wang, D. Li, J. Gao, X. Li, R. Zhang, X. Jin, Z. Hu, B. Zheng, S. Persson, P. Chen, The 2'-O-methyladenosine nucleoside modification gene OsTRM13 positively regulates salt stress tolerance in rice, *Journal of Experimental Botany*. 68 (2017). <https://doi.org/10.1093/jxb/erx061>.
- [117] M. Ghildiyal, P.D. Zamore, Small silencing RNAs: an expanding universe, *Nature Reviews Genetics*. 10 (2009) 94–108. <https://doi.org/10.1038/nrg2504>.
- [118] G. Hutvágner, J. McLachlan, A.E. Pasquinelli, É. Bálint, T. Tuschl, P.D. Zamore, A cellular function for the RNA-interference enzyme dicer in the maturation of the let-7 small temporal RNA, *Science* (1979). 293 (2001). <https://doi.org/10.1126/science.1062961>.
- [119] Y.S. Lee, K. Nakahara, J.W. Pham, K. Kim, Z. He, E.J. Sontheimer, R.W. Carthew, Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways, *Cell*. 117 (2004). [https://doi.org/10.1016/S0092-8674\(04\)00261-2](https://doi.org/10.1016/S0092-8674(04)00261-2).
- [120] B.P. Lewis, I.H. Shih, M.W. Jones-Rhoades, D.P. Bartel, C.B. Burge, Prediction of Mammalian MicroRNA Targets, *Cell*. 115 (2003). [https://doi.org/10.1016/S0092-8674\(03\)01018-3](https://doi.org/10.1016/S0092-8674(03)01018-3).
- [121] S. Yekta, I.H. Shih, D.P. Bartel, MicroRNA-Directed Cleavage of HOXB8 mRNA, *Science* (1979). 304 (2004). <https://doi.org/10.1126/science.1097434>.
- [122] M.W. Rhoades, B.J. Reinhart, L.P. Lim, C.B. Burge, B. Bartel, D.P. Bartel, Prediction of plant microRNA targets, *Cell*. 110 (2002). [https://doi.org/10.1016/S0092-8674\(02\)00863-2](https://doi.org/10.1016/S0092-8674(02)00863-2).
- [123] E. Bonnet, Y. He, K. Billiau, Y. van de Peer, TAPIR, a web server for the prediction of plant microRNA targets, including target mimics, *Bioinformatics*. 26 (2010). <https://doi.org/10.1093/bioinformatics/btq233>.
- [124] C. Llave, K.D. Kasschau, M.A. Rector, J.C. Carrington, Endogenous and silencing-associated small RNAs in plants, *Plant Cell*. 14 (2002). <https://doi.org/10.1105/tpc.003210>.
- [125] G. Tang, B.J. Reinhart, D.P. Bartel, P.D. Zamore, A biochemical framework for RNA silencing in plants, *Genes and Development*. 17 (2003). <https://doi.org/10.1101/gad.1048103>.
- [126] B. Yu, H. Wang, Translational inhibition by microRNAs in plants., *Prog Mol Subcell Biol*. 50 (2010). https://doi.org/10.1007/978-3-642-03103-8_3.
- [127] C. Kuscu, P. Kumar, M. Kiran, Z. Su, A. Malik, A. Dutta, tRNA fragments (tRFs) guide Ago to regulate gene expression post-transcriptionally in a Dicer-independent manner, (2018). <https://doi.org/10.1261/rna.066126>.

- [128] S. Luo, F. He, J. Luo, S. Dou, Y. Wang, A. Guo, J. Lu, *Drosophila* tsRNAs preferentially suppress general translation machinery via antisense pairing and participate in cellular starvation response, *Nucleic Acids Research*. 46 (2018) 5250–5268. <https://doi.org/10.1093/nar/gky189>.
- [129] R.L. Maute, C. Schneider, P. Sumazin, A. Holmes, A. Califano, K. Basso, R. Dalla-Favera, TRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma, *Proc Natl Acad Sci U S A*. 110 (2013) 1404–1409. <https://doi.org/10.1073/pnas.1206761110>.
- [130] I. Olovnikov, K. Chan, R. Sachidanandam, D.K. Newman, A.A. Aravin, Bacterial Argonaute Samples the Transcriptome to Identify Foreign DNA, *Molecular Cell*. 51 (2013). <https://doi.org/10.1016/j.molcel.2013.08.014>.
- [131] M.T. Couvillion, G. Bounova, E. Purdom, T.P. Speed, K. Collins, A Tetrahymena Piwi Bound to Mature tRNA 3' Fragments Activates the Exonuclease Xrn2 for RNA Processing in the Nucleus, *Molecular Cell*. 48 (2012) 509–520. <https://doi.org/10.1016/j.molcel.2012.09.010>.
- [132] S.P. Keam, P.E. Young, A.L. McCorkindale, T.H.Y. Dang, J.L. Clancy, D.T. Humphreys, T. Preiss, G. Hutvagner, D.I.K. Martin, J.E. Cropley, C.M. Suter, The human Piwi protein Hiwi2 associates with tRNA-derived piRNAs in somatic cells, *Nucleic Acids Research*. 42 (2014). <https://doi.org/10.1093/nar/gku620>.
- [133] H.K. Kim, J. Xu, K. Chu, H. Park, H. Jang, P. Li, P.N. Valdmanis, Q.C. Zhang, M.A. Kay, A tRNA-Derived Small RNA Regulates Ribosomal Protein S28 Protein Levels after Translation Initiation in Humans and Mice, *Cell Reports*. 29 (2019) 3816–3824.e4. <https://doi.org/10.1016/j.celrep.2019.11.062>.
- [134] J. Gebetsberger, L. Wyss, A.M. Mleczko, J. Reuther, N. Polacek, A tRNA-derived fragment competes with mRNA for ribosome binding and regulates translation during stress, *RNA Biology*. 14 (2017) 1364–1373. <https://doi.org/10.1080/15476286.2016.1257470>.
- [135] S. Krishna, D.G. Yim, V. Lakshmanan, V. Tirumalai, J.L. Koh, J.E. Park, J.K. Cheong, J.L. Low, M.J. Lim, S.K. Sze, P. Shivaprasad, A. Gulyani, S. Raghavan, D. Palakodeti, R. DasGupta, Dynamic expression of tRNA-derived small RNAs define cellular states, *EMBO Rep*. 20 (2019). <https://doi.org/10.15252/embr.201947789>.
- [136] R. Fricker, R. Brogli, H. Luidalepp, L. Wyss, M. Fasnacht, O. Joss, M. Zywicki, M. Helm, A. Schneider, M. Cristodero, N. Polacek, A tRNA half modulates translation as stress response in *Trypanosoma brucei*, *Nature Communications*. 10 (2019). <https://doi.org/10.1038/s41467-018-07949-6>.
- [137] P. Ivanov, M.M. Emara, J. Villen, S.P. Gygi, P. Anderson, Angiogenin-Induced tRNA Fragments Inhibit Translation Initiation, *Molecular Cell*. 43 (2011) 613–623. <https://doi.org/10.1016/j.molcel.2011.06.022>.

- [138] S.M. Lyons, D. Gudanis, S.M. Coyne, Z. Gdaniec, P. Ivanov, Identification of functional tetramolecular RNA G-quadruplexes derived from transfer RNAs, *Nature Communications*. 8 (2017) 1127. <https://doi.org/10.1038/s41467-017-01278-w>.
- [139] S.M. Lyons, C. Achorn, N.L. Kedersha, P.J. Anderson, P. Ivanov, YB-1 regulates tiRNA-induced Stress Granule formation but not translational repression, *Nucleic Acids Research*. 44 (2016) 6949–6960. <https://doi.org/10.1093/nar/gkw418>.
- [140] N. Guzzi, M. Cieřła, P.C.T. Ngoc, S. Lang, S. Arora, M. Dimitriou, K. Pimková, M.N.E. Sommarin, R. Munita, M. Lubas, Y. Lim, K. Okuyama, S. Soneji, G. Karlsson, J. Hansson, G. Jönsson, A.H. Lund, M. Sigvardsson, E. Hellström-Lindberg, A.C. Hsieh, C. Bellodi, Pseudouridylation of tRNA-Derived Fragments Steers Translational Control in Stem Cells, *Cell*. 173 (2018) 1204-1216.e26. <https://doi.org/10.1016/j.cell.2018.03.008>.
- [141] I. Avcilar-Kucukgoze, H. Gamper, C. Polte, Z. Ignatova, R. Kraetzner, M. Shtutman, Y.-M. Hou, D.W. Dong, A. Kashina, tRNAArg-Derived Fragments Can Serve as Arginine Donors for Protein Arginylation, *Cell Chemical Biology*. 27 (2020) 839-849.e4. <https://doi.org/10.1016/j.chembiol.2020.05.013>.
- [142] A.J. Schorn, M.J. Gutbrod, C. LeBlanc, R. Martienssen, LTR-Retrotransposon Control by tRNA-Derived Small RNAs, *Cell*. 170 (2017). <https://doi.org/10.1016/j.cell.2017.06.013>.
- [143] M.L. Yeung, Y. Bennasser, K. Watashi, S.Y. Le, L. Houzet, K.T. Jeang, Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: Evidence for the processing of a viral-cellular double-stranded RNA hybrid, *Nucleic Acids Research*. 37 (2009). <https://doi.org/10.1093/nar/gkp707>.
- [144] K. Ruggero, A. Guffanti, A. Corradin, V.K. Sharma, G. de Bellis, G. Corti, A. Grassi, P. Zanovello, V. Bronte, V. Ciminale, D.M. D'Agostino, Small Noncoding RNAs in Cells Transformed by Human T-Cell Leukemia Virus Type 1: a Role for a tRNA Fragment as a Primer for Reverse Transcriptase, *Journal of Virology*. 88 (2014). <https://doi.org/10.1128/jvi.02823-13>.
- [145] Q. Chen, M. Yan, Z. Cao, X. Li, Y. Zhang, J. Shi, G.H. Feng, H. Peng, X. Zhang, Y. Zhang, J. Qian, E. Duan, Q. Zhai, Q. Zhou, Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder, *Science* (1979). 351 (2016). <https://doi.org/10.1126/science.aad7977>.
- [146] Y. Zhang, X. Zhang, J. Shi, F. Tuorto, X. Li, Y. Liu, R. Liebers, L. Zhang, Y. Qu, J. Qian, M. Pahima, Y. Liu, M. Yan, Z. Cao, X. Lei, Y. Cao, H. Peng, S. Liu, Y. Wang, H. Zheng, R. Woolsey, D. Quilici, Q. Zhai, L. Li, T. Zhou, W. Yan, F. Lyko, Y. Zhang, Q. Zhou, E. Duan, Q. Chen, Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs, *Nature Cell Biology*. 20 (2018). <https://doi.org/10.1038/s41556-018-0087-2>.
- [147] H. Gu, B. Lian, Y. Yuan, C. Kong, Y. Li, C. Liu, Y. Qi, A 5' tRNA-Ala-derived small RNA regulates anti-fungal defense in plants, *Science China Life Sciences*. 65 (2022). <https://doi.org/10.1007/s11427-021-2017-1>.

- [148] G. Loss-Morais, P.M. Waterhouse, R. Margis, Description of plant tRNA-derived RNA fragments (tRFs) associated with argonaute and identification of their putative targets, *Biology Direct*. 8 (2013). <https://doi.org/10.1186/1745-6150-8-6>.
- [149] B. Ren, X. Wang, J. Duan, J. Ma, Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation, (2019) 1–4. <http://science.sciencemag.org/>.
- [150] A. Trolet, P. Baldrich, M.C. Criqui, M. Dubois, M. Clavel, B.C. Meyers, P. Genschik, Cell cycle-dependent regulation and function of ARGONAUTE1 in plants, *Plant Cell*. 31 (2019) 1734–1750. <https://doi.org/10.1105/tpc.19.00069>.
- [151] S. Lalande, R. Merret, T. Salinas-Giegé, L. Drouard, Arabidopsis tRNA-derived fragments as potential modulators of translation, *RNA Biology*. 17 (2020) 1137–1148. <https://doi.org/10.1080/15476286.2020.1722514>.
- [152] E. Glawischnig, Camalexin, *Phytochemistry*. 68 (2007). <https://doi.org/10.1016/j.phytochem.2006.12.005>.
- [153] B.T. Werner, A. Koch, E. Šečić, J. Engelhardt, L. Jelonek, J. Steinbrenner, K.H. Kogel, *Fusarium graminearum* DICER-like-dependent sRNAs are required for the suppression of host immune genes and full virulence, *PLoS ONE*. 16 (2021). <https://doi.org/10.1371/journal.pone.0252365>.
- [154] W. Luan, Y. Dai, X.-Y. Li, Y. Wang, X. Tao, C.-X. Li, P. Mao, X.-R. Ma, Identification of tRFs and phasiRNAs in tomato (*Solanum lycopersicum*) and their responses to exogenous abscisic acid, *BMC Plant Biology*. 20 (2020) 320. <https://doi.org/10.1186/s12870-020-02528-8>.
- [155] S. Zhang, L. Sun, F. Kragler, The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation1[W][OA], *Plant Physiology*. 150 (2009). <https://doi.org/10.1104/pp.108.134767>.
- [156] A. v. Savelyeva, D.N. Baryakin, E.D. Chikova, E. v. Kuligina, V.A. Richter, D. v. Semenov, Vesicular and extra-vesicular RNAs of human blood plasma, *Advances in Experimental Medicine and Biology*. 924 (2016). https://doi.org/10.1007/978-3-319-42044-8_23.
- [157] F. Gámbaro, M. Li Calzi, P. Fagúndez, B. Costa, G. Greif, E. Mallick, S. Lyons, P. Ivanov, K. Witwer, A. Cayota, J.P. Tosar, Stable tRNA halves can be sorted into extracellular vesicles and delivered to recipient cells in a concentration-dependent manner, *RNA Biology*. 17 (2020) 1168–1182. <https://doi.org/10.1080/15476286.2019.1708548>.
- [158] J.M. Dhahbi, S.R. Spindler, H. Atamna, A. Yamakawa, D. Boffelli, P. Mote, D.I.K. Martin, 5' tRNA halves are present as abundant complexes in serum, concentrated in blood cells, and modulated by aging and calorie restriction, *BMC Genomics*. 14 (2013). <https://doi.org/10.1186/1471-2164-14-298>.
- [159] J.P. Tosar, A. Cayota, Extracellular tRNAs and tRNA-derived fragments, *RNA Biology*. 17 (2020) 1149–1167. <https://doi.org/10.1080/15476286.2020.1729584>.
- [160] K.Y. Cao, T.M. Yan, J.Z. Zhang, T.F. Chan, J. Li, C. Li, E. Lai-Han Leung, J. Gao, B.X. Zhang, Z.H. Jiang, A tRNA-derived fragment from Chinese yew suppresses ovarian

- cancer growth via targeting TRPA1, *Molecular Therapy - Nucleic Acids*. 27 (2022) 718–732. <https://doi.org/10.1016/j.omtn.2021.12.037>.
- [161] R. Merret, M.C. Carpentier, J.J. Favory, C. Picart, J. Descombin, C. Bousquet-Antonelli, P. Tillard, L. Lejay, J.M. Deragon, Y.Y. Charng, Heat shock protein HSP101 affects the release of ribosomal protein mRNAs for recovery after heat shock, *Plant Physiology*. 174 (2017) 1216–1225. <https://doi.org/10.1104/pp.17.00269>.
- [162] R. Merret, J. Descombin, Y. ting Juan, J.J. Favory, M.C. Carpentier, C. Chaparro, Y. yung Charng, J.M. Deragon, C. Bousquet-Antonelli, XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress, *Cell Reports*. 5 (2013) 1279–1293. <https://doi.org/10.1016/j.celrep.2013.11.019>.
- [163] H. Tsutsui, T. Higashiyama, PKAMA-ITACHI vectors for highly efficient CRISPR/Cas9-mediated gene knockout in *Arabidopsis thaliana*, *Plant and Cell Physiology*. 58 (2017) 46–56. <https://doi.org/10.1093/pcp/pcw191>.
- [164] M. Karimi, D. Inzé, A. Depicker, GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation, *Trends in Plant Science*. 7 (2002). [https://doi.org/10.1016/S1360-1385\(02\)02251-3](https://doi.org/10.1016/S1360-1385(02)02251-3).
- [165] M. Haeussler, K. Schönig, H. Eckert, A. Eschstruth, J. Mianné, J.B. Renaud, S. Schneider-Maunoury, A. Shkumatava, L. Teboul, J. Kent, J.S. Joly, J.P. Concordet, Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR, *Genome Biology*. 17 (2016). <https://doi.org/10.1186/s13059-016-1012-2>.
- [166] E. Meijering, M. Jacob, J.C.F. Sarria, P. Steiner, H. Hirling, M. Unser, Design and Validation of a Tool for Neurite Tracing and Analysis in Fluorescence Microscopy Images, *Cytometry Part A*. 58 (2004). <https://doi.org/10.1002/cyto.a.20022>.
- [167] J. Mutterer, E. Zinck, Quick-and-clean article figures with FigureJ, *Journal of Microscopy*. 252 (2013). <https://doi.org/10.1111/jmi.12069>.
- [168] S. Lalande, M. Chery, E. Ubrig, G. Hummel, J. Kubina, A. Geldreich, L. Drouard, Transfection of small noncoding RNAs into *Arabidopsis thaliana* protoplasts, in: *Methods in Molecular Biology*, Humana Press Inc., 2020: pp. 413–429. https://doi.org/10.1007/978-1-0716-0712-1_24.
- [169] L. Kuhn, T. Vincent, P. Hamman, H. Zuber, Exploring Protein Interactome Data with IPinquiry: Statistical Analysis and Data Visualization by Spectral Counts, in: Burger Thomas (Ed.), *Statistical Analysis of Proteomic Data*, *Methods Mol Biol*, 2022. https://doi.org/https://doi.org/10.1007/978-1-0716-1967-4_11.
- [170] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data, *Bioinformatics*. 26 (2009). <https://doi.org/10.1093/bioinformatics/btp616>.
- [171] Y. Zuo, M.P. Deutscher, Exoribonuclease superfamilies: Structural analysis and phylogenetic distribution, *Nucleic Acids Research*. 29 (2001). <https://doi.org/10.1093/nar/29.5.1017>.

- [172] Z. Li, M.P. Deutscher, Exoribonucleases and Endoribonucleases, *EcoSal Plus*. 1 (2004). <https://doi.org/10.1128/ecosalplus.4.6.3>.
- [173] G.R. Marshall, J.A. Feng, D.J. Kuster, Back to the future: Ribonuclease A, *Biopolymers - Peptide Science Section*. 90 (2008) 259–277. <https://doi.org/10.1002/bip.20845>.
- [174] C.N. Pace, U. Heinemann, U. Hahn, W. Saenger, Ribonuclease T1: Structure, Function, and Stability, *Angewandte Chemie International Edition in English*. 30 (1991). <https://doi.org/10.1002/anie.199103433>.
- [175] G.C. MacIntosh, RNase T2 Family: Enzymatic Properties, Functional Diversity, and Evolution of Ancient Ribonucleases, in: 2011: pp. 89–114. https://doi.org/10.1007/978-3-642-21078-5_4.
- [176] M. Irie, Structure-Function Relationships of Acid Ribonucleases, *Pharmacology & Therapeutics*. 81 (1999) 77–89. [https://doi.org/10.1016/S0163-7258\(98\)00035-7](https://doi.org/10.1016/S0163-7258(98)00035-7).
- [177] N. Luhtala, R. Parker, T2 Family ribonucleases: ancient enzymes with diverse roles, *Trends in Biochemical Sciences*. 35 (2010) 253–259. <https://doi.org/10.1016/j.tibs.2010.02.002>.
- [178] M. Irie, H. Watanabe, K. Ohgi, Y. Minami, H. Yamada, G. Gunki, F. Funatsu, Base Specificity of Two Plant Seed Ribonucleases from *Momordica charantia* and *Luffa cylindrica*, *Bioscience, Biotechnology and Biochemistry*. 57 (1993). <https://doi.org/10.1271/bbb.57.497>.
- [179] N. Tanaka, J. Arai, N. Inokuchi, T. Koyama, K. Ohgi, M. Irie, K.T. Nakamura, Crystal structure of a plant ribonuclease, RNase LE, *Journal of Molecular Biology*. 298 (2000). <https://doi.org/10.1006/jmbi.2000.3707>.
- [180] R.A. Deshpande, V. Shankar, Ribonucleases from T2 Family, *Critical Reviews in Microbiology*. 28 (2002) 79–122. <https://doi.org/10.1080/1040-840291046704>.
- [181] T. Uchida, Purification and properties of RNase T2, *Journal of Biochemistry*. 60 (1966). <https://doi.org/10.1093/oxfordjournals.jbchem.a128410>.
- [182] M.S. Hillwig, L. Rizhsky, Y. Wang, A. Umanskaya, J.J. Essner, G.C. MacIntosh, Zebrafish RNase T2 genes and the evolution of secretory ribonucleases in animals, *BMC Evolutionary Biology*. 9 (2009). <https://doi.org/10.1186/1471-2148-9-170>.
- [183] M.S. Hillwig, A.L. Contento, A. Meyer, D. Ebany, D.C. Bassham, G.C. MacIntosha, RNS2, a conserved member of the RNase T2 family, is necessary for ribosomal RNA decay in plants, *Proc Natl Acad Sci U S A*. 108 (2011) 1093–1098. <https://doi.org/10.1073/pnas.1009809108>.
- [184] P.F. Spahr, B.R. Hollingworth, Purification and Mechanism of Action of Ribonuclease from *Escherichia coli* Ribosomes, *Journal of Biological Chemistry*. 236 (1961). [https://doi.org/10.1016/s0021-9258\(18\)64315-7](https://doi.org/10.1016/s0021-9258(18)64315-7).
- [185] G.C. MacIntosh, M.S. Hillwig, A. Meyer, L. Flagel, RNase T2 genes from rice and the evolution of secretory ribonucleases in plants, *Molecular Genetics and Genomics*. 283 (2010) 381–396. <https://doi.org/10.1007/s00438-010-0524-9>.

- [186] C.B. Taylor, P.J. Green, Genes with homology to fungal and S-gene RNases are expressed in *Arabidopsis thaliana*, *Plant Physiology*. 96 (1991) 980–984. <https://doi.org/10.1104/pp.96.3.980>.
- [187] C.B. Taylor, P.A. Bariola, S.B. delCardayré, R.T. Raines, P.J. Green, RNS2: a senescence-associated RNase of *Arabidopsis* that diverged from the S-RNases before speciation., *Proceedings of the National Academy of Sciences*. 90 (1993) 5118–5122. <https://doi.org/10.1073/pnas.90.11.5118>.
- [188] P.A. Bariola, C.J. Howard, C.B. Taylor, M.T. Verburg, V.D. Jaglan, P.J. Green, The *Arabidopsis* ribonuclease gene RNS1 is tightly controlled in response to phosphate limitation, *The Plant Journal*. 6 (1994) 673–685. <https://doi.org/10.1046/j.1365-313X.1994.6050673.x>.
- [189] P.A. Bariola, G.C. MacIntosh, P.J. Green, Regulation of S-like ribonuclease levels in *Arabidopsis*. Antisense inhibition of RNS1 or RNS2 elevates anthocyanin accumulation, *Plant Physiology*. 119 (1999) 331–342. <https://doi.org/10.1104/pp.119.1.331>.
- [190] L.C. Hsieh, S.I. Lin, A.C.C. Shih, J.W. Chen, W.Y. Lin, C.Y. Tseng, W.H. Li, T.J. Chiou, Uncovering small RNA-mediated responses to phosphate deficiency in *Arabidopsis* by deep sequencing, *Plant Physiology*. 151 (2009). <https://doi.org/10.1104/pp.109.147280>.
- [191] B.E. Floyd, Y. Mugume, S.C. Morriss, G.C. MacIntosh, D.C. Bassham, Localization of RNS2 ribonuclease to the vacuole is required for its role in cellular homeostasis, *Planta*. 245 (2017) 779–792. <https://doi.org/10.1007/s00425-016-2644-x>.
- [192] G. Morelle, Les ARN de transfert, une nouvelle source de petits ARN non-codants chez *Arabidopsis thaliana*, 2015.
- [193] T. Gaj, C.A. Gersbach, C.F. Barbas, ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering, *Trends in Biotechnology*. 31 (2013). <https://doi.org/10.1016/j.tibtech.2013.04.004>.
- [194] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity, n.d. <http://science.sciencemag.org/>.
- [195] D. Zhang, Z. Zhang, T. Unver, B. Zhang, CRISPR/Cas: A powerful tool for gene function study and crop improvement, *Journal of Advanced Research*. 29 (2021). <https://doi.org/10.1016/j.jare.2020.10.003>.
- [196] S. Heuer, R. Gaxiola, R. Schilling, L. Herrera-Estrella, D. López-Arredondo, M. Wissuwa, E. Delhaize, H. Rouached, Improving phosphorus use efficiency: a complex trait with emerging opportunities, *The Plant Journal*. 90 (2017) 868–885. <https://doi.org/10.1111/tpj.13423>.
- [197] G. Huang, D. Zhang, The plasticity of root systems in response to external phosphate, *International Journal of Molecular Sciences*. 21 (2020) 1–12. <https://doi.org/10.3390/ijms21175955>.

- [198] F. Chevalier, M. Pata, P. Nacry, P. Doumas, M. Rossignol, Effects of phosphate availability on the root system architecture: Large-scale analysis of the natural variation between *Arabidopsis* accessions, *Plant, Cell and Environment*. 26 (2003) 1839–1850. <https://doi.org/10.1046/j.1365-3040.2003.01100.x>.
- [199] M. Gellert, M.N. Lipsett, D.R. Davies, Helix formation by guanylic acid, *Proc Natl Acad Sci U S A*. 48 (1962). <https://doi.org/10.1073/pnas.48.12.2013>.
- [200] D. Sen, W. Gilbert, A sodium-potassium switch in the formation of four-stranded G4-DNA, *Nature*. 344 (1990). <https://doi.org/10.1038/344410a0>.
- [201] V.T. Mukundan, A.T. Phan, Bulges in G-Quadruplexes: Broadening the Definition of G-Quadruplex-Forming Sequences, *J Am Chem Soc*. 135 (2013) 5017–5028. <https://doi.org/10.1021/ja310251r>.
- [202] D. Bhattacharyya, G.M. Arachchilage, S. Basu, Metal cations in G-quadruplex folding and stability, *Frontiers in Chemistry*. 4 (2016). <https://doi.org/10.3389/fchem.2016.00038>.
- [203] S. Neidle, Quadruplex nucleic acids as targets for anticancer therapeutics, *Nature Reviews Chemistry*. 1 (2017). <https://doi.org/10.1038/s41570-017-0041>.
- [204] L. Dumas, P. Herviou, E. Dassi, A. Cammas, S. Millevoi, G-Quadruplexes in RNA Biology: Recent Advances and Future Directions, *Trends in Biochemical Sciences*. 46 (2021) 270–283. <https://doi.org/10.1016/j.tibs.2020.11.001>.
- [205] S.K. Mishra, A. Tawani, A. Mishra, A. Kumar, G4IPDB: A database for G-quadruplex structure forming nucleic acid interacting proteins, *Scientific Reports*. 6 (2016). <https://doi.org/10.1038/srep38144>.
- [206] V. Brázda, J. Cerveň, M. Bartas, N. Mikysková, J. Coufal, P. Pečinka, The amino acid composition of quadruplex binding proteins reveals a shared motif and predicts new potential quadruplex interactors, *Molecules*. 23 (2018). <https://doi.org/10.3390/molecules23092341>.
- [207] Z.L. Huang, J. Dai, W.H. Luo, X.G. Wang, J.H. Tan, S. bin Chen, Z.S. Huang, Identification of G-Quadruplex-Binding Protein from the Exploration of RGG Motif/G-Quadruplex Interactions, *J Am Chem Soc*. 140 (2018). <https://doi.org/10.1021/jacs.8b09329>.
- [208] A. Volná, M. Bartas, J. Nezval, V. Špunda, P. Pečinka, J. Červeň, Searching for G-quadruplex-binding proteins in plants: New insight into possible G-quadruplex regulation, *BioTech*. 10 (2021). <https://doi.org/10.3390/biotech10040020>.
- [209] K.K.P. Yan, I. Obi, N. Sabouri, The RGG domain in the C-terminus of the DEAD box helicases Dbp2 and Ded1 is necessary for G-quadruplex destabilization, *Nucleic Acids Research*. 49 (2021). <https://doi.org/10.1093/nar/gkab620>.
- [210] T. Masuzawa, T. Oyoshi, Roles of the RGG Domain and RNA Recognition Motif of Nucleolin in G-Quadruplex Stabilization, *ACS Applied Materials and Interfaces*. (2020). <https://doi.org/10.1021/acsomega.9b04221>.

- [211] B. Herdy, C. Mayer, D. Varshney, G. Marsico, P. Murat, C. Taylor, C. D'Santos, D. Tannahill, S. Balasubramanian, Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts, *Nucleic Acids Research*. 46 (2018) 11592–11604. <https://doi.org/10.1093/nar/gky861>.
- [212] M. Caterino, K. Paeschke, Action and function of helicases on RNA G-quadruplexes, *Methods*. (2021). <https://doi.org/10.1016/j.ymeth.2021.09.003>.
- [213] M. Sauer, K. Paeschke, G-quadruplex unwinding helicases and their function in vivo, *Biochemical Society Transactions*. 45 (2017) 1173–1182. <https://doi.org/10.1042/BST20170097>.
- [214] K.A. Hossain, M. Jurkowski, J. Czub, M. Kogut, Mechanism of recognition of parallel G-quadruplexes by DEAH/RHAU helicase DHX36 explored by molecular dynamics simulations, *Computational and Structural Biotechnology Journal*. 19 (2021). <https://doi.org/10.1016/j.csbj.2021.04.039>.
- [215] B. Chang-Gu, D. Bradburn, P.M. Yangyuoru, R. Russell, The DHX36-specific-motif (DSM) enhances specificity by accelerating recruitment of DNA G-quadruplex structures, *Biological Chemistry*. 402 (2021) 593–604. <https://doi.org/10.1515/hsz-2020-0302>.
- [216] M.C. Chen, R. Tippana, N.A. Demeshkina, P. Murat, S. Balasubramanian, S. Myong, A.R. Ferré-D'amaré, Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36, *Nature*. 558 (2018) 465–483. <https://doi.org/10.1038/s41586-018-0209-9>.
- [217] X. Yang, J. Cheema, Y. Zhang, H. Deng, S. Duncan, M.I. Umar, J. Zhao, Q. Liu, X. Cao, C.K. Kwok, Y. Ding, RNA G-quadruplex structures exist and function in vivo in plants, *Genome Biology*. 21 (2020). <https://doi.org/10.1186/s13059-020-02142-9>.
- [218] M.A. Mullen, K.J. Olson, P. Dallaire, F. Major, S.M. Assmann, P.C. Bevilacqua, RNA G-Quadruplexes in the model plant species *Arabidopsis thaliana*: prevalence and possible functional roles, *Nucleic Acids Research*. 38 (2010) 8149–8163. <https://doi.org/10.1093/nar/gkq804>.
- [219] B.D. Griffin, H.W. Bass, Review: Plant G-quadruplex (G4) motifs in DNA and RNA; abundant, intriguing sequences of unknown function, *Plant Science*. 269 (2018) 143–147. <https://doi.org/10.1016/j.plantsci.2018.01.011>.
- [220] H. Cho, H.S. Cho, H. Nam, H. Jo, J. Yoon, C. Park, T.V.T. Dang, E. Kim, J. Jeong, S. Park, E.S. Wallner, H. Youn, J. Park, J. Jeon, H. Ryu, T. Greb, K. Choi, Y. Lee, S.K. Jang, C. Ban, I. Hwang, Translational control of phloem development by RNA G-quadruplex-JULGI determines plant sink strength, *Nature Plants*. 4 (2018) 376–390. <https://doi.org/10.1038/s41477-018-0157-2>.
- [221] C.K. Kwok, Y. Ding, S. Shahid, S.M. Assmann, P.C. Bevilacqua, A stable RNA G-quadruplex within the 5'-UTR of *Arabidopsis thaliana* ATR mRNA inhibits translation, *Biochemical Journal*. 467 (2015) 91–102. <https://doi.org/10.1042/BJ20141063>.
- [222] N.K. Tanner, P. Linder, Review DExD/H Box RNA Helicases: From Generic Motors to Specific Dissociation Functions, 2001.

- [223] F. de Bortoli, S. Espinosa, R. Zhao, DEAH-Box RNA Helicases in Pre-mRNA Splicing, *Trends in Biochemical Sciences*. 46 (2021). <https://doi.org/10.1016/j.tibs.2020.10.006>.
- [224] A.L. Mallam, D.J. Sidote, A.M. Lambowitz, Molecular insights into RNA and DNA helicase evolution from the determinants of specificity for a DEAD-box RNA helicase, *Elife*. 3 (2014). <https://doi.org/10.7554/eLife.04630>.
- [225] C.F. Bourgeois, F. Mortreux, D. Auboeuf, The multiple functions of RNA helicases as drivers and regulators of gene expression, *Nature Reviews Molecular Cell Biology*. 17 (2016) 426–438. <https://doi.org/10.1038/nrm.2016.50>.
- [226] R. Xu, S. Zhang, J. Huang, C. Zheng, Genome-Wide Comparative In Silico Analysis of the RNA Helicase Gene Family in *Zea mays* and *Glycine max*: A Comparison with *Arabidopsis* and *Oryza sativa*, *PLoS ONE*. 8 (2013) e78982. <https://doi.org/10.1371/journal.pone.0078982>.
- [227] N.K. Singh, E. Paz, Y. Kutsher, M. Reuveni, A. Lers, Tomato T2 ribonuclease LE is involved in the response to pathogens, *Molecular Plant Pathology*. 21 (2020) 895–906. <https://doi.org/10.1111/mpp.12928>.
- [228] K. Hugot, M. Ponchet, A. Marais, P. Ricci, E. Galiana, A tobacco S-like RNase inhibits hyphal elongation of plant pathogens, *Molecular Plant-Microbe Interactions*. 15 (2002). <https://doi.org/10.1094/MPMI.2002.15.3.243>.
- [229] K. Kim, D. Yadav, M. Cho, Multi-phased internalization of murine norovirus (MNV) in *Arabidopsis* seedlings and its potential correlation with plant defensive responses, *Microbial Pathogenesis*. 135 (2019) 103648. <https://doi.org/10.1016/j.micpath.2019.103648>.
- [230] Y. Hu, A. Cai, J. Xu, W. Feng, A. Wu, R. Liu, W. Cai, L. Chen, F. Wang, An emerging role of the 5' termini of mature tRNAs in human diseases: Current situation and prospects, *Biochimica et Biophysica Acta - Molecular Basis of Disease*. 1868 (2022). <https://doi.org/10.1016/j.bbadis.2021.166314>.
- [231] Q. Pan, T. Han, G. Li, Novel insights into the roles of tRNA-derived small RNAs, *RNA Biology*. 18 (2021) 2157–2167. <https://doi.org/10.1080/15476286.2021.1922009>.
- [232] X. Yu, Y. Xie, S. Zhang, X. Song, B. Xiao, Z. Yan, tRNA-derived fragments: Mechanisms underlying their regulation of gene expression and potential applications as therapeutic targets in cancers and virus infections, *Theranostics*. 11 (2020) 461–469. <https://doi.org/10.7150/thno.51963>.
- [233] T. Zong, Y. Yang, H. Zhao, L. Li, M. Liu, X. Fu, G. Tang, H. Zhou, L.H.H. Aung, P. Li, J. Wang, Z. Wang, T. Yu, tsRNAs: Novel small molecules from cell function and regulatory mechanism to therapeutic targets, *Cell Proliferation*. 54 (2021). <https://doi.org/10.1111/cpr.12977>.
- [234] S. Yamasaki, P. Ivanov, G.F. Hu, P. Anderson, Angiogenin cleaves tRNA and promotes stress-induced translational repression, *Journal of Cell Biology*. 185 (2009). <https://doi.org/10.1083/jcb.200811106>.

- [235] M.M. Emara, P. Ivanov, T. Hickman, N. Dawra, S. Tisdale, N. Kedersha, G.F. Hu, P. Anderson, Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly, *Journal of Biological Chemistry*. 285 (2010) 10959–10968. <https://doi.org/10.1074/jbc.M109.077560>.
- [236] S.M. Lyons, P. Kharel, Y. Akiyama, S. Ojha, D. Dave, V. Tsvetkov, W. Merrick, P. Ivanov, P. Anderson, eIF4G has intrinsic G-quadruplex binding activity that is required for tRNA function, *Nucleic Acids Research*. 48 (2020) 6223–6233. <https://doi.org/10.1093/nar/gkaa336>.
- [237] X. Ma, F. Si, X. Liu, W. Luan, PRMdb: A Repository of Predicted RNA Modifications in Plants, *Plant and Cell Physiology*. 61 (2020). <https://doi.org/10.1093/pcp/pcaa042>.
- [238] Q. Chen, X. Zhang, J. Shi, M. Yan, T. Zhou, Origins and evolving functionalities of tRNA-derived small RNAs, *Trends in Biochemical Sciences*. 46 (2021) 790–804. <https://doi.org/10.1016/j.tibs.2021.05.001>.
- [239] Z. Su, B. Wilson, P. Kumar, A. Dutta, Noncanonical Roles of tRNAs: TRNA Fragments and beyond, *Annual Review of Genetics*. 54 (2020) 47–69. <https://doi.org/10.1146/annurev-genet-022620-101840>.
- [240] S.M. Lyons, M.M. Fay, P. Ivanov, The role of RNA modifications in the regulation of tRNA cleavage, *FEBS Letters*. 592 (2018). <https://doi.org/10.1002/1873-3468.13205>.
- [241] T. Peng, M. Qiao, H. Liu, S. Teotia, Z. Zhang, Y. Zhao, B. Wang, D. Zhao, L. Shi, C. Zhang, B. Le, K. Rogers, C. Gunasekara, H. Duan, Y. Gu, L. Tian, J. Nie, J. Qi, F. Meng, L. Huang, Q. Chen, Z. Wang, J. Tang, X. Tang, T. Lan, X. Chen, H. Wei, Q. Zhao, G. Tang, A Resource for Inactivation of MicroRNAs Using Short Tandem Target Mimic Technology in Model and Crop Plants, *Molecular Plant*. 11 (2018). <https://doi.org/10.1016/j.molp.2018.09.003>.
- [242] P.L.T. Tran, A. de Cian, J. Gros, R. Moriyama, J.L. Mergny, Tetramolecular quadruplex stability and assembly, *Topics in Current Chemistry*. 330 (2013). <https://doi.org/10.1007/128-2012-334>.

Les figures ont été réalisées en partie en utilisant les images de Servier Medical Art (License Creative Commons Attribution 3.0 (<https://creativecommons.org/licenses/by/3.0/>)) et de BioRender (<https://biorender.com/library/>).

Étude de la biogenèse et des fonctions des petits ARN non codants dérivant d'ARN de transfert chez *Arabidopsis thaliana*

Résumé

Les petits ARN non codants dérivant des ARN de transfert (tRFs) ont été identifiés dans tous les domaines de la vie et sont impliqués dans des mécanismes biologiques variés. Mon travail de thèse, mené sur la plante modèle *Arabidopsis thaliana*, a eu pour but de mieux comprendre la biogenèse des tRFs et leurs fonctions chez les plantes. Des lignées stables n'exprimant pas les trois principales RNases T2, RNS1, 2 et 3, acteurs majeurs de la biogenèse des tRFs, ont été réalisées avec la technologie CRISPR/Cas9. De plus, le tRF-5D Ala (AGC) de 20 nucléotides, montré précédemment comme inhibant la traduction *in vitro*, a été utilisé comme modèle pour déterminer sa fonction. Des protoplastes d'*A. thaliana* ont été transfectés avec ce tRF pour observer sa localisation subcellulaire et identifier ses partenaires protéiques. Celui-ci est retrouvé accumulé dans des foci cytosoliques qui ne s'apparentent pas à des mitochondries, P-bodies ou granules de stress. Une hélicase ARN, DEXH1, interagissant avec des G-quadruplex, a été retrouvée associée avec le tRF Ala et la présence de structure en G-quadruplex sur le tRF Ala a été confirmée *in vitro*. Les résultats obtenus suggèrent que le tRF Ala structuré en G-quadruplex peut être stocké dans la cellule et que DEXH1 interviendrait pour le relâcher sous forme non structurée.

Mots-clés : ARNt, tRF, RNase T2, RNS, hélicase, DEXH1 ; G-quadruplex

Abstract

Small non-coding RNA derived from transfer RNAs (tRFs) have been identified in all domains of life and are implicated in various biological process. Based on the *Arabidopsis thaliana* model plant, my thesis work aimed to contributed to the understanding of tRFs biogenesis and their functions in plants. Stable knock out lines of the three main RNases T2, RNS1, 2, and 3, the main actors of tRFs biogenesis, were created with CRISPR/Cas9 technology. In addition, the tRF-5D Ala (AGC) of 20 nucleotides, previously shown to inhibit translation *in vitro*, was used as a model to determine its function. *A. thaliana* protoplasts were transfected with this tRF to determine its subcellular localization and to identify its protein partners. The tRF Ala is found accumulated in cytosolic foci that are neither mitochondria, P-bodies nor stress granules. Additionally, the tRF Ala that we showed able to create G-quadruplex *in vitro*, associates with DEXH1, an RNA helicase interacting with G-quadruplex. These results suggest that the tRF Ala, structured in G-quadruplex, can be stored in the cell and that DEXH1 would release it in an unstructured form.

Keywords: tRNA, tRF, RNase T2, RNS, helicase, DEXH1, G-quadruplex