

# THÈSE DE L'UNIVERSITÉ DE STRASBOURG

Présentée à la faculté des Sciences de la Vie

En vue d'obtenir le grade de

Docteur de l'Université de Strasbourg

Spécialité : Aspects moléculaire et cellulaire de la biologie

---

*Geoffrey Morelle*

*Les ARN de transfert,*

*une nouvelle source de petits ARN non-codants chez Arabidopsis thaliana*

Soutenue le 17 Mars 2015 devant la commission d'examen :

Dr. Michel DRON	—	Rapporteur externe
Dr. Michel GOLDSCHMIDT-CLERMONT	—	Rapporteur externe
Dr. Magali FRUGIER	—	Examineur interne
Dr. Laurence DROUARD	—	Directrice de thèse
Dr. Ian SMALL	—	Directeur de thèse

---

Institut de Biologie Moléculaire des Plantes (UPR-CNRS 2357)

12, rue du Général Zimmer

67084 Strasbourg cedex



**ibmp**



THE UNIVERSITY OF  
WESTERN AUSTRALIA



*À mon ange*



# Remerciements

Je tiens tout d'abord à remercier les docteurs Michel Dron ainsi que, Michel Goldschmidt-Clermont et Magali Frugier pour avoir accepté de juger mon travail de thèse.

Je tiens aussi à remercier mes directeurs de thèse Laurence Drouard et Ian Small pour m'avoir accepté dans leurs laboratoires respectifs et m'avoir donné l'opportunité de cette expérience unique qu'est le doctorat.

J'adresse aussi mes plus grands remerciements à Elodie Ubrig, Thalia Salinas et Cyrille Megel pour avoir participé si activement dans ce projet. J'adresse aussi à Bernard Gutmann, Peter Kindgren, Joanna Melonek ainsi que Hannes Ruwe ma plus grande gratitude pour ces excellents moments passés ensemble, mais aussi pour ces discussions qui se sont révélées si importantes pour ce projet.

Je tiens enfin à remercier toute ma famille pour avoir été si patient. Sans le savoir ils ont participé plus que de raison à cette thèse.



# Table des matières

<b>Abbréviations.....</b>	<b>11</b>
<b>Introduction .....</b>	<b>15</b>
<b>1- Les grands ARN .....</b>	<b>15</b>
1.1 – Les grands ARN codants .....	15
1.1.1 – Les ARN messagers .....	15
1.2 – Les grands ARN non-codants .....	19
1.2.1 – Les rétro-transposons .....	19
1.2.1 – Les lncRNA .....	23
1.2.2 – Les ARN ribosomiques .....	23
<b>2-Les petits ARN.....</b>	<b>25</b>
2.1 – Les petits ARN codants.....	25
2.2 – Les petits ARN non-codants.....	27
2.2.1 – Les snoRNAs.....	27
2.2.2 - Les snRNA .....	29
2.2.3 – Les petits ARN et l’interférence ARN.....	31
2.2.4 – piRNAs.....	39
2.2.5 - Les SINEs.....	39
2.2.6 – Les tRNA.....	41
<b>3 - Caractéristiques et fonctions des tRNA .....</b>	<b>41</b>
3.1 - Organisation des gènes de tRNA chez les plantes.....	41
3.1.1 – Génome nucléaire .....	43
3.1.2 - Génome des organites .....	45
3.2- Structure des tRNA .....	47
3.2.1 – Les tRNA cytosoliques .....	47
3.2.2 – Les tRNA d’organites.....	49
3.3 - Transcription des tRNA.....	49
3.3.1 – Dans le noyau.....	49
3.3.2 – Dans les organites.....	53
3.4 – Maturation des tRNA.....	55
3.4.1 – Maturation des extrémités 5’ et 3’ .....	55
3.4.2 – La queue CCA .....	57
3.4.3 – Épissage des pré-tRNA.....	57
3.4.5 – Autres modifications des tRNA .....	59
3.5 –tRNA et utilisation des codons.....	63
3.6 –Contrôle qualité des tRNA .....	63
3.7 –Fonctions des tRNA.....	65

3.7.1 – Fonctions communes aux eucaryotes et procaryotes.....	65
3.7.3 – Fonctions spécifiques aux procaryotes.....	69
3.7.2 – Fonctions spécifiques aux eucaryotes .....	69
<b>4 - Revue “Surveillance et clivage des tRNA eucaryotiques” .....</b>	<b>71</b>
<b>5 - Travaux de thèse.....</b>	<b>95</b>
5.1 - Analyse de la population en tRFs chez <i>A. thaliana</i> .....	95
5.2 - Biogenèse des tRFs chez <i>A. thaliana</i> .....	95
5.3 - Les tRFs et le séquençage haut débit .....	97
<b>Matériels et méthodes.....</b>	<b>101</b>
<b>1 - Matériels .....</b>	<b>101</b>
1.1 - Matériels vivants.....	101
1.1.1 - Bactéries.....	101
1.1.2 - Plantes.....	103
1.2 - ADN.....	103
1.2.1 - Plasmides.....	103
1.2.2 - Oligonucléotides.....	109
1.3 – Anticorps.....	114
1.3 – Outils informatiques.....	115
<b>2 – Méthodes.....</b>	<b>117</b>
2.1 – ADN.....	117
2.1.1 - Extraction des plasmides .....	117
2.1.2 - Extraction d’ADN génomique et génotypage .....	117
2.1.3 - PCR (Polymerase Chain Reaction).....	119
2.1.4 - Mutagenèse dirigée.....	123
2.1.5 - Génération d’amiRNA.....	125
2.1.6 - Séquençage d’ADN .....	127
2.1.7 – Synthèse d’ADNc .....	127
2.1.8 – qRT-PCR .....	127
2.1.9 – Gateway .....	129
2.1.10 - Préparation et transformation d’ <i>E. coli</i> thermocompétentes.....	131
2.1.11 - Préparation d’ <i>A. tumefaciens</i> électrocompétentes .....	133
2.1.12 - Transformation d’ <i>A. tumefaciens</i> .....	135
2.1.13 - Électrophorèse de fragments d’ADN en gel d’agarose.....	135
2.2 - ARN.....	135
2.2.1 - Extraction des ARN de petites tailles .....	135
2.2.2 - Électrophorèse des ARN sur gel de polyacrylamide.....	137
2.2.3 - Transfert des ARN .....	139
2.2.4 - Marquage radioactif des sondes.....	139

2.2.5 - Northern blot.....	141
2.2.6 - Cinétique de clivage des transcrits tRNA.....	141
2.2.7 - Préparation de banques de petits ARN et séquençage Illumina®.....	143
2.3 - Protéines.....	147
2.3.1 - Expression et induction protéique dans <i>E. coli</i> .....	147
2.3.2 - Purification de protéines en conditions natives.....	147
2.3.3 - SDS-PAGE.....	149
2.3.4 - Détection des protéines.....	151
2.3.5 - Western blot.....	151
2.4 - Techniques relatives aux plantes et à leurs cultures.....	155
2.4.1- Culture d' <i>A. thaliana</i> et <i>N. benthamiana</i> .....	155
2.4.2 - Transformation d' <i>A. thaliana</i> et <i>N. benthamiana</i> .....	155
2.4.3 - Purification de chloroplastes d' <i>A. thaliana</i> .....	157
2.4.4 - Analyse GFP en microscopie confocale.....	159
<b>Résultats.....</b>	<b>165</b>
<b>1 - Analyse de la population en tRFs chez <i>A. thaliana</i>.....</b>	<b>165</b>
<b>2 - Biogenèse des tRFs.....</b>	<b>201</b>
2.1 – Etude de lignées mutantes <i>dcl</i> et <i>rdr</i> d' <i>A. thaliana</i> .....	201
2.2 – Identification <i>in silico</i> d'homologues de <i>rny1</i> de <i>S. cerevisiae</i> chez <i>A. thaliana</i> .....	205
2.3 – Caractéristiques des RNases T2.....	205
2.4 - Données bibliographiques des RNS d' <i>A. thaliana</i> .....	213
2.5 – Clonage des cDNA codant pour les RNS.....	215
2.6 – Production des RNS chez <i>E. coli</i> .....	217
2.7 – Activité de clivage des RNS recombinantes sur des tRNA transcrits <i>in vitro</i>	219
2.8 – Mutagenèse dirigée des RNS et activité de clivage des protéines recombinantes mutantes.....	223
2.9 – Localisation <i>in vivo</i> des RNS.....	225
2.10 – Mutants KO ou lignées inhibant la production des RNS d' <i>A. thaliana</i> .....	227
<b>3 – Les tRFs et le séquençage haut débit.....</b>	<b>229</b>
3.1 – RNase T2, séquençage haut débit.....	229
3.1.1 – Rétablissement des extrémités 5' et 3' normales.....	231
3.1.2 – Comparaison de la banque traitée avec la T4 polynucléotides kinase avec la banque non traitée.....	231
3.1.3 – Clivage des tRFs dans l'anticodon dans la banques non traitée.....	239
<b>Discussion et perspectives.....</b>	<b>249</b>
<b>1 – Conclusion générale : récapitulatif des résultats.....</b>	<b>249</b>

1.1 – Analyse des banques de tRFs .....	249
1.2 – Biogenèse des tRFs.....	251
1.3 – Clonage des tRFs, activité des RNS.....	255
<b>2 – Discussion .....</b>	<b>255</b>
2.1 - Les tRFs .....	255
2.2 – Biogenèse des tRFs et localisation des RNS .....	261
2.3 – T4 polynucléotides kinase et tRFs.....	267
<b>3 – Perspectives .....</b>	<b>269</b>
<b>Bibliographie .....</b>	<b>273</b>

## Abbreviations

<b>ARN</b>	Acide Ribo Nucléique
<b>ADN</b>	Acide Désoxyribo Nucléique
<b>mRNA</b>	“messenger RNA“
<b>rRNA</b>	“ribosomal RNA“
<b>NEP</b>	“Nucleus Encoded Polymerase“
<b>PEP</b>	“Plastidial Encoded Polymerase“
<b>CPSF</b>	“Cleavage and Polyadenylation Specificity factor“
<b>CStF</b>	“Cleavage Stimulation Factor“
<b>PABP</b>	“Poly Adenosine Binding Protein“
<b>hnRNP</b>	“heterogeneous nuclear RiboNucleoProtein
<b>CBC</b>	“Cap-Binding Complex“
<b>TE</b>	“Transposable Elements“
<b>LTR</b>	“ Long Terminal Repeat“
<b>LINES</b>	“ Long INterspersed Elements“
<b>SINEs</b>	“ Short INterspersed Elements“
<b>lncRNA</b>	“long non-coding RNA“
<b>tmRNA</b>	“transfer messenger RNA“
<b>smpB</b>	“small protein B“
<b>snoRNA</b>	“small nucleolar RNA“
<b>snoRNP</b>	“small nucleolar RiboNucleoProtein“
<b>snRNA</b>	“small nuclear RNA“
<b>PTGS</b>	“Post-Transcriptional Gene Silencing“
<b>TGS</b>	“Transcriptional Gene Silencing“
<b>DCL</b>	“Dicer Like“
<b>miRNA</b>	“micro RNA“
<b>pri-miRNA</b>	“primary miRNA“
<b>pre-miRNA</b>	“precursor miRNA“
<b>HEN1</b>	“ Hua Enhancer 1“
<b>HST</b>	“Hasty“
<b>RISC</b>	“RNA-Induced Silencing Complex“
<b>siRNA</b>	“Small Interferent RNA“
<b>ta-siRNA</b>	“trans-acting siRNA“

<b>nat-siRNA</b>	“natural antisens transcript-derived siRNA“
<b>endo-siRNA</b>	“endogenous siRNA“
<b>NERD siRNA</b>	“NEeded for RDR2 independent DNA methylation“
<b>Rdr</b>	“RNA dependent RNA polymerase“
<b>casRNA</b>	“cis-acting siRNA“
<b>RdDM</b>	“RNA directed DNA methylation“
<b>piRNA</b>	“piwi-interacting RNA“
<b>tRNA</b>	“transfer RNA“
<b>tDNA</b>	“transfer DNA“
<b>pre-tRNA</b>	“precursor tRNA“
<b>tRFs</b>	“tRNA fragments“
<b>RNS</b>	Ribonucléase
<b>LB, RB</b>	“Left Border, Right Border“
<b>Tm</b>	“melting temperature“
<b>amiRNA</b>	“artificial miRNA“
<b>PCR</b>	“Polymerase Chain Reaction“
<b>qPCR</b>	“quantitative PCR“
<b>EDTA</b>	“ Ethylene di-amine tetra-acetic acid“
<b>LB</b>	“Lysogeny Broth“
<b>TAE</b>	Tris-Acétate EDTA
<b>BET</b>	Bromure d'éthidium
<b>TBE</b>	Tris-Borate EDTA
<b>PNK</b>	Phospho-Nucléotides Kinase
<b>SSC</b>	“Saline Sodium Citrate“
<b>SDS</b>	Sodium dodécyl sulfate
<b>IPTG</b>	Isopropyl $\beta$ -D-1-thiogalactopyranoside
<b>APS</b>	“Ammonium Persulfate“
<b>ECL</b>	“Enhanced Chemical Luminescence“
<b>D.O</b>	Densité Optique
<b>GFP</b>	“Green Fluorescent Protein“
<b>RFP</b>	“Red Fluorescent Protein“
<b>PBS</b>	“Phosphate Buffer Saline“
<b>mGFP<sub>er</sub></b>	“mutateg GFP endoplasmic reticulum“



---

# *INTRODUCTION*



# Introduction

On suppose que dès les balbutiements de la vie les ARN étaient les uniques porteurs de l'information génétique ainsi que les effecteurs du génome, à l'image de certaines répliques décrites par le passé (Mills et al., 1967). Cependant, au cours de l'évolution, ceux-ci furent remplacés dans leurs fonctions de porteurs de l'information par l'ADN, plus stable et plus résistant aux conditions environnementales et par les protéines en tant qu'effecteurs (Kawamura, 2012). Ainsi, à l'exception de certains virus (*e.g.* virus à ARN, rétrovirus) les ARN, chez les autres organismes, ont changé de rôles et sont devenus les porteurs transitoires de l'information génétique, tandis que certains ont gardé leurs rôles d'effecteurs (*e.g.* ribosomes, ribozymes).

En changeant de rôles les ARN se sont aussi diversifiés jusqu'à occuper de multiples fonctions dans la cellule. Parmi ces rôles, nous pouvons citer les plus connus comme les ARN messagers (mRNA) qui servent d'intermédiaires pour la synthèse des protéines lors de la traduction, les ARN ribosomiaux (rRNA) composants essentiels du ribosome, ou encore les ARN de transfert (tRNA) molécules essentielles aux transports des acides aminés sur les chaînes peptidiques naissantes. De nombreux autres types d'ARN existent, ils sont brièvement décrits ci-dessous.

## 1- Les grands ARN

### 1.1 – Les grands ARN codants

#### 1.1.1 – Les ARN messagers

Les seuls ARN de grandes tailles codants pour des protéines sont les ARN messagers, ou mRNA. Le rôle des mRNA est de servir d'intermédiaire pour la synthèse des protéines. Nous pouvons distinguer les mRNA en fonction de l'organisme qui les produit, et notamment selon qu'ils soient d'origine procaryote ou eucaryote.



### 1.1.1.1 – Chez les procaryotes

Chez les procaryotes l'étape de transcription s'effectue en tandem avec la traduction. Les mRNA y sont polycistroniques, ils possèdent donc plusieurs codons d'initiations et codons-stop sur un même ARN. Ceci permet la synthèse de plusieurs protéines à partir du même mRNA. Ils sont souvent transcrits à partir de gènes organisés en opérons c'est-à-dire une série de gènes (souvent associés à la même voie métabolique) sous le contrôle d'un même promoteur.

L'enzyme responsable de la transcription chez les procaryotes est composé de 5 sous-unités dont notamment la sous-unité  $\sigma$  qui sera la première à reconnaître l'ADN et la séquence initiatrice TATAAT appelée boîte de Pribnow (Pribnow, 1975). Il existe plusieurs types de facteurs  $\sigma$ , chacun est spécifique d'un type de promoteur (*e.g.*,  $\sigma 70$  transcription de gène standard,  $\sigma 54$  transcription de gène nécessaire au métabolisme de l'azote) (Zhang et al., 2014; Zhao et al., 2005; Zhou et al., 1988).

La terminaison de la transcription s'effectue généralement de deux manières différentes. La première dite rho indépendante fait intervenir une tige-boucle suivie d'une série d'uridines. Le second type de terminaison fait intervenir un facteur protéique de terminaison nommé rho (Pour revue : Adhya et al., 1979; Bear et Peabody, 1988; Howard et de Crombrugghe, 1976).

Les mRNA sont dégradés par plusieurs endonucléases qui vont cliver les mRNA dans des régions simples brins riches en AU (RNase E, G, J et Y), ou dans des régions doubles brins (RNase III). Ils sont aussi pris en charge par des exonucléases (RNase II, RNase R et la PNPase : Polynucleotide phosphorylase) qui dégradent les mRNA dans le sens 5'-3' ou 3'-5'. Divers éléments participent à la stabilisation des mRNA (*e.g.* présence d'une extrémité 5' triphosphate et de structures secondaires), ou à leur déstabilisation (*e.g.* présence d'une extrémité 5' monophosphate et de queues permettant le recrutement d'exonucléases) (Pour une revue : Evguenieva-Hackenberg et Klug, 2011).

### 1.1.1.2 – Chez les eucaryotes

Chez les eucaryotes, la transcription des mRNA est réalisée par la polymérase II dans le noyau, par des NEP ("Nucleus-Encoded Polymerase") dans la mitochondrie auxquelles s'ajoute une PEP ("Plastid-Encoded Polymerase")



dans le chloroplaste (Pour revue, Arnold et al., 2012; Braun et Young, 2014; Yu et al., 2014). Dans le noyau la polymérase II nécessite plusieurs facteurs d'initiation, ou TF, (Transcriptional Factor) dont certains vont se fixer à la TATA box en jouant un rôle similaire au facteur  $\sigma$  bactérien. Le mRNA est ensuite transcrit de l'extrémité 5' vers l'extrémité 3'. La transcription est stoppée aux environs du signal polyadénylation (5'-AAUAAA-3'). Ce signal permet le recrutement des facteurs CPSF (Cleavage and Polyadenylation Specificity factor), CStF (Cleavage Stimulation Factor) ainsi que de la poly-A polymérase (Pour revue : Buratowski, 2009; Cheung et Cramer, 2012; Kim et Buratowski, 2013; Shandilya et Roberts, 2012; Munroe et Jacobson, 1990).

Peu de temps après le début de sa synthèse, le mRNA est coiffé par l'ajout à son extrémité 5' d'une guanosine méthylée (Ghosh et Lima, 2010). Puis il est épissé via l'action de plusieurs facteurs (Pour revue : Brown et al., 2014; Galej et al., 2014; Matera et Wang, 2014). Enfin le mRNA est exporté vers le cytoplasme *via* l'interaction avec un complexe ribonucléique nommé hnRNP (heterogeneous nuclear RiboNucleoProtein). Ce complexe est composé de la CBC (Cap-Binding Complex), la PABP (Poly Adenine Binding Protein) et du complexe de liaison au pore nucléaire Nxt1-Nxf1 (Björk et Wieslander, 2014).

La cellule possède des mécanismes de contrôle qualité des mRNA (Pour revue : Kilchert et Vasiljeva, 2013; Walters et Parker, 2014; Zhai et Xiang, 2014). Ainsi, lorsqu'un mRNA aberrant est détecté on distingue trois processus de dégradation le "No-Go Decay", pour un mRNA sans codon start, le "Non-Stop Decay", pour un mRNA sans codon stop et enfin le "Nonsense Mediated mRNA Decay" pour des mRNA possédant un codon stop prématuré (Pour revue : Parker, 2012; Roy et Jacobson, 2013).

## 1.2 – Les grands ARN non-codants

### 1.2.1 – Les rétro-transposons

Il existe deux types d'éléments transposables (TE), les rétrotransposons, ou transposons de classe 1 sont des transposons avec un intermédiaire ARN. Tandis que les transposons à ADN, dit de classe 2, ne font pas appel à ces intermédiaires ARN.



Ils sont particulièrement abondants chez les plantes. Entre 49 et 78% du génome du maïs en est composé (Sanmiguel et Bennetzen, 1998), contre 68% du génome du blé (Li et al., 2004). Afin de se répliquer et de s'intégrer de nouveau dans le génome, les rétrotransposons sont transcrits en ARN, puis ils sont rétrotranscrits en ADN à l'aide d'une transcriptase inverse pour laquelle ils codent eux-mêmes (Dombroski et al., 1994).

Il existe 2 classes de rétrotransposons les LTR (Long Terminal Repeat) et les non-LTR.

#### 1.1.2.1 – Les rétrotransposons LTR

Les rétrotransposons LTR dérivent probablement de rétrovirus qui se sont définitivement intégrés au génome de leurs hôtes. Leur taille varie de 1 à 10kpb. Ils sont subdivisés en 3 classes les Ty1-copia-like, Ty3-gypsy-like et les BEL-Pao-Like. Les 2 premiers types sont présents en très grand nombre chez les animaux, les champignons et les plantes, tandis que les BEL-Pao-Like n'ont été caractérisés que chez les animaux (Copeland et al., 2005; Schulman, 2013; Wicker et al., 2007).

#### 1.1.2.2 – Les rétrotransposons non-LTR

Les rétrotransposons non-LTR sont divisés en 2 sous-groupes les LINEs (Long INterspersed Elements) et les SINEs (Short INterspersed Elements). Les LINEs codent pour tous les éléments nécessaires à leurs rétrotranscription et leurs transposition tandis que les SINEs dépendent des éléments codés par les LINEs et ne sont donc pas autonomes (McCarthy et McDonald, 2004; Schulman, 2013).

Les LINEs forment un groupe d'éléments génétiques allant de 6 à 8kpb, transcrit par la polymérase II. Ils représentent 17% du génome humain mais sont bien moins nombreux chez les plantes (e.g, 4% du génome de *Lilium speciosum*) (Wenke et al., 2009). Peu d'entre eux sont encore capables de se répliquer seuls (Doucet et al., 2010). Ils possèdent deux cadres de lecture ouverts. Le premier code pour une protéine de liaison à l'ARN tandis que le second code pour une enzyme possédant une activité de transcription inverse et une activité de type RNase H (Ohshima et Okada, 2005; Schulman, 2013).



Les SINEs sont plus petits, entre 100 et 300pb, et ils ne codent pour aucune protéine. C'est pourquoi ce type de rétrotransposon appartient plus à la classe des petits ARN non codants (cf. 2.2.5).

### 1.2.1 – Les lncRNA

Les lncRNA (long non coding RNA) sont définis comme des séquences d'ARN supérieures à 200 nucléotides, ne codant pour aucune protéine. Les lncRNA se trouvent dans les introns, en antisens de gènes ou dans des régions intergéniques, nombres d'entres eux proviennent d'insertions de rétrotransposons.

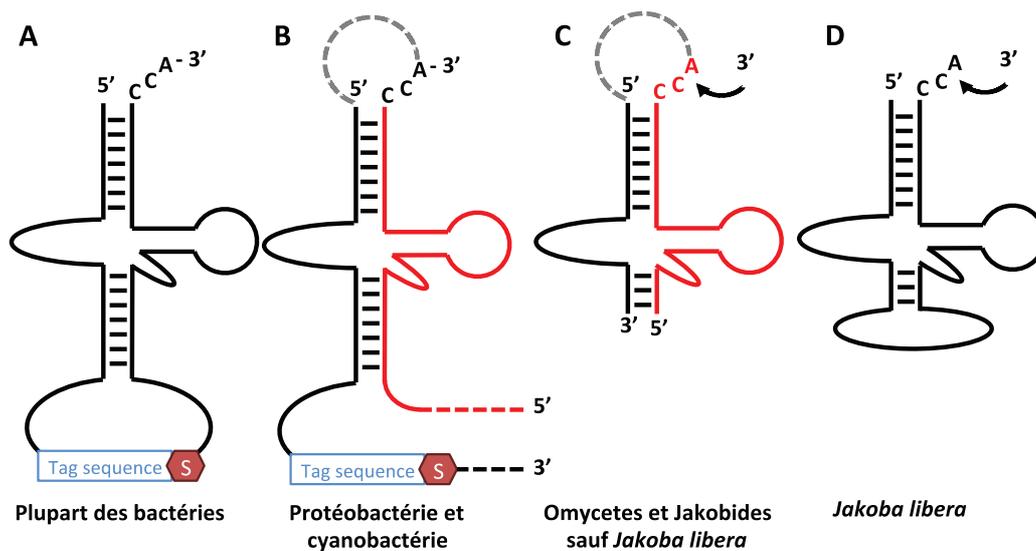
La plupart des lncRNA analysés jusqu'à aujourd'hui ont un rôle dans la régulation de la différenciation cellulaire, dans le développement de l'embryon et dans certaines réactions épigénétiques (guidage des enzymes modifiant la chromatine). Ils guident et aident aussi au rapprochement des régions "enhancers" et promotrices de la transcription, à l'épissage alternatif, à la localisation des protéines ainsi qu'à la stabilité et la traduction mRNA (Pour une revue, Morris et Mattick, 2014).

### 1.2.2 – Les ARN ribosomiques

Les ARN ribosomiques (rRNA) sont des molécules essentielles à la traduction. Chez les eucaryotes, la transcription des rRNA est réalisée par les polymérases I et III dans le noyau, par des NEP dans la mitochondrie et une PEP dans le chloroplaste (Pour revue, Arimbasseri et al., 2013; Braun et Young, 2014; Németh et al., 2013; Yu et al., 2014).

Chez les procaryotes, la mitochondrie et le chloroplaste, la petite sous-unité dite 30S (Melnikov et al., 2012) est composée d'un rRNA nommé 16S tandis que la grande sous-unité, la 50S, est composée de 2 rRNA de 5 et 23S. (Alberts, 2002).

Chez les eucaryotes, le ribosome est composé de 2 sous-unités 60S et 40S, le rRNA 18S se trouve dans la petite sous-unité du ribosome (40S) tandis que la grande sous-unité contient 3 rRNA, le 5S, 5,8S et le 28S (chez les mammifères, 25S chez les plantes), tous sont transcrits par l'ARN polymérase I à l'exception du rRNA 5S qui est transcrit par la polymérase III (Melnikov et al., 2012).



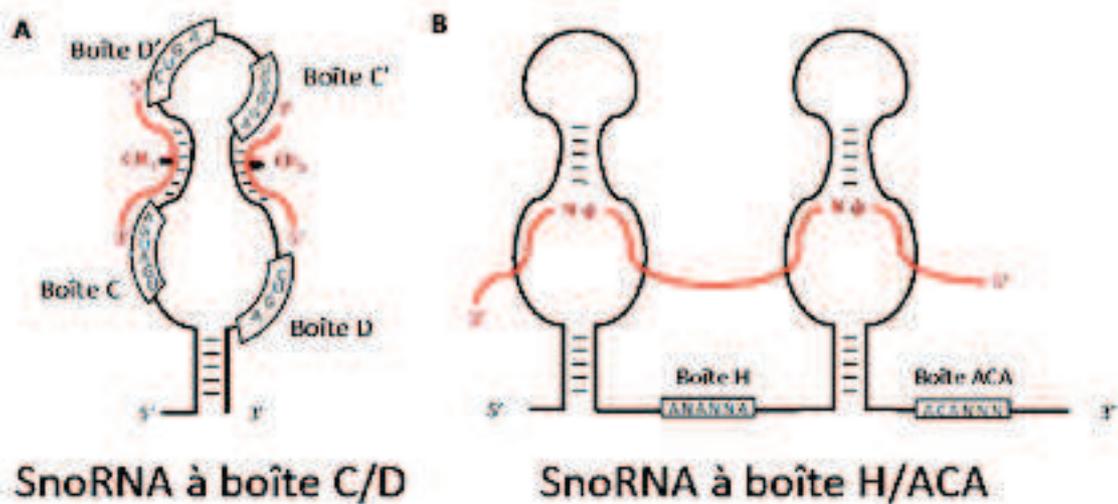
**Figure 1 : Comparaison des différentes structures de tmRNA de bactéries et mitochondries. A et B, tmRNA bactérien. Le CCA est ici encodé par le génome chez la plupart des bactéries. Les tmRNA comprennent deux domaines ayant des fonctions distinctes: le domaine de type tRNA (noir) et le domaine de type mRNA (bleu). B et C, en rouge figure les domaines de type tRNA codés par un second gène. C et D, tmRNA mitochondrial, chez lesquels il manque le domaine de type mRNA. La flèche indique que le CCA est ajouté post-transcriptionnellement (figure d'après Hafez et al., 2013).**

## 2-Les petits ARN

### 2.1 – Les petits ARN codants

Les seuls ARN connus faisant partie de cette catégorie se nomment les tmRNA (abréviation de “transfer-messenger RNA”) (figure 1). Ce type d’ARN se trouve principalement chez les bactéries. Ils possèdent des caractéristiques propres aux mRNA ainsi qu’aux tRNA. Les tmRNAs forment des complexes ribonucléoprotéiques avec SmpB (Small protein B), le facteur d’élongation Tu (Ef-Tu), ainsi que la protéine ribosomale S1. C’est lors de l’étape de trans-translation que les tmRNA et ces protéines se lieront à un ribosome bactérien bloqué en plein milieu de la synthèse d’une protéine, comme par exemple lors de la traduction d’un mRNA auquel il manque un codon stop. Le tmRNA permet le recyclage du ribosome bloqué, et le marquage du peptide en cours de synthèse de façon à ce qu’il soit dégradé tout comme le mRNA aberrant (Keiler, 2008). Chez la plupart des bactéries ces fonctions sont assurées par des tmRNA transcrits d’un seul tenant. Cependant, chez quelques espèces, les tmRNA sont transcrits à partir de 2 gènes permutés dont les transcrits s’hybrident l’un à l’autre de manière à former le tmRNA complet (figure 1D) (Hafez et al., 2013). La structure canonique des tmRNA consiste en 2 domaines. Le premier domaine, similaire au tRNA<sup>Ala</sup>, est aminoacylable par l’alanine et permet l’entrée du tmRNA dans le site A du ribosome. Le second domaine, similaire aux mRNA, code pour un peptide qui étiquètera le polypeptide néosynthétisé de façon à ce qu’il soit dégradé.

Du fait de leur origine commune avec les bactéries, les mitochondries de certains organismes, possèdent aussi des gènes codants pour les tmRNA (*ssrA*). La présence de ces ARN dans les mitochondries a été mise en évidence, pour la première fois, dans des organismes eucaryotes de l’ordre des jakobida (Burger et al., 2013; Jacob et al., 2004; Keiler et al., 2000), depuis ils ont aussi été décrits chez des oomycètes (Hafez et al., 2013).



**Figure 7 : Structure des 2 classes connues de snoRNA.** En noir figure les snoRNA et en rouge les brins d'ARN à modifier. **A** : Structure d'un snoRNA à boîte C/D, utilisés pour la méthylation des transcrits : CH<sub>2</sub> (en rouge). **B** : Structure d'un snoRNA à boîte H/ACA servant à l'isomérisation d'une uridine en pseudouridine : Ψ (en rouge).

## 2.2 – Les petits ARN non-codants

### 2.2.1 – Les snoRNAs

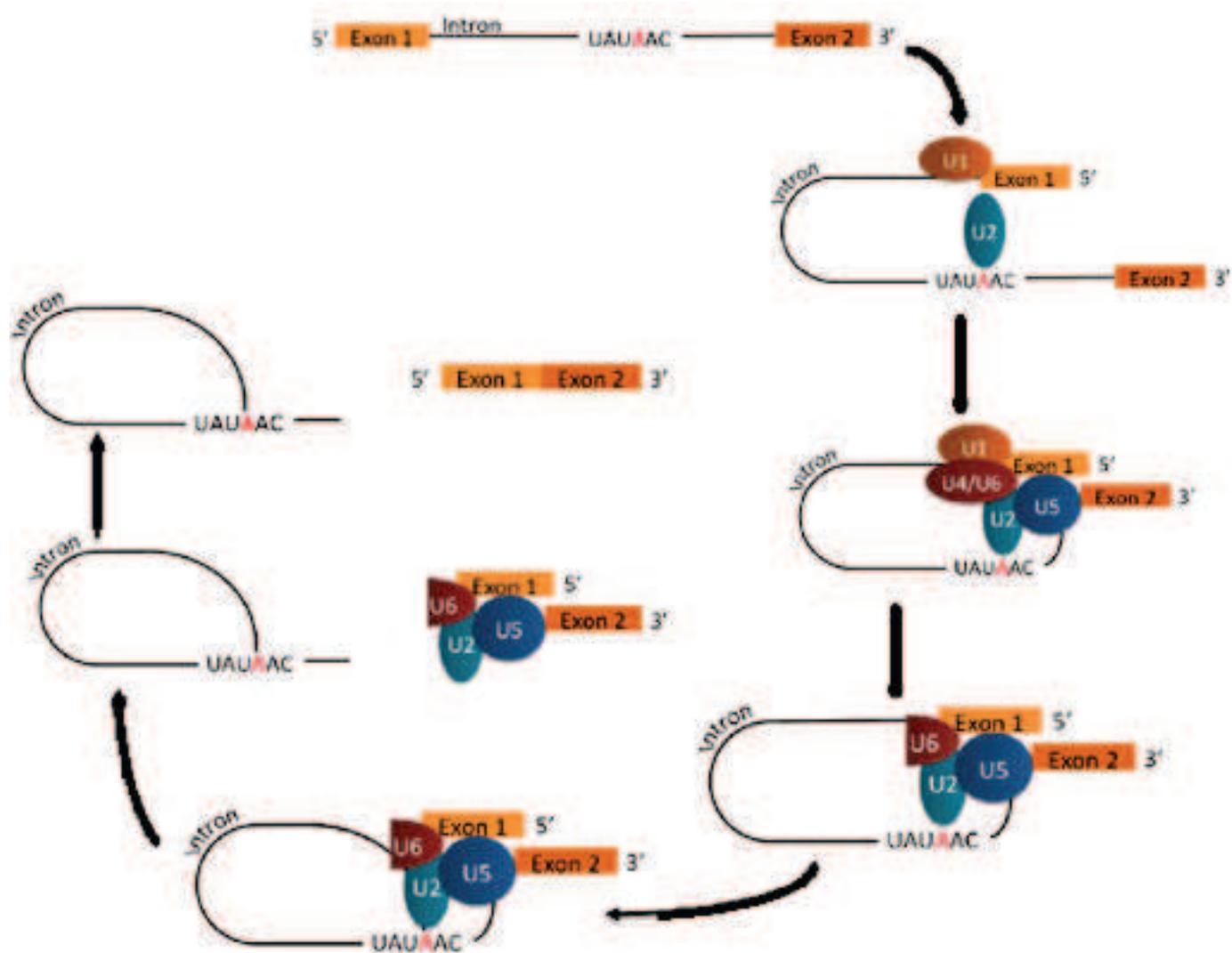
Les snoRNA (Small nucleolar RNAs) sont une classe de petits ARN d'environ 50 à 100nt que l'on trouve codés dans les introns, chez les eucaryotes. Ils servent de guide à la modification d'autres ARN, notamment les ARN ribosomiaux. Il existe 2 classes de snoRNA. La classe contenant une boîte C/D qui est associée à la méthylation des transcrits et celle contenant une boîte H/ACA qui est associée à la pseudouridylation des ARN (Figure 2).

La majorité des méthylations a lieu au niveau du ribose, en position 2', tandis que la pseudouridylation consiste en l'isomérisation de l'uridine en pseudouridine (Bratkovič et Rogelj, 2014).

Chaque snoRNA ne peut participer qu'à la modification d'un site. Cela nécessite le recrutement d'un complexe d'environ 4 protéines, l'ensemble formant alors un snoRNP (« small nucleolar RiboNucleoProtein »). Ce complexe reconnaît le site d'édition de l'ARN grâce au snoRNA qui possède une séquence de 10 à 20nt complémentaire au site d'édition. Dès lors les différentes protéines du complexe vont pouvoir jouer leurs rôles selon la modification visée.

#### 2.2.1.1 - SnoRNA à boîte C/D et méthylation

La séquence de cette classe de snoRNA contient deux séries de deux séquences conservées, la première série est constituée des séquences C et C' (AGUAGU) et la seconde série est constituée des séquences D et D' (CUGA). De petites régions complémentaires permettent la formation d'une structure rapprochant les domaines C et D ainsi que C' et D' respectivement. Entre chaque couple de domaine C/D et C'/D' se trouve la région complémentaire au site de modification. Ce site permet au snoRNA de former un duplex ARN dont l'un des nucléotides sera modifié par l'ajout d'un groupement méthyl (figure 2A) (Samarsky et al., 1998).



**Figure 3 : Schéma des snRNP et de leur rôle dans l'épissage.** L'épissage se divise en plusieurs étapes d'interaction, avec les snRNP (small nucleolar RiboNucleo Protein). Tout d'abord, le snRNP U1 reconnaît la jonction 5' de l'intron et U2 reconnaît le A de branchement (rouge). Ensuite U4 et U6 s'associent et rapprochent U1 et U2, permettant la jonction en 5' entre l'intron et la boîte de branchement. Dans le même temps, U5 rapprochera les extrémités des 2 introns. Après que U1 et U4 aient quitté le complexe, le 2'-OH de l'adénosine de la boîte de branchement attaque l'extrémité 5' de l'intron et s'y lie. Puis le 3'-OH de l'exon 1 va couper la liaison retenant l'exon 2 à l'intron, reliant ainsi les 2 exons.

### 2.2.1.2 - SnoRNA à Boîte H/ACA et pseudo-urydiation

Les snoRNA à boîte H/ACA sont structurés en une série de 2 tiges boucles ainsi que 2 domaines d'ARN simple brin. Le tout organisé de la façon suivante : tige-boucle/charnière/tige-boucle/queue (figure 2B). Les snoRNA H/ACA contiennent deux régions conservées, la boîte H dont la séquence consensus est ANANNA, et la boîte ACA. Chacune des 2 boucles contient les séquences nécessaires à la reconnaissance de l'ARN cible, formant alors une structure en "pseudo-nœud" incluant l'ARN cible ainsi que 4 protéines (Cbf5p, Gar1p, Nhp2p, et Nop10p). Le SnoRNA convertira alors en pseudo-uridines, par isomérisation d'un de leurs cycles, les uridines de transcrits tels que les rRNA, ou les snRNA (Bachellerie et al., 2002; Bratkovič et Rogelj, 2014; Ganot et al., 1997).

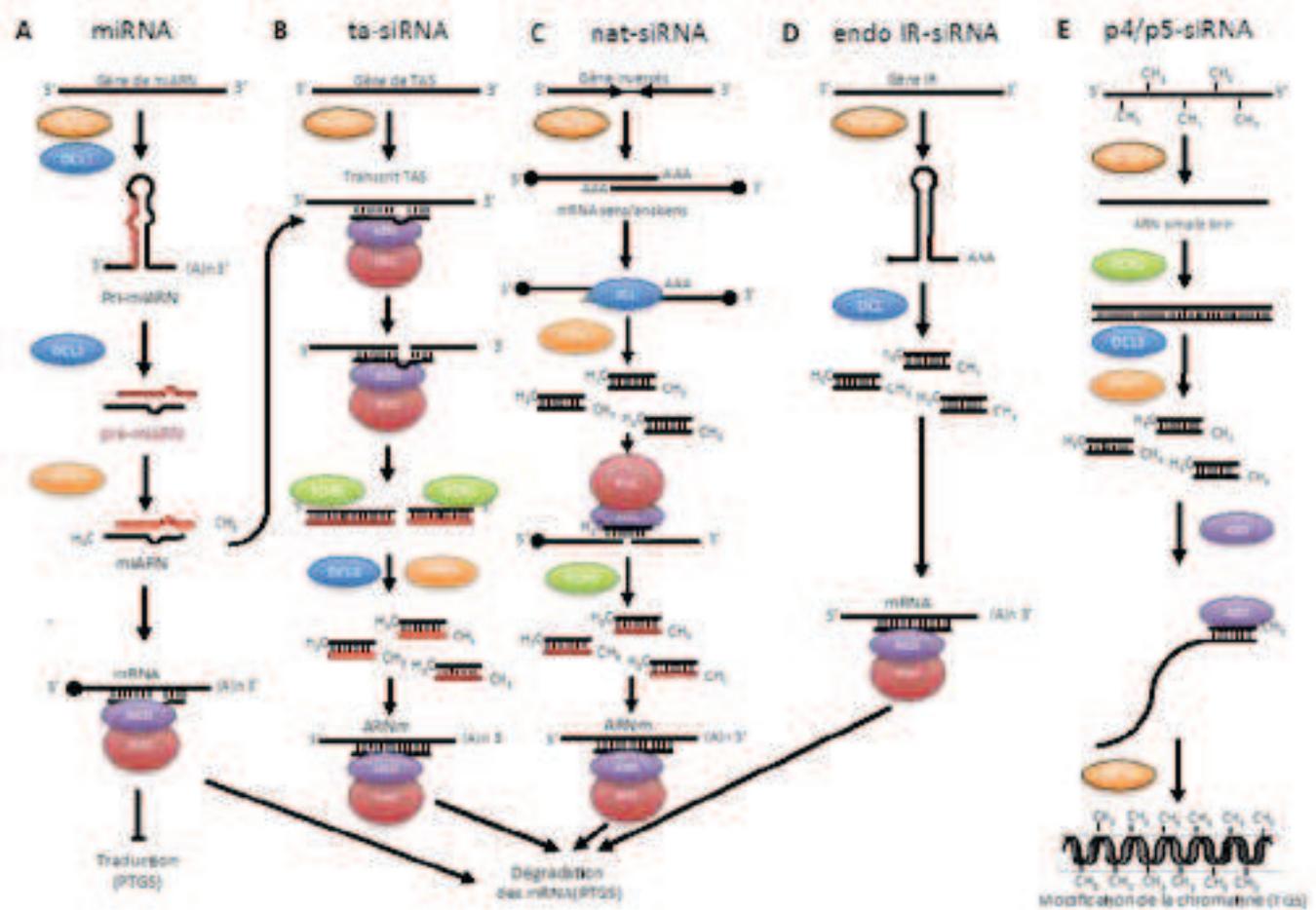
### 2.2.2 - Les snRNA

Les snRNA (Small nuclear RNA), occupent un rôle-clé dans la cellule. Souvent appelés U-RNA, ils constituent une classe de petits ARN non codants trouvés dans le noyau des eucaryotes. Ils participent à l'épissage des introns des mRNA (figure 3). Ils ont une taille moyenne de 150nt et ils peuvent être transcrits soit par la polymérase II, soit par la polymérase III (Matera et al., 2007).

Les snRNA sont toujours associés à un jeu de protéines, ce complexe est alors nommé snRNP (small nuclear RiboNucleo Protein). Chaque snRNP est composé de plusieurs protéines ainsi que d'un snRNA. Ils sont divisés en 2 classes (Matera et al., 2007).

La première classe, la classe Sm, est la plus décrite. Les snRNA de cette classe sont transcrits par l'ARN polymérase II. Ils sont ensuite coiffés dans le noyau d'une 5' monométhylguanosine. Puis ils sont rallongés en 3' de manière à former une tige boucle. Ils sont aussi hyperméthylés au niveau de la coiffe en 5'. La classe Sm des snRNA compte les ARN U1, U2, U4, U4atac, U5, U7, U11, et U12. Tous ces ARN riches en uridine (hormis U7) forment le cœur du spliceosome. Cette structure participe à l'excision des introns des mRNA (figure 3).

Dans la seconde classe, la classe Lsm, les snRNA possèdent une coiffe monométhylphosphate ainsi qu'une tige boucle en 3' se terminant par un polyU.



**Figure 4 : Schéma des différentes voies de l'interférence ARN chez *A. thaliana*.** **A**, Après transcription par la polymérase II les miRNA sont maturés par DCL 1, méthylés par HEN1 et chargés par AGO1 au sein du complexe RISC via Hst. Ainsi chargés les complexes RISC vont soit dégrader les mARN ciblés par le miRNA, soit inhiber leurs traductions (PTGS). **B**, Les fragments de mRNA dégradés sont utilisés afin d'induire des clivages dans un transcrit TAS dont les fragments serviront de matrice à Rdr6. Ils participeront alors au PTGS. **C**, Les nat-siRNA sont produits à partir de 2 gènes dont les transcrits s'hybrident l'un avec l'autre, la zone complémentaire est ensuite utilisée par DCL 1 et 2 afin de produire des siRNA qui serviront à dégrader d'autres transcrits dont les produits de dégradation serviront dans le PTGS. **D**, Les endo IR-siRNA sont transcrit à partir de loci dispersés dans le génome. Après transcription ils sont maturés par DCL 2 et intégrés au PTGS. **E**, Les p4/p5-siRNA sont transcrits à partir de gènes méthylés par les polymérases 4 et 5. Après synthèse du brin complémentaire par RDR2, DCL 3 produira de petit ARN qui serviront à modifier la chromatine afin d'inhiber la transcription de certains gènes (TGS) (cf. Introduction - 2.2.3.2) (figure d'après Martínez de Alba et al., 2013).

Ce polyU constitue le site de liaison de l'hétérohaptamère de la protéine Lsm. Cette classe de snRNA est transcrite par l'ARN polymérase III. Les snRNA ne quittent jamais le noyau. À ce jour, seuls 2 ARN de ce type sont connus : U6 et U6atac (Matera et al., 2007).

L'ensemble de ces petits ARN participe à la formation de deux spliceosomes. U1, U2, U4, U5 et U6 participent à la formation du spliceosome majeur, intervenant dans l'épissage des introns communs dit de type U2. Tandis que U11, U12, U4atac, U6atac et U5 participent à la formation du spliceosome mineur intervenant dans l'épissage d'intron plus rare dit de type U12. U7 quant à lui est impliqué dans la modification de la tige boucle située à l'extrémité 3' des pré-mRNA d'histones (Schümperli et Pillai, 2004).

### 2.2.3 – Les petits ARN et l'interférence ARN

#### 2.2.3.1 – Les miRNA

Les miRNA (microRNA) sont codés dans des régions intergéniques ou inclus dans des introns de mRNA (figure 4). Chez les plantes, contrairement aux animaux, les miRNA s'hybrident parfaitement dans les régions codantes des mRNA dont ils induisent la dégradation. Ce processus est appelé PTGS ("Post-Transcriptional Gene Silencing") (Bartel, 2009; Lewis et al., 2003; Wang et al., 2004).

Les gènes de miRNA sont transcrits par l'ARN polymérase II (Lee et al., 2004) en une série de tiges-boucles nommée pri-miRNA (primary miRNA). Puis le transcrit est coiffé, polyadénylé, et mûri à la manière d'un mRNA (Cai et al., 2004). Il arrive parfois que certains pri-miRNA possèdent des introns, et ceux-ci sont épissés de la même manière que chez les mRNA (Jones-Rhoades et al., 2006). Les tiges-boucles sont ensuite séparées les unes des autres par DCL1 ("Dicer Like 1") (figure 4) pour former un pre-miRNA ("precursor miRNA") (tige-boucle isolée) dont la boucle sera clivée, dans un second temps, par la même enzyme (Cai et al., 2004; Lee et al., 2004; Martínez de Alba et al., 2013). Cependant, DCL1 ne prend en charge que la maturation de miRNA conservés, issus de tiges boucles courtes et imparfaites (*e.g.* pas totalement hybridées), d'autres miRNA, moins conservés, sont mûris par DCL4 à partir de tiges boucles longues et parfaites (Martínez de Alba et al., 2013).



Les miRNA sont ensuite méthylés dans leur extrémité 3' par HEN1 ("Hua Enhancer 1") puis tous les ARN sont exportés du noyau vers le cytoplasme par un homologue de l'exportine 5 de drosophile, "Hasty" (HST) (Lund et Dahlberg, 2006; Murchison et Hannon, 2004; Park et al., 2005). Enfin, après ces étapes de maturation l'ARN est chargé sur Ago I qui est lui-même intégré au complexe RISC ("RNA-induced silencing complex"). C'est ce complexe qui clive le transcrite cible ou en inhibe la traduction *via* l'hybridation parfaite ou quasi-parfaite du miRNA au mRNA (Seto et al., 2007). Certains des fragments ainsi produits sont alors utilisés dans la voie des ta-siRNA (cf. 2.2.3.2) (figure 4) afin d'amplifier le phénomène d'interférence ARN aux cellules ou tissus voisins.

#### 2.2.3.2 – Les siRNA

Chez les plantes, les siRNA (small interfering RNA) (90%) sont bien plus représentés que les miRNA (10%) (Kasschau et al., 2007). On distingue 5 grandes catégories de siRNA. Les "trans-acting siRNAs"(ta-siRNA), les "natural antisense transcript-derived siRNAs" (nat-siRNA), les "endogenous siRNAs (endo-siRNA)", les p4/p5-siRNA et les siRNA nécessaires à la méthylation de l'ADN *via* RDR2 (NERD siRNAs) (figure 4) (Martínez de Alba et al., 2013).

##### Les ta-siRNA

Les ta-siRNA dérivent de longs transcrits non-codants contenant les sites d'hybridation de quelques miRNA spécifiques. Ils sont transcrits à partir de gènes spécifiques nommés TAS ("trans-acting siRNA"). Quatre *loci* "TAS" et 3 miRNA sont impliqués dans la biogenèse des ta-siRNA. Les TAS sont transcrits par la polymérase III, puis clivés grâce à l'un des 3 miRNA en générant alors des fragments d'ARN qui seront amplifiés par RDR6 ("RNA-dependent RNA polymerase 6") (figure 4). Ces fragments d'ARN doubles brins sont ensuite maturés par DCL4 qui clivera leurs extrémités (Xie et al., 2005). Ces fragments, appelés ta-siRNA pourront se propager à travers la plante et aller jusqu'à diffuser le phénomène d'interférence ARN dans tout l'organisme (Martínez de Alba et al., 2013). On distingue 4 types de ta-siRNA chez *A. thaliana*, définis en fonction du *locus* dont ils proviennent (TAS1, 2, 3 et 4) et des miRNA provoquant leurs clivage (les miRNA 173, 390 et 828) (Martínez de Alba et al., 2013) (figure 4).



Les ta-siRNA 1 et 2 sont produits *via* le clivage des transcrits *TAS1* et *2* par AGO1 portant le miRNA 173. Les produits de clivage sont ensuite stabilisés par SGS3 (“Supressor of Gene Silencing 3”) tandis que RDR6 en synthétise le brin complémentaire. Enfin, DCL4 génère des siRNA de 21nt complémentaires au mRNA cible (Allen et Howell, 2010; Martínez de Alba et al., 2013). Les ta-siRNA *TAS4* sont, quant à eux, produits par le clivage d’un transcrit *TAS4* par le complexe composé de AGO1 et du miRNA 828. Ils sont ensuite maturés selon le même procédé que les ta-siRNA 1 et 2. Les ta-siRNA 3 possèdent 2 sites de fixations AGO7-miRNA 390, après clivage par ces 2 complexes, les fragments suivront le même processus que les ta-siRNA 1, 2 et 4 (Allen et Howell, 2010). L’ensemble des ta-siRNA joue un rôle dans l’interférence ARN, ils induisent donc la dégradation du transcrit qu’ils ciblent (PTGS) (Martínez de Alba et al., 2013).

#### Les nat-siRNA

Le nat-siRNA (“natural antisense transcript-derived siRNA”) sont de petits ARN de 21nt maturés par DCL 1 ou 2 et impliqués dans la régulation de gènes de réponse aux stress (Sarkies and Miska, 2014). Ce type de siRNA dérive d’ARN doubles brins formés par l’hybridation de deux transcrits NAT (“Natural Antisens Transcript”). Les cis-NAT sont codés par le génome à partir du même *locus*, tandis que les trans-NAT sont portés par 2 *loci* différents. Ils sont tout d’abord chargés par une protéine AGO encore non-identifiée à ce jour, et servent à la dégradation du transcrit qu’ils ciblent. Puis les produits de ce clivage sont convertis en dsRNA via l’action de la polymérase IV, RDR6 et SGS3 dont les extrémités sont méthylés par HEN1 afin de former de nouveau des siRNA (figure 4). On estime qu’environ 10% des gènes d’*Arabidopsis* sont des cis-NAT (Martínez de Alba et al., 2013), bien que très peu aient effectivement été décrits. Les nat-siRNA compteraient parmi les plus grands régulateurs de l’expression du génome des plantes, notamment lors de réponses adaptatives aux stress biotiques ou abiotiques (Budde et al., 2008).

#### Les endoIR-siRNAs

Les endoIR-siRNAs (Endogenous IR-derived siRNAs) dérivent de tiges-boucles transcrites à partir de *loci* dispersés à travers tout le génome d’*A.*



*thaliana* (figure 4). Les tiges boucles générées sont plus grandes que celles décrites pour des gènes de miRNA. Cependant, comme les pri-miRNA, les endoIR-siRNAs se replient et forment une tige-boucle parfaite ou quasi-parfaite, ce qui en fait des candidats pour la maturation *via* DCL2, 3 et 4, mais pas DCL1 (Martínez de Alba et al., 2013).

Deux de ces tiges boucles ont été décrites, IR71 et IR2039. Elles produisent des siRNA de 21, 22 et 24nt qui sont tous capables de se répandre dans la plante *via* le tissu vasculaire. La fonction des endoIR-siRNAs est actuellement inconnue, cependant on suppose que ces siRNA sont des acteurs de l'interférence ARN (PTGS), et qu'ils jouent un rôle dans l'adaptation à l'environnement et la mémoire trans-générationnelle (Dunoyer et al., 2010).

#### NERD siRNA

Ce sont des siRNA de 21 et 22nt dont la formation est indépendante des enzymes du TGS ("Transcriptionnal Gene Silencing") (*e.g* RDR2, AGO4) mais dépendante des enzymes du PTGS comme RDR6 et AGO2 (figure 4). Pourtant, ils jouent un rôle dans le TGS, car ils servent à la méthylation de *loci* ADN grâce à la protéine NERD ("Needed for RDR2 independent DNA methylation") qui méthylera la lysine 9 de l'histone H3, rendant alors accessible les *loci* pour le siRNA et le complexe de méthylation (Pontier et al., 2012).

#### Les p4/p5-siRNAs (DNA-dependent RNA polymerases IV/V-derived siRNAs)

Anciennement appelés *casiRNA* (cis-acting siRNA), les p4/p5-siRNAs sont de petits ARN de 24nt impliqués dans le remodelage de la chromatine. Ils sont codés par le génome et sont transcrits par les polymérase IV et V qui transcrivent le premier brin d'ARN, puis RDR2 (RNA Dependent RNA polymerase) synthétisera le brin complémentaire. Enfin DCL3 convertit ces doubles brins d'ARN en siRNA (figure 4) (Vazquez et al., 2004). Ils sont la classe de siRNA la plus abondante chez la plante et ils agissent à de nombreux niveaux pour modeler la chromatine et induire la TGS dans un processus nommé la RdDM (RNA directed DNA methylation) (Pour revue, Martínez de Alba et al., 2013).

	Embranchement	Espèces	Nombres de tDNA	Taille du génome (Mgb)	
Eucaryotes	Vertébrés	<i>Homo sapiens</i>	505 513 *	3 400	
		<i>Mus musculus</i>	430 432 *	3 400	
	Insectes	<i>Drosophila melanogaster</i>	298	128	
	Echinodermes	<i>Strongylocentrotus purpuratus</i>	1265	89	
	Hemosporidés	<i>Plasmodium falciparum</i>	35	21,8	
	Leishmanie	<i>Leishmania major</i>	82	32,8	
	Champignons	<i>Schizosaccharomyces pombe</i>	186	12,5	
		<i>Saccharomyces cerevisiae</i>	286	12	
	Ostreophytes	<i>Ectocarpus siliculosus</i>	144	214	
		<i>Phaeodactylum tricornutum</i>	52	27	
	Glaucophytes	<i>Cyanoptera paradoxa</i>	106	16,5	
		<i>Ostreococcus tauri</i>	44	12,5	
	Chlorophytes	<i>Chlamydomonas reinhardtii</i>	259	111	
		<i>Arabidopsis thaliana</i>	598	128	
	Trachéophytes	<i>Brachypodium distachyon</i>	477	272	
		<i>Oryza sativa</i>	738	430	
<i>Vitis vinifera</i>		550	475		
Archéo-bactéries	Thaumarchaeota	<i>Candidatus</i> <i>Symbiocaldococcus</i>	46	2,05	
	Euryarchaeota	<i>Pyrococcus furiosus</i>	46	1,9	
Bactéries	Gammaproteobactéries	<i>Escherichia coli</i>	79-101 *	4,64	
		Firmicutes	<i>Staphylococcus aureus</i>	52-61 *	2,8
			<i>Clostridium botulinum</i>	70-81 *	4

**Figure 5 : Nombres de tDNA présents dans le génome nucléaire comparé à la taille de ce génome chez divers organismes.** Lorsque l'on compare le nombre de tDNA à la taille du génome, les hémospodidés ainsi que les *Leishmania* sont ceux qui possèdent le moins de tDNA (données d'après <http://getnadb.ucsc.edu> et Michaud et al., 2011). \* : Plusieurs annotations étant disponibles, le nombre de tDNA minimum et maximum identifié chez chacun de ces organismes figure dans ce tableau.

#### 2.2.4 – piRNAs

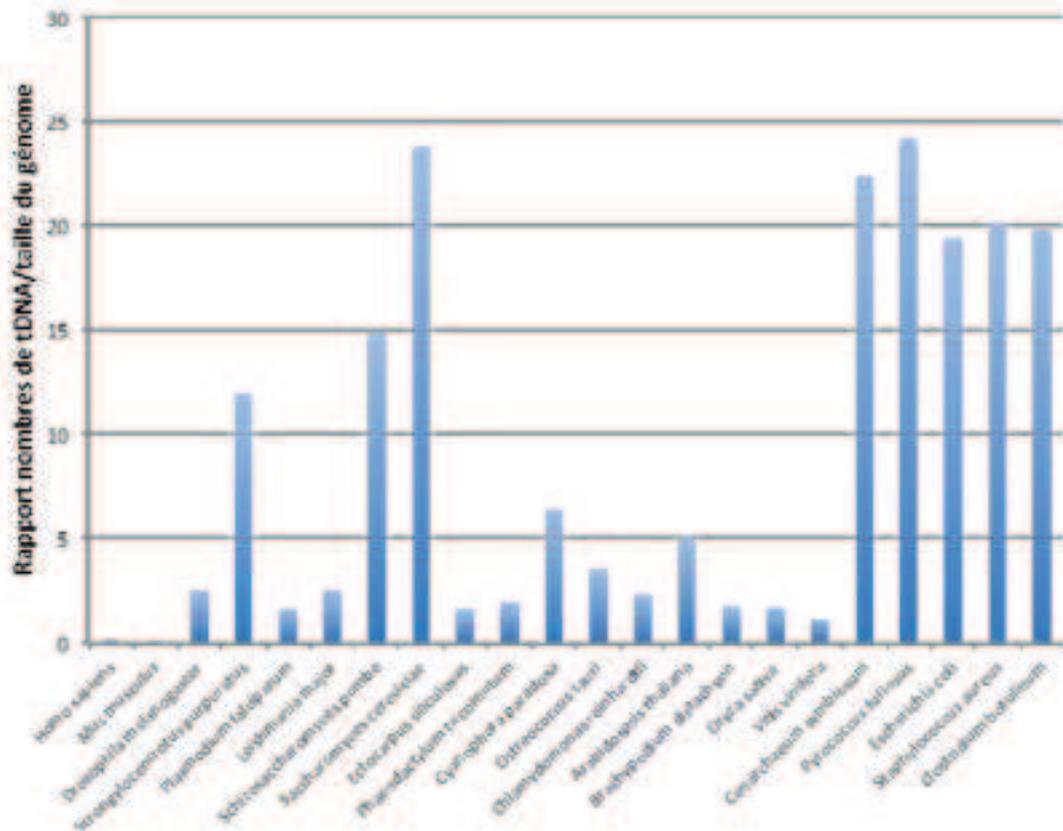
Les piRNA (PIWI-interacting ARN) représentent la classe d'ARN non codant la plus abondante. Ils ne sont présents que dans les cellules animales ainsi que les protozoaires. Les piRNA sont appelés ainsi car ils forment des complexes avec des protéines dénommées PIWI. Ces complexes sont impliqués dans des mécanismes de modifications épigénétiques de l'ADN régulant l'expression des rétro-transposons (Siomi et al., 2011). On distingue les piRNA des autres petits ARN car ils sont légèrement plus grand, 21-24nt pour les miRNA, 26-31nt pour les piRNA, mais aussi parce que leurs séquences sont moins conservées que celles des miRNA (Seto et al., 2007).

Les piRNA se trouvent sous forme de "clusters" de tailles très diverses (de 100bp à 100kbp) pouvant coder jusqu'à plusieurs milliers de piRNA (O'Donnell et Boeke, 2007). Leur biogenèse est assez mal connue. On sait cependant qu'ils jouent un rôle dans l'interférence ARN, où ils participent à la formation du complexe RISC *via* les protéines PIWI dont l'activité est similaire aux protéines argonautes. Il existe 3 familles de protéines PIWI chez la souris : MIWI, MIWI2 et MILI (O'Donnell et Boeke, 2007), toutes participent à la spermatogenèse. Les piRNA dirigent le complexe RISC vers sa cible, souvent un transposon et induisent alors sa sous-expression (Aravin et al., 2008). Certains piRNAs peuvent être transmis par la mère (Aravin et al., 2008) et d'après certains résultats obtenus chez la drosophile, ceux-ci pourraient intervenir dans des processus de modifications épigénétiques (Brennecke et al., 2008).

#### 2.2.5 - Les SINEs

Les SINEs de plantes, comme chez d'autres organismes, sont un groupe de rétrotransposons non-autonomes qui sont amplifiés et propagés par les éléments codés par les LINEs (Schulman, 2013). Les SINEs incluent un grand nombre de transcrit de la polymérase III tel que les tRNA ou les rRNA qui ont été rétro-transcrits puis intégrés au génome. Leurs tailles varient entre 75 et 662pb.

Les SINEs dérivant de tRNA en possèdent l'extrémité 5' mais sont homologues au LINEs en 3'. Ils sont très répandus chez les plantes avec toutefois une prévalence chez les *Solanaceae* (Schulman, 2013; Wenke et al., 2011).



**Figure 6 : Rapport entre le nombre de tDNA et la taille du génome.** Lorsque l'on compare le nombre de tDNA à la taille du génome, les vertébrés sont ceux qui possèdent le moins de tDNA par Mpb, après quoi viennent les hémisporidés ainsi que les *Leishmania*, puis les ochrophytes (algues brunes) et enfin les trachéophytes, à l'exception d'*A. thaliana* (données d'après <http://etrnadb.ucsc.edu>) (Michaud et al., 2011).

### 2.2.6 – Les tRNA

Les tRNA représentent l'élément central de mon sujet de thèse et font l'objet d'une partie distincte.

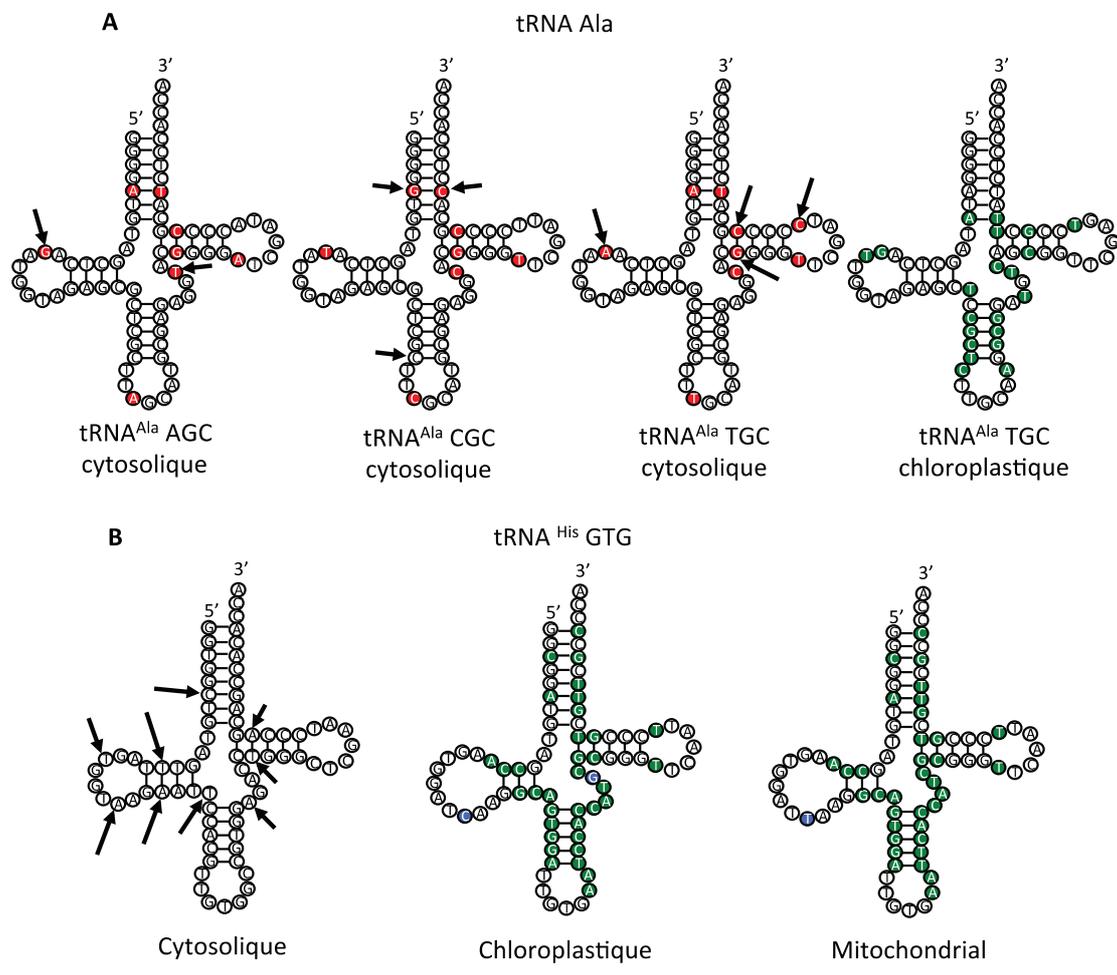
## 3 - Caractéristiques et fonctions des tRNA

Les tRNA sont, avec les rRNA, les molécules les plus transcrites et produites par la cellule. Leurs tailles varient généralement entre 75 et 90 nucléotides. Cependant des tRNA beaucoup plus petits ont été identifiés dans les mitochondries de métazoaires (*e.g. Caenorhabditis elegans* ou *Ascaris suum*), (43-50nt) chez lesquels les tiges-boucles D et T sont réduites à une boucle de 5 à 8 nucléotides seulement, ou dans certains cas plus extrêmes totalement absentes (*e.g. R. culicivora*) (Wende et al., 2014).

Les tRNA ont pour rôle d'apporter au ribosome l'acide aminé nécessaire à la synthèse des protéines, tout en respectant la séquence, une série de codons, imposée par le mRNA. C'est grâce à l'anticodon, série de trois nucléotides portés par les tRNA, que ceux-ci peuvent identifier l'acide aminé à ajouter en s'hybridant au codon porté par le mRNA. Ainsi chaque codon ne code que pour un acide aminé, et l'ensemble des couples codon/acide aminé définit le code génétique universel. Si un codon ne code que pour un acide aminé, l'acide aminé peut par contre correspondre à plusieurs codons. Il existe ainsi des tRNA qui portent le même acide aminé, mais pas le même anticodon. On parlera alors de tRNA isoaccepteurs. Il existe aussi des tRNA qui ne possèdent pas la même séquence mais possèdent le même anticodon, on parlera alors de tRNA isodécodateurs.

### 3.1 - Organisation des gènes de tRNA chez les plantes

Les eucaryotes sont parmi les organismes présentant le plus grand nombre de gène de tRNA (tDNA). On compte ainsi 599 tDNA nucléaires chez *A. thaliana* (Cognat et al., 2013) contre 513 chez l'homme (Chan et Lowe, 2009). Pour comparaison, le génome d'*E.coli* possède entre 79 et 101 tDNA selon les sources. Des exceptions existent cependant, comme chez le champignon, *Encephalitozoon cuniculi* qui ne possède que 46 tDNAs, les parasites *P.*



**Figure 7 : Les tRNA<sup>Ala</sup> et tRNA<sup>His</sup> chez *A. thaliana*.** Les nucléotides différents entre les isoaccepteurs sont indiqués en **rouge**. Les positions différentes entre deux isoaccepteurs (un cytosolique et l'autre d'origine plastidiale ou mitochondriale) sont affichées en **vert** tandis que les positions n'étant pas similaires entre les différents isodécodateurs sont indiquées par une flèche **noire**. Chez *A. thaliana* certains isoaccepteurs, à l'image des tRNA<sup>Ala</sup> sont très semblables. D'autres isoaccepteurs sont très différents selon le compartiment. Ainsi les tRNA<sup>His</sup> GTG cytosoliques sont très différents de leurs homologues plastidiaux et mitochondriaux. Ces deux derniers sont par ailleurs très similaires, car seul 2 nucléotides les différencient (**bleu**) (Cognat et al. 2013).

*falciparum* et *L. Major* chez qui le nombre de tDNA est respectivement de 35 et 82 gènes ou encore des microalgues vertes comme *Osteococcus tauri* (43 tDNA) ou algues brunes comme *Phaeodactylum tricornutum* (52 tDNA) (figure 5 et 6) (Chan and Lowe, 2009; Cognat et al., 2013). L'augmentation du nombre de tDNA et de la complexité de leurs transcrits tout au long de l'évolution est concomitante à l'apparition d'enzyme de modification des tRNA, comme les adénosines désaminases chez les eucaryotes. Ces modifications augmentent souvent les capacités de décodage d'un même tRNA (Grosjean et al., 2010; Spieth and Lawson, 2006).

Chez les plantes, les tRNA peuvent être codés par 3 génomes différents, le génome nucléaire à l'origine des tRNA localisés dans le cytosol et de certains tRNA importés dans la mitochondrie, le génome mitochondrial comme chez la quasi totalité des eucaryotes et le génome plastidial spécifique aux organismes photosynthétiques. Nombre de ces tRNA possède une grande homologie de séquences que ce soit pour des tRNA cytosoliques, plastidiaux ou mitochondriaux. Ainsi, il arrive que seulement 1 ou 2 nucléotides différencient des tRNA provenant d'un même compartiment (dans le cas de tRNA isodécodateurs, et dans une moindre mesure de tRNA isoaccepteurs) (figure 7), ou de compartiments différents (*e.g.* tRNA mitochondriaux provenant de l'expression de gènes chloroplastiques insérés dans le génome mitochondrial) (figure 7).

### 3.1.1 – Génome nucléaire

Chez les angiospermes, le nombre de tDNA varie entre 500 et 600 d'une espèce à une autre (Michaud et al., 2011). Ainsi, le génome nucléaire d'*A. thaliana* possède 599 tDNA codant pour 45 tRNA isoaccepteurs différents. Ils sont moins nombreux chez les bryophytes ou les algues (figure 5 et 6).

Les tDNA sont généralement dispersés à travers le génome des plantes et il est rare qu'ils soient organisés en clusters. Quelques exceptions existent cependant. Ainsi chez *A. thaliana*, le chromosome 1 porte deux clusters contenant respectivement 27 répétitions en tandem du tDNA proline, ainsi que 27 répétitions en tandem de tDNA tyrosine et sérine organisés de la façon suivante, Tyr-Tyr-Ser (Theologis et al., 2000). Ces régions en tandem entraînent

Isoaccepteur	Anticodon	Nucléaires	Plastidiaux	Mitochondriaux
Ala	AGC	16	-	(i)
	CGC	7	-	(i)
	TGC	10	2	(i)
Arg	ACG	9	2	(i)
	CCG	4	-	(i)
	CCT	8	-	(i)
	TCG	6	-	(i)
	TCT	9	1	(i)
Asn	GTT	15	2	1
Asp	GTC	25	1	1
Cys	GCA	14	1	1
Gln	CTG	9	-	-
	TTG	7	1	1
Glu	CTC	13	-	-
	TTC	10	1	1
Gly	CCC	5	-	-
	GCC	22	1	1
	TCC	12	1	-
His	GTG	10	1	1
Ile	AAT	17	-	-
	CAT	-	2	1
	GAT	-	2	-
	TAT	5	-	-
Leu	AAG	11	-	(i)
	CAA	10	2	(i)
	CAG	3	-	(i)
	TAA	6	1	(i)
	TAG	9	1	(i)
Lys	CTT	17	-	-
	TTT	13	1	2
Met_e	CAT	11	1	1 *
Met_i	CAT	10	1	1
Phe	GAA	16	1	-
Pro	AGG	15	-	-
	CGG	5	-	-
	TGG	44	1	1
Ser	AGA	37	-	-
	CGA	4	-	-
	GGA	-	1	1
	GCT	11	1	2
	TGA	7	1	2
Thr	AGT	10	-	(i)
	CGT	5	-	(i)
	GGT	-	1	-
Trp	TGT	8	1	(i)
	CCA	13	1	1(i) *
Tyr	GTA	70	1	4
Val	AAC	15	-	(i)
	CAC	8	-	(i)
	GAC	-	2	-
	TAC	7	1	(i)

**Figure 8 : Nombres de tDNA présents dans les différents compartiments d'*A. thaliana*.** Les symboles (-) indiquent des tRNA qui ne sont pas codés par le génome du compartiment. Le symbole (i) indique les tRNA importés dans le compartiment (figure d'après Cognat *et al.*, 2013). \*: Dans le cas des gènes mitochondriaux codants pour les tRNA<sup>Trp</sup> et <sup>Met-e</sup> des expériences de Northern blot suggèrent qu'ils ne sont pas exprimés (Duchêne and Maréchal-Drouard, 2001).

une augmentation importante des tDNA proline et sérine dans le génome d'*A. thaliana* (Michaud et al., 2011). Chez *A. thaliana*, quatre-vingt un tDNA possèdent un intron. Ils codent essentiellement pour les tRNA<sup>Tyr</sup> et <sup>Mete</sup> (Méthionine élongateur) (figure 8). Le pourcentage et l'identité des tDNA possédant un intron est beaucoup plus important chez les algues vertes et rouges ainsi que les microalgues (Cognat et al., 2013).

### 3.1.2 - Génome des organites

#### 3.1.2.1 - Génome mitochondrial

Le génome mitochondrial de plantes supérieures possède entre 17 et 29 tDNA (Cognat et al., 2013) répartis sur l'ensemble du génome. Ainsi, chez *A. thaliana* on dénombre 23 gènes de tRNA dans le génome mitochondrial, codant pour seulement 17 tRNA isoaccepteurs différents (Michaud et al., 2011). Onze de ces tRNA sont dits "natifs", c'est à dire codés par un tDNA provenant du génome de l'alpha-protéobactérie ancestrale à l'origine des mitochondries. Quatre de ces tRNA proviennent de l'expression de fragments d'ADN chloroplastiques insérés dans le génome mitochondrial au cours de l'évolution. Il s'agit des tRNA<sup>Asp, Asn, His</sup> et <sup>Ser(GGA)</sup>. Ces tRNA dits de type chloroplastique sont exprimés de la même manière que les tRNA mitochondriaux natifs (Duchêne and Maréchal-Drouard, 2001). De plus, deux tDNA correspondant aux tRNA<sup>Mete</sup> et <sup>Trp</sup> chloroplastiques sont également présents sur le génome mitochondrial. Des expériences de Northern blot suggèrent qu'ils ne sont pas (ou du moins très peu) exprimés dans les mitochondries d'*Arabidopsis* (Duchêne and Maréchal-Drouard, 2001).

En supposant l'existence d'appariements wobble entre la première position de l'anticodon et la troisième d'un codon, on estime que pour assurer la lecture de l'ensemble des codons du code génétique universel, un organisme a besoin d'un minimum de 23 tRNA (Phizicky et Hopper, 2010). Le nombre de tRNA codé par le génome mitochondrial est donc insuffisant, ainsi pour compenser ce manque, des tRNA codés par le génome nucléaire sont importés du cytosol (figure 8) (Michaud et al., 2011). C'est pourquoi les tRNA<sup>Ala, Arg, Leu, Thr</sup> et <sup>Val</sup> sont importés du cytosol vers la mitochondrie chez toutes les angiospermes étudiés à ce jour (Michaud et al., 2011). De plus, les gènes de tRNA ayant disparu du génome mitochondrial pouvant différer d'une espèce végétale à l'autre, le

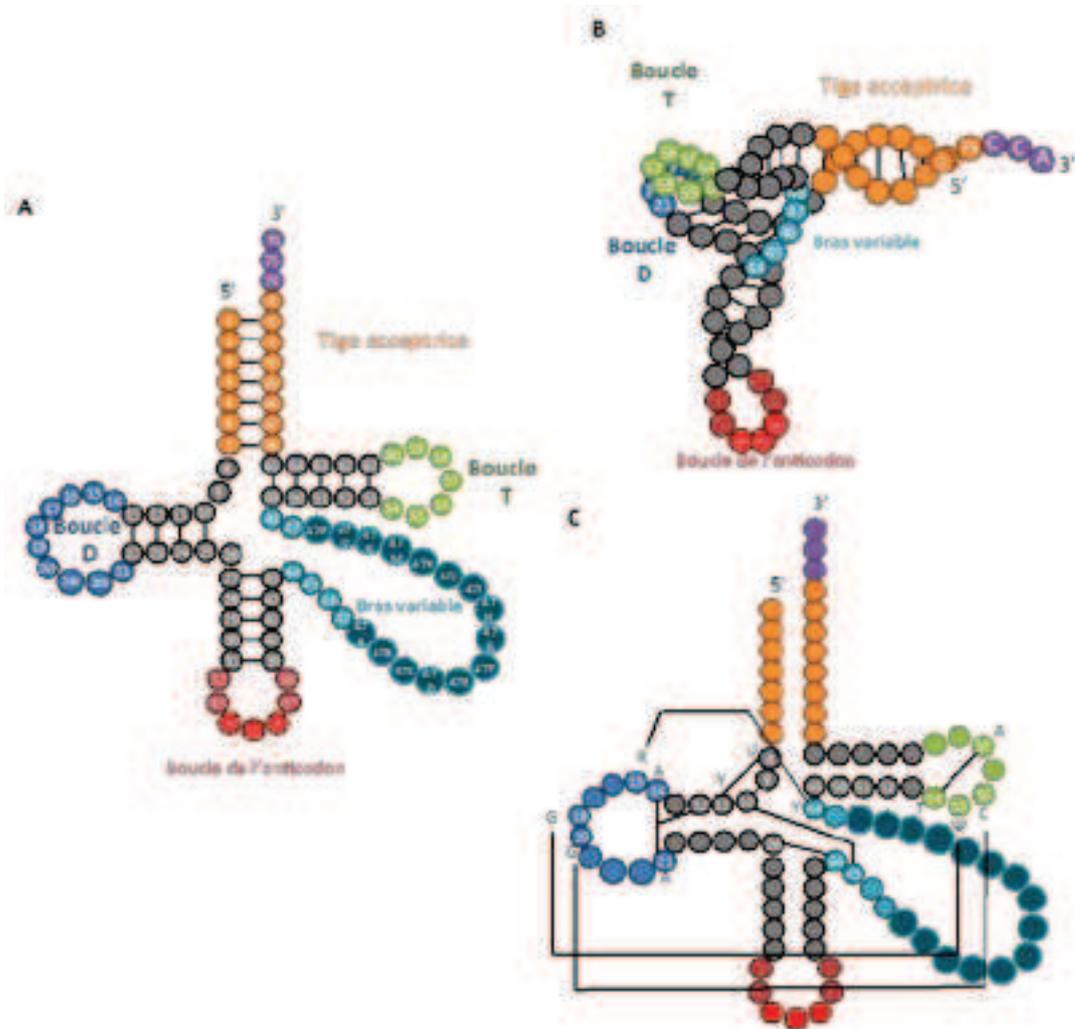


Figure 9 : Schéma de la structure secondaire (A) et tertiaire (B) d'un tRNA et des interactions permettant son repliement correct (C). Les tRNA ont généralement une taille d'environ 75 à 90 nucléotides, et on différencie 5 principales structures. Chez les eucaryotes c'est principalement la taille de la boucle variable qui est responsable de la diversité de taille des tRNA. La figure C représente les interactions intervenant dans la structure tertiaire des tRNA (figures d'après Sprinzl et al., 1998).

nombre et l'identité des tRNA importés fluctuent en fonction des espèces, c'est notamment le cas du tRNA<sup>Trp</sup> chez *A. thaliana* (figure 8), (Duchêne et al., 2009). Chez certains organismes photosynthétiques le nombre de gènes de tRNA mitochondriaux est très réduit comme l'algue verte *C. reinhardtii* où le génome mitochondrial ne code que pour 3 tRNA (Cognat et al., 2013).

### 3.1.2.2 – Génome plastidial

Chez la majorité des plantes supérieures, le chloroplaste a conservé un jeu complet de tDNA pour assurer une synthèse protéique active. Chez *O. sativa* et *P. trichocarpa*, on dénombre respectivement 38 et 39 tDNA codant pour 30 isoaccepteurs. Chez *A. thaliana* le génome plastidial contient 37 tDNA répartis sur l'ensemble du génome et codant pour 30 tRNA isoaccepteurs (figure 8), (Michaud et al., 2011).

Il est toutefois intéressant de noter que chez certains végétaux, le génome plastidial a perdu un nombre important de tDNA suggérant que l'importation de tRNA du cytosol existe également chez ces organismes, même si aucune preuve n'a été apportée à ce jour. Il s'agit en particulier de végétaux parasites d'arbre comme *Epifagus virginiana* qui ne possède que 15 tDNA plastidiaux (Wolfe et al., 1992) ou encore l'orchidée parasite *Rhizanthella garneri* dont le génome plastidial ne code que pour 9 tDNAs (Delannoy et al., 2011).

## 3.2- Structure des tRNA

### 3.2.1 – Les tRNA cytosoliques

La structure secondaire des tRNA est souvent comparée à un trèfle, tandis que leur structure tertiaire est en forme de « L ». La structure secondaire des tRNA peut être divisée en 6 domaines, la tige acceptrice, l'extrémité CCA, les régions D et T et le domaine de l'anticodon (figure 9). Les tiges-boucles D et T participent à la conformation correcte de la structure tertiaire des tRNA *via* leurs boucles qui vont se superposer l'une à l'autre ("stacking") (figure 9) (Giegé et al., 2012). Plusieurs modifications post-transcriptionnelles stabilisent la structure des tRNA. Certains tRNA possèdent un bras variable plus grand (*e.g.* tous les tRNA<sup>Leu</sup> et <sup>Ser</sup> chez *A. thaliana*) (Cognat et al., 2013).

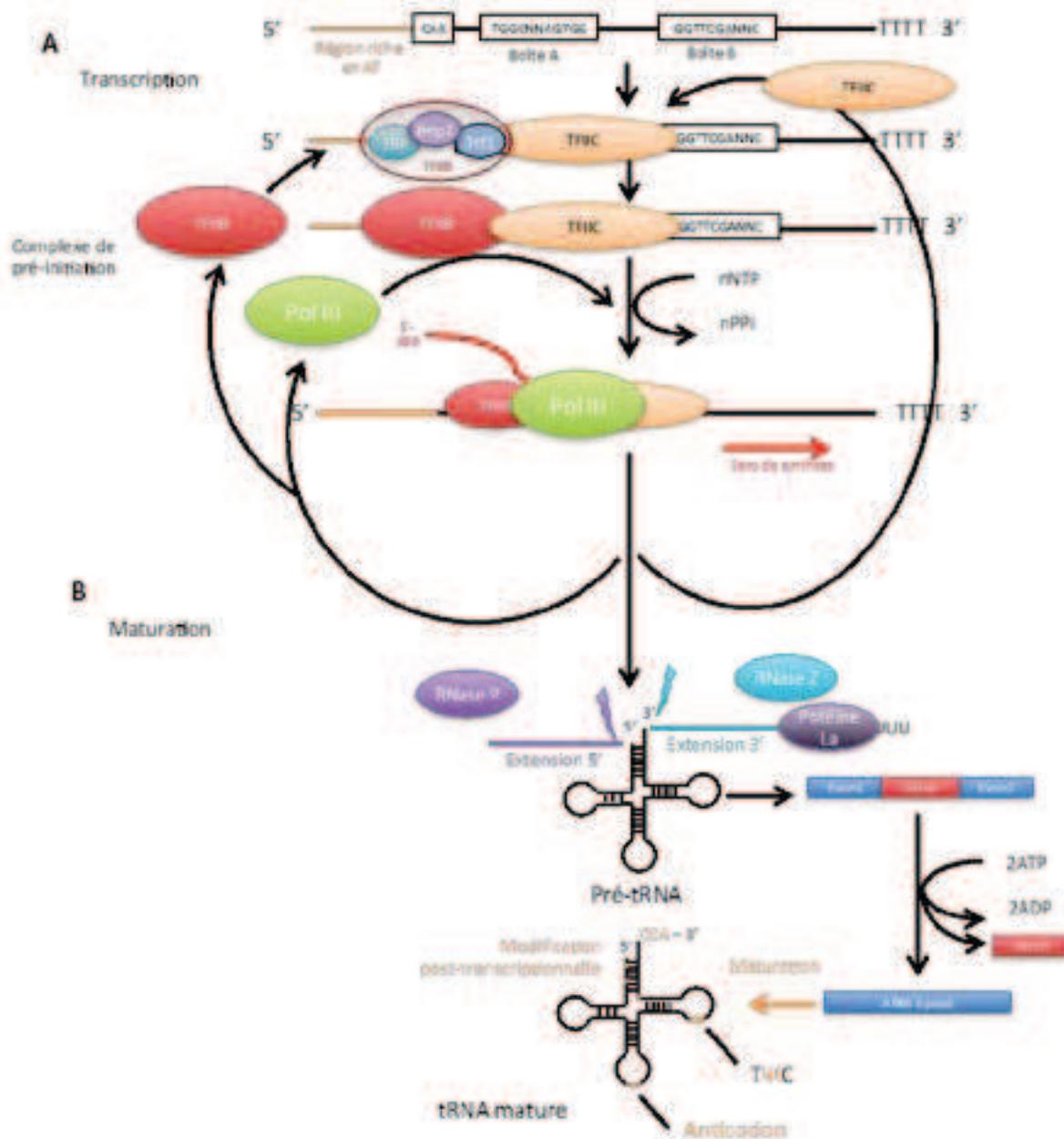


Figure 10 : Schéma de la biogenèse des tRNA dans le noyau. A, Les différentes étapes de la transcription (cf. Introduction - 3.3). B, Les différentes étapes de la maturation (cf. Introduction - 3.4).

### 3.2.2 – Les tRNA d'organites

Dans le règne végétal, les tRNA natifs du génome mitochondrial et les tRNA chloroplastiques sont très similaires (de 65% à 80% d'homologie de séquence) aux tRNA procaryotiques (Maréchal-Drouard et al., 1993). Ils possèdent la même structure générale que les tRNA cytosoliques. Les tRNA d'organites contiennent environ deux fois moins de modifications post-transcriptionnelles comparés aux tRNA cytosoliques (Chen et al., 2010). Tout comme les tRNA nucléaires, certains tRNA d'organites possèdent une grande boucle variable. Il s'agit dans ce cas des tRNA<sup>Ser, Leu</sup> et Tyr.

## 3.3 - Transcription des tRNA

La transcription des tRNA est un processus faisant appel à plusieurs séquences régulatrices et facteurs protéiques.

### 3.3.1 – Dans le noyau

Chez les eucaryotes, les tRNA sont transcrits par l'ARN polymérase III en tant que précurseur de tRNA (pré-tRNA) possédant de courtes séquences en 5' et 3' (figure 10B), (pour une revue voir Phizicky et Hopper, 2010). Dans quelques cas le pré-tRNA peut posséder un intron.

#### 3.3.1.1 – Les promoteurs internes, les boîtes A et B

La polymérase III reconnaît 2 régions promotrices internes A et B très conservées chez les eucaryotes. Ces régions sont situées à l'intérieur du gène au niveau des boucles D et T (Galli et al., 1981; Hofstetter et al., 1981) (figure 10A). La boîte A s'étend des positions 8 à 18 du tRNA mature tandis que la boîte B se situe de 30 à 60 nucléotides en aval de la boîte A. Les séquences consensus chez *A. thaliana* sont respectivement 5'-Tg/aGC/tNNAGTg/tG-3' et 5'-G/aGTTTCGANN-3'.

#### 3.3.1.2 – Séquence riche en AT et séquence "TATA-like"

Chez les plantes supérieures, des régions très riches en AT existent dans les 50 nucléotides en amont de la séquence des tDNA. On peut noter aussi une différence entre, d'une part, les tDNA de plantes mono et dicotylédones chez qui l'enrichissement en nucléotides AT atteint les 70% (e.g 64% chez *A. thaliana*) et



d'autre part les algues vertes comme *C. reinhardtii* chez qui cet enrichissement ne dépasse pas les 50%. Chez les angiospermes, ces éléments se trouvent principalement entre les positions -25 et -35.

C'est à cette position que l'on trouve fréquemment une séquence ressemblant à la "TATA box", dont il a été démontré qu'elle contribuait à l'efficacité de transcription des gènes transcrits par la polymérase III grâce à la fixation de la TBP ("TATA-box binding protein") qui aidera au recrutement de la polymérase III (Dieci et al., 2006, 2007; Giuliadori et al., 2003; Hamada et al., 2001; Schramm et Hernandez, 2002). Chez les algues vertes, comme chez les métazoaires, ces motifs TATA-like ne sont pratiquement jamais trouvés.

#### 3.3.1.3 – Le motif CAA

Environ 65% des tDNA d'angiospermes sont précédés d'un triplet CAA directement en amont (-1 à -10) de leurs séquences. Chez *A. thaliana*, 78% des tDNA présentent au moins un motif CAA dans les 50 nucléotides en amont. La présence d'un tel motif est beaucoup plus rare chez les algues (*e.g.* 13% chez *C. reinhardtii*). C'est chez *N. benthamiana*, qu'il a été montré que ces triplets CAA étaient les sites d'initiation de la transcription de la polymérase III (Yukawa et al., 2000).

#### 3.3.1.4 – La région polyU

Chez les eucaryotes, la transcription des tRNA se termine au niveau d'une courte queue d'au moins 4U codés par le génome. C'est à ce niveau que la polymérase III est décrochée et la protéine La est recrutée permettant ainsi la stabilisation du transcrit néosynthétisé.

Plusieurs exceptions ont été trouvées dans le monde végétal. La première concerne l'algue verte *C. reinhardtii*, chez qui beaucoup de séquences de tDNA sont co-transcrites et ne se terminent pas par une extension polyT. Le second chez la mousse *Physcomitrella patens*, chez qui seulement 40% des tDNA possèdent une extension polyT dans les 25nt en aval. Chez *A. thaliana*, l'absence de queue polyU est caractéristique d'un tRNA co-transcrit avec un snoRNA à boîte C/D (cf. 2.2.1.1) (Kruszka et al., 2003).



### 3.3.1.5 – Les facteurs protéiques

Afin d'initier la transcription, la polymérase III utilise des promoteurs internes hautement conservés nommés boîtes A et B (Galli et al., 1981). Ces promoteurs permettent le recrutement de TFIIB et TFIIC, TFIIB étant composés des Bdp1, Trf1 ainsi que de la TBP ("TATA box Binding Protein") qui reconnaît les séquences riches en AT. Le tout forme alors le complexe de pré-initiation. Une fois ce complexe formé, la polymérase III sera recrutée, l'ensemble formera le complexe d'initiation. La transcription pourra alors débiter à partir du motif CAA. Une fois la transcription terminée, la polymérase III laissera une série d'uridine que la protéine La reconnaîtra stabilisant le tRNA qui devra subir plusieurs étapes de maturation. Il devra notamment être débarrassé de ses extensions 5' et 3' par les RNases P et Z, respectivement, puis épissé s'il contient un intron. Des modifications post-transcriptionnelle seront ensuite ajoutées ainsi que l'extrémité CCA terminale (figure 10A) (Michaud et al., 2011; Phizicky and Hopper, 2010).

D'autres éléments situés en amont et en aval des tDNA ont aussi été caractérisés comme essentiels à l'efficacité de transcription, notamment chez les plantes supérieures. Une description de ces différents éléments (pour une revue voir Michaud et al., 2011) ainsi que des différentes étapes et facteurs protéiques nécessaires à la transcription des tDNA est présentée ci-dessous.

## 3.3.2 – Dans les organites

### 3.3.2.1 - Dans les mitochondries

Il n'existe pas d'ARN polymérase spécifique pour la transcription des gènes de tRNA mitochondriaux. Chez les plantes, l'ensemble du génome mitochondrial est transcrit par deux ARN polymérases de type NEP ("Nuclear-Encoded Polymerase"). Ces polymérases reconnaissent des régions promotrices dispersées tout le long du génome. Plusieurs régions polycistroniques existent et de nombreux gènes comportent plusieurs régions promotrices. La séquence promotrice est caractérisée par un motif nucléotidique CRTA mais de nombreuses variantes à ce motif existent (Fey and Maréchal-Drouard, 1999; Kühn et al., 2005).



### 3.3.2.2- Dans les chloroplastes

Comme pour les mitochondries, il n'existe pas de promoteurs spécifiques pour la transcription des gènes de tRNA. Beaucoup de gènes sont transcrits sous forme polycistronique et beaucoup d'entre eux possèdent plusieurs régions promotrices. Les chloroplastes possèdent deux types d'ARN polymérase. Premièrement, l'ARN polymérase PEP ("Plastidial Encoded Polymérase") reconnaît les promoteurs de type PEP qui ressemblent fortement aux promoteurs d'*E. coli* avec des régions promotrices en position -35 (TTGACA) et -10 (TATAAT) (Maréchal-Drouard et al., 1993). Deuxièmement, comme chez les mitochondries végétales, on retrouve deux NEP ARN polymérases. Il existe deux types de promoteurs NEP-dépendants. Le type I possède un motif YRTA, tandis qu'aucun motif n'a encore été défini pour le type II. Les précurseurs des tRNA plastidiaux peuvent être mono ou polycistroniques et dépendent des deux types d'ARN polymérases (Hanaoka et al., 2003; Leal-Klevezas et al., 2000; Shiina et al., 2005).

## 3.4 – Maturation des tRNA

### 3.4.1 – Maturation des extrémités 5' et 3'

Afin de produire un tRNA mature, la cellule doit éliminer les séquences en 5' et en 3' des pré-tRNA.

Chez *A. thaliana* le clivage en 5' est catalysé par une famille d'enzyme de type RNase P appelée PRORP ("Proteinaceous RNase P"). Ces RNases P sont uniquement protéique et ne possèdent pas d'ARN catalytique contrairement aux RNases P ribonucléoprotéiques initialement découverte par Altman *et al.* (Stark et al., 1978). Trois protéines PRORP sont présentes, PRORP1 localisée dans le chloroplaste et la mitochondrie, PRORP2 et 3 localisées dans le noyau (Gutmann et al., 2012). Ces RNases P font partie de la famille des protéines PPR ("Pentatrico Peptide Repeat").

Le clivage du côté 3' quand à lui est catalysé par la RNase Z (Vogel et al., 2005). Une famille de quatre RNases Z existe chez *Arabidopsis*. L'une d'elle clive les pré-tRNA dans le noyau, deux sont localisées dans les mitochondries, la localisation de la quatrième RNase Z est incertaine (Canino et al., 2009).



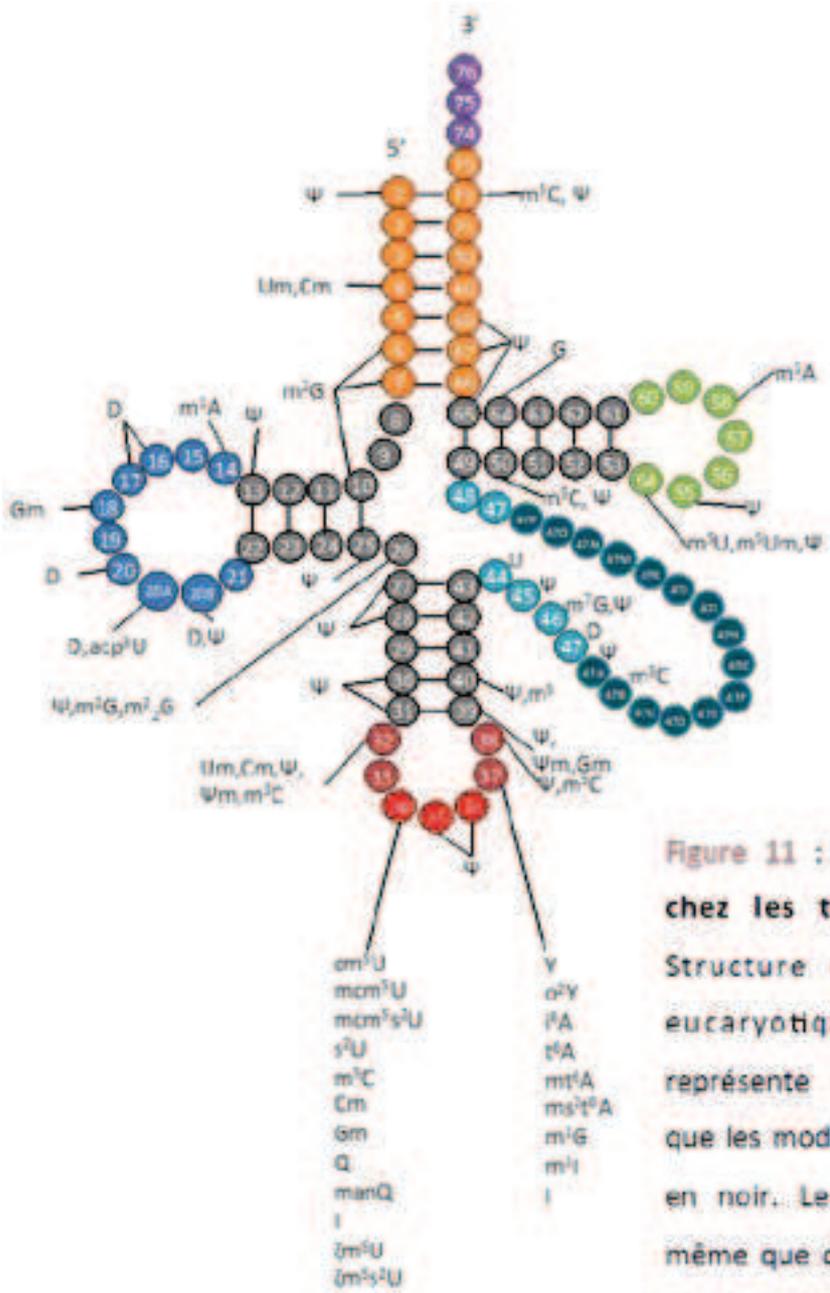
### 3.4.2 – La queue CCA

Une autre étape importante de la modification des tRNA concerne l'ajout du triplet CCA qui est effectué par une tRNA-nucléotidyl transférase (Deutscher, 1982). Chez les plantes, la même enzyme est localisée dans les trois compartiments cellulaires, noyau, mitochondrie et chloroplaste (von Braun et al., 2007). Lors de cette étape, le contrôle qualité des tRNA peut avoir lieu. Si par exemple la tige acceptrice porte des mutations, le tRNA sera marqué d'un second CCA, pour ensuite être pris en charge et dégradé par l'exosome (Wilusz et al., 2011). Cette étape de contrôle qualité a été démontrée chez *S. cerevisiae*. Elle est décrite plus en détail dans la revue située en section 4. L'existence d'un processus identique chez les plantes n'est pas démontrée à ce jour.

### 3.4.3 – Épissage des pré-tRNA

À l'heure actuelle, seul le mécanisme d'épissage des introns de levure a été caractérisé en détail. Il nécessite l'action de trois enzymes. Deux enzymes vont tout d'abord cliver les 2 extrémités de l'intron menant à la formation de 2 moitiés de tRNA possédant une extrémité 5'-OH et une autre 3'-PO<sub>4</sub> cyclique (Budde et al., 2008; Trotta et al., 1997; Xue et al., 2006). Une ligase (Trl1) va alors modifier puis liguer les 2 extrémités de chaque exon, menant le phosphate, originalement situé à l'extrémité 5' du site d'épissage, en position 2'. Afin de terminer l'épissage la même enzyme viendra supprimer ce groupement 2' résiduel. L'épissage des introns est réalisé à la surface des mitochondries (Hopper and Shaheen, 2008; Huh et al., 2003; Yoshihisa et al., 2007). Ceci implique que les pré-tRNA doivent sortir du noyau pour subir cette étape de maturation, où là aussi, le contrôle qualité des pré-tRNA peut être réalisé.

Concernant les tRNA d'origine nucléaire chez les plantes supérieures, seuls les tRNA<sup>Tyr</sup> et <sup>Met-e</sup> possèdent des introns. Ces introns font entre 11 et 25 nucléotides de long et se situent entre les positions 37 et 38 du tRNA mature. L'épissage des introns de pré-tRNA de plantes a lieu dans le noyau (Michaud et al., 2011) et nécessite l'action de 3 enzymes. Une endonucléase, composée des sous-unité AtSen1 et AtSen2, la ligase AtRL ainsi que la phosphotransférase AtPt (Englert et al., 2007). L'efficacité d'épissage dépend fortement de la séquence



**Figure 11 : Nucléosides modifiés chez les tRNA eucaryotiques.** Structure en trèfle du tRNA eucaryotique. Chaque cercle représente un nucléotide tandis que les modifications sont inscrites en noir. Le code couleur est le même que celui utilisé en figure 8 (figure d'après Chen et al., 2010).

nucléotidique et de la structure du pré-tRNA (Akama et al., 2000; Stange et al., 1991). Chez les algues vertes, le nombre de tRNA possédant un intron est beaucoup plus important et les caractéristiques structurales nécessaires à l'épissage semble différentes (Michaud et al., 2011).

Aucun tRNA d'origine mitochondriale ne possède un intron. Par contre plusieurs tRNA codés par le génome plastidial possèdent de très grands introns de groupe II (Longevialle et al., 2010). L'épissage de ces introns nécessite de nombreux facteurs protéiques (Longevialle et al., 2010).

### 3.4.5 – Autres modifications des tRNA

Les ARN sont des molécules généralement modifiées (pour une revue Machnicka et al., 2013). Les tRNA sont ceux possédant le plus de modifications. En 2005, un total de 92 types de modifications post-transcriptionnelles de tRNA a été dénombré (<http://www.biochem.ncsu.edu/RNAmods>), (figure 11). Les tRNA eucaryotiques sont au moins deux fois plus modifiés que les tRNA procaryotiques ou d'organites.

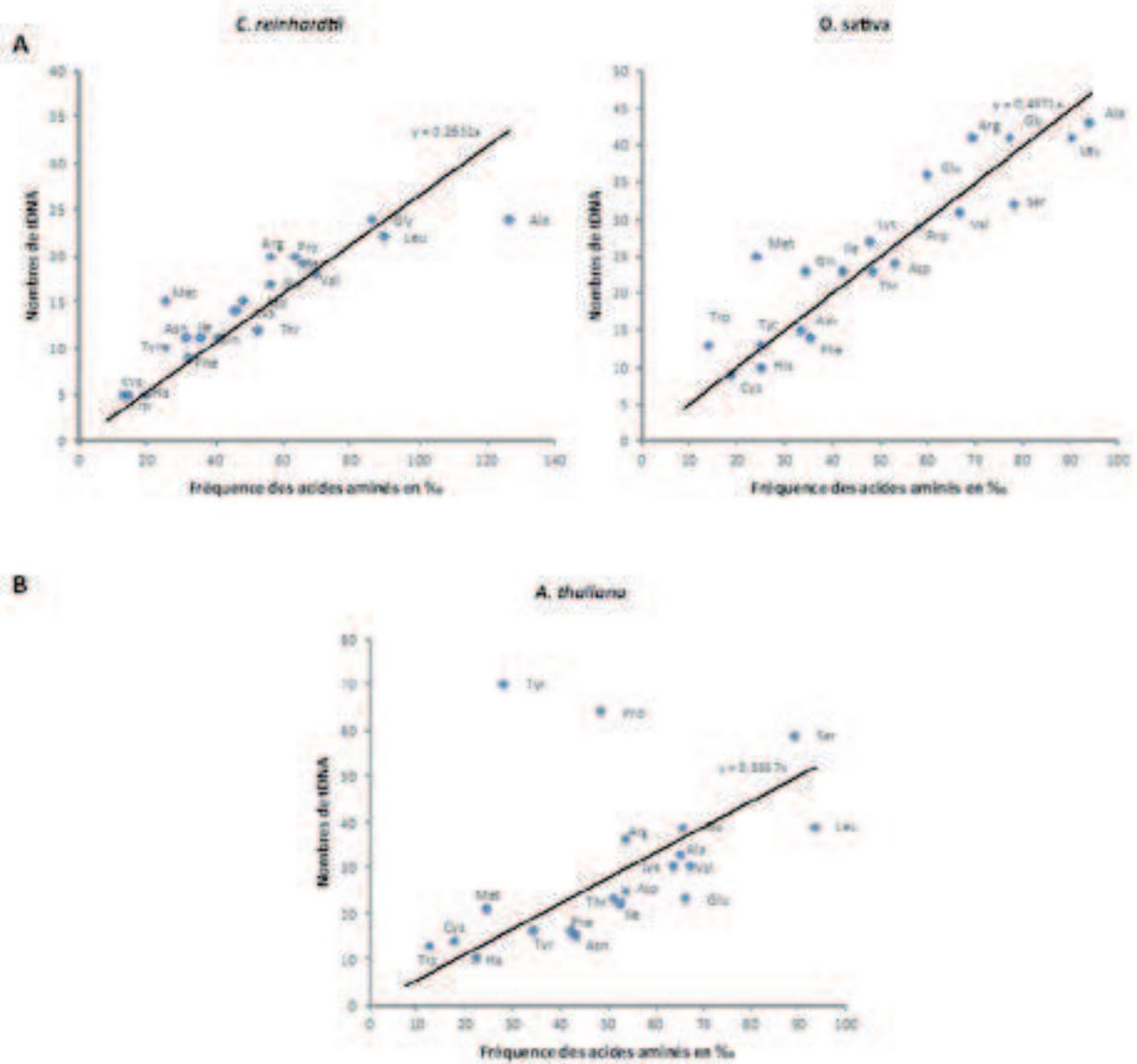
Chez la levure et *A. thaliana*, une grande variété d'enzymes impliquées dans la modification post-transcriptionnelle des tRNA a été identifiée certaines d'entre elles aident aussi au recrutement de la polymérase (Chen et al., 2010; Grosjean et al., 2010; Phizicky and Hopper, 2010), (Données disponibles sur les plantes sur le site : <http://plantrna.ibmp.cnrs.fr>). Des données supplémentaires concernant d'autres organismes sont disponibles sur le site : <http://modomics.genesilico.pl>.

Bien qu'encore mal connues, les modifications post-transcriptionnelles des tRNA ont des fonctions importantes (pour une revue voir Phizicky et Hopper, 2010). Les modifications dans la région de l'anticodon peuvent affecter la traduction, par exemple en ne permettant pas une reconnaissance codon-anticodon correct. Il s'agit souvent de la première position de l'anticodon ou de la position 37. C'est notamment le cas de souche de levure chez qui l'adénine en position 34 n'est pas modifiée en inosine, ou chez qui la guanine 37 n'est pas méthylée (Gerber and Keller, 1999; Björk et al., 2001; El Yacoubi et al., 2009). Des modifications dans le « corps » des tRNA modifient souvent leurs stabilités et/ou leurs structures. Ainsi chez la levure l'absence de thymine en position 54



abaisse la température de fusion des tRNA<sup>Phe</sup> et <sup>Met</sup> de 2 à 6°C (Davanloo et al., 1979; Sengupta et al., 2000). Chez l'Homme, l'absence de méthylation sur l'adénine 9 du tRNA<sup>Lys</sup> entraîne une modification de sa structure (Helm et al., 1999). La modification peut aussi altérer à la fois la reconnaissance du codon et l'identité du tRNA. Ainsi, chez les bactéries et les organites de plantes, l'addition d'un résidu lysine sur une cytosine (modification appelée lysidine) en première position de l'anticodon d'un tRNA avec un anticodon CAU entraîne à la fois un changement d'identité du tRNA de méthionine vers isoleucine et de reconnaissance de codon de AUG vers AUA (Muramatsu et al., 1988; Weber et al., 1990). Certains tRNA, en particulier dans les mitochondries végétales, peuvent être édités. Ce processus d'édition entraîne le remplacement de résidus cytosine en uridine. Il y a donc changement de la séquence du tRNA par rapport à la séquence du tDNA. Ces modifications sont nécessaires à la maturation correcte du tRNA ou à leur stabilisation (pour une revue, voir Fey et al., 2002). Ces exemples ne sont pas exhaustifs, d'autres exemples peuvent être trouvés dans (Grosjean et al., 2010).

Enfin, un autre exemple de modification post-transcriptionnelle concerne l'addition d'une guanine en position -1 (G<sub>-1</sub>) des tRNA<sup>His</sup> eucaryotiques. Tous les tRNA<sup>His</sup>, à l'exception de quelques alpha-protéobactéries et de celui d'*Acanthamoeba catelanii* (Wang et al., 2007) possède un tel résidu en position -1. Le tRNA<sup>His</sup> a de ce fait 8 paires de base dans le bras accepteur de l'acide aminé. Ce G<sub>-1</sub> est essentiel à la reconnaissance par l'histidyl-tRNA synthétase. Chez la plupart des bactéries et quelques archéobactéries (ainsi que dans le chloroplaste et la plupart des mitochondries), cette guanine est codée par le génome, et elle est conservée grâce à une maturation anormale de l'extrémité 5' du tRNA par la RNase P (Orellana et al., 1986). Chez les eucaryotes, il est ajouté grâce à l'activité d'une tRNA<sup>His</sup> guanylyl transférase (Cooley et al., 1982). Il est à noter que dans les mitochondries et les chloroplastes de plantes supérieures, ce G<sub>-1</sub> est codé par les génomes d'organites mais il peut également être ajouté post-transcriptionnellement par une tRNA<sup>His</sup> guanylyl transférase qui reste à caractériser (Placido et al., 2010).



**Figure 12:** Fréquence des acides aminés comparée au nombre de tDNA. (Michaud et al., 2011).

### 3.5 –tRNA et utilisation des codons

Un très bon ajustement de la population en tRNA à l'utilisation des codons et a été démontré chez quelques organismes comme *S. cerevisiae*, *E. coli*, *C. elegans* ou *C. reinhardtii* ainsi que dans les chloroplastes (Bulmer, 1987; Pfitzinger et al., 1987; Duret, 2000; Cognat et al., 2008). D'autre part, il a également été montré dans certains cas comme chez *E. coli*, *C. elegans* ou encore *C. reinhardtii*, que le nombre de tDNA est corrélé à l'utilisation des codons (Novoa and Ribas de Pouplana, 2012; Novoa et al., 2012). Chez les plantes supérieures, une telle corrélation semble exister pour les tDNA d'origine nucléaire (figure 12A, Michaud et al., 2011). Cette corrélation existe chez *A. thaliana* sauf pour les tDNA Tyr et Pro qui sont fortement amplifiés (figure 12B). S'il existe une signification biologique à cette amplification, celle-ci est à ce jour inconnue. Très globalement, il semble donc que au moins dans certains cas une relation directe en nombre de tDNA transcrits, abondance des tRNA stables dans la cellule et utilisation des codons existe. Toutefois, il est bon de noter que cette interprétation relève d'un nombre limité d'expériences et qu'il est encore à l'heure actuelle difficile de mesurer avec précision l'abondance de chaque tRNA dans la cellule. De plus, une analyse détaillée de l'abondance en tRNA en lien avec l'utilisation des codons à différents stades de développement ou dans différents tissus n'a jamais été réalisée. Enfin chez les organites comme les chloroplastes de plantes, plusieurs tRNA sont co-transcrits mais ne sont présents en quantité similaire. La part jouée par la régulation transcriptionnelle et celle jouée au niveau post-transcriptionnelle pour établir le taux de chaque tRNA reste à déterminer.

### 3.6 –Contrôle qualité des tRNA

Les tRNA, comme les rRNA, sont des molécules extrêmement stables. Ces RNA très structurés et leurs précurseurs peuvent cependant être dégradés en particulier lorsque leur structure est incorrecte et/ou leur fonctionnalité est affectée. Deux voies principales de dégradation ont été identifiées chez la levure *S. cerevisiae*. Elles sont décrites dans la revue présentée en section 4.

tRNA	<i>E. coli</i>	<i>S. cerevisiae</i>	Eucaryotes supérieurs	<i>Arabidopsis thaliana</i>
Ala	A73, G2-C71, G3-U70, G4-C69, G20	G3-U70	G3-U70	G3-U70
Arg	A ou G73, C35, U ou G36, A20	C35, U ou G36	?	?
Asn	G73, G34, U35, U36	?	?	?
Asp	G73, G2-C71, G34, U35, C36, C38, G20	G73, G34, U35, C36, C38, G10-U25	?	?
Cys	U73, G2-C71, G3-G70, G34, C35, A36, G15-G46, A13-A22	U73	?	?
Gln	G73, U3, A72, G2-C71, C1-G70, F34, U35, G36, A37, U38, G10	?	?	?
Glu	G1-C72, U7-A73, A4-U34, U35, A37, U11-A24, U13-G22-A46, G47	?	?	?
Gly	U73, G1-C72, C2-G71, G3-C70, C35, C36	A73, G1-C72, C2-G71, G3-C70, C35, C36	A73, G2-C71	?
His	C73, U-1, anticodon	A73, G-1, G14, U35	?	?
Ile	A73, C4-G69, U ou G34	U4, A35, U36	?	?
	A35, U36, U6A37, A38, U12-A73, C29-G41			
Leu	A73, U8-A19	A73, A35, G37	A73, C3-G70, A4-B5, G5-C68, C70	?
Lys	A73, U ou mm 5, C2-U34, U35, U36	?	?	?
Met (Met)	A73, G2-C71, C3-G70, U4-A69, A5-U68, C34, A35, U36, C32, U33, A37	A73, C34, A35, U36, 4 nt AC boucle anticodon	?	?
Phe	A73, G34, A35, A36, G27-C43, G28-C42, U20, G44, U45, U55, U60	A73, G34, A35, A36, U6A37, G20	A73, G34, A35, A36, C37-C43, A11-U39, G20	G34, A35, A36, G1-C70
Pro	A73, G72, G35, C36, G15-C48	?	G35, G36	?
Ser	G73, C72, G2-C71, A3-U70, A4-Y63, C11-G34, C11-G24, boucle variable	boucle variable	G73, boucle variable	?
Thr	G1-C72, C2-G71, G34, G35, U36	G1-C72, G35, G36	?	?
Tip	G73, A1-U72, G2-C71, G3-C70, C34, C35, A36	C34, C35	?	A73, C34, C35, A36
Tyr	A73, U35	A73, C1-G72, G34, U35	?	?
Val	A73, G3-C70, U4-A69, A35, C36	A73, A35	?	anticodon

**Figure 13 : Éléments d'identité des tRNA.** Tableau récapitulatif des éléments d'identité impliqués dans l'aminocyclation des tRNA (d'après Giegé et al., 1998). Dans le cas des nucléotides appariés : (:) paires Watson-Crick, (") paires non Watson-Crick (Appariements Wobble), (-) paires tertiaires ou Hoogsteen (pouvant faire participer jusqu'à 3 nucléotides)

R : purine, Y : Pyrimidine.

Les résidus sont numérotés selon le modèle de (Sprinzl et al., 1998).

## 3.7 – Fonctions des tRNA

### 3.7.1 – Fonctions communes aux eucaryotes et procaryotes

Les tRNA jouent leur rôle central dans la traduction. Ils sont les molécules adaptatrices formant un lien entre les mRNA et les protéines en cours de synthèse. On découpe le processus de traduction en 3 parties, initiation, élongation et terminaison. Les tRNA y jouent 2 rôles en permettant tout d'abord l'initiation par l'apport du premier acide aminé (méthionine), ainsi que l'élongation en apportant les acides aminés correspondant à la séquence imposée par le mRNA en cours de traduction. Pour assurer sa fonction, le tRNA doit être chargé par un acide aminé, c'est la réaction d'aminacylation. L'aminacylation est une des étapes clés dans la vie du tRNA, la cellule doit s'assurer que le bon acide aminé est attaché au bon tRNA. Chez les eucaryotes, celle-ci est effectuée par un jeu complet d'enzyme (Ibba and Söll, 2000), les aminoacyl-tRNA synthétases (aaRS). Chacune de ces enzymes reconnaît un acide aminé et un ou plusieurs tRNA isoaccepteurs. Chez les procaryotes et dans les organites, les asparaginyl- et/ou les glutaminyl-tRNA synthétases manquent. Dans ce cas les tRNA correspondant sont dans un premier temps chargé par une aspartyl ou une glutamyl-tRNA synthétase non discriminante pour former des Asp-tRNA<sup>Asn</sup> et Glu-tRNA<sup>Gln</sup>. Dans un second temps, une réaction de transamidation mettant en œuvre une amidotransférase tRNA-dépendante permet l'obtention de Asn-tRNA<sup>Asn</sup> et de Gln-tRNA<sup>Gln</sup> (Feng et al., 2004; Frechin et al., 2009).

Des éléments d'identités au sein du tRNA permettent une reconnaissance spécifique entre le tRNA et l'aaRS responsable de son aminoacylation. Ces éléments d'identités ont principalement été étudiés chez les bactéries et la levure *S. cerevisiae*, peu de données sont disponibles concernant les eucaryotes supérieurs (pour une revue Giegé et al., 1998), (figure 13). On distingue deux types d'éléments d'identités, les éléments majeurs et les éléments mineurs, selon l'importance qu'ils ont lors du processus d'aminacylation. Chez la plupart des tRNA, deux éléments d'identités majeurs ont été identifiés. Ils sont situés aux deux extrémités distales du tRNA à savoir la boucle de l'anticodon (notamment la position 35) et les acides aminés du bras accepteur ainsi que le nucléotide



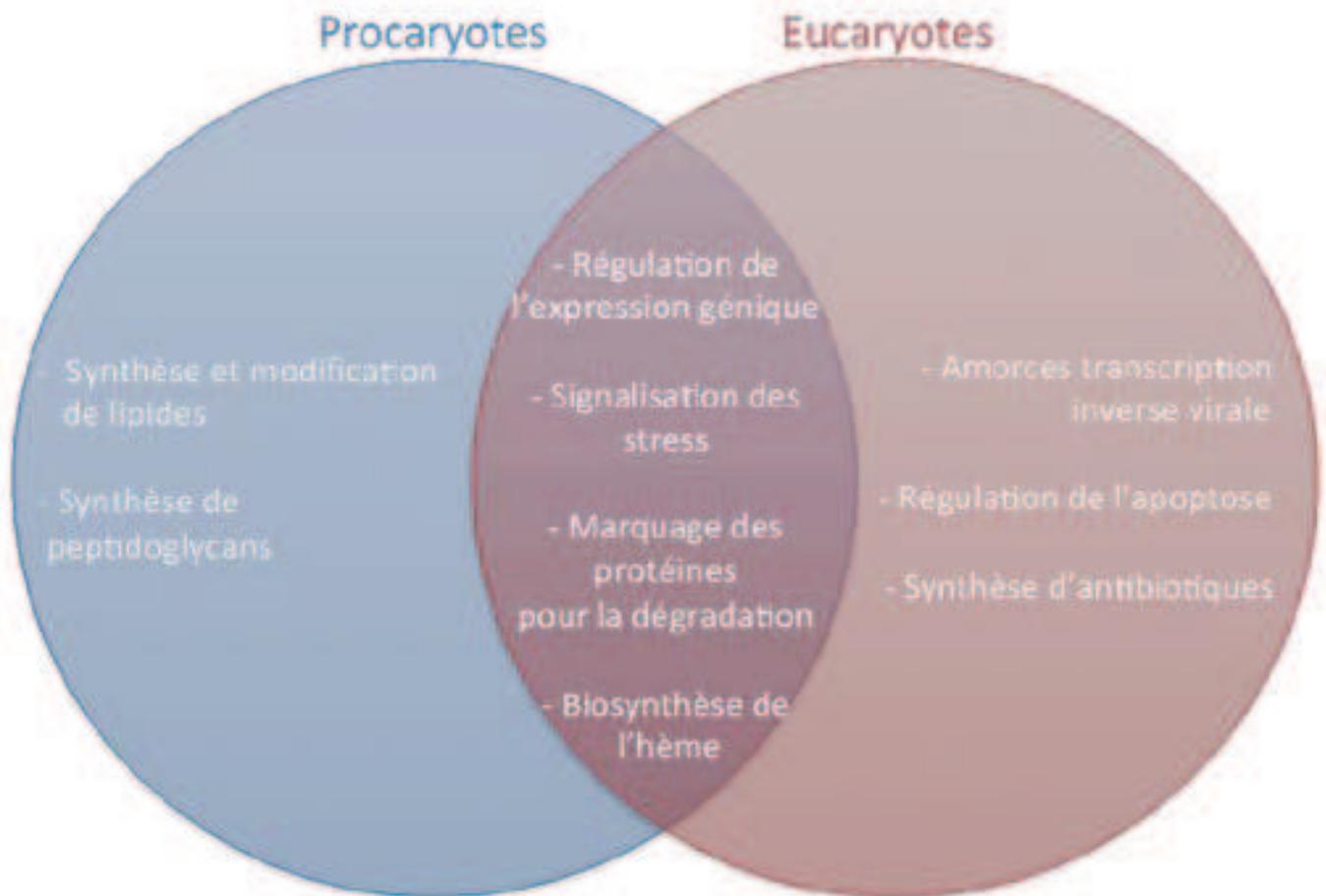
discriminant N73. Chez *A. thaliana*, les éléments d'identités pour l'aminacylation des tRNA alanine (Carneiro et al., 1994; Dietrich et al., 1996), phenylalanine (Carneiro et al., 1994), tryptophane (Ulmasov et al., 1998) et valine (Delage et al., 2003) cytosoliques sont les seuls à avoir été caractérisés.

Les tRNA occupent aussi d'autres fonctions non conventionnelles chez les eucaryotes ainsi que les procaryotes. Ils peuvent ainsi servir de donneur d'acides aminés afin de marquer des protéines en N-terminal pour leur dégradation (Mogk et al., 2007). Découvert par Bachmair et al en 1986, chez la levure (Bachmair et al., 1986), ce processus fait intervenir une aminoacyl tRNA transférase qui, après avoir reconnu un signal en N-terminale (*n.b.* un résidu spécifique nommé pro-N degron), marque la protéine pour la dégradation à l'aide d'un acide aminé (N-degron) porté par un tRNA (Mogk et al., 2007). Depuis, ce processus a été identifié chez les bactéries (Tobias et al., 1991), les plantes (Potuschak et al., 1998) et les mammifères (Gonda et al., 1989) (figure 14).

Chez les eucaryotes, le N-degron est généralement un résidu arginine. Il agit comme un précurseur à l'ubiquitylation entraînant ainsi l'enzyme au protéasome (Tasaki et al., 2012). Le N-degron est ajouté en N-ter par une arginyl (R)-transférase, qui utilise l'arginine se trouvant sur un tRNA afin de le fixer sur une cystéine, une asparagine ou une glutamine oxydées. Ce sont ces acides aminés endommagés qui constituent les "pro-N-degrons" chez les eucaryotes (Graciet et al., 2006; Rai and Kashina, 2005).

Chez les procaryotes, les N-degrons sont souvent des résidus Leu et Phe tandis que les pro-N-degron sont des Lys, des Met ou des Arg. La protéine ainsi marquée est reconnue par ClpS ("Cyclic lipopeptides") qui recrutera la protéase ClpAP (Mogk et al., 2007; Shrader et al., 1993).

Les tRNA participent aussi à la synthèse de l'hème *via* la conversion d'un tRNA<sup>Glu</sup> en L-glutamyl-tRNA<sup>Glu</sup> par la glutamyl-tRNA-reductase. Le produit de cette réaction servira ensuite à la synthèse de l'hème (Pour revue : Layer et al., 2010).



**Figure 14 : Fonctions secondaires des tRNA.** La cellule, et les fonctions des tRNA chez les procaryotes sont indiquées en bleu. Les fonctions des tRNA chez les eucaryotes sont indiqués en rouge. Les fonctions communes aux 2 sont affichées au centre.

### 3.7.3 – Fonctions spécifiques aux procaryotes

Chez les procaryotes, les tRNA non-aminoacylés peuvent réguler l'expression génique en réponse à des changements en disponibilité des acides aminés dans le milieu.

L'un de ces modes de régulations a été découvert en 1969 par Cashel et Gallant (Cashel and Gallant, 1969) chez *E. coli* en réponse à une carence en acide aminé. Ce processus repose sur la production de 2 hormones nommées ppGpp (5'-diphosphate 3'-diphosphate guanosine) et pppGpp (5'-triphosphate 3'-diphosphate guanosine). Leurs synthèses dépendent des protéines RelA et SpoT. RelA est une (p)ppGpp synthase associée au ribosome qui détecte la présence de tRNA non chargés s'accumulant au site A du ribosome (Haseltine and Block, 1973; Sy and Lipmann, 1973). La protéine SpoT quant à elle, est une enzyme bi-fonctionnelle pouvant synthétiser et hydrolyser la (p)ppGpp, à l'heure actuelle les mécanismes d'activation de cette enzyme sont inconnus (Magnusson et al., 2005). Une fois produite la (p)ppGpp agit sur une sous-unité de l'ARN polymérase II en diminuant la transcription tout en augmentant l'expression des gènes liés à la synthèse des acides aminés (Ross et al., 2013) (figure 14).

Un autre mécanisme grâce auquel les bactéries régulent l'expression des gènes à l'aide de tRNA non chargés a été démontré chez *B. subtilis* et d'autres bactéries gram positive impliquant le système "T-box riboswitch" (pour revue Green et al., 2010). Chez ces organismes, durant une carence en acides aminés, un tRNA est recruté sur une séquence nommée anti-terminateur empêchant la formation de la tige boucle permettant la terminaison de la transcription (Grundy et al., 2005) (figure 14).

Les tRNA aminoacylés occupent aussi un rôle essentiel dans la production de peptidoglycanes ainsi que de lipides (pour revue Raina and Ibba, 2014).

### 3.7.2 – Fonctions spécifiques aux eucaryotes

Certains des rôles secondaires des tRNA sont spécifiques aux eucaryotes. Ainsi, il a été montré que des champignons unicellulaire (*e.g. S. cerevisiae, S. noursei* ou *S. coeruleorubiduil*) étaient capables de se servir de tRNA aminoacylés afin de produire des antibiotiques (Pour revue Raina and Ibba, 2014).



Un autre de ces rôles concerne cette fois-ci des tRNA non chargés qui participent à la transduction du signal durant des carences en nutriments. Ainsi, chez la levure, lors de carences en acides aminés, la cellule utilise les tRNA non chargés afin d'interagir avec la protéine kinase Gcn2p qui phosphoryle le facteur eIF2 (eukaryotic Initiation Factor 2). Cette phosphorylation sur eIF2 induit une diminution du niveau de la traduction. Pourtant, en parallèle, le facteur Gcn4 ("general control nonderepressible 2") est traduit plus activement, promouvant la transcription de nombreux gènes impliqués dans la synthèse d'acides aminés ou d'acides nucléiques (Raina and Ibba, 2014).

Chez les mammifères, les tRNA peuvent être impliqués dans la régulation de l'apoptose (Mei et al., 2010) (figure 14). Il a été montré que des tRNA se lient au cytochrome *c*, empêchant son activation par la caspase Apaf-1. Enfin, certains tRNA ont été détournés par des virus et sont utilisés par les transcriptases inverses virales pour convertir leur génome ARN en génome ADN. C'est, entre autres, le cas du VIH, qui utilise le tRNA<sup>Lys</sup> comme amorce pour sa transcription inverse (Marquet et al., 1995; Piekna-Przybylska et al., 2010, 2011).

Finalement, au cours de ces 10 dernières années, de nouvelles fonctions ont été attribuées aux tRNA à travers de petits fragments d'ARN générés par leur clivage. Ces fragments d'ARN dérivant, de tRNA ont été appelés tRFs pour « tRNA-derived RNA Fragments ». Ils ont été détectés dans des organismes eucaryotiques évolutivement différents tels les protozoaires (Couvillion et al., 2012), les champignons (Thompson et al., 2008), les animaux (Lee et al., 2009) et les plantes (Hsieh et al., 2009). Nos connaissances actuelles sur la biogenèse et les fonctions potentielles des tRFs sont présentées dans la revue ci-dessous (section 4).

## 4 - Revue "Surveillance et clivage des tRNA eucaryotiques"

La revue ci-dessous a été écrite dans le cadre d'un numéro spécial sur les tRNA du journal "International Journal of Molecular Sciences". Elle traite de la dégradation et du clivage des tRNA, des processus essentiels pour contrôler la



qualité des tRNA tout au long de leur genèse, pour recycler les nucléotides et le phosphate ou encore pour générer de nouvelles fonctions. Je suis co-premier auteur de cette revue avec Cyrille Megel également doctorant dans l'équipe du Dr. L. Maréchal-Drouard. Cyrille Megel a rédigé la partie concernant la biogenèse et le rôle des tRFs et pour ma part, j'ai rédigé la partie concernant le « turnover » et le contrôle qualité des tRNA ainsi que la description des différentes classes de tRFs.

Review

## Surveillance and Cleavage of Eukaryotic tRNAs

Cyrille Mège<sup>1</sup>†, Geoffrey Morelle<sup>1,2</sup>†, Stéphanie Lalande<sup>1</sup>, Anne-Marie Duchêne<sup>1</sup>, Ian Small<sup>2</sup>  
and Laurence Maréchal-Drouard<sup>1\*</sup>

<sup>1</sup> Institut de Biologie Moléculaire des Plantes, CNRS, Associated with University of Strasbourg,  
12 rue du Général Zimmer, F-67084 Strasbourg Cedex, France;  
E-Mail: cyrille.megel@ibmp-cnrs.unistra.fr (C.M.), geoffrey.morelle@ibmp-cnrs.unistra.fr (G.M.),  
stephanie.lalande@ibmp-cnrs.unistra.fr (S.L.), anne-marie.duchene@ibmp-cnrs.unistra.fr (A.-M.D.)

<sup>2</sup> Australian Research Council Centre of Excellence in Plant Energy Biology,  
The University of Western Australia, Crawley WA6009, Australia; E-Mail: ian.small@uwa.edu.au

† These authors contributed equally to this work.

\* Author to whom correspondence should be addressed; E-Mail: laurence.drouard@ibmp-cnrs.unistra.fr;  
Tel: +33-3-6715-5398; Fax: +33-3-8861-4442

Academic Editor: Michael Eblin

Received: 19 December 2014 / Accepted: 9 January 2015 / Published:

**Abstract:** Beyond their central role in protein synthesis, transfer RNAs (tRNAs) have many other crucial functions. This includes various roles in the regulation of gene expression, stress responses, metabolic processes and priming reverse transcription. In the RNA world, tRNAs are, with ribosomal RNAs, among the most stable molecules. Nevertheless, they are not eternal. As key elements of cell function, tRNAs need to be continuously quality-controlled. Two tRNA surveillance pathways have been identified. They act on hypo-modified or mis-processed pre-tRNAs and on mature tRNAs lacking modifications. A short overview of these two pathways will be presented here. Furthermore, while the exonucleases acting in these pathways ultimately lead to complete tRNA degradation, numerous tRNA-derived fragments (tRFs) are present within a cell. These cleavage products of tRNAs now potentially emerge as a new class of small non-coding RNAs (sncRNAs) and are suspected to have important regulatory functions. The tRFs are evolutionarily widespread and created by cleavage at different positions by various endonucleases. Here, we review our present knowledge on the biogenesis and function of tRFs in various organisms.

**Keywords:** tRNA decay; tRNA quality control; asRNAs; tRNA stress response; gene expression regulation; tRNA endonucleases.

## 1. Introduction

Transfer RNAs (tRNAs) are among the most ancient RNAs in the world. Since their discovery by Hoagland and Zamenick in 1957 and the historical hypothesis of Francis Crick postulating the existence of an adapter molecule at the interface between the genetic code and protein, tRNA molecules are still “à la mode”. The first primary tRNA sequence [1] was followed by the determination of the characteristic cloverleaf secondary structure, and the L-shaped three dimensional tertiary structure was obtained in the seventies [2,3]. The sizes of tRNAs vary classically between 73 and 90 nucleotides. However, tRNAs from metazoan mitochondria can be much smaller. The world’s smallest aimless tRNAs (42 nucleotides) were recently identified in the mitochondria of nematodes [4]. The first step in the production of a mature tRNA molecule is the synthesis of a primary transcript. In eukaryotes, RNA polymerase III transcribes tRNA genes and transcription stops after a series of U residues present a few markoticks T of the tRNA sequence. The 5' leader and the 3' trailer sequences are removed by RNase P and RNase Z, respectively. The CCA 3' extremity is added by the tRNA nucleotidyl transferase. For a subset of tRNAs, introns located between positions 37 and 38 are removed through tRNA splicing. Splicing of tRNAs is a process including mainly two steps, the excision of the intron by an endonuclease and the ligation of two tRNA half molecules. In yeast and plants the second step involves a single enzyme (tRI) catalyzing phosphodiesterase, kinase and ligase activities. In vertebrates, in addition to the yeast like ligation pathway, a second pathway exists with a ligase that still needs to be identified (for a review see e.g., [5]). Multiple post-transcriptional modifications are also added all along the maturation of tRNA. To function in protein synthesis, tRNA is finally charged with an amino acid by its cognate aminoacyl-tRNA synthetase (e.g., [5]).

Transfer RNA molecules are among the most stable RNAs in a cell. Compared to mRNAs, with a half-life varying from minutes to a few hours, tRNA half-lives are more in the range of several days [5]. This extreme stability is due to a highly folded structure and the presence of numerous modifications. Nevertheless, quality control mechanisms exist to avoid the accumulation of non-functional tRNAs. In contrast to the many studies of mRNA decay [6–9], there has been less attention to the degradation of stable RNAs such as tRNAs or rRNAs, although they are the most abundant RNAs in the cell. Two major surveillance pathways have been uncovered that act in tRNA turnover in eukaryotes [10]. The degradation can occur at the level of pre-tRNA or mature tRNA species. The first tRNA degradation pathway involves the TRAMP (Trf1 (topoisomerase 1-related 1)/Air2 (Arginine methyltransferase-interacting RING finger protein 2)/Mtrp (mRNA transport regulator 4 protein) Polyadenylation) complex and the nuclear exosome. The second tRNA degradation pathway is called RTD for Rapid tRNA Decay and involves cytoplasmic exonucleases. Our knowledge on the two pathways will be the topic of the first part of this review.

Beside their canonical role during protein synthesis, novel additional functions are regularly assigned to tRNAs (for reviews see [5,11–13]). In particular, they are involved in many processes of cellular metabolism. They can deliver amino acids not only for protein synthesis but also for cell envelope remodeling, antibiotic synthesis or tagging for proteolysis. Regulation of cell death by tRNA represents another important role. They can also be used as primers for viral reverse transcription. Moreover, uncharged tRNAs can act as stress sensors and as regulators of gene expression. Finally, small non-coding RNAs deriving from tRNAs potentially represent a new class of regulatory RNAs. Indeed, while in the tRNA surveillance pathways tRNA degradation is complete and no degradation products are expected to persist, relatively stable tRNA-derived RNA fragments (tRFs) of different sizes have been found in many different organisms. Half tRNAs are primarily created upon different stresses but shorter tRNA fragments of about 15–25 nucleotides are also abundant. Endonucleases responsible for tRNA or pre-tRNA slicing have been identified. Slicing can occur at different sites, in particular within the anticodon, D and T loops of the tRNA molecule. Potentially important functions have been ascribed to tRFs, including the inhibition of translation, or the induction of RNA silencing or apoptosis. In the second part of the review, we summarize our present knowledge on the population, biogenesis and function of tRFs.

## 2. Eukaryotic Transfer RNA (tRNA) Turnover and Quality Control

### 2.1. The Nuclear Surveillance Pathway

The nuclear surveillance pathway (Figure 1A) was discovered in temperature-sensitive yeast (*Saccharomyces cerevisiae*) strains lacking the tRNA methyl transferases (Tm) Tm6 or Tm61 (also named Gcd10 or Gcd14 for General control non-depressible proteins 10 and 14) the two essential subunits of the methyltransferase responsible for tRNA m<sup>1</sup>A modification. In these strains, the precursor of initiator tRNA<sup>Met</sup> (pre-tRNA<sup>Met</sup>) lacks m<sup>1</sup>A and is subject to rapid turnover [14,15]. *In vivo* genetic analyses provided evidence that as a first step towards its degradation, the pre-tRNA<sup>Met</sup> is polyadenylated at its 3' end by the TRAMP complex. This complex was first shown to be required for polyadenylation and degradation of tRNA and snoRNA (small nucleolar RNA) precursors as part of a post-transcriptional quality control mechanism [16]. It is constituted by Trf4, a poly(A) polymerase, Mtr4, a RNA helicase, and Air2, a zinc knuckle protein interacting with Rps6 (Ribosomal RNA Processing subunit 6), the 3' exonuclease of the nuclear exosome. Following polyadenylation by Trf4, pre-tRNA<sup>Met</sup> is degraded from the 3' end by Rps6 [16,17–21]. Both the TRAMP and exosome complexes are localized in the nucleus. As described above in yeast, the pre-tRNA nuclear surveillance pathway is implicated in the degradation of defective hypomodified pre-tRNA<sup>Met</sup>. However, how a tRNA is recognized or not as a substrate to be adenylated and degraded remains largely unknown. The lack of m<sup>1</sup>A is *per se* not sufficient to explain why the pre-tRNA<sup>Met</sup> is rapidly degraded as other tRNAs lacking the same modification are not degraded [18]. The most reasonable hypothesis is that the nuclear surveillance pathway recognizes aberrant mis-folded tRNA tertiary structure. *In vitro* data suggest that other tRNAs can be TRAMP substrates, but this remains to be demonstrated *in vivo* [16,22].

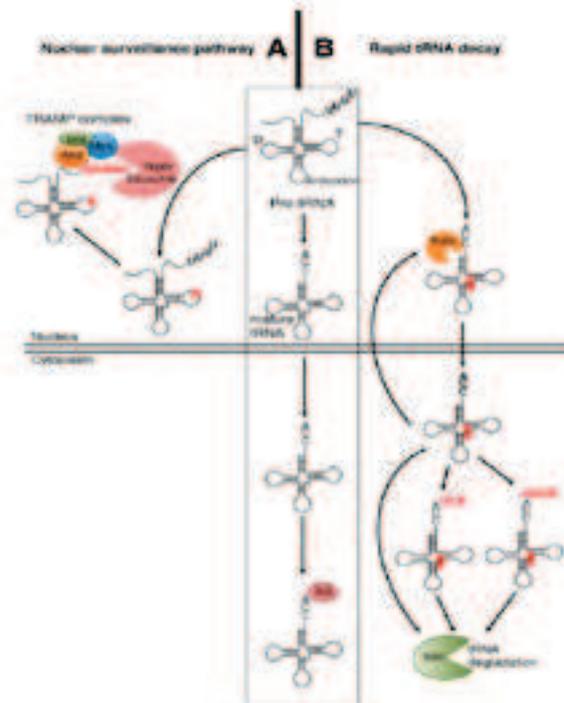
Different end-natural but unspliced intron-containing pre-tRNAs were also shown to be TRAMP targets [23]. Indeed, the authors showed that such targets are enriched in *trf4* cells, a phenotype

accentuated in *Rp60* cells. These results are linked to the complexity of the 3'-end maturation of pre-tRNAs, maturation that can be achieved by two different pathways [21,24]. In the first major pathway, the La protein (Lhp1p in yeast) binds the short stretch of uridines (Us), (at least 3) present at the 3' extremity of eukaryotic pre-tRNA, stabilizes the conformation and protects the extremity from exonucleases [25]. After removal of the 5' extension by RNAse P, the 3' tail is removed by an endonuclease such as RNAse Z. In the second pathway, Lhp1p does not bind or is released from the poly(U) tail. This allows Rex1p, a non-exosomic endonuclease, to trim the 3' trailer up to a base-paired region between the 5' leader and the 3' trailer. Following cleavage of the 5' extremity by RNAse P, Rex1p or other exonucleases mature the 3' extremity and the CCA sequence can be added. Considering these two pathways, it is likely that a fraction of intron-containing pre-tRNAs folding into alternative abnormal structure fails to bind Lhp1p. This aberrant structure may slow the rate of end maturation, and the competition between Lhp1p and Rex1p will allow additional exonucleolytic nibbling and/or TRAMP polyadenylation [26]. Altogether there is a strong interplay between maturation pathways that functional tRNAs undergo and quality control pathways that degrade abnormal RNAs. This also implies strong competition between the La-protein Lhp1p, the endonuclease Rex1p and the TRAMP complex [23].

## 2.2. The Rapid tRNA Decay (RTD) Pathway

While the nuclear surveillance pathway acts primarily on pre-tRNAs, mature tRNAs also undergo quality control surveillance. As for the nuclear surveillance pathway, the first evidence of the existence of a pathway called rapid tRNA decay (RTD) came from yeast mutant strains lacking enzymes involved in tRNA modifications. The temperature sensitive *trm1/trm2* yeast mutant strain presents a severe growth defect. In this strain, the mature tRNA<sup>45(Ser)</sup> lacks both 7-methylguanine (m<sup>7</sup>G<sub>46</sub>) and 5-methylcytosine (m<sup>5</sup>C) and is rapidly deacylated and degraded at 37 °C [27]. Similarly, degradation of several other mature tRNAs lacking different combinations of modifications was also observed. For instance, in *trm1/trm2* or in *trm1/trm2* yeast cells lacking 2'-O-methyluridine (Um<sub>2</sub>) and 4-acetylcytosine (ac<sup>4</sup>C<sub>47</sub>) or di-methylguanine (m<sub>2</sub>G<sub>46</sub>) and 5-methylcytosine (m<sup>5</sup>C) modification enzymes respectively, both mature tRNA<sup>50(Glu)</sup> and tRNA<sup>50(Ser)</sup> are progressively lost at 37 °C [28,29]. As modifications are known to be involved in thermal stability of tRNAs (e.g., [30,31]), this supports the idea that this pathway may act on a wide variety of unstable or non-functional tRNA species.

The hypothesis is that, in strains lacking tRNA modification enzymes, under normal conditions most tRNAs are correctly folded and functional, but after the temperature switch the hypo-modified tRNAs are destabilized and trigger RTD. It has also been shown that the RTD pathway can act on tRNAs bearing mutations that are deleterious for the stability of the acceptor and T-stems. Fully modified tRNAs bearing these mutations are degraded, thus demonstrating that surveillance by the RTD pathway does not necessarily depend on tRNA modifications but rather on the structural integrity of the RNA [32].



**Figure 1.** Overview of the two tRNA degradation pathways of *S. cerevisiae*. This figure is adapted from [10]. **(A)** The tRNA nuclear surveillance pathway was primarily shown to act on pre-tRNAs lacking modifications (e.g., pre-tRNA<sup>Met</sup>). Hypo-modified pre-tRNA is first polyadenylated by the TRAMP complex and then degraded by the exosome; **(B)** Degradation of mature tRNAs through the rapid tRNA decay (RTD) pathway. RTD was shown to act on hypo-modified tRNAs (marked with red crosses) (e.g., tRNA<sup>ValAAAC</sup>). The major components are depicted. The implication of the tRNA nucleotidyl transferase (CCAase) enzyme in the RTD pathway is also presented [13]. A CCA triplet or a short poly(A) tail (AAAA) depicted in red can be added by the CCAase enzyme.

The loss of hypo-modified tRNAs is not dependent on Tif1 or Rtp6, implying that the TRAMP complex and the nuclear exosome are not implicated. Rather, other protein components belonging to the RTD pathway were genetically identified (Figure 1B). In the *trm113/trm115* yeast mutant strain, the degradation and loss of aminoacylation of tRNA<sup>ValAAAC</sup> is mediated by Met22 (Methionine-requiring protein 22) and the two 5'-3' exonucleases Rnt1 (Ribonucleic-acid-trafficking protein 1) and Xrn1 (Exonuclease 1) [34]. Deletion of Met22 or Rnt1/Xrn1 in a *trm113/trm115* or *trm114/trm115* strain rescues the temperature-sensitive phenotype entirely by preventing the degradation of several tRNAs in yeast [34], and in humans [35]. The degradation by the nuclease Xrn1 is likely due to the instability of the acceptor stem, which exposes the 5' end of the tRNA molecule [32]. Rnt1 is localized in the

nucleus, thus suggesting the need to re-import deleterious tRNAs into this compartment to allow their degradation by this endonuclease. However, the requirement of the retrograde tRNA nuclear import pathway using importin Mtr10 (mRNA transport regulator 10) [36] needs to be demonstrated. Mtr22 is a phosphatase not directly implicated in the RTD pathway. Rather, when function of Mtr22 is lost, accumulation of its metabolic substrate adenosine 5',5'-diphosphate (pAp) inhibits Rat1 and Xrn1 [27,34,37,38]. Interestingly, the inhibition of plant exonucleases by pAp has also been observed and it is the plastidial enzyme SAL1 (SAL for salt tolerance) which plays the role of Mtr22. Whether the degradation of tRNAs is prevented in *sef1* mutants, remains to be established [39].

Furthermore, in the *trm68trm69* strain, the stability of hypo-modified tRNA<sup>Val</sup><sup>68/69</sup> is restored when the genes *TEF1* and *VAS1* encoding the elongation factor eEF1A and valyl-tRNA synthetase respectively are over-expressed [29,40]. These two proteins interact directly with the tRNA molecules: tRNA<sup>Val</sup> is aminoacylated by valyl-tRNA synthetase, bound to eEF1A and delivered to ribosomes. Thus, it is likely that an increase in the levels of proteins directly interacting with imperfectly folded hypo-modified tRNAs can protect them from degradation by competing with ribonucleases such as Rat1 and Xrn1. This shows that components of the RTD and of the translation machinery actively interact to either deliver charged tRNA molecules to ribosomes or degrade them. Another category of proteins able to limit the RTD process corresponds to proteins affecting global tRNA transcription. In the *trm68trm69* strain, over expression of Maf1, the global negative regulator of polymerase III (*polIII*) [41,42] results in lower degradation of the hypo-modified tRNA<sup>Val</sup><sup>68/69</sup> [40]. Similarly, in the same mutant strain, point mutations in the RNA *polIII* Rpe178 subunit (RNA closed complex subunit 178) or decreased expression of Rpe17, another *polIII* subunit, leads to a decrease in tRNA transcription and to tRNA<sup>Val</sup><sup>68/69</sup> stabilization.

The presence of a CCA triplet at the 3' extremity of tRNAs is essential for their charging with cognate amino acids. The tRNA nucleotidyltransferase, also called CCA-adding enzyme or CCAase, post-transcriptionally adds the CCA triplet at the 3' end of tRNAs, a sequence not genome-encoded in eukaryotes [43–45].

In addition, this enzyme can also add CCA to the 3' ends of tRNA-like transcripts (e.g., [46]). Wilusz and collaborators [33] provided evidence that the CCA adding enzyme is able to selectively mark unstable tRNAs (e.g., hypo-modified tRNAs or tRNAs lacking integrity of the acceptor stem) or tRNA-like small RNAs for degradation by adding the CCACCA sequence. The CCACCA addition requires an isomerization step and the presence of two Gs at the 5' extremity of the RNA molecule. This creates a long 5' single stranded RNA tail and likely allows access by the exonucleases Rps44 (Ribosomal RNA processing subunit 44) and Xrn1 (Exonuclease 1), which then initiate tRNA degradation. The proposed model also explains why some tRNAs are sensitive to RTD whereas others, similarly hypo-modified, are not. The latter do not contain two guanosines (Gs) at their 5' extremity, thus precluding the isomerization step required for the double CCA addition. Finally, the authors also observed an increased number of tRNAs with short poly(A) tails added to the CCA motifs, another way to flag unstable tRNAs for degradation.

In other eukaryotes, tRNA degradation pathways have not been studied yet. Many of the enzymes involved in either the nuclear surveillance pathway or the RTD pathway in yeast exist in plants and mammals. Therefore, it is likely the same tRNA quality control pathways are present in organisms from yeast to vertebrates [47], although it remains to be demonstrated.

### 3. tRNA-Derived Fragments (tRFs): A New Class of Small Non-Coding RNAs

Since the discovery in 1993 of small non-coding RNAs (snRNAs) in *Caenorhabditis elegans* [48], many classes of snRNAs (15–40 nt long) with various regulatory functions have been identified in a multitude of organisms of the eukaryotic kingdom [49]. The most well known are the siRNAs (small interfering RNAs), miRNAs (microRNAs) and piRNAs (pilot interacting RNAs).

Thanks to the increasing use of high-throughput sequencing technologies many more classes of snRNAs are emerging [50,51]. Among them, short tRNA-derived RNA fragments (tRFs) of 15–28 nt have been found in snRNA libraries of various organisms. Eukaryotic tRFs were first characterized in the human HCT (Human Lymphocyte T) 116 cell line [52]. Since then, they have been described in all domains of life (e.g., [53–64]). Longer tRFs of about 30–40 nt in size are also reported in the literature. In eukaryotes, they were discovered in evolutionarily divergent organisms such as yeast, protozoans, plants and metazoans (e.g., [54,56,60,61,65–72]). Most of the time, RNA fragments with a length exceeding 30 nt are excluded from snRNA libraries and this second category of tRFs was mainly identified by northern blots.

tRFs are present in very low amounts within cells as compared to the high steady-state levels of tRNAs from which they derive. Nevertheless, it is now suspected that tRFs are not just degradation products but may be implicated in important regulatory and biological processes. In the following paragraphs, the different classes of tRFs will be presented in more detail.

tRFs are classified according to the region of the tRNA or of the pre-tRNA from which they derive. Three classes of short tRFs have been described. They originate from (i) the 5' extremity of mature tRNA after cleavage in the D region, (ii) the 3' extremity of the mature tRNA (meaning the CCA end is present) after cleavage in the T region and (iii) the 3' trailer of pre-tRNA after cleavage by RNase Z.

As the data were obtained by several teams concurrently, various nomenclatures have been proposed (Table 1). For example, tRFs corresponding to RNA fragments derived from the 5' end of mature tRNA after cleavage in the D region were named 5' tRF or tRF5 (e.g., [73]); those deriving from the 3' end were named 3' CCA tRF or tRF3 [61,74]. These tRFs were recovered in snRNA libraries because of their size. The remaining RNA fragments (i.e., fragments of around 40–60 nt) would be considered as long tRFs. Due to their size, these RNA fragments cannot be retrieved from the snRNA libraries, but their existence cannot be excluded. Indeed, using northern experiments, such RNA fragments are present in total *Arabidopsis thaliana* RNA extract (personal communication, Laurence Marchal-Drouard). Taking into account the existence of a more complex population of tRFs than presently described, we propose a novel nomenclature where each tRF is named according to the extremity of the mature tRNA and the cleavage region from which it is generated (Table 1). For instance, tRF-5D represents a tRF generated from the 5' extremity of a tRNA after cleavage in the D region. The complementary fragment is then called tRF-3D. To this nomenclature it is possible to add the amino acid and the anticodon specifying the corresponding tRNA. For example, tRF-TT (His-GHG) means tRF generated from 3' extremity of tRNA<sup>His(GHG)</sup> after cleavage in the T region. Most tRFs described so far originate from nuclear-encoded cytosolic tRNAs. However a few examples of tRFs deriving from plastidial or mitochondrial tRNAs have been described [75,76]. It is thus also important for the future to add this information by adding c (cytosolic), p (plastidial) or m (mitochondrial) when necessary. In mammals, tRFs deriving from

the 5' trailer of pre-tRNAs were also identified. These tRFs, called 5' tRF or tRF-1, match to the 5' trailer of precursor tRNA molecules [52,57,77]. They are released upon RNase Z cleavage and end at the level of the short stretch of U residues where polymerase III is released. They do not derive from mature tRNAs but rather from precursor transcripts, and we propose to rename them pre-tRF-3H.

**Table 1.** General nomenclature of tRNA derived fragments (tRFs). Mature tRNAs and pre-tRNAs can be cleaved at different positions to produce various tRFs. Different names have been proposed according to their size, function and position [13,50,66,78]. We now propose a general nomenclature where the letter corresponds to the extremity of the tRNA and the number to the cleavage site. The localization of the tRFs are drawn either in red or in black on the tRNA molecule and their names have the same color code.

tRFs	Yamasaki, 2009; Anderson, 2014	Raina, 2014	Loe, 2009	General Nomenclature
	5' tRF	5' tRF	tRF5	tRF-5B
	3' tRF	3' tRF	tRF3	tRF-3B
	5' tRF (5' tRF3a)	5' tRF (3' tRF)	tRF5	pre-tRF-5B
	5' tRF (5' tRF3a)	5' tRF (3' tRF)	tRF5	tRF-5A
	3' tRF (3' tRF3a)	3' tRF (5' tRF)	tRF3	tRF-3A

Another category of tRFs corresponds to mature tRNA halves generated through cleavage in the region of the anticodon. From a single tRNA, two fragments of about 30–40 nt are thus generated, and correspond to the 5' and 3' tRNA halves respectively (Table 1). This category of tRFs is generally produced in response to a multitude of stresses (e.g., amino acid starvation, phosphate starvation, oxidative stress) (e.g., [61,72,78–80]). In mammals, they were therefore called tRNAs (tRNA derived stress induced RNAs) or also 5' tRNA halves (for reviews see [13,81]). This category of tRFs has been found in all kingdoms of life: bacteria, protozoa, fungi, plants and animals and seem to be a universal marker of stress (e.g., [56,69,61,65–67,69–72]). Nevertheless, it is worth to mention that not all stresses induce tRNA cleavage. For instance, in yeast cells undergoing amino acid or phosphate starvation or UV irradiation, no increase of this category of tRFs was observed [61,69,82]. Furthermore, tRFs corresponding to tRNA halves were also shown to be present *in vivo* in the absence of stress. [61,72,80]. The anticodon region is the most accessible part of a tRNA molecule for either mechanical or enzymatic cleavages. The presence of these RNA fragments is thus not unexpected and whether some of them play regulatory functions still remain to be established. Finally, as the 5' tRNA half and the 3' tRNA half do not have similar functions (see below for more details), it is essential

to distinguish them. Thus, as above for short tRFs, we propose a similar nomenclature where tRF-5A and tRF-3A will represent the 5' tRNA half and the 3' tRNA half of a mature tRNA after cleavage in the region of the anticodon.

Finally, the last known example of tRFs corresponds to 5' leader-exon tRNA fragments. In mice lacking CLP1 (Cleavage and Polyadenylation factor I subunit 1, an RNA kinase implicated in tRNA splicing in mammals), accumulation of such tRFs deriving from an aberrant splicing of intron-containing pre-tRNAs<sup>Pro</sup> was observed [83]. Interestingly the authors showed that this category of tRFs sensitizes cells to oxidative stress induced p53 activation and p53 dependent cell death but the exact molecular mechanism remains to be characterized.

#### 4. Biogenesis of tRF: Which Endonuclease for Which Job?

Among the first identified tRNA endonucleases were bacterial Colicin E5 and Colicin D (Table 2). In *E. Coli*, both enzymes cleave specific tRNAs in the anticodon loop. The major consequence is the arrest of protein synthesis leading to cell death [65,66]. Colicin E5 endonuclease recognizes tRNAs containing a guanine (G) residue at the first position of the anticodon and cleaves between position 34 and 35. Colicin D endonuclease cleaves the four tRNA<sup>Pro</sup> isoacceptors between positions 38 and 39. Another study illustrates an anticodon endonuclease activity. During phage T4 infection, suppression of EcoRI DNA restriction activity allows the PrrC (Pre mRNA cleavage complex II) protein subunit to cleave tRNA<sup>Leu</sup> between position 13 and 34 [84,85].

In eukaryotes, tRNA endonucleases have also been identified. Zymocin ( $\gamma$ -subunit) of the dairy yeast *Kluyveromyces fragilis* induces G1 phase arrest of sensitive yeast cells like *Saccharomyces cerevisiae* by cleavage in the anticodon loop between position 34 and 35 of tRNA<sup>Leu</sup> and tRNA<sup>Val</sup> [86]. These tRNAs are unusual in that they carry a 5-methoxycarbonylmethyl 2-thiouridine (mcm<sup>2</sup>) residue at the wobble position. The absence of mcm<sup>2</sup> group drastically reduces the cleavage efficiency suggesting that post-transcriptional nucleotide modifications are important for a proper enzyme-substrate interaction and cleavage. *S. cerevisiae* also possesses its own tRNA endonuclease called Rny1 (Re-associated Y1). Rny1 belongs to the RNase T2 family and is able to cleave tRNAs within the anticodon loop [87,88]. Rny1 is normally localized in the vacuole [87]. Where this ribonuclease cleaves tRNAs is still under debate. Thompson and Parker showed that, after oxidative stress, Rny1 can be released from the vacuole to the cytoplasm and that this release correlates with tRNA cleavage [87]. In contrast, Luthala and Parker presented evidence that tRNAs are targeted to vacuoles and that Rny1 cleaves tRNAs in this compartment. They proposed that this enzyme might participate in a form of tRNA ribophagy [87,88]. Recently, in agreement with these data, work from Huang *et al.* demonstrated that vacuole-localized Rny1 is responsible for the degradation of unidentified RNA in nitrogen starving conditions [89]. RNase T2 enzymes are also implicated in tRNA cleavage in the protozoan *Tetrahymena thermophila* [90]. Three endonucleases induced during amino acid starvation are expressed from homologous genes (*RNY2A*, *RNY2B* and *RNY2C*) and contribute to the production of tRNA halves by cleavage in the anticodon loop. In mammals, several endonucleases have been identified as capable of tRNA cleavage. They are involved at different stages of tRNA processing and exert their activities at different sites. A study performed by Lee *et al.* [52] showed that the RNase Z, HLAZ2 (cDC ribonuclease Z2), produces pre-tRF-3D from tRNA precursors. For instance,

the pre-tRF-3D called tRF1001 is generated by cleavage at the 3' end of the precursor of tRNA<sup>SupT22</sup>. Dicer, a double-stranded RNA-specific endonuclease playing a central role in RNA silencing pathway, is implicated in the production of a specific tRF-5D deriving from tRNA<sup>SupT22</sup> after cleavage in the single-stranded D loop [73]. Other studies also revealed the role of Dicer in the production of pre-tRF-3U and tRF-3T deriving respectively from either the 3' extremity of pre-tRNA<sup>Sup</sup> [91] or from the 3' extremity of mature tRNA<sup>Sup</sup> [57]. These results suggest that some tRNAs, tRNA halves or pre-tRNAs can adopt alternative secondary structures to form a dsRNA enabling Dicer cleavage. However, Dicer does not seem to be implicated in the production of all tRF-5D or tRF-3T and other still unknown endonucleases are likely required. Finally, a vertebrate-specific endonuclease called angiogenin (ANG) has been shown to cleave tRNAs in the anticodon loop generating the long tRFs, tRF-5A and tRF-3A [69]. ANG belongs to the RNase A family and is induced by a large variety of stresses such as oxidative stress, UV irradiation, heat shock or viral infection. The particularity of ANG is that it was first identified as a tumor angiogenic factor acting in the nucleus during cell growth and proliferation. Under normal conditions, ANG localizes primarily to the nucleus and to a lesser extent to the cytoplasm where it is associated with RNH1 (RNase H inhibitor 1), its inhibitor [92]. Under stress conditions, it has been proposed that either the nucleus-localized ribonuclease is exported to the cytoplasm or that the cytoplasmic-localized ANG is dissociated from RNH1 to produce tRNA halves from mature tRNAs [93]. This enzyme seems to play a strong role in a cytoprotective stress response program. ANG catalytic activity is also essential to promote stress granule assembly through tRF production [94]. Importantly, beside the ability to cleave in the anticodon region of tRNAs, ANG plays a crucial role in cleaving the 3'-CCA extremity of tRNAs upon oxidative stress [95]. Indeed, the authors showed that the cleavage by ANG at the level of the CCA extremity occurs much faster than the cleavage in the anticodon region. Eliminating the CCA ends of tRNA molecules represents a very rapid and efficient way to inhibit translation. This process is reversible as the CCA extremities can be rapidly repaired by tRNA-nucleotidyl transferase. This cleavage at the 3' ends of tRNAs reflects the substrate specificity of ANG which preferentially targets single-stranded RNA at the level of a CA motif. While ANG can cleave all tRNAs at their 3' end, only a subset of tRNAs (*i.e.*, those with a CA motif in the anticodon region) will be cleaved in the anticodon loop to generate tRFs. These data suggest that the primary function of ANG is the cleavage of the CCA 3' ends of tRNAs.

**Table 2.** List of endonucleases capable of cleaving tRNAs or pre-tRNAs.

Organisms	Endonuclease	tRNA Specificities	Cleavage Sites	References
	Colicin E5	Queuine-containing tRNA	Anticodon loop	[65]
<i>E. coli</i>	Colicin D	tRNA <sup>Sup</sup>	Anticodon region	[66]
	PreC	tRNA <sup>Sup</sup>	Anticodon loop	[84]
<i>K. lactis</i>	Zymocin	tRNA <sup>Sup</sup> , tRNA <sup>Sup</sup>	Anticodon loop	[86]

Table 2. Cont.

Organism	Endonuclease	tRNA Specificities	Cleavage Sites	References
<i>S. cerevisiae</i>	Rny1	No	Anticodon loop	[87]
<i>T. brucei</i>	Rnt2A-C	No	Anticodon loop	[90]
	ELAC2	pre-tRNA <sup>Asp</sup>	3' end	[52]
		tRNA <sup>Asp</sup>	D loop	[73]
Mammals	Dicer	tRNA <sup>Asp</sup>	T loop	[55]
		pre-tRNA <sup>Asp</sup>	3' end	[91]
	Ango2ts1	No	Anticodon loop	[69]

Concerning short tRFs, with the exception of Dicer in a very few instances, mostly nothing is known concerning the identification and the location of the endonucleases responsible for the cleavage in the D or T regions of mature tRNAs. Similarly, the subcellular location of these tRFs is largely unknown. In the eukaryote form of *Trypanosoma cruzi*, the subcellular location of tRFs has been reported by [55]. Using Fluorescent *In Situ* Hybridization (FISH) experiments, the authors not only observed the recruitment of tRNA halves to specific cytoplasmic granules but also showed that depending on the nature of the tRFs (tRF-Asp or tRF-3A), the RNA fragments localize in distinct cytoplasmic granules. Furthermore, Reifar and collaborators [96] reported a differential location of tRFs depending on the developmental stage of *T. cruzi*. In the epimastigote form, tRFs localize to posterior cytoplasmic granules whereas in the metacyclic form, they are evenly distributed in the cytoplasm. These studies suggest that location of short tRFs may have an important biological significance since they localize in specific structures or compartments where they may interact with particular partners for as yet undescribed functions.

### 5. The Emerging Roles of tRFs

Under nutrient starvation (e.g. phosphate or nitrogen starvation), it is essential to recycle constituents present in stable macromolecules for the cell to survive. For instance, RNA degradation constitutes a very efficient way to deliver phosphate and nitrogen upon starvation. Ribosomal RNAs and tRNAs are the most abundant RNAs and represent, *via* their degradation, major sources for the recycling of these components. Indeed, in *S. cerevisiae* the existence in vacuoles of a nitrogen starvation induced ribophagy pathway involving Rny1 has been identified (Figure 2A). [89]. Thus tRFs are likely to be products of the earliest stages of ribophagy, on the route to providing needed phosphate and nitrogen to the cell. In plants, in response to phosphate deficiency, tRF-3s accumulate at a high level [58,71]. Whether the only fate of this pool of ancRNAs is to provide nutrients to the cell via their complete degradation or whether some of them play other molecular functions is an open question. Evidence of additional functions for some tRFs is emerging in many organisms. Our present knowledge is summarized below and Figure 2 likely represents only the tip of the iceberg.



Three-generated tRF-5Ds are only poorly associated with AGO complexes (Figure 2D). [73]. In contrast, pre-tRF-3Ds were found to be preferentially associated with AGO1 and AGO4 in HEK293 cell lines [57]. Moreover, AGO1-associated tRFs deriving from several human tRNAs [76] and from the 5' end of tRNA<sup>Leu</sup> in *Schistosoma mansoni parvite* were also uncovered [104]. So far, the only identified target for a tRF-AGO complex was shown in mature B cells (Figure 2E). In this case, a Dicer-dependent tRF-3T deriving from tRNA<sup>Arg</sup> (tRF-3T), is bound to AGO and represses RPA1 (Replication Protein A 1), a gene implicated in DNA replication and repair. This is the first evidence that a tRF can play the role of a miRNA [74]. In plants, association of tRFs with several AGO proteins was also observed [105]. At least two hypotheses can be given concerning tRF functions in AGO complexes: (i) Loaded in RISC the tRFs can target a mRNA bearing the antisense sequence to modulate gene expression; (ii) In cooperation with AGO proteins, tRFs compete with the original small RNAs for loading in the RISC complex, repressing therefore the whole RNA silencing pathway. Still related to silencing, work in *Drosophila melanogaster* reported that the siRNA pathway can be affected by heat-shock induced tRFs. Indeed, the authors showed that tRFs are able to inhibit Dicer-2 activity on long dsRNA in a *dicer2* mutant background [106].

In *T. thermophila* (Figure 2F), the implication of tRFs in RNA metabolism has also been demonstrated. A detailed study by Couvillion *et al.* showed that tRF-3Ts interact with Fwi12, a Prw1 protein required for vegetative growth. This is the only Prw1 Argonaute protein (out of eight) essential for cell survival. It is specialized in the loading of tRFs and these tRFs are essential for the nuclear RNA decay pathway. Indeed, the tRF-Fwi12 complexes recruit Tard (Two-associated novel 1) and Xrn2 (exonuclease 2) ribonucleases forming a TXT (Fwi12/Xrn2/Tard) complex. The TXT complex is then imported into the nucleus and stimulates Xrn2 exonuclease activity engaged in RNA processing and degradation pathways [98,107,108].

Further research has established a link between ANG-induced tRNA halves and particular functions. In the case of RSV (Respiratory Syncytial Virus) infection, Wang and collaborators found that an ANG-dependent tRNA half generated from tRNA<sup>Arg</sup> represses target mRNA in the cytoplasm and stimulates RSV replication [62]. Work from Saika *et al.* [99] demonstrated that hyperosmotic stress leads to the accumulation of ANG induced tRNA halves in the cytoplasm of mouse embryonic fibroblasts (Figure 2G). They showed that specific tRF-5As or tRF-3As interact with cytochrome c released from mitochondria to form a ribonucleoprotein complex. This complex interferes with the formation of the apoptosome by preventing the association of cytochrome c with Apaf-1, the apoptotic protease activating factor 1. Altogether this study demonstrates the cytoprotective effect of argipogenin via the production and accumulation of tRNA halves. Another recent function assigned to a tRF is as a primer for reverse transcriptase (Figure 2H). A tRF-3T (Pro) was shown to be capable of priming HTLV-1 (Human T-cell leukemia/lymphoma virus type 1) reverse transcriptase, thus suggesting an important role of some tRFs in viral infection [100].

Not only tRFs deriving from mature tRNAs may have important biological functions, but also those deriving from the 5' end of pre-tRNAs. A pre-tRF-3U called tRF-1001 deriving from the pre-tRNA<sup>Met</sup> was found to be increased in various cancer cell lines and shown to be required for cell viability [52]. Several pre-tRFs-3U, including tRF-1001, were found associated with Argonaute proteins. This supports the idea that this category of tRFs plays a role in the global regulation of RNA silencing [57].

## 6. Conclusions and Perspectives

As key elements of the protein synthesis machinery, tRNAs and rRNAs must be, and are, highly stable macromolecules. Indeed, a rapid turnover of these structural non-coding RNAs would not be beneficial for the cell [32]. Nevertheless, and as described above, there are at least three main reasons to destroy a tRNA.

The first reason is linked to RNA quality control. If tRNA is unable to accurately and efficiently play its role in protein synthesis by correctly charging its cognate amino acid or to perform other additional functions more recently uncovered, its presence may be harmful for the cell. As described above, two major RNA degradation pathways involving exonucleases and acting at the level of either pre-tRNAs or mature tRNAs have evolved to avoid the presence of deleterious misfolded or hypo-modified tRNAs. Most of the work to unravel these pathways has been achieved in the yeast *S. cerevisiae* but we can speculate that these tRNA surveillance pathways are likely conserved among evolutionarily divergent eukaryotes.

The second reason is to recycle key nutrients within the cell. It has been demonstrated that when eukaryotic cells are deprived of various nutrients, there is a retrograde nuclear import of tRNAs which accumulate in the nucleus. The reduction of tRNA abundance in the cytoplasm induces translational repression (for a review, see [109]). Another way for the cell to respond to nutrient deprivation is to degrade its own proteins and RNAs. Many data on tRNAs report their accumulation under stress conditions. It is thus easily conceivable that stable RNAs such as rRNAs and tRNAs are degraded to provide resources to the cell. Quite surprisingly, there is little data in the literature on this aspect of tRNA turnover, and only a few studies on ribophagy are available.

The third reason to cut tRNAs into pieces is to use the pieces for other purposes. First assumed to be simply degradation products, there is now some evidence that specific tRNAs play a role in regulating gene expression. The involvement of tRNAs in a number of human diseases such as cancer or chronic diseases and in various environmental stresses is also emerging (see e.g., [13]). The functions attributed so far to tRNAs are likely only the tip of the iceberg. Many questions still remain regarding both their biogenesis and their functions. From an evolutionary point of view, we can also wonder whether the same tRNAs play identical roles in different eukaryotes. Finally, the role of modified nucleotides in producing functional tRNAs also remains elusive and needs to be addressed in the future.

## Acknowledgments

The work was supported by the French National Program "Investissement d'Avenir" administered by the "Agence Nationale de la Recherche" (ANR) ("Microcosm" Laboratory of Excellence (Labex), funded as ANR-10-IDEX-0002-02), the University of Strasbourg and the CNRS. Stéphanie Lalande is recipient of a fellowship from the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie. Geoffrey Morelle is recipient of a doctoral fellowship supported by the "region Alsace" and by the Australian Research Council Centre of Excellence in Plant Energy Biology (CE0561493). Cyrille Megeil is recipient of a doctoral fellowship supported by the ANR-10-IDEX-0002-02 "Investissements d'Avenir" program.

#### Author Contributions

Cyrille Meyl and Geoffrey Morelle contributed equally in writing the manuscript. Stéphanie Lalancle, Anne-Marie Duchêne and Ian Small contributing equally in improving the manuscript. Laurence Marechal-Drouard was involved in all steps required for the elaboration of the manuscript.

#### Conflicts of Interest

The authors declare no conflict of interest.

#### References

- Holley, R.; Appar, J.; Everett, G.A.; Madison, J.T.; Marqusee, M.; Merrill, S.H.; Penwick, J.R.; Zamir, A. Structure of a ribonucleic acid. *Science* **1965**, *147*, 1462–1465.
- Kim, S.H.; Quigley, G.J.; Suckuth, F.L.; McPherson, A.; Sroden, D.; Kim, J.J.; Weisz, J.; Rich, A. Three-dimensional structure of yeast phenylalanine transfer RNA: Folding of the polynucleotide chain. *Science* **1973**, *179*, 285–288.
- Ladner, J.E.; Jark, A.; Roberts, J.D.; Brown, R.S.; Blocks, D.; Clark, B.F.C.; Klug, A. Structure of yeast phenylalanine transfer RNA at 2.5 angstrom resolution. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4414–4418.
- Wende, S.; Platzer, E.G.; Jähling, F.; Patz, J.; Florentz, C.; Stadler, P.F.; Modl, M. Biological evidence for the world's smallest tRNAs. *Biochimie* **2013**, *100*, 151–158.
- Phizicky, E.M.; Hopper, A.K. tRNA biology changes to the front. *Genes Dev.* **2010**, *24*, 1812–1860.
- Housley, J.; Tollervey, D. The many pathways of RNA degradation. *Cell* **2009**, *136*, 761–776.
- Reznik, B.; Lykke Andersen, J. Regulated and quality control mRNA turnover pathways in eukaryotes. *Biochem. Soc. Trans.* **2010**, *38*, 1506–1510.
- Parker, R. RNA degradation in *Saccharomyces cerevisiae*. *Genetics* **2012**, *191*, 671–707.
- Schmidt, D.R.; Magout, L.B. Regulation of cytoplasmic mRNA decay. *Nat. Rev. Genet.* **2012**, *13*, 246–259.
- Phizicky, E.M.; Grayhack, E.J.; Chernyakov, I.; Whipple, J.M. Roles of tRNA modifications in tRNA turnover. In *DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution*; Grosjean, H., Ed.; Landes Bioscience: Austin, TX, USA, 2009; pp. 564–576.
- Durr, K.; Ibbn, M. Roles of tRNA in cell wall biosynthesis. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 247–264.
- Gebetsberger, J.; Polacek, N. Sheng tRNAs to boost functional ncRNA diversity. *RNA Biol.* **2013**, *10*, 1–9.
- Rana, M.; Ibbn, M. tRNAs as regulators of biological processes. *Front. Genet.* **2014**, *5*, 1–14.
- Anderson, J.; Phan, L.; Casata, R.; Carlson, B.A.; Pak, M.; Asano, K.; Hsok, G.R.; Tamame, M.; Hinnebusch, A.G. The essential God10p-God11p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev.* **1998**, *12*, 3650–3662.
- Anderson, J.; Phan, L.; Hinnebusch, A.G. The God10p/God11p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5173–5178.

16. LaCava, J.; Housley, J.; Savicini, C.; Petlibski, F.; Thompson, F.; Jaouani, A.; Tollervey, D. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **2005**, *121*, 713–724.
17. Kadaba, S.; Krueger, A.; Trice, T.; Krcic, A.M.; Hinnebusch, A.G.; Anderson, J. Nuclear surveillance and degradation of hypomodified initiator tRNA<sup>Met</sup> in *S. cerevisiae*. *Genes Dev.* **2004**, *18*, 1227–1240.
18. Anderson, J.T. RNA turnover: Unexpected consequences of being tailed. *Curr. Biol.* **2005**, *15*, R635–R638.
19. Kadaba, S.; Wang, X.; Anderson, J.T. Nuclear RNA surveillance in *Saccharomyces cerevisiae*: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5s rRNA. *FNA* **2006**, *12*, 508–521.
20. Wang, X.; Ju, H.; Jankowsky, B.; Anderson, J.T. Degradation of hypomodified tRNA<sup>Met</sup> *in vivo* involves RNA-dependent ATPase activity of the DExH helicase Mtr4p. *FNA* **2008**, *14*, 107–116.
21. Mariani, R.J.; Lamichhane, T.N. *F* Processing of eukaryotic precursor tRNAs. *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 162–175.
22. Varasova, S.; Wolf, J.; Martin, G.; Blank, D.; Dettwiler, S.; Imboden, A.; Langen, H.; Keith, G.; Keller, W. A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol.* **2005**, *3*, e189.
23. Cepeda, I.A.; LeRNandez, C.F.; Shemet, R.L.; Wolin, S.L. Competition between the Rex1 exonuclease and the La protein affects both Trf4p-mediated RNA quality control and pre-tRNA maturation. *FNA* **2008**, *14*, 1214–1227.
24. Morf, M.; Marchfelder, A. The final cut: The importance of tRNA 3' processing. *EMBO Rep.* **2001**, *2*, 17–20.
25. Yoo, C.J.; Wolin, S.L. The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell* **1997**, *89*, 393–402.
26. Oranick, S.G.; Wang, X.; Costanzo, M.; Brost, R.L.; Heere, C.; Anderson, J.T. Rex1p deficiency leads to accumulation of precursor initiator tRNA<sup>Met</sup> and polyadenylation of substrate RNAs in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **2009**, *37*, 298–308.
27. Alexandrov, A.; Chernyakov, I.; Gu, W.; Hukey, S.J.; Hughes, T.R.; Grayback, E.J.; Phizicky, E.M. Rapid tRNA decay can result from lack of nonessential modifications. *Mol. Cell* **2006**, *21*, 87–96.
28. Kotclawala, I.; Grayback, E.J.; Phizicky, E.M. Identification of yeast tRNA Ura5 2'-O-methyltransferase (Trm44) and demonstration of a Trm44 role in sustaining levels of specific tRNA<sup>Met</sup> species. *FNA* **2006**, *12*, 158–169.
29. Dewe, J.M.; Whipple, J.M.; Chernyakov, I.; Jaramillo, I.N.; Phizicky, E.M. 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. *FNA* **2012**, *18*, 1886–1896.
30. Sampson, J.R.; Uhlenbeck, O.C. Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed *in vitro*. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 1033–1037.
31. Nobles, K.N.; Yaman, C.S.; Liu, G.; Guenther, R.H.; Apris, P.F. Highly conserved modified nucleosides influence Mg<sup>2+</sup>-dependent tRNA folding. *Nucleic Acids Res.* **2002**, *30*, 4751–4760.

12. Whipple, J.M.; Lane, F.A.; Chernyakov, I.; D'Silva, S.; Phizicky, E.M. The yeast rapid tRNA decay pathway primarily monitors the structural integrity of the acceptor and T-stems of mature tRNA. *Genes Dev.* **2011**, *25*, 1173–1184.
13. Wilusz, J.E.; Whipple, J.M.; Phizicky, E.M.; Sharp, P.A. tRNAs marked with CCACCA are targeted for degradation. *Science* **2011**, *334*, 817–821.
14. Chernyakov, I.; Whipple, J.M.; Kotchirwala, L.; Gnybank, H.J.; Phizicky, E.M. Degradation of several hypomodified mature tRNA species in *Saccharomyces cerevisiae* is mediated by Mei22 and the 5'–3' exonucleases Rat1 and Xrn1. *Genes Dev.* **2008**, *22*, 1369–1380.
15. Watanabe, K.; Miyagawa, R.; Tomikawa, C.; Mizuno, R.; Takahashi, A.; Hori, H.; Ijiri, K. Degradation of initiator tRNA<sup>Met</sup> by Xrn1/2 via its accumulation in the nucleus of heat-treated HeLa cells. *Nucleic Acids Res.* **2013**, *41*, 4671–4685.
16. Shaheen, H.H.; Hopper, A.K. Retrograde movement of tRNAs from the cytoplasm to the nucleus in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 11290–11295.
17. Dichil, B.; Stevens, A.; Tollervey, D. Lithium toxicity in yeast is due to the inhibition of RNA processing enzymes. *EMBO J.* **1997**, *16*, 7144–7155.
18. Pelfalski, E.; Dandekar, T.; Henry, Y.; Tollervey, D. Processing of the precursors to small nuclear RNAs and tRNAs requires common components. *Mol. Cell Biol.* **1998**, *18*, 1181–1189.
19. Estavillo, G.M.; Crisp, P.A.; Pornsinwong, W.; Wirtz, M.; Collinge, D.; Curie, C.; Giraud, E.; Whelan, J.; David, P.; Jawot, H., et al. Evidence for a SAL1-PAP chloroplast retrograde pathway that functions in drought and high light signaling in *Arabidopsis*. *Plant Cell* **2011**, *23*, 3997–4012.
20. Turowski, T.W.; Karkusiewicz, I.; Kowal, J.; Boguta, M. Mal1-mediated repression of RNA polymerase III transcription inhibits tRNA degradation via RTD pathway. *RNA* **2012**, *18*, 1821–1832.
21. Boguta, M.; Graczyk, D. RNA polymerase III under control: Repression and de-repression. *Trends Biochem. Sci.* **2011**, *36*, 451–456.
22. Karkusiewicz, I.; Turowski, T.W.; Graczyk, D.; Towpik, J.; Dhargiel, N.; Hopper, A.K.; Boguta, M. Mal1 protein, repressor of RNA polymerase III, indirectly affects tRNA processing. *J. Biol. Chem.* **2011**, *286*, 39478–39488.
23. Xiong, Y.; Sartz, T.A. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* **2004**, *430*, 640–645.
24. Xiong, Y.; Sartz, T.A. A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. *Curr. Opin. Struct. Biol.* **2006**, *16*, 12–17.
25. Lizano, E.; Scheibe, M.; Rammelt, C.; Betat, H.; Møll, M. A comparative analysis of CCA-adding enzymes from human and *E. coli*: Differences in CCA addition and tRNA 3'-end repair. *Biochemistry* **2008**, *47*, 762–772.
26. Betat, H.; Rammelt, C.; Møll, M. tRNA nucleotidyltransferases: Ancient catalysts with an unusual mechanism of polymerization. *Cell Mol. Life Sci.* **2010**, *67*, 1417–1463.
27. Swaszczyk, A.; Okleja, M.; Dziembowski, A. Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. *RNA Biol.* **2014**, doi:10.4161/rna.34406.
28. Lee, R.C.; Feinbaum, R.L.; Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **1993**, *75*, 843–854.
29. Czech, T.R.; Sartz, T.A. The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* **2014**, *157*, 77–94.

50. Kawap, H., Nakamura, M., Takahashi, Y., Sankeln, A., Katayama, S., Fukuda, S., Daulty, C.O., Kai, C., Kawan, J., Yasuda, J., *et al.* Hidden layers of human small RNAs. *BMC Genomics* **2008**, *9*, 157.
51. Farazi, T.A.; Jarman, S.A.; Tuschl, T. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* **2008**, *135*, 1201–1214.
52. Lee, Y.S.; Shibata, Y.; Malboin, A.; Datta, A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev* **2009**, *23*, 2649–2649.
53. Aravin, A.A.; Lagos Quintana, M.; Yalcin, A.; Zavolan, M.; Marks, D.; Snyder, B.; Gasterland, T.; Meyer, J.; Tuschl, T. The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell* **2003**, *5*, 337–350.
54. Galizi, R.; Spino, F.; Grubisic, M.A.; Capucini, B.; Magini, A.; Urlandi, L.; Oprowski, T.; Doherty, J.P.; Sparacelo, R.; Emiliiani, C., *et al.* Evidence of tRNA cleavage in apicomplexan parasites: tRFs as new potential regulatory molecules of *Toxoplasma gondii* and *Plasmodium berghei*. *Mol. Biochem. Parasitol.* **2013**, *188*, 99–108.
55. Garcia-Silva, M.R.; Fragata, M.; Tovar, J.P.; Correa-Dominguez, A.; Ronalde-Alves, L.; Parodi-Jalce, A.; Rovira, C.; Robello, C.; Goldenberg, S.; Cayota, A. A population of tRNA-derived small RNAs is actively produced in *Trypanosoma cruzi* and recruited to specific cytoplasmic granules. *Mol. Biochem. Parasitol.* **2010**, *171*, 64–73.
56. Gebetsberger, J.; Zywicki, M.; Kunzi, A.; Polacek, N. tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in *Halobacter volcanii*. *Archives* **2012**, *2012*, 260909.
57. Hainsworth, D.; Huang, Y.; Liu, A.; Paramaswaran, P.; Fire, A.Z.; Kay, M.A. Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* **2010**, *16*, 673–695.
58. Büsch, L.C.; Lin, S.L.; Shih, A.C.; Chen, J.W.; Lin, W.Y.; Tseng, C.Y.; Li, W.H.; Chiu, T.J. Uncovering small RNA-mediated responses to phosphate deficiency in *Arabidopsis* by deep sequencing. *Plant Physiol.* **2009**, *151*, 2120–2132.
59. Joehli, C.; Reberstorff, M.; Herdel, J.; Stoffer, P.F.; Hofacker, H.; Schrettel, M.; Haas, H.; Huttenhofer, A. Small ncRNA transcriptome analysis from *Aspergillus fumigatus* suggests a novel mechanism for regulation of protein synthesis. *Nucleic Acids Res.* **2008**, *36*, 2677–2689.
60. Li, Y.; Liao, J.; Zhou, H.; Luo, J.Y.; Ma, J.M.; Chen, Y.Q.; Qu, L.H. Stress-induced tRNA-derived RNAs: A novel class of small RNAs in the primitive eukaryote *Giardia lamblia*. *Nucleic Acids Res.* **2008**, *36*, 6048–6055.
61. Thompson, D.M.; Lu, C.; Green, P.J.; Parker, R. tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA* **2008**, *14*, 2095–2103.
62. Wang, Q.; Lee, I.; Ren, J.; Ajay, S.S.; Lee, Y.S.; Bao, X. Identification and functional characterization of tRNA-derived RNA fragments (tRFs) in respiratory syncytial virus infection. *Mol. Ther.* **2013**, *21*, 368–379.
63. Kawap, H.; Hayashizaki, Y. Exploration of small RNAs. *PLoS Genet.* **2008**, *4*, e22.
64. Peng, H.; Shi, J.; Zhang, Y.; Zhang, H.; Luo, S.; Li, W.; Lei, L.; Han, C.; Ning, L.; Cao, Y., *et al.* A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Res.* **2012**, *22*, 1609–1617.

65. Ooyawa, T.; Tomita, K.; Ueda, T.; Watanabe, K.; Hozumi, T.; Masaki, H. A cytotoxic ribonuclease targeting specific transfer RNA anticodons. *Science* **1999**, *283*, 2097–2100.
66. Tomita, K.; Ooyawa, T.; Hozumi, T.; Watanabe, K.; Masaki, H. A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8278–8283.
67. Lee, S.R.; Collins, K. Starvation-induced cleavage of the tRNA anticodon loop in *Tetrahymena thermophila*. *J. Biol. Chem.* **2005**, *280*, 42744–42749.
68. Lee, S.R.; Collins, K. Two classes of endogenous small RNAs in *Tetrahymena thermophila*. *Genes Dev.* **2006**, *20*, 28–33.
69. Yamasaki, S.; Ivanov, P.; Hu, G.F.; Anderson, P. Argonemin cleaves tRNA and promotes stress-induced translational repression. *J. Cell Biol.* **2009**, *185*, 35–42.
70. Hauser, H.J.; Karpov, I.V.; Hannon, G.J.; Elliot, M.A. Developmentally regulated cleavage of tRNAs in the bacterium *Streptomyces coelicolor*. *Nucleic Acids Res.* **2008**, *36*, 732–741.
71. Esich, L.C.; Lin, S.L.; Kuo, H.F.; Chou, T.J. Abundance of tRNA-derived small RNAs in phosphate-starved *Arabidopsis* roots. *Plant Signal. Behav.* **2010**, *5*, 537–539.
72. Dhabhi, J.M.; Spindler, S.R.; Atarua, H.; Yamakawa, A.; Hoffelli, D.; Mose, P.; Martin, D.J. tRNA halves are present as abundant complexes in serum, concentrated in blood cells, and modulated by aging and caloric restriction. *BMC Genomics* **2013**, *14*, 298.
73. Cole, C.; Sobala, A.; Lu, C.; Fletcher, S.R.; Bowman, A.; Brown, J.W.S.; Green, P.J.; Barton, G.J.; Hutvagner, G. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* **2009**, *15*, 2147–2160.
74. Manic, R.I.; Schneider, C.; Samarin, P.; Holmes, A.; Calvano, A.; Haas, K.; Dalla-Favera, R. tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1404–1409.
75. Wang, L.; Yu, X.; Wang, H.; Lu, Y.Z.; de Raeter, M.; Prins, M.; He, Y.S. A novel class of heat-responsive small RNAs derived from the chloroplast genome of Chinese cabbage (*Brassica rapa*). *BMC Genomics* **2013**, *12*, 289.
76. Burroughs, A.M.; Ando, Y.; de Hoon, M.J.; Tomaru, Y.; Suzuki, H.; Hayashizaki, Y.; Daub, C.O. Deep-sequencing of human Argonemin-associated small RNAs provides insight into miRNA sorting and reveals argonemin association with RNA fragments of diverse origin. *RNA Biol.* **2011**, *8*, 158–177.
77. Luo, J.Y.; Ma, F.M.; Gao, Y.H.; Zhang, Y.C.; Zhou, H.; Shao, P.; Chen, Y.Q.; Qi, F.H. Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 5' trailers. *PLoS One* **2010**, *5*, e10563.
78. Saikia, M.; Krolowski, D.; Guan, H.J.; Ivanov, P.; Pansien, M.; Hu, G.F.; Anderson, P.; Pan, T.; Hatzeoglou, M. Genome-wide identification and quantitative analysis of cleaved tRNA fragments induced by cellular stress. *J. Biol. Chem.* **2012**, *287*, 42708–42715.
79. Dhabhi, J.M. Circulating small noncoding RNAs as biomarkers of aging. *Ageing Res Rev.* **2014**, *17*, 86–98.
80. Nowacka, M.; Stumyczki, P.M.; Jackowski, P.; Hojka-Omanska, A.; Szymanski, M.; Figlerowicz, M. Identification of stable, high copy number, medium-sized RNA degradation intermediates that accumulate in plants under non-stress conditions. *Plant Mol. Biol.* **2013**, *83*, 191–204.

81. Anderson, P.; Ivanov, P. tRNA fragments in human health and disease. *FEBS Lett.* **2014**, *568*, 4797–4804.
82. Fu, H.; Feng, J.; Liu, Q.; Sun, F.; Tie, Y.; Zhu, J.; Xing, R.; Sun, Z.; Zheng, X. Stress induces tRNA cleavage by angiotensin in mammalian cells. *FEBS Lett.* **2009**, *583*, 437–442.
83. Hanada, T.; Wentzer, S.; Maier, B.; Bernreuther, C.; Wainger, B.J.; Ichiki, J.; Hanada, R.; Ordoñez, M.; Cronin, S.J.; Komnenovic, V., et al. C1P1 links tRNA metabolism to progressive motor neuron loss. *Nature* **2013**, *495*, 474–480.
84. David, M.; Borasio, G.D.; Kaufmann, G. Bacteriophage T4 induced anticodon loop nuclease detected in a host strain restrictive to RNA ligase mutants. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 7097–7101.
85. Blampa-Kamfi, S.; Amisur, M.; Azem, A.; Kaufmann, G. ProC-anticodon nuclease: Functional organization of a prototypical bacterial restriction RNase. *Nucleic Acids Res.* **2006**, *34*, 3209–3219.
86. Lu, J.; Esberg, A.; Huang, B.; Hymowitz, A.S. *Kluyveromyces fragilis* γ-toxin, a ribonuclease that recognizes the anticodon stem loop of tRNA. *Nucleic Acids Res.* **2008**, *36*, 1077–1080.
87. Thompson, D.M.; Parker, R. The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stress in *Saccharomyces cerevisiae*. *J. Cell Biol.* **2009**, *185*, 43–50.
88. Lahtala, S.; Parker, R. Structure-function analysis of Rny1 in tRNA cleavage and growth inhibition. *PLoS One* **2012**, *7*, e41111.
89. Huang, H.; Kawamura, T.; Horie, T.; Tsugawa, H.; Nakayama, Y.; Ohsami, Y.; Fukunaki, E. Bulk tRNA depletion by nitrogen starvation-induced autophagy in yeast. *EMBO J.* **2014**, doi:10.15252/embo.201489083.
90. Anderson, K.L.; Collins, K. Several RNase T7 enzymes function in induced tRNA and rRNA turnover in the cilium *Tetrahymena*. *Mol. Biol. Cell* **2011**, *23*, 36–44.
91. Babiarz, J.E.; Ruby, J.G.; Wang, Y.; Bartel, D.P.; Blelloch, R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.* **2008**, *22*, 2773–2785.
92. Li, S.; Hu, G.F. Emerging role of angiotensin in stress response and cell survival under adverse conditions. *J. Cell. Physiol.* **2012**, *227*, 2822–2826.
93. Thompson, D.M.; Parker, R. Stressing out over tRNA cleavage. *Cell* **2009**, *138*, 215–219.
94. Emami, M.M.; Ivanov, P.; Hickman, T.; Bawa, N.; Tisdale, S.; Kerkela, N.; Hu, G.F.; Anderson, P. Angiotensin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *J. Biol. Chem.* **2010**, *285*, 10959–10968.
95. Czech, A.; Wende, S.; Mehl, M.; Pan, T.; Ignatova, Z. Reversible and rapid transfer-RNA deactivation as a mechanism of translational repression in stress. *PLoS Genet.* **2013**, *9*, e1003067.
96. Reifur, L.; Garcia-Silva, M.R.; Paschel, S.H.; Alves, I.R.; Amico, P.; Buar, D.K.; Gokkerberg, S.; Cayota, A.; Dallagiovanna, B. Distinct subcellular localization of tRNA-derived fragments in the infective metacyclic forms of *Trypanosoma cruzi*. *Mem. Inst. Oswaldo Cruz* **2012**, *107*, 816–819.
97. Ivanov, P.; Emami, M.M.; Villen, J.; Gygi, S.P.; Anderson, P. Angiotensin-induced tRNA fragments inhibit translation initiation. *Mol. Cell* **2011**, *43*, 613–623.

98. Coscillaro, M.T.; Boumova, G.; Pando, E.; Speed, T.P.; Collins, K. A *Tetrahymena* Pwi1 bound to mature tRNA<sup>Y</sup> fragments activates the exonuclease Xrn2 for RNA processing in the nucleus. *Mol. Cell* **2012**, *46*, 509–520.
99. Sakai, M.; Jobava, R.; Parsian, M.; Putnam, A.; Krokowski, D.; Gao, X.H.; Guan, H.J.; Yuan, Y.; Jankowski, E.; Feng, Z., et al. Angiogenin-cleaved tRNA halves interact with cytochrome c, protecting cells from apoptosis during osmotic stress. *Mol. Cell Biol.* **2014**, *34*, 2450–2463.
100. Ruggiero, K.; Guffanti, A.; Comodin, A.; Sharma, V.K.; De Bellis, G.; Conti, G.; Grassi, A.; Zancovello, P.; Bronte, V.; Canniale, V., et al. 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. *J. Virol.* **2014**, *88*, 3612–3622.
101. Ivanov, P.; O'Day, E.; Emara, M.M.; Wagner, G.; Lieberman, J.; Anderson, P. G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 18201–18206.
102. Wang, H.; Zhang, X.; Liu, J.; Kiba, T.; Woo, J.; Ojo, T.; Hafner, M.; Tuschl, T.; Chao, N.H.; Wang, X.J. Deep sequencing of small RNAs specifically associated with *Arabidopsis* AGO1 and AGO4 uncovers new AGO functions. *Plant J.* **2011**, *67*, 292–304.
103. Kumar, P.; Araya, J.; Madhavi, S.B.; Datta, A. Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol.* **2014**, *12*, 78.
104. Bulter, M.; Spies, N.; Bartel, D.P.; Mougel, D. TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the *Schistosoma* *parvum* siRNA pathway. *Nat. Struct. Mol. Biol.* **2008**, *15*, 1015–1021.
105. Loss Morán, G.; Waterhouse, P.M.; Marjot, R. Description of plant tRNA-derived RNA fragments (TRFs) associated with Argonaute and identification of their putative targets. *Biol. Direct* **2013**, *8*, 6.
106. Dunckvic, Z.; Mohan, M.B.; Hanna, K.; Lyko, F.; Schaefer, M. The RNA methyltransferase Dnm2 is required for efficient Dicer 2 dependent siRNA pathway activity in *Drosophila*. *Cell Rep.* **2013**, *4*, 931–937.
107. Coscillaro, M.T.; Lee, S.R.; Hogstad, B.; Malone, C.D.; Toekin, L.A.; Sachidanandan, R.; Harrison, G.J.; Collins, K. Sequence, biogenesis, and function of diverse small RNA classes bound to the Pwi1 family proteins of *Tetrahymena thermophila*. *Genes Dev.* **2009**, *23*, 2016–2032.
108. Coscillaro, M.T.; Sachidanandan, R.; Collins, K. A growth-essential *Tetrahymena* Pwi1 protein carries tRNA fragment cargo. *Genes Dev.* **2010**, *24*, 2742–2747.
109. Hoppe, A.K. Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*. *Genetics* **2013**, *194*, 43–67.

## 5 – Travaux de thèse

Mon travail de thèse s'est organisé autour de deux grands axes d'études. Le premier a porté sur l'analyse des tRFs *in vivo* (Parties I et III) et le second sur la détermination des voies de biogenèse de ces tRFs (Parties II).

### 5.1 - Analyse de la population en tRFs chez *A. thaliana*

Si les tRFs ont été identifiés chez de nombreux organismes (*i.e* homme, *S. cerevisiae*, *T. thermophila*, *T. cruzi*), peu d'information était disponible chez *A. thaliana* au moment où j'ai débuté ma thèse. Ainsi, la nature des tRFs (cytosolique, plastidial et mitochondrial), leur séquence (de quel tRNA proviennent t'ils ?), leur abondance, ou le domaine dans lequel les tRNA sont clivés étaient mal connus. Nous avons donc choisi, dans un premier temps, d'analyser différentes banques de données de petits ARN issus d'*A. thaliana* dans le but d'en dresser une liste quasi exhaustive.

### 5.2 - Biogenèse des tRFs chez *A. thaliana*

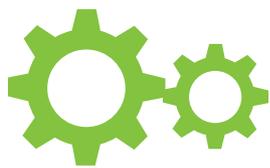
La seconde partie de mes travaux a porté sur la caractérisation des enzymes responsables de la biogenèse des tRFs chez *A. thaliana*. Nous nous sommes servis des connaissances déjà disponibles dans la littérature afin de déterminer si des homologues aux protéines clivant les tRNA dans d'autres organismes existaient chez *A. thaliana*. À partir de ces données nous avons étudié différents mutants d'insertion d'*A. thaliana*. Ceux-ci, KO pour les homologues à la protéine DICER humaine (dcl : dicer-like) nous ont alors été fournis par l'équipe de P. Dunoyer (IBMP). En parallèle de cette approche, pour les homologues chez *A. thaliana* dont nous n'avions pas les lignées d'insertion à disposition (*e.g* RNY1 de levure, angiogénine A humaine ou Rnt2 de *T. thermophila*), ces enzymes ont été clonées à partir de leurs cDNA afin de tester leurs activités *in vitro*.



### 5.3 - Les tRFs et le séquençage haut débit

Cette dernière partie de ma thèse n'a vu le jour qu'après l'identification des enzymes responsables de la biogenèse des tRFs chez *A. thaliana* : les RNS. Ces enzymes clivent les tRNA en libérant une extrémité 5'-OH et 3'-P ce qui pourrait empêcher leur séquençage. J'ai donc eu pour objectif de préparer des banques de petits ARN de Col 0 et de les traiter de manière à améliorer la qualité du clonage des tRFs dans ces banques. La seconde partie de cette étude a ensuite visé à comprendre, *via* l'analyse bioinformatique de ces banques, les changements qui se sont opérés entre les différentes classes de tRFs (*i.e* tRF-5 tRF-3 et les autres tRFs).





# *MATÉRIEL ET MÉTHODES*



# Matériels et méthodes

## 1 - Matériels

### 1.1 - Matériels vivants

#### 1.1.1 - Bactéries

##### **Escherichia coli**

Trois souches *d'E. coli* ont été utilisées au cours de mes travaux. Les deux premières ont été utilisées afin d'amplifier les vecteurs recombinants. Toutes deux présentent une mutation dans le gène *endA1* inactivant une endonucléase intracellulaire, et une seconde mutation dans le gène *recA* éliminant les recombinaisons homologues.

TOP 10 : F- *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\phi$ 80*lacZ* $\Delta$ M15  $\Delta$ *lacX74* *nupG* *recA1* *araD139*  $\Delta$ (*ara-leu*)7697 *galE15* *galK16* *rpsL*(StrR) *endA1*  $\lambda$ -

DH5 $\alpha$  : F- *endA1* *glnV44* *thi-1* *recA1* *relA1* *gyrA96* *deoR* *nupG*  $\Phi$ 80*dlacZ* $\Delta$ M15  $\Delta$ (*lacZYA-argF*)U169, *hsdR17*(rK- mK+),  $\lambda$ -

La troisième souche possède le gène codant pour la polymérase T7 sous contrôle d'un promoteur *lacUV5* inductible à l'IPTG. Elle a été utilisée pour l'expression et la production de protéines recombinantes en système bactérien.

BL21 (DE3) : F- *ompT* *gal* *dcm* *lon* *hsdSB*(rB- mB-)  $\lambda$ (DE3 [*lacI* *lacUV5*-T7 gene 1 *ind1* *sam7* *nin5*])

##### **Agrobacterium tumefaciens** :

Les souches *d'A. tumefaciens* C58C1 et GV3101 ont été utilisées afin de transformer les plantes *d'A. thaliana* et *N. benthamiana* par un vecteur binaire contenant les constructions que nous souhaitons exprimer. La souche GV3101



est une souche produite par l'université de Gand, obtenue par guérison de la souche sauvage C58 du plasmide inducteur de tumeur pTic58. Ces 2 souches, résistantes à la rifampicine, ont été utilisées pour des transformations stables d'*A. thaliana*, en trempant les fleurs dans une solution contenant les bactéries. Elles ont aussi servi à la transformation transitoire de feuilles de *N. benthamiana* par injection d'une solution contenant les bactéries.

### 1.1.2 - Plantes

#### **Arabidopsis thaliana**

*A. thaliana* est l'organisme modèle de notre étude. L'écotype Columbia 0 a été utilisé au cours de nos expériences.

#### **Nicotiana benthamiana**

Les plantes de *N. benthamiana* ont été utilisées pour l'expression transitoire de différentes constructions (fusion avec la GFP) afin d'étudier la localisation des ribonucléases (RNS).

## 1.2 - ADN

### 1.2.1 - Plasmides

#### **Classiques**

**-pGEM :** commercialisé sous le nom pGEM-T Easy Vector par PROMEGA®. Ce plasmide linéaire a la particularité de posséder une extrémité sortante en 3' constituée d'un T. Il est donc plus facile d'y cloner des fragments d'ADN amplifiés par PCR par des Taq polymérase qui ajoutent un A en 3' de la séquence amplifiée.

**-pBIN19 :** est un vecteur binaire qui permet le transfert, par *A. tumefaciens*, du matériel génétique se trouvant entre les « Left » Border (LB) et



« Right » Border (RB). Il contient un ADN-T ayant une cassette de clonage en aval du promoteur fort 35S. Il confère une résistance à la kanamycine aux plantes, comme aux bactéries.

- **pTMV-MP:RFP** : a été obtenu par insertion de la mRFP1 dans le plasmide pTf5-nx2 (Ashby et al., 2006). Il permet la production de la RFP couplée à un signal HDEL en C-ter de rétention dans le réticulum endoplasmique.

- **pRS300** : correspond au vecteur pBSK dans lequel a été cloné le précurseur du miR319a ce vecteur est utilisé dans le clonage des amiRNA (Schwab et al., 2006).

### **Plasmides pour transcription *in vitro***

- **pUC18-tRNA ALA** : vecteur pUC18 dans lequel à été cloné la séquence du tRNA Ala UGC cytosolique afin de produire des tRNA transcrits *in vitro* sous contrôle d'un promoteur T7. En aval de la séquence du tRNA, se trouve le site de restriction BstNI qui permettra le dégagement d'une extrémité CCA en 3' de la séquence du tRNA.

- **pUC19-tRNA HIS** : vecteur pUC18 dans lequel à été cloné la séquence du tRNA His GUG cytosolique afin de produire des tRNA transcrits *in vitro* sous contrôle d'un promoteur T7. En aval de la séquence du tRNA, se trouve le site de restriction BstNI qui permettra le dégagement d'une extrémité CCA en 3' de la séquence du tRNA.

### **Gateway**

La plupart des plasmides utilisés dans cette étude proviennent de l'université de Gand (Belgique). Ils sont disponibles à cette adresse : <http://gateway.psb.ugent.be>.

-**pDONR207** : ce vecteur commercialisé par INVITROGEN® est un vecteur dit "accepteur", c'est-à-dire qu'il n'a pour vocation que le clonage et la multiplication dans *E. coli* d'un insert recombiné entre les sites attP1 et attP2.



Il confère une résistance à la gentamycine aux bactéries transformées.

**-pHGWA** : est un vecteur dit de “destination”. Il porte les sites attR1 et attR2 qui permettent le clonage de séquence *via* une seconde étape de recombinaison. Il est utilisé pour l’expression inductible de protéine dans la bactérie. Ce vecteur possède deux étiquettes 6xHis possibles, une en N-terminal, ou une en C-terminal de sa cassette de clonage. Il confère une résistance à l’ampicilline aux bactéries.

**-pB7WGF2.0** : le vecteur pB7WGF2.0 est un vecteur binaire de destination. Il permet l’expression *in planta* sous contrôle d’un promoteur 35S d’une protéine en fusion N-terminal avec la GFP. Il confère une résistance à la spectinomycine aux bactéries ainsi qu’une résistance au glufosinate chez les plantes.

**-pB7FWG2.0** : ce vecteur binaire est similaire au précédent, si ce n’est que la fusion avec la GFP est réalisée en C-terminal.

**-pB2GW7** : ce vecteur binaire ne possède aucune étiquette. Il possède néanmoins un promoteur 35S, il est donc utilisé dans le cadre de l’expression de protéines *in planta*. Il confère les mêmes résistances que les deux vecteurs précédents.

**-pGWB615** : est un vecteur binaire de surexpression protéique chez les plantes, avec un tag 3xHA en N-terminal. Il confère les résistances spectinomycine aux bactéries et glufosinate aux plantes.

**-pGWB14** : est un vecteur binaire de surexpression protéique chez les plantes avec un tag 3xHA du côté C-terminal. Il confère les résistances kanamycine aux bactéries et hygromycine aux plantes.

**-pH2GW7** : le vecteur pH2GW7 est un vecteur d’expression binaire chez les plantes. Il permet l’expression d’une construction sous contrôle d’un



promoteur 35S. Il confère les résistances spectinomycine aux bactéries ainsi que l'hygromycine aux plantes.

**-pB2GW7** : même vecteur que précédemment mis à part que ce vecteur rend la plante résistante au glufosinate.

**-pK2GW7** : même vecteur que précédemment mis à part que ce vecteur rend la plante résistante à la kanamycine.

### 1.2.2 - Oligonucléotides

Les oligonucléotides ont été synthétisés soit par Eurofins (France), soit par Sigma-Aldrich (Australie). Les oligonucléotides ont été utilisés pour des clonages, de la mutagenèse dirigée ou des expériences d'hybridations. Le  $T_m$  est calculé selon la formule :  $T_m(^{\circ}C) = 2 \times (A+T) + 4 \times (C+G)$ .

Pour les expériences d'hybridations, les sondes utilisées ont été marquées radioactivement au laboratoire.

#### Clonage des RNS

Rôle de l'oligo	Nom de l'oligo nucléotide	Séquence (5' à 3')
Clonage des RNS à partir d'ADNc	RNS1-5'	ATGAAGATTCTTCTAGCATCATTG
	RNS1-3'	CTAAAAAGAAGGGAATTCGATCTC
	RNS2-5'	ATGGCGTCACGTTTATGTCTTCTC
	RNS2-3'	TCAAAGAGCTTCTCTTTCTGTTGG
	RNS3-5'	ATGAAATTCTTCATTTTTATTCTAGCG
	RNS3-3'	TTAGAACTTGGGAAATTGAACTC
	RNS4-5'	ATGAGAGGAATCATCATCGTAAGC
	RNS4-3'	TTAGGTAGTTGAAACAGAGGACATC
	RNS5-5'	ATGGGAGCAAAGGGTTGTGTTAACG
	RNS5-3'	TTAGAACTTAGAAACTGGATTTGAG



Clonage des RNS en système gateway avec fusion N -ter	N-ter RNS1-5' site attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAGATTCTTCTAGCATCATTGTGTTTG
	N-ter RNS1-3' site attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAAAAAGAAGGGAATTCGATCTCA
	N-ter RNS2-5' site attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGCGTCACGTTTATGTCTTCTCC
	N-ter RNS2-3' site attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAAGAGCTTCTCTTCTGTTGG
	N-ter RNS3-5' site attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAATTCCTATTTTTATTCTAGCGTTA
	N-ter RNS3-3' site attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGAAGCTGGGAAATTGAACTCG
	N-ter RNS4-5' site attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAGAGGAATCATCATCGTAAGCT
	N-ter RNS4-3' site attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGTAGTTGAAACAGAGGACAT
Mutagenèse dirigée du premier site actif des RNS	RNS1 H65 en A sens	TTAGGCCAAAGACCAGCAATGCCAAATCAGC
	RNS1 H65 en A antisens	GCTGATTTTGGCATTGCTGGTCTTTGGCCTAA
	RNS2 H71 en A sens	TCAGGCCATAACCCAGCAATTGTGAATTGAGT
	RNS2 H71 en A antisens	ACTCAATTCACAATTGCTGGTTATGGCCTGA
	RNS3 H57 en A sens	TTAGGCCAAAGACCGCAATCCAAATCTGC
	RNS3 H57 en A antisens	GCAGATTTTGAATTGCCGGTCTTTGGCCTAA
	RNS4 H61 en A sens	TGAGGCCAGAGTCCAGCGATGATAAAATCCGA
RNS4 H61 en A antisens	TCGGATTTTATCATCGCTGGACTCTGGCCTCA	
Mutagenèse dirigée du second site actif des RNS	RNS1 H124 en A sens	TCAGAGCAAGTACCAGCCTTCTCCACTCGTG
	RNS1 H124 en A antisens	CACGAATGGGAGAAGGCTGGTACTTGCTCTGA
	RNS2 H133 en A sens	GAAGAACAAGTCCAGCTTTCTCCACTCGTG
	RNS2 H133 en A antisens	CACGAGTGGGAGAAAGCTGGGACTTGTCTTC
	RNS3 H115 en A sens	TCAGGCGACGTACCGGCTTTCTCCACTCATG
	RNS3 H115 en A antisens	CATGAGTGGGAGAAAGCCGGTACGTGCGCTGA
	RNS4 H121 en A sens	TGGACACAGGTGCCAGCCTTATCCACTCGTG
	RNS4 H121 en A antisens	CACGAGTGAATAAGGCTGGCACCTGTGTCCA
Clonage des RNS en système gateway avec fusion C-ter	C-ter RNS1-5' site attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAAGATTCTTCTAGCATCATTGTGTTTG
	C-ter RNS1-3' site attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTAAAAAGAAGGGAATTCGATCTCAGCTCCAC
	C-ter RNS2-5' site attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGCGTCACGTTTATGTCTTCTCCTCTC
	C-ter RNS2-3' site attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTAAAGAGCTTCTCTTCTGTTGGCATCTTCA
	C-ter RNS3-5' site attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAAATTCCTATTTTTATTCTAGCGTTA

Clonage des RNS en système gateway	C-ter RNS3-3' site attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTTGAACTTGGGAAATTGAACTCGAGAGTCGC
	C-ter RNS4-5' site attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAGAGGAATCATCATCGTAAGCTTGTTA
	C-ter RNS4-3' site attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTTGGTAGTTGAAACAGAGGACATCACACTTA

## Clonage des amiRNA

Rôle de l'oligo	Nom de l'oligo nucléotide	Séquence (5' à 3')
Clonage des amiARN dans le pRS300	RNS1-I	GATTATGGTATATACTGTACCTGCTCTCTTTTGTATTCCA
	RNS1-II	AGCAGGTACAGTATATACCATAATCAAAGAGAATCAATGA
	RNS1-III	AGCAAGTACAGTATAAACCATATTCACAGGTCGTGATATG
	RNS1-IV	GAATATGGTTTATACTGTACTTGCTACATATATATTCCTA
	RNS1-I'	GATTATGGTATATACTGTACCCCTCTCTTTTGTATTCCA
	RNS1-II'	AGGGGTACAGTATATACCATAAATCAAAGAGAATCAATGA
	RNS1-III'	AGGGATACAGTATATTCATAATTCACAGGTCGTGATATG
	RNS1-IV'	GAATTATGGAATATACTGTATCCCTACATATATATTCCTA
	RNS2-I	GATTGCCGGAATACACGCCATTACTCTCTTTTGTATTCCA
	RNS2-II	AGTAATGGCGTGTATTCCGGCAATCAAAGAGAATCAATGA
	RNS2-III	AGTACTGGCGTGTATACCGGCATTCACAGGTCGTGATATG
	RNS2-IV	GAATGCCGGTATACACGCCAGTACTACATATATATTCCTA
	RNS2-I'	GATAGAGCGAAATAATCGACCTCCTCTCTTTTGTATTCCA
	RNS2-II'	AGGAGGTCGATTATTTTCGCTCTATCAAAGAGAATCAATGA
	RNS2-III'	AGGAAGTCGATTATTACGCTCTTTCACAGGTCGTGATATG
	RNS2-IV'	GAAAGAGCGTAATAATCGACTTCCTACATATATATTCCTA
	RNS3-I	GATTTGTTACATTCAAGGCCCGCTCTCTTTTGTATTCCA
	RNS3-II	AGCCGGGCCTTGAATGTAACAAATCAAAGAGAATCAATGA
	RNS3-III	AGCCAGGCCTTGAATCTAACAATTCACAGGTCGTGATATG
	RNS3-IV	GAATTGTTAGATTCAAGCCTGGCTACATATATATTCCTA
	RNS3-I'	GATTTTTACCATGCAAATCACAGCTCTCTTTTGTATTCCA
	RNS3-II'	AGCTGTGATTGTCATGGTAAAAATCAAAGAGAATCAATGA
	RNS3-III'	AGCTATGATTGTCATCGTAAAATTCACAGGTCGTGATATG
	RNS3-IV'	GAATTTTACGATGCAAATCATAGCTACATATATATTCCTA
	RNS4-I	GATTGGATGATAAAAACCGACGCTCTCTTTTGTATTCCA
	RNS4-II	AGGCGTCGGTTTTTATCATCCAATCAAAGAGAATCAATGA
	RNS4-III	AGGCATCGGTTTTTAAACATCCATTCACAGGTCGTGATATG
	RNS4-IV	GAATGGATGTTAAAACCGATGCCTACATATATATTCCTA
	RNS4-I'	GATTATCCGGAACATACGCGCATCTCTCTTTTGTATTCCA
	RNS4-II'	AGATGCGCGTATGTTCCGGATAATCAAAGAGAATCAATGA
	RNS4-III'	AGATACGCGTATGTTGCGGATATTCACAGGTCGTGATATG
	RNS4-IV'	GAATATCCGCAACATACGCGTATCTACATATATATTCCTA
	pRS300-A	CTGCAAGGCGATTAAGTTGGGTAAC

Clonage des amiARN en système gateway	pRS300-A-GW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTGCAAGGCGATTAAGTTGGGTAAC
	pRS300-B-GW	GGGGACCACTTTGTACAAGAAAGCTGGGTGCGGATAACAATTCACACAGGAAACAG
	AttB1	GGGGACAAGTTTGTACAAAAAAGCAGGCT
	AttB2	GGGGACCACTTTGTACAAGAAAGCTGGGT

### Northern Blots

Les sondes utilisées pour le northern blot sont nommées de la façon suivante : tRN-code à une lettre des acides aminés - séquence de l'anticodon du tRNA - extrémité du tRNA à laquelle la sonde s'hybride - génome portant le tRNA ciblé.

Rôle de l'oligo	Nom de l'oligo nucléotide	Séquence (5' à 3')	Numéro d'accension
Sonde spécifique de l'extrémité 5' des tRNA	tRNAA- AGC/UGC/CGC-5'- cyto	ACCATCTGAGCTACATCCCC	At5G55505, At5G43455, At5G61445, At1G06610, At1G07580, At2G02600, At2G02900, At3G28695, At3G28685, At4G26675, At4G12065, At4G11355, At5G39895, At5G01365, At5G07675
	tRND-GUC-5'-cyto	GGGAATACTTACCACTATACTACAAC	At5G40545, At5G21378, At5G15815, At5G59055, At5G05945, At5G14495, At3G02335, At3G13335, At3G59926, At3G51265, At3G27555, At3G52765, At2G33660, At2G33650, At2G39860, At1G49020, At1G76000, At1G73900, At1G75070, At1G18430, At1G17660, At1G20210, At1G03515, At5G32017
	tRNR-UCG-5'-cyto	ATCCACTGCGCTATGCGGTC	At5G02615, At3G09505, At1G11030, At1G52350, At1G61020, At1G79980
	tRNR-ACG-5'-plaste	ATCCTCTGAGCTACAAGCCC	AtCG00980, AtCG01150
	tRNV-UAC-5'-plaste	TGGTAGGGCTATACGGACTC	AtCG00450
	tRNL-UAA-5'-plaste	ACCAATTCGCCATATCCCC	AtCG00400
	tRNG-GCC-5' plaste	TACCATTGACTATATCCGC	AtCG00310
	tRNQ-UUG-5' plaste	TACCGCTTGGCTACGCCCA	AtCG00060
	tRNE-UUC-5' plaste	AACCACTAGACGATGGGGGC	AtCG00250
tRNH-GUG-5' plaste	TCCACTTGGCTACATCCGCC	AtCG00010	
Sonde spécifique de l'extrémité 3' des tRNA	tRNI-GAU-3'-cyto	TGGTGGGCCATCCTGGATTT	AtCG01200, AtCG00930
	tRNA-UGC-3'-cyto	TGGTGGAGATAAGCGGACTC	AtCG01190, AtCG00940
	tRNK-UUU-3'-cyto	TGGTGGGTTGCCCGGACTC	AtCG00030

## qPCR

Rôle de l'oligo	Nom de l'oligo nucléotide	Séquence (5' à 3')	Numéro d'accèsion
Quantification par qPCR du transcrit RNS1	RNS1 sens	TCTTCTCGTAATCTTGCCTTCTG	At2G02990
	RNS1 antisens	ACAGTATGATCCTGGCCATTG	
Quantification par qPCR du transcrit RNS2	RNS2 sens	TGGGAAAGGGTCATTTTGG	At2G39780
	RNS2 antisens	TGCTTCAAGTAGAGATTAAGTGTTGG	
Quantification par qPCR du transcrit RNS3	RNS3 sens	TCGCCCAAGATTTTCGATT	At1G26820
	RNS3 antisens	GCAACAACATATGTCTTGAATCACA	
Quantification par qPCR du transcrit RNS4	RNS4 sens	TTCAATTTCTTCTACTGGGTCAACT	At1G14210
	RNS4 antisens	GGCAACATCCTTTTGGCTA	
Quantification par qPCR du transcrit Tip 41-Like	Tip 41 sens	GTGAAAAGTGTGGAGAGAAGCAA	At4G34270
	Tip 41 antigens	TCAACTGGATACCCTTTTCGCA	
Quantification par qPCR du transcrit Exp	Exp sens	GAGCTGAAGTGGCTTCAATGAC	At4G26410
	Exp antisens	GGTCCGACATACCCATGATCC	
Quantification par qPCR du transcrit FBox	FBox sens	TTTCGGCTGAGAGGTTTCGAGT	At5G15710
	FBox antisens	GATTCCAAGACGTAAGCAGA	
Quantification par qPCR du transcrit GAPDH	GAPDH sens	CAGTCACTGTTTTCGGCATC	At3G04120
	GAPDH antisens	AAGACACCAGTAGACTCAACAACG	

## 1.3 – Anticorps

Rôle de l'anticorp	Nom de l'anticorp	Origine
Anticorp primaire	@RNS1	Anticorp polyclonal de lapin dirigé contre RNS1 (Produit par le laboratoire)
	@RNS2	Anticorp polyclonal de lapin dirigé contre RNS2 (Produit par le laboratoire)
	@RNS3	Anticorp polyclonal de lapin dirigé contre RNS3 (Produit par le laboratoire)
	@RNS4	Anticorp polyclonal de lapin dirigé contre RNS4 (Produit par le laboratoire)
	@GFP	Anticorp polyclonal de lapin dirigé contre la GFP (Produit par l'équipe de A. Dietrich, IBMP)
	@ HA	Anticorp monoclonal de souris dirigé contre un épitope HA (Sigma)
	@ 5xHis	Anticorp monoclonal de souris couplé à la HRP dirigé contre un épitope 5xHis (Qiagen®)
Anticorp secondaire	@ Souris	Anticorp polyclonal de chèvre couplé à la HRP et dirigé contre des anticorps de souris (Sigma)
	@ Lapin	Anticorp polyclonal de chèvre couplé à la HRP et dirigé contre des anticorps de lapin (Life Technologies)

## 1.3 – Outils informatiques

- Alignements de séquences: les alignements de séquences acides nucléiques et protéiques ont été réalisés sur le site de l'EBI (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) (logiciel ClustalOmega)

- Blast: les recherches de séquences nucléotidiques et protéiques par homologie ont été réalisées à l'aide du logiciel Blast (Basic Local Alignment Search Tool) développé par le NCBI (National Center of Biotechnology Information) (<http://blast.st.va.ncbi.nlm.nih.gov/Blast.cgi>).

- Prédiction de structure: les structures des protéines étudiées ont été réalisées à l'aide du logiciel Robetta disponible à l'adresse : <http://rosetta.bakerlab.org>.

- Prédiction de localisation subcellulaire: la localisation des protéines étudiées a été prédite à l'aide du logiciel SUBA3 développé par l'ARC Center of Excellence In Plant Energy Biology disponible à l'adresse : <http://suba.plantenergy.uwa.edu.au>.

- Banques de données utilisées:

TAIR : <http://www.arabidopsis.org>

plantRNA : <http://plantrna.ibmp.cnrs.fr>

NCBI : <http://www.ncbi.nlm.nih.gov>

SAIL : <http://arabidopsis.info>

SALK : <http://signal.salk.edu/cgi-bin/tdnaexpress>

RIKEN : <https://www.brc.riken.jp/lab/epd/catalog/cdnaclone.html>



## 2 – Méthodes

### 2.1 – ADN

#### 2.1.1 - Extraction des plasmides

En Australie les plasmides ont été purifiés à l'aide du kit "QIAprep Spin Miniprep Kit" commercialisé par Qiagen© qui utilise le principe de la lyse alcaline. En France, les plasmides ont été purifiés à l'aide du kit "Nucleospin Plasmid" de Macherey-Nagel.

#### 2.1.2 - Extraction d'ADN génomique et génotypage

Des extractions d'ADN génomique (ADNg) ont été effectuées à chaque étape de sélection des plantes transformées.

Pour chaque échantillon, un morceau de feuille d'*A. thaliana* d'environ 1 cm de diamètre est placé dans un tube PreCellys® contenant 400µL de **tampon d'extraction** et l'équivalent d'une spatule de bille de 0,5mm de diamètre (Roth), Après avoir effectué 2 séquences de lyse 3 minutes à 5500 rpm dans le PreCellys, avec une pause de 10 secondes entre chaque séquence, les tubes sont centrifugés 5 minutes à 16 000g à température ambiante. Deux cents microlitres de surnageant sont alors mélangés à 150µL d'isopropanol et incubés 10 minutes à température ambiante afin de précipiter les acides nucléiques. Après une seconde centrifugation de 5 minutes à 16 000g, 1mL d'éthanol 70% est ajouté aux culots. Après centrifugation, de nouveau 5 minutes à 16 000g, les culots sont séchés et les acides nucléiques dissous dans 50µL de **tampon T**, puis traités à la RNase A à une concentration finale de 40µg/mL durant 1h à température ambiante.

Enfin, après une précipitation éthanolique l'ADN est dissous dans 30 à 50µL de **tampon T**.



**Tampon d'extraction** : 200mM Tris HCl pH 7,5 ; 250mM NaCl ; 25mM EDTA

**Tampon T** : 5mM Tris HCl pH8,5

### 2.1.3 - PCR (Polymerase Chain Reaction)

La PCR sert à amplifier des séquences d'ADN spécifiques à l'aide d'une ADN polymérase ADN dépendante ainsi que d'amorces oligonucléotides.

La PCR a 2 rôles en biologie moléculaire : soit elle sert à confirmer la présence d'une construction dans un organisme transformé (criblage, génotypage, etc), soit elle sert à créer ou cloner une séquence d'ADN d'intérêt.

Deux types d'ADN polymérase ADN dépendante ont été utilisées au cours de cette étude. On les discrimine notamment à cause de leurs rapidités ainsi que de leurs fidélités. Ainsi une polymérase telle que la Taq (Promega®) enzyme assez processive, mais ayant tendance à faire des erreurs, sera privilégiée pour des criblages de bactéries transformées, ou des génotypages. À l'inverse, d'autres enzymes telles la Pfu (*Pyrococcus furiosus*, Thermo Scientific®) sont privilégiées pour les clonages et les expériences requérant une grande fidélité de polymérisation.

#### Conditions de PCR :

#### Clonage :

Le mélange réactionnel a été effectué de façon à obtenir un tampon de réaction 1X (New England Biolabs), une concentration de 250µM de chaque dNTP, 5pmol de chaque amorce, environ 35ng de matrice et 1U de polymérase (Pfu : New England Biolabs). Suite à quoi les mélanges sont placés dans un thermocycleur, où ils subissent différentes étapes.



- 95°C – 2 minutes
  - 95°C – 30 secondes
  - 55°C – 1 minute par kb de plasmide
  - 72°C – 40 secondes
  - 72°C – 7 minutes
- } x15-30

### **Génotypage :**

Les conditions du génotypage sont les mêmes que pour l'amplification de fragments PCR cités plus haut. La seule exception consiste en l'enzyme utilisée qui est la Taq polymérase de *Thermophilus aquaticus*.

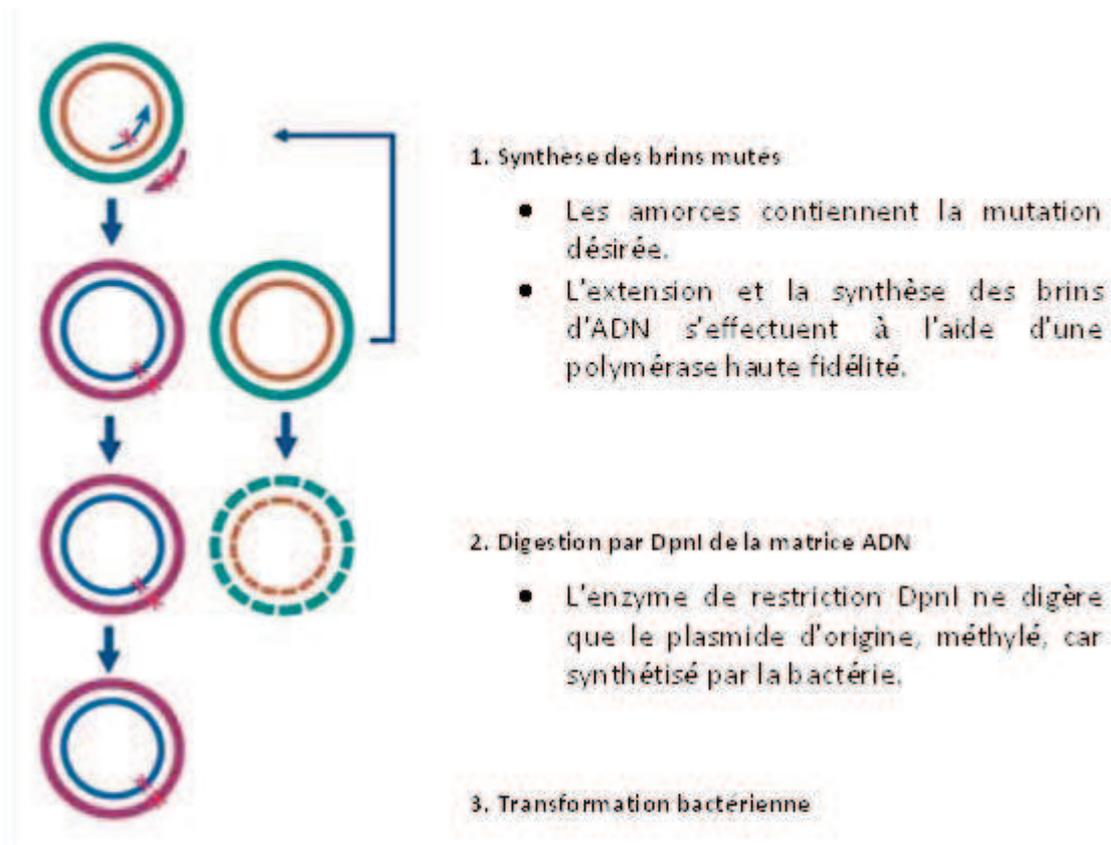
### **Mutagenèse dirigée :**

Les réactions ont été effectuées en mélangeant 1µL de dNTP fournis dans le kit QuikChange II Site-Directed Mutagenesis Kit (d'Agilent technologies©), 125ng de plasmide, 5µL de tampon de PCR 10X et enfin 2,5 U de PfuUltra HF DNA polymerase (NEB), pour un volume final de 50µL. Enfin, le mélange a été placé dans un appareil PCR pour être soumis au cycle suivant :

- 95°C – 30 secondes
  - 95°C – 30 secondes
  - 55°C – 1 minute par kb de plasmide
  - 68°C – 1 minute
  - 68°C – 2 minutes
- } x16

### **Synthèse des amiRNA :**

Les trois premières PCR sont effectuées en mélangeant 5µL de dNTP 2mM, 2µL de chaque amorce à 10µM, environ 10 à 20ng de plasmide pRS300, 2,5U d'enzyme Pfu (Sigma Aldrich) ainsi que 5µL de tampon Pfu (Sigma-Aldrich) pour un volume de 50µL. S'ensuit une incubation dans un thermocycleur avec le programme suivant :



**Figure 15 :** Représentation schématique du principe de mutagenèse dirigée. En vert et en orange sont représentés les brins d'ADN non mutés, et méthylés produit par *E. coli*. En mauve et en bleu figurent les amorces ainsi que les brins d'ADN néosynthétisés. La croix rouge symbolise la mutation introduite par les amorces.

- 95°C – 2 minutes
  - 95°C – 30 secondes
  - 55°C – 30 secondes
  - 72°C – 40 secondes
  - 72°C – 7 minutes
- } x24

Les différents produits PCR sont purifiés sur gel à l'aide du kit "QIAquick Gel Extraction Kit" de chez Qiagen, afin d'en extraire des bandes de 278 (PCR a), 170 (PCR b) et 305pb (PCR c).

Un demi microlitre de chaque éluat est mélangé selon les mêmes conditions que pour les PCR a, b et c, avec les amorces pRS300 A et B. Puis le tout est incubé comme suit :

- 95°C – 2 minutes
  - 95°C – 30 secondes
  - 55°C – 30 secondes
  - 72°C – 1 minute 30 secondes
  - 72°C – 7 minutes
- } x24

Les fragments PCR de 699 pb sont purifiés sur gel et clonés dans un vecteur d'entrée gateway.

#### 2.1.4 - Mutagenèse dirigée

La mutagenèse dirigée permet d'introduire une mutation au sein d'une séquence. Au cours de cette étude nous avons utilisé le kit d'Agilent technologies© : QuikChange II Site-Directed Mutagenesis Kit qui permet d'introduire des mutations dans des séquences déjà clonées dans un vecteur (Figure 15).

Cette technique consiste en une synthèse par PCR de la totalité d'un plasmide à l'aide de deux amorces complémentaires portant la mutation choisie. Les plasmides synthétisés *de novo* porteront ainsi la mutation introduite par les amorces PCR (les conditions de mutagenèse dirigée se trouvent dans la section 2.1.3).

Étapes	Oligo sens	Oligo antisens	Matrice
(a)	pRS300-A	RNS(1-4) IV (')	pRS300
(b)	RNS(1-4) III (')	RNS(1-4) II (')	pRS300
(c)	RNS(1-4) I (')	pRS300-B	pRS300
(d)	pRS300-A	pRS300-B	(a)+(b)+(c)

**Figure 16** : Liste des oligonucléotides et matrices nécessaires aux différentes étapes de clonage des amiRNA. Les séquences des différents oligonucléotides sont disponibles dans le tableau "Clonage des amiRNA". Afin de créer chaque amiRNA, cette technique nécessite 4 amorces spécifiques, numérotées de I à IV, ainsi que 2 amorces communes à tous les amiRNA (pRS300A et B). Deux amiRNA ont été créés par RNS, et donne lieu à un nouveau set d'amorces, annotés " ' ".

L'étape suivante consiste à éliminer les plasmides matrices. Ces derniers proviennent d'*E. coli* et sont donc méthylés. Il est possible de les éliminer à l'aide d'une enzyme de restriction capable de cliver spécifiquement les ADN méthylés. Un microlitre (10 U/ $\mu$ L) de l'enzyme de digestion Dpn I est ajouté dans la totalité du volume réactionnel de la PCR, le tout est ensuite incubé 1 heure à 37°C.

Enfin 1 $\mu$ L de la solution de plasmide traité par Dpn I est transformé dans *E. coli* où les plasmides non dégradés sont recircularisés grâce au mécanisme de réparation de l'ADN bactérien.

### 2.1.5 - Génération d'amiRNA

Le « design » des amiRNA (artificial microRNA) a été réalisé grâce aux logiciels disponibles sur le site internet : <http://wmd3.weigelworld.org/> (Schwab et al., 2006). Le principe de cette technique consiste en la modification d'une séquence d'un amiRNA déjà présent dans le vecteur pRS300 (Figure 16).

La première étape de ce protocole consiste à choisir les amorces nécessaires au clonage des amiRNA. Pour ce faire, il faut rentrer la séquence des gènes que l'on cherche à inhiber dans le site internet puis sélectionner un amiRNA parmi ceux proposés, en suivant 5 critères.

- Aucun mésappariement ne doit être présent entre les positions 2 et 12.

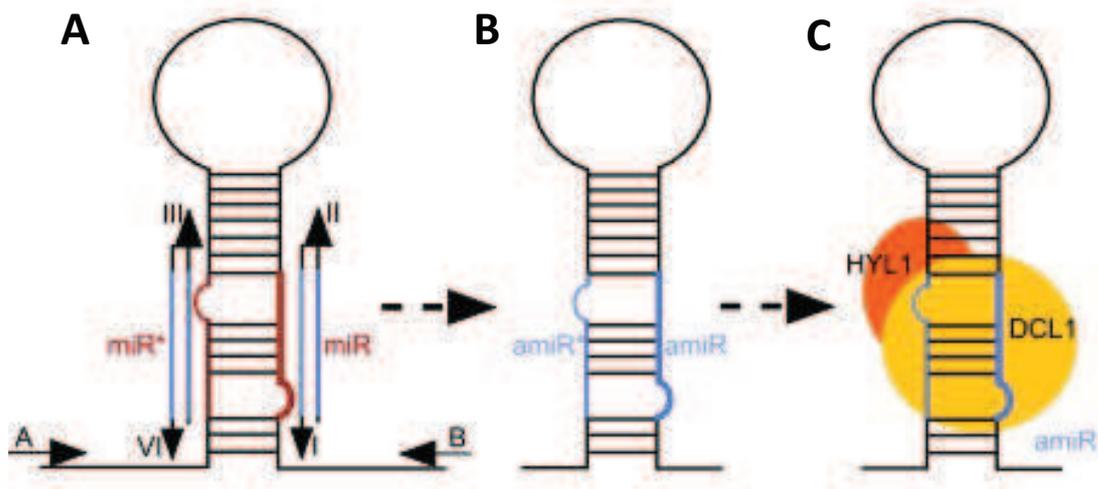
- 1 à 2 mésappariements à l'extrémité 3' de l'amiRNA (position 18 à 21).

- Tous les mésappariements doivent être les mêmes sur chacune des cibles choisies.

- L'énergie d'hybridation absolue doit se trouver entre -35 et -38 kcal/mole.

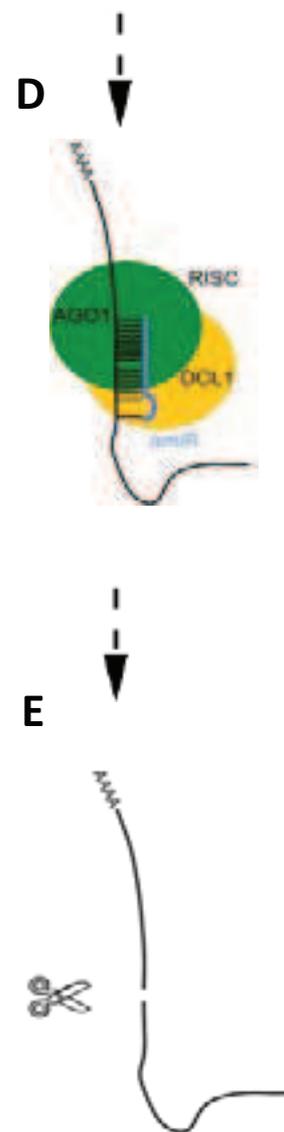
- Le site d'hybridation des amiRNA doit se trouver préférentiellement du côté 3' du transcrit cible, et ne posséder aucune autre cible dans le génome.

Une fois l'amiRNA sélectionné le jeu d'amorces PCR nécessaire à son clonage est proposé. Le clonage de l'amiRNA se divise en 3 étapes de PCR successives, la première permet d'amplifier différentes parties de l'amiRNA en utilisant les amorces A et B spécifiques au plasmide ainsi que les amorces spécifiques de l'amiRNA (I, II, III, IV) soit les étapes a+b+c (Figure 16 et 17). La



**Figure 17 : Principe de clonage, et mode de fonctionnement des amiRNA.** **A :** En rouge figure le miRNA d'origine du vecteur pRS300. En bleu sont représentés les amorces qui contiennent la séquence du nouvel amiRNA que l'on souhaite créer (amorces I, II, III, IV). Les amorces A et B servent à l'amplification de la totalité du construct. **B :** Structures finalisées du nouvel amiRNA. **C :** Après transformation et transcription, le construct sera reconnu par HYL1, et clivé par DCL1, qui en dégagera un fragment d'ARN simple brin. **D et E :** Ce fragment d'ARN simple brin permettra à AGO1, et au complexe RISC, de reconnaître l'ARNm cible, permettant dès lors sa destruction et l'inhibition de la synthèse de la protéine correspondante.

(Figure d'après le site internet <http://wmd3.weigelworld.org/cgi-bin/webapp.cgi> )



deuxième étape permet d'amplifier l'ensemble de l'amiRNA en n'utilisant que les amorces pRS300-A et B, ainsi que les 3 amplicons issues des PCR a + b + c. La dernière étape permet d'incorporer des sites gateway en 5' et 3' de l'amiRNA afin de pouvoir cloner l'insert dans différents vecteurs d'expressions chez la plante.

Après recombinaison des inserts dans des vecteurs d'expressions gateway *in planta* (cf. 2.1.5) la plante est transformée (se référer à la section 2.4.4) puis les amiRNA seront transcrits et maturés par les différentes enzymes de la voie du « silencing » inhibant alors la production des protéines issues du/des gènes ciblés (Figure 4).

### 2.1.6 - Séquençage d'ADN

Les séquençages ont été réalisés par M. Malek Alioua en France (plateforme de séquençage de l'IBMP) et par l'entreprise MACROGEN en Australie. Dans les deux cas, la technique utilisée consiste en une PCR asymétrique basée sur la méthode de Sanger.

### 2.1.7 – Synthèse d'ADNc

Cette technique consiste à synthétiser des brins d'ADN complémentaires (cDNA) aux ARN. Ces cDNA sont obtenus grâce à une ADN polymérase ARN dépendante.

Pour synthétiser un mélange de cDNA représentant les ARN totaux, le volume réactionnel initial de 13µl contient (concentrations ajustées pour un volume réactionnel de 20 µl) 1,25µM d'oligo(dT)20, 0,5 mM dNTP, 0,5 ng/µl d'amorces aléatoires et 1µg d'ARN. La solution est incubée 5 min à 65°C, puis 5 min dans la glace. Puis sont ajoutés 5 mM de DTT, 1x tampon "First-Strand", 40 U RNaseOUT (Invitrogen), 200 U SuperScript III RT (Invitrogen) pour un volume final de 20µL. Le milieu réactionnel est alors incubé 5 min à 25°C, 1 h à 50°C puis l'enzyme est inactivée 15 min à 70°C. Un microlitre d'ADNc obtenu est utilisé pour l'amplification par PCR.

### 2.1.8 – qRT-PCR

Cette technique consiste à mesurer l'expression d'un gène, (*i.e.* à mesurer



la quantité d'un ARN cible par l'intermédiaire d'un cDNA). Le volume réactionnel de 10µl comporte 1x SYBR Green I Master, 0,25µM d'amorces sens et antisens et 1µl d'ADNc. Un triplicat technique est réalisé pour chaque triplicat biologique.

La qRT-PCR est réalisée dans un LightCycler 180 Real-Time PCR System (Roche). La PCR commence par une dénaturation initiale de 5 min à 95°C, suivie de 40 cycles à 95°C de 10 sec, à 60°C de 15 sec et à 72°C de 15 sec. À chaque cycle la quantité de produit est mesurée, permettant de mesurer le Ct (Crossing threshold), qui est le point pour lequel le signal fluorescent est significativement supérieur au bruit de fond. Il apparaît au début de la phase exponentielle.

À la fin des cycles d'amplifications, une courbe de fusion est obtenue en chauffant les produits PCR de 55 à 95°C, permettant ainsi de vérifier qu'un seul produit a été amplifié.

Les résultats obtenus sont normalisés en utilisant des gènes représentant des contrôles endogènes (Tip 41-Like, Exp, FBox, GAPDH, séquences en section 1.2.2) et analysés en utilisant les programmes geNorm et NormFinder. L'efficacité de PCR est calculée par le programme LinRegPCR. Les coefficients de corrélation obtenus pour chaque paire d'amorces excèdent 0,99.

### 2.1.9 – Gateway

Le gateway est un système de clonage commercialisé par Invitrogen®. Il se base sur la capacité de recombinaison du bactériophage lambda dans *E. coli*. Le système gateway consiste en une succession de 2 recombinaisons. La première fait intervenir les sites attB clonés par PCR sur le fragment d'intérêt, les sites attP présents sur un vecteur dit "donneur" et enfin la BP clonase qui ne recombinera que les sites attB et P entre eux. Une fois les amplicons clonés, et séquencés, la seconde recombinaison est effectuée. Elle fait intervenir le vecteur "donneur" recombiné possédant désormais les sites attL, un vecteur "accepteur" possédant les sites attR, et enfin la LR clonase.

Une fois cloné dans ce dernier vecteur l'insert pourra être exprimé soit *in planta* soit dans les bactéries, en fonction du plasmide accepteur.

Les réactions de recombinaisons sont réalisées en additionnant 150ng de fragments PCR et/ou de plasmides donneurs/accepteurs, 2µL de mix BP ou LR



clonase (tampon inclus dans ce mix) le mélange est complété à 10µL avec de l'eau. Le tout est incubé 1h à température ambiante, ou sur la nuit à 4°C.

Après recombinaison, il est nécessaire de supprimer les protéines pouvant inhiber la transformation en restant liées à l'ADN, les mélanges sont donc incubés 10 minutes à 37°C avec 4U de protéinase K.

### 2.1.10 - Préparation et transformation d'*E. coli* thermocompétentes

#### **Préparation :**

Une pré-culture de 2mL de LB des bactéries DH5alpha ou TOP10 est incubée à 37°C et 200 rpm sur la nuit. Un millilitre de cette pré-culture est transféré dans un erlenmeyer de 500mL contenant 100mL de LB. La culture est incubée suivant les mêmes conditions jusqu'à une D.O comprise entre 0,3 et 0,4 ( $\lambda=595\text{nm}$ ). Pour stopper la croissance, les bactéries sont incubées 10 minutes sur la glace, puis culotées en centrifugeant à 3000g durant 10 minutes, à 4°C. Après élimination du surnageant, 25mL de **tampon CCMB** est ajouté et les cellules sont resuspendues. Incubées sur glace durant 20 minutes, les bactéries sont de nouveau centrifugées et resuspendues dans 4mL de **tampon CCMB**. Après une nouvelle incubation de 20 minutes, les bactéries sont aliquotées par fractions de 100µL et congelées immédiatement dans de l'azote liquide.

**Tampon CCMB (competent cell making buffer) :** 10mM PIPES, 55mM  $\text{MnCl}_2$ , 15mM  $\text{CaCl}_2$ , 50mM KCl

#### **Transformation :**

Vingt-cinq microlitres des bactéries précédemment préparées sont transformées avec 10 à 100ng d'ADN plasmidique ou la moitié du volume réactionnel de recombinaison gateway.

Après incubation 30 minutes à 4°C sur la glace, les bactéries subissent un choc thermique en étant placées 30 secondes à 42°C puis de nouveau 3 minutes sur la glace. Puis, après resuspension dans 200µL de milieu **LB**, les bactéries sont incubées 1h à 37°C, puis étalées sur du milieu **LB-agar** possédant l'antibiotique de sélection adéquat.



**LB (Lysogeny Broth)** : 10% tryptone, 5% extrait de levure, 10%NaCl

**LB-Agar** : LB + 1,5% (p/v) Agar (Sigma)

En fonction des vecteurs et souches bactériennes, voici la liste des antibiotiques utilisés et leur concentration finale :

- Ampicilline (Sigma) : 100µg/mL, dilué dans de l'eau
- Carbénicilline (Duchefa) : 100µg/mL, dilué dans de l'eau
- Spectinomycine (Duchefa) : 100µg/mL, dilué dans de l'eau
- Gentamycine (Duchefa) : 15µg/mL, dilué dans de l'eau
- Chloramphénicol (Sigma) : 30µg/mL, dilué dans du méthanol
- Kanamycine (Sigma) : 50µg/mL, dilué dans de l'eau
- Rifampicine (Sigma) : 50µg/mL, dilué dans du DMSO (Di-méthylsulfoxyde)
- Tétracycline (Sigma) : 10µg/mL, dilué dans de l'eau

#### 2.1.11 - Préparation d'*A. tumefaciens* électrocompétentes

Une pré-culture de 2mL de LB d'agrobactéries est incubée à 28°C et 200 rpm sur la nuit. Puis, les 2mL de préculture sont transférés dans 250mL de milieu LB contenant de la tétracycline ainsi que de la rifampicine. La culture est incubée à 28°C jusqu'à ce que la densité optique atteigne 0,4-0,6 ( $\lambda=595\text{nm}$ ). Les bactéries sont alors placées à 4°C durant 30 minutes.

Les bactéries sont ensuite culotées par centrifugation durant 10 minutes à 5 000rpm à 4°C puis resuspendues dans 250ml d'eau et de nouveau, elles sont incubées à 4°C durant 30 minutes. Puis, après une deuxième centrifugation, les bactéries sont resuspendues dans 125mL d'eau.

Enfin, après une nouvelle centrifugation les bactéries sont resuspendues dans 50mL de glycérol 10% (p/v) stérile, puis de nouveau centrifugées et resuspendues dans 2,5mL de glycérol 10% pour être congelées rapidement dans de l'azote liquide par aliquote de 50µL.



### 2.1.12 - Transformation d'*A. tumefaciens*

Afin de transformer les bactéries compétentes, 50µL sont incubées 30 minutes dans la glace, en présence de 30 à 50ng de plasmide.

Les bactéries sont alors transférées dans une cuve à électroporation (Gene Pulser/MicroPulser Cuvettes, BioRad®) dans laquelle est appliqué un courant électrique de 2.5V, une capacitance de 125µFD et une résistance de 200Ω. Après le choc électrique, 750µL de LB sont ajoutées aux bactéries qui sont ensuite incubées 40 minutes à 28°C, puis étalées sur milieu sélectif.

### 2.1.13 - Électrophorèse de fragments d'ADN en gel d'agarose

L'électrophorèse consiste en la séparation des acides nucléiques selon leurs tailles. On mélange ainsi 100ng à 1µg d'ADN à un **tampon de charge** qui contient des marqueurs de migrations (bleu de bromophénol, xylène cyanol) ainsi que du glycérol comme agent de charge. Le tout est chargé sur un gel composé de 0,8 à 2% d'agarose (p/v), de **tampon TAE 1X**, ainsi que de 0,5µg/ml de bromure d'éthidium (BET).

Le gel est placé dans une cuve contenant le même **tampon (TAE)**, puis l'électrophorèse est effectuée sous une tension constante de 100V durant un temps variable (de 20 minutes à 1h) selon la taille des fragments à observer. Les acides nucléiques sont révélés sous UV grâce au bromure d'éthidium un agent intercalant de l'ADN.

**Tampon TAE 0,5X** : Tris-Acétate 40 mM pH 8,0, EDTA 1 mM

**Tampon de charge** : Glycérol 50% (v/v), SDS 1% (p/v), EDTA 1 mM, xylène cyanol 0.1% (p/v), bleu de bromophénol 0.1% (p/v)

## 2.2 - ARN

### 2.2.1 - Extraction des ARN de petites tailles

Cent à 150mg de matériel végétal sont broyés dans un mortier froid contenant de l'azote liquide. Après ajout d'un millilitre de Tri Reagent (Molecular



Research Center, Trizol®) le mélange est de nouveau broyé jusqu'à obtenir une pâte qui se liquéfie peu à peu.

Le mélange est transféré dans un tube Eppendorf de 1,5mL et incubé 3 à 4 minutes à température ambiante, 250µL de chloroforme est ensuite ajouté et le tout est vigoureusement agité durant 2 minutes. Les échantillons sont centrifugés à 4°C durant 10 minutes à 20 000g. La phase aqueuse est prélevée et mélangée à 500µL d'isopropanol puis incubée 30 minutes à température ambiante afin de précipiter les acides nucléiques en centrifugeant le mélange 30 minutes à 20 000g. Les culots sont séchés, lavés avec de l'éthanol 80%, puis repris dans 180µL d'eau et mélangés à 120µL de chlorure de lithium 5M pour être incubés sur la nuit à 4°C. Dans ces conditions, les grands ARN sont précipités tandis que les tRNA et petits ARN non codants restent solubles. Les acides nucléiques contenus dans le surnageant sont ensuite précipités en ajoutant 30µL d'acétate de sodium 3M pH 5,8 puis 750µL d'éthanol absolu refroidi à -20°C. Le mélange est ensuite incubé une nuit à -20°C, après quoi il est centrifugé durant 30 minutes à 4°C et 16 000g pour enfin être resuspendu dans 20µL d'eau.

### 2.2.2 - Électrophorèse des ARN sur gel de polyacrylamide

La migration sur gel de polyacrylamide est utilisée pour la séparation des molécules d'ARN. Le gel est préparé à partir d'une solution d'acrylamide/N,N' méthylène bisacrylamide (19/1) de 15% ou de 6% et 7 M urée (dénaturant) dans du tampon TBE 1X. La polymérisation du gel est catalysée par l'addition d'APS (0,1% (p/v)) et de TEMED (0.035% (v/v)). Avant dépôt, les acides nucléiques sont additionnés d'un volume de **tampon de charge**. La migration s'effectue sous 25 mA dans le **tampon TBE 1X**. Le bleu de bromophénol et le xylène cyanol constituent les témoins de migration (l'équivalence en nucléotides (nt) dépend de la concentration du gel : 15 et 60 nt sur un gel 15% ou 26 et 106 nt sur un gel 6%).

**Tampon TBE 1X:** Tris 90 mM, EDTA pH 8,3 , 2,5 mM, borate 90 mM



**Tampon de charge:** Formamide désionisée 95% (v/v), EDTA 20 mM, xylène cyanol 0,05% (p/v), bleu de bromophénol 0,05% (p/v)

### 2.2.3 - Transfert des ARN

Après séparation sur un gel dénaturant de polyacrylamide, les ARN sont transférés sur une membrane de nylon de type Hybond N+ (Amersham). Le gel est placé sur la membrane et l'ensemble est disposé en "sandwich" entre deux feuilles de papier Whatmann et une éponge de part et d'autres. L'électrotransfert est effectué dans du tampon **TBE 0,5X** pendant 1h15 minutes sous une intensité de 300 mA. La membrane est brièvement séchée et les ARN sont fixés par irradiation en lumière UV (366 nm) pendant 3 à 5 min.

**Tampon TBE 0,5X :** 45 mM Tris, 1,25 mM EDTA pH8,0 , 45 mM borate

### 2.2.4 - Marquage radioactif des sondes

Les oligonucléotides synthétisés chimiquement sont marqués radioactivement à leur extrémité 5' par incubation 30 min à 37°C dans le milieu réactionnel suivant :

- 200 nM d'oligonucléotides
- 2,5 µL [ $\gamma$ -<sup>32</sup>P]ATP 10 µCi/µl (activité spécifique 3000 Ci/mmol)
- 1 µL **Tampon PNK 10X** (Fermentas)
- 10 U de T4 polynucléotide kinase (Fermentas)
- H<sub>2</sub>O q.s.p. 10 µL

**Tampon PNK 10X :** Tris-HCl pH 7,6 500 mM, MgCl<sub>2</sub> 100 mM, DTT 50 mM, spermidine 1 mM, EDTA 1mM pH8,

Le volume est ajusté à 50 µL d'eau et la radioactivité non incorporée est éliminée par chromatographie d'exclusion sur une colonne de Séphadex G50 (seringue de 1 mL). La colonne est séchée par centrifugation 30 secondes à 200g. La solution est déposée et éluée dans les mêmes conditions.



### 2.2.5 - Northern blot

La sonde radioactive est placée dans un tube d'hybridation avec la membrane ainsi que 3 à 6mL de tampon PerfectHyb (Sigma-Aldrich) et incubée à 42°C sur la nuit.

Après incubation, la membrane est lavée en augmentant progressivement la stringence par 3 lavages successifs : 2 de 5 minutes dans du **SSC 2X à 42°C**, et un de 30 minutes avec du **SSC 2X**, 0,1% (p/v) SDS (Sodium Dodecyl Sulfate) à la même température.

Les signaux sont révélés par exposition de la membrane à l'aide du phosphorimager ou par autoradiographie.

**Tampon SSC 2X** : Citrate-trisodique 30 mM pH 7,0 ; NaCl 0,3 M.

### 2.2.6 - Cinétique de clivage des transcrits tRNA

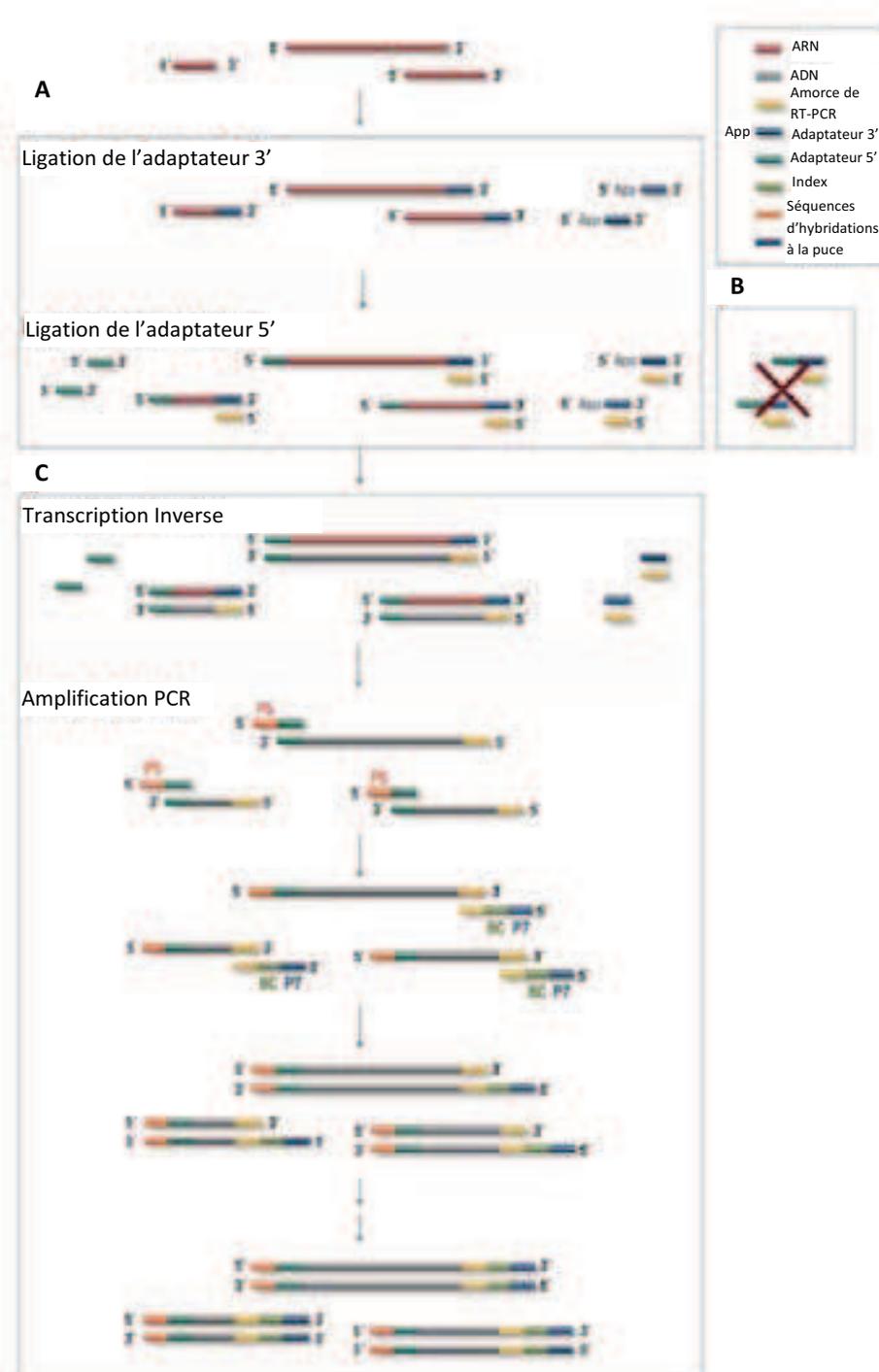
#### **Réalisation d'un extrait brut bactérien surexprimant une RNS :**

Cinq millilitres de culture de bactéries induites à l'IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside) sont centrifugés à 13000g pendant 5 minutes. Le culot de bactéries est repris dans 10 $\mu$ L de **tampon d'extraction protéique** et maintenu à 4°C sur la glace. Les bactéries sont soniquées 8 fois 10 secondes, puis l'extrait protéique brut est congelé dans de l'azote.

**Tampon extraction protéique** : 10% glycérol (v/v) ; 80mM Tris-HCl pH 6,8 ; cocktail d'inhibiteur de protéase (1 tablette pour 50mL de solution cOmplete protease inhibitor cocktail, Roche)

#### **Cinétique de clivage des ARN :**

Les cinétiques de clivage ont été effectuées en mélangeant 10 $\mu$ L d'extrait brut bactérien avec 20 $\mu$ L de tRNA (environ 10 $\mu$ g) pour un volume total ajusté à 100 $\mu$ L avec de l'eau. Divers prélèvements de 5 $\mu$ L sont effectués à différents temps d'incubation (5, 10, 15 minutes *etc...*). Chaque prélèvement est immédiatement transféré dans du Trisol afin d'en extraire les ARN (cf. 2.2.1). Ces ARN sont ensuite dissous dans 5 $\mu$ L d'eau puis séparés par électrophorèse (cf.



**Figure 18 : Principe du clonage de petits fragments d'ARN selon la méthode décrite dans le kit NEBNext Multiplex Small RNA Library Prep Set For Illumina .** Le kit utilise pour matériel de base de l'ARN. Le but de ce kit est d'amplifier au maximum les séquences afin de pouvoir les détecter. **A :** ligation des adaptateurs. **B :** l'hybridation de l'amorce de RT-PCR empêche la formation de dimère d'amorce. **C :** transcription inverse et amplification des séquences. (Illustration d'après le site internet du fabricant : <https://www.neb.com/products/e7300-nebnext-multiplex-small-rna-library-prep-set-for-illumina-set-1>)

2.2.2) et transférés sur membrane (cf. 2.2.3) dans le but de les analyser par northern blot (cf 2.2.5).

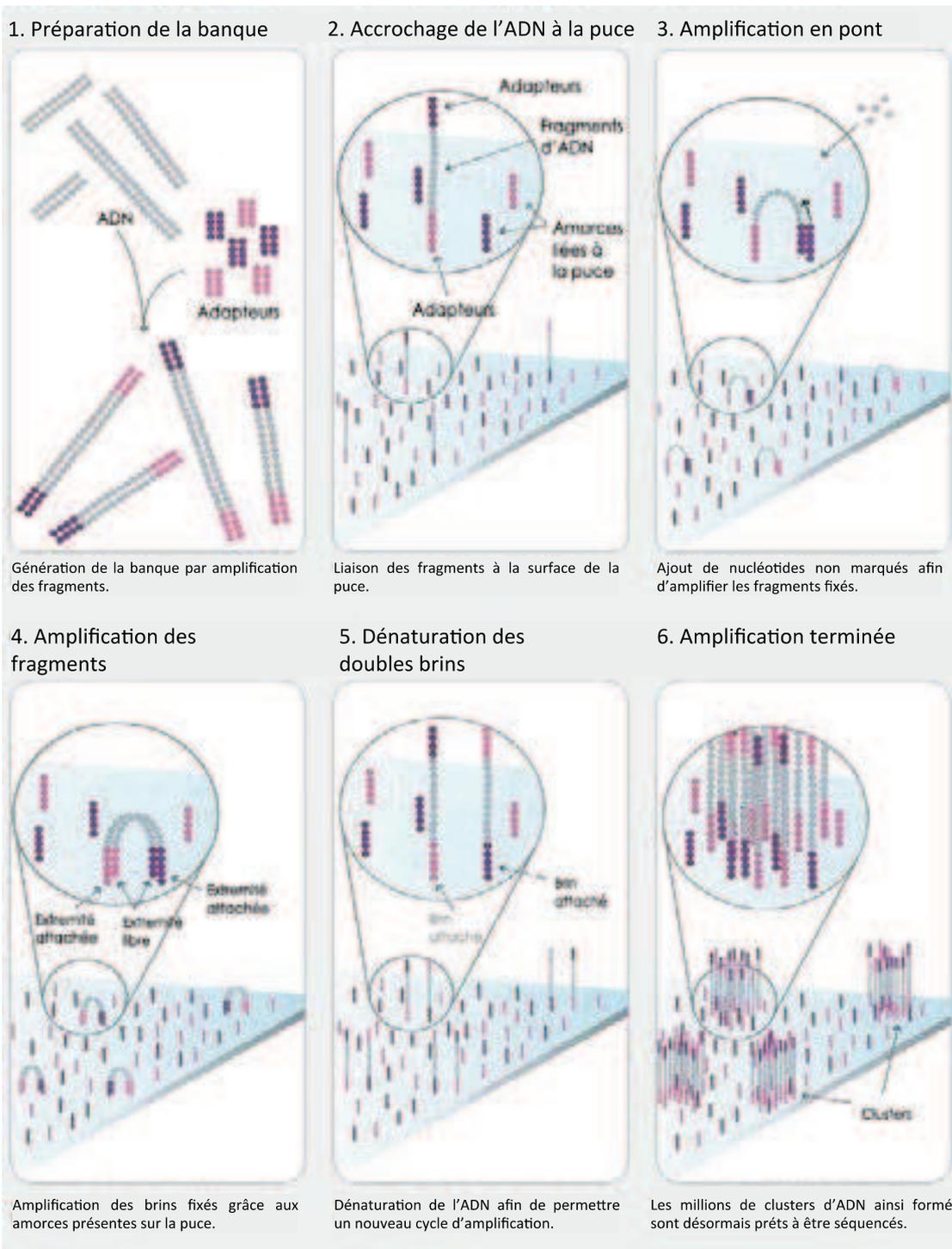
### 2.2.7 - Préparation de banques de petits ARN et séquençage Illumina®

Les tRF-5 possèdent un 5'-P et les tRF-3 un 3'-OH, leurs autres extrémités respectives ne possèdent pas d'extrémité normale puisque les enzymes que nous étudions sont des RNases de type T2. Or ces enzymes libèrent des extrémités d'ARN 5'-OH et 3'-P. Il est donc impossible de cloner ces fragments sans rétablir auparavant un 5'-P, ainsi qu'un 3'-OH à chaque extrémité des ARN. C'est pourquoi, avant de débiter la préparation des banques, environ 3µg d'ARN ont été incubés avec 2 unités de T4 polynucléotides kinase (NEB), en suivant le protocole du fabricant, qui possède les activités kinase et déphosphorylase nécessaires à l'établissement des extrémités compatibles avec le kit. L'enzyme ayant ces 2 activités dans les mêmes conditions, une seule incubation est nécessaire.

Afin de préparer les banques de petits ARN nous avons utilisé un kit distribué par NEB dont nous avons suivi le protocole : NEBNext® Multiplex Small RNA Library Prep Set For Illumina®. Très peu des composants utilisés dans ce kit sont précisés (enzymes, compositions des tampons) cependant son mode de fonctionnement est assez simple et il peut se diviser en 5 étapes.

La première consiste en la ligation d'adaptateurs en position 5' et 3' des séquences d'ARN à cloner (figure 18A). Afin d'éviter la dimérisation des adaptateurs, le protocole contient une étape intermédiaire. Elle consiste en l'hybridation de l'amorce de RT-PCR aux adaptateurs 3' se trouvant encore libres dans la solution. Puis lors de l'ajout de l'adaptateur du côté 5', la ligase sera incapable de former des dimères d'amorces car celle-ci ne fonctionne pas sur les fragments doubles brins (figure 18B).

La seconde étape vise en la transcription inverse des ARN clonés (figure 18C), puis en leurs amplifications via PCR. C'est d'ailleurs au cours de cette étape que sont intégrés les différents index. Les index sont des séquences de 6 nucléotides portées par les amorces PCR. Ils sont ajoutés en amont du fragment



**Figure 19 : Principe du clonage de petits fragments d'ARN selon la méthode décrite dans le kit NEBNext® Multiplex Small RNA Library Prep Set For Illumina®.** Le kit utilise pour matériel de base de l'ARN. Le but de ce kit d'amplifier au maximum les séquences afin de pouvoir les détecter. **A** : ligation des adaptateurs. **B** : L'hybridation de l'amorce de RT-PCR empêche la formation de dimère d'amorce. **C** : Transcription inverse et amplification des séquences. (Illustration d'après le site <http://seqanswers.com/forums/showthread.php?t=21>)

cloné lors de cette PCR. Ces index permettent l'identification de chaque banque de données, un peu à la manière d'un code barre. Ils permettent le « multiplexing », c'est-à-dire le mélange de plusieurs banques sur une même puce. Les index permettront ensuite de déterminer la banque dont chaque fragment provient, même si ceux-ci ont été séquencés en même temps et sur la même puce.

L'étape finale de préparation de la banque consiste en la sélection des tailles des fragments que nous cherchons à séquencer. Nous avons sélectionné à l'aide d'un gel d'acrylamide dénaturant de 6% les fragments d'environ 15 à 65nt.

Une fois les banques prêtes, celles-ci sont chargées sur le séquenceur très haut débit, Illumina HiSeq®. Le principe de fonctionnement de ce séquençage se base sur la méthode de Sanger. Cependant, afin d'augmenter fortement la sensibilité de détection nécessaire à l'identification de fragments très peu abondants, les fragments sont adsorbés sur une puce et y sont amplifiés de nouveau par PCR jusqu'à former des clusters (figure 19). Ce sont ces clusters qui seront finalement séquencés.

### **Analyse bio-informatique :**

Le but de ces banques est d'obtenir la population respective en tRF-5, 3 et de connaître la taille de chacun d'entre eux. La distinction est faite entre tRNA d'origine mitochondriale, chloroplastique et nucléaire. Une fois séquencées, les banques de données sont donc filtrées par Mme Valérie Cognat, bio-informaticienne à l'IBMP, à l'aide d'une banque de données identifiant tous les gènes de tRNA connus dans le génome d'*A. thaliana* (Cognat et al., 2013) disponibles sur le site <http://plantrna.ibmp.cnrs.fr/>.

Les adaptateurs sont extraits des données de séquençage grâce au programme FastX-toolkit. Les séquences obtenues sont alignées aux tRNA avec le logiciel PatMaN 1.2.1 contre l'ensemble des tRNA épissés et possédant le CCA, en autorisant 2 mésappariements dont 1 insertion/délétion. Les tRNA annotés comme étant des insertions mitochondriales ou plastidiales sont exclus. Les tRNA dans les régions répétées sont inclus. Enfin, les graphiques présentés ont été générés avec le logiciel R.



## 2.3 - Protéines

### 2.3.1 - Expression et induction protéique dans *E. coli*

Les bactéries BL21 transformées avec des vecteurs d'expression bactériens (pHGWA) ont notamment été utilisées pour exprimer les différentes RNS ainsi que leurs mutants.

Une préculture est tout d'abord effectuée sur la nuit à 37°C avec 10mL de LB liquide contenant 1% de glucose et l'agent de sélection (ampicilline), à partir d'une colonie d'*E. coli* isolée sur boîte petri.

Après centrifugation et resuspension des bactéries dans le même volume de pré-culture, afin d'éliminer le milieu de pré-culture, 5mL sont transférés dans 50mL de LB contenant le même agent de sélection. Après une heure de culture à 37°C la production de protéines est induite par l'ajout d'1mM d'IPTG. Après 6h à 37°C, les bactéries sont centrifugées 30 minutes à 3 000g puis congelées dans de l'azote liquide. Les bactéries congelées sont lysées dans du **tampon Laemmli** (cf. 2.3.3) et chargées sur un SDS-Page afin de procéder à un western blot (cf. 2.3.5) pour vérifier si l'induction de la production de protéines a bien fonctionné. Où bien elles ont été utilisées afin de produire des protéines recombinantes.

### 2.3.2 - Purification de protéines en conditions natives

Le principe de cette purification est basé sur l'affinité de la queue poly-His au Nickel. L'intégralité de la réaction a lieu à 4°C afin de prévenir la dégradation des protéines. Le culot de bactéries induites est lysé en soniquant 10 fois durant 8 secondes dans 1,8mL de **tampon de lyse**. Le lysat est ensuite centrifugé 30 min à 4 000 g.

Puis 200µL de résine de nickel (Ni-NTA, Qiagen) sont lavés 3 fois dans 500µL de **tampon de lavage**. La résine est récupérée après chaque lavage par centrifugation des tubes 5 minutes à 9 000g.

Le surnageant du lysat centrifugé est ajouté à la résine de nickel lavée et le tout est incubé durant une heure sous agitation lente. La résine est ensuite culotée par centrifugation à 9 000g durant 5 minutes, puis lavée 3 fois avec 300µL de **tampon de lavage**.



Les protéines fixées à la résine sont ensuite éluées avec 100 $\mu$ L de **tampon d'éluion** durant 15 minutes, sous agitation lente. Cette éluion est répétée 2 fois et les éluats sont congelés dans l'azote liquide.

**Tampon de lyse** : Tris malate 50mM pH 8.2, NaCl 0.6M, MgCl<sub>2</sub> 5mM, Glycérol 1% (v/v), Imidazole 20mM, Triton X-100 1% (v/v)

**Tampon de lavage** : Tris malate 50mM pH 8.2, NaCl 0.6M, MgCl<sub>2</sub> 5mM, Glycérol 1% (v/v), Imidazole 20mM

**Tampon d'éluion** : Hepes 25mM pH 7.5, NaCl 0.6M, MgCl<sub>2</sub> 5mM, imidazole 250mM

### 2.3.3 - SDS-PAGE

La séparation des protéines est réalisée par migration électrophorétique en gel de polyacrylamide en conditions dénaturantes (Laemmli, 1970). Les mailles de la matrice sont définies par le rapport polyacrylamide/N,N' méthylène bisacrylamide (37.5/1 dans le cas d'un gel de protéines). Le gel comporte deux parties : un **gel de concentration** (hauteur 2 cm) et un **gel de séparation** (hauteur 4 cm). La polymérisation de l'acrylamide est catalysée par l'ajout d'APS 0,1% (p/v) et de TEMED 0.05% (v/v).

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

**Gel de séparation** : Acrylamide 12 à 15% selon la taille des protéines à fractionner, Tris- HCl pH 8,8 100 mM, SDS 0.1% (p/v)

Les protéines additionnées à 0,2 volume de **tampon de charge** sont chargées sur le gel, et la migration s'effectue dans du **tampon de migration** pendant environ 1h sous ampérage constant de 25mA. La migration est contrôlée par la visualisation du front de migration, et du ladder.



**Tampon de charge (Laemmli) :** Tris-HCl pH 6,8 100 mM, glycérol 10%, SDS 2% (p/v), 0,1% bleu de bromophénol (p/v),  $\beta$ -mercaptoéthanol 3% (v/v)

**Tampon de migration :** Tris-HCl pH 8,3 25 mM, glycine 250 mM, SDS 0,1% (p/v)

### 2.3.4 - Détection des protéines

Après migration, les protéines sont fixées au gel et colorées au bleu de Coomassie pendant une heure dans une **solution de coloration** sous agitation lente et constante. Le gel est ensuite décoloré par plusieurs bains dans une **solution de décoloration**.

**Solution de coloration :** acide acétique 10% (v/v); éthanol 25% (v/v); bleu de Coomassie R250 0.125% (p/v)

**Solution de décoloration :** acide acétique 10% (v/v), éthanol 20% (v/v)

### 2.3.5 - Western blot

Cette technique consiste à révéler une protéine particulière à l'aide d'anticorps spécifiquement dirigés contre cette protéine. Elle nécessite la fixation des protéines sur une membrane.

#### **Transfert des protéines sur membrane**

Après séparation sur un gel de polyacrylamide (cf. 2.3.3), les protéines sont transférées sur une membrane Immobilon-P 0,45 $\mu$ m (Millipore) que l'on active par le méthanol 100%. L'ensemble est disposé en « sandwich » entre 4 feuilles de papier Whatman imbibées avec du **tampon de transfert** et placé dans une cuve à électrotransfert (Bio-Rad). Gel et membrane sont placés de sorte que les protéines se fixent sur la membrane en migrant de la cathode vers l'anode. Ce transfert est effectué à 80mA pendant 45 min.



**Tampon de transfert :** Tris 50 mM, borate 50 mM

### **Coloration des protéines après électrotransfert sur membrane**

Les protéines transférées sont observées par coloration avec une **solution de coloration** à base de bleu de Coomassie puis décolorées à l'aide de la **solution de décoloration**. Le colorant est éliminé totalement par lavage prolongé au méthanol 100%.

**Solution de coloration :** bleu de Coomassie R250 0,25% (p/v), méthanol 50% (v/v), acide acétique 7% (v/v)

**Solution de décoloration :** méthanol 50% (v/v), acide acétique 7% (v/v)

### **Fixation des anticorps**

La membrane est saturée pendant 1 heure sous agitation constante, dans du **tampon bloquant**. L'anticorps primaire spécifique de lapin ou de souris est ajouté à la solution, puis l'ensemble est incubé toujours sous agitation constante, une nuit à 4°C ou 1 à 2 h à température ambiante.

La membrane est ensuite lavée 3 fois 10 min dans du **tampon bloquant**, puis les anticorps secondaires (anticorps de chèvre dirigés contre les immunoglobulines de lapin ou de souris et couplés à la peroxydase (Life Technologies ou Sigma) sont ajoutés à du tampon bloquant. L'ensemble est agité une heure à température ambiante, puis lavé une fois 5 min dans le **tampon bloquant** et 3 fois 10 min avec le **tampon de lavage** avant révélation.

**Tampon bloquant :** Tris-HCl pH 7,4 20 mM, NaCl 150 mM, Tween 20 0,2% (v/v), lait écrémé en poudre 5% (p/v)

**Tampon de lavage :** Tris-HCl pH 7,4 20 mM, NaCl 150 mM, Tween 20 0,2% (v/v)



### Révélation de la fixation des anticorps

Les anticorps secondaires utilisés sont couplés à la peroxydase. On utilise une révélation par kit ECL (pour « Enhanced Chemical Luminescence ») qui consiste en une réaction de chimioluminescence catalysée par la peroxydase et produisant une émission de photons au lieu de fixation des anticorps secondaires. Dans la pratique, les conditions préconisées par le fournisseur (Amersham ou Roche) sont suivies et l'émission lumineuse est révélée par autoradiographie en exposant entre 10 secondes et 10 minutes les membranes.

## 2.4 - Techniques relatives aux plantes et à leurs cultures

### 2.4.1- Culture d'*A. thaliana* et *N. benthamiana*

Dans le cadre de la sélection *in vitro* d'*A. thaliana* transformées, les plantes ont auparavant été semées sur un milieu de gélose tamponné à pH5.7 contenant 0,8% d'agar (p/v), 1% (p/v) de sucrose et du Murashige & Skoog (Duchefa) contenant les micro et macro éléments nécessaires à la croissance des plantes.

Une fois les plantes germées (soit en milieu terre soit en semis *in vitro*) les jeunes plants d'une vingtaine de jours sont transférés en pot de culture contenant un terreau à base de tourbe (La Terra Standard Pikiererde). Ces plantes sont cultivées en serre avec une alternance jour-nuit de 16h/8h et une température comprise entre 18 et 22°C.

### 2.4.2 - Transformation d'*A. thaliana* et *N. benthamiana*

Les agrobactéries d'un glycérol stock sontensemencées dans 250mL de milieu LB contenant les agents de sélection, et le tout est incubé à 28°C à 200rpm jusqu'à ce que la densité optique atteigne 0.6-0.8 ( $\lambda=595\text{nm}$ ). Après centrifugation et élimination du surnageant, les bactéries sont resuspendues dans le même volume de **tampon d'infiltration** et incubées une heure à température ambiante. On procède ensuite à la transformation des plantes.



**Tampon d'infiltration** : 0,24% de Murashige and Skoog mélange micro et macroélément (MS) (Duchefa), 5% sucrose (p/v), 0,1% acétosyringone 200mM (v/v), 0,05% de silwet (v/v).

**Transformation d'*A. thaliana* en "flower dip", transformation stable :**

Lorsque la plante produit suffisamment de bourgeons non éclos (5 à 6 semaines) les hampes florales sont trempées dans le tampon d'infiltration contenant les bactéries. Les plantes sont ensuite placées à l'obscurité durant 1 journée, puis replacées environ 4 semaines dans des conditions de cultures décrites en 2.4.1. Les graines sont récoltées une fois sèches.

**Transformation de *N. benthamiana* par infiltration, transformation transitoire :**

Une fois les plants arrivés à maturité pour l'infiltration (stade 4 à 6 feuilles), une incision superficielle est pratiquée sur l'épiderme inférieur des feuilles à transformer. Ayant accès au parenchyme et à l'aide d'une seringue, environ 500µL de tampon d'infiltration contenant les bactéries à une D.O d'environ 0,6 ( $\lambda=595\text{nm}$ ) est alors injecté entre les 2 épidermes, dans le parenchyme.

Après 2 jours d'incubation, afin d'atteindre le pic d'expression maximale des constructions, les feuilles sont analysées (microscopie, western blot).

### 2.4.3 - Purification de chloroplastes d'*A. thaliana*

Des plantes de 5 semaines (Stade rosette, 6-8 feuilles) sont placées durant une à deux journées dans l'obscurité avant la récolte des feuilles, afin d'en éliminer l'amidon. Le jour précédant l'extraction des chloroplastes, des gradients de Percoll (Sigma-Aldrich®) sont préparés en additionnant un volume de 15mL de **tampon de resuspension 2X** à 15mL de Percoll, l'ensemble est ensuite centrifugé durant 30 minutes à 4 300g à 4°C et réservé à la même température jusqu'au lendemain.



Environ 20g de jeunes feuilles sont broyées dans un mixeur contenant 200mL de **tampon d'extraction** froid (4°C) en effectuant 2 séquences de 3 secondes de broyages, rapides puis lente. Après filtration des débris à l'aide de 2 couches de Miracloth, les chloroplastes sont culottés par centrifugation à 1 000g à 4°C durant 10 minutes. Les chloroplastes sont délicatement resuspendus à l'aide d'un pinceau dans 5mL dans du **tampon de resuspension 1X**.

Les chloroplastes sont ensuite déposés sur le gradient de Percoll préformés le jour précédent, et centrifugé à 13 200g à 4°C durant 10 minutes. La bande correspondant aux chloroplastes intacts se trouve vers le milieu du gradient. Ils sont prélevés et lavés avec 10mL de **tampon de suspension 1X**, puis centrifugés 10 minutes à 1000g à 4°C. Après un second lavage, les chloroplastes sont culottés et congelés dans de l'azote liquide afin d'en extraire les ARN ultérieurement.

**Tampon d'extraction** : 350mM Sorbitol, 50mM TrisHCl pH8, 0.5mM EDTA pH8, 0,1% BSA (m/v), 15mM  $\beta$ -mercaptoéthanol.

**Gradient de Percoll préformé** : 15mL Percoll (Sigma) + 15 mL tampon de suspension 2X. Centrifugation 30 minutes à 43 000g à 4°C.

**Tampon de suspension 2X** : 0,6M Sorbitol, 40mM Tricine pH7,6 , 10mM MgCl<sub>2</sub>, 5mM EDTA pH8

#### 2.4.4 - Analyse GFP en microscopie confocale

Les plantes transformées (*N. benthamiana*) avec des constructions GFP et/ou RFP-HDEL sont analysées au microscope confocal. Des échantillons sont prélevés sur les plantes vivantes, puis placés dans une cloche sous vide entre lame et lamelle en présence de PBS. Un vide de -2bar est appliqué aux échantillons afin d'en chasser l'air pour le remplacer par du PBS. Les différentes molécules fluorescentes sont excitées et observées au microscope (Zeiss LSM510), aux longueurs d'ondes suivantes :



- eGFP (fusionnée aux différentes protéines) : excitation 395nm, détection 509nm.
- mGFP<sub>er</sub> (modified GFP endoplasmic reticulum) : excitation 395 nm, détection 509 nm, localisée dans le réticulum endoplasmique.
- RFP-HDEL : excitation 558nm, détection 583nm, localisée dans le réticulum endoplasmique.





# *RÉSULTATS*



# Résultats

## 1 - Analyse de la population en tRFs chez *A. thaliana*

Lorsque j'ai débuté mon doctorat en 2011, peu de données étaient disponibles concernant les tRFs chez les plantes, cette situation n'est d'ailleurs pas très différente aujourd'hui alors que je rédige ma thèse (figure 15). Actuellement les tRFs ont été identifiés chez diverses plantes sous diverses conditions. Zhang et al., ont démontré la présence de fragments de tRNA de 31 à 68 nucléotides de long chez la citrouille (*Cucurbita maxima*) correspondant à des clivages des tRNA dans les boucles D et de l'anticodon (Hsieh et al., 2010; Zhang et al., 2009). De la même manière des études de séquençage haut débit ont montré chez *Brassica rapa* que des tRF-5A de 24nt issus de tRNA chloroplastiques étaient produits (à l'exception des tRNA<sup>Ser,Thr</sup> et Trp) en réponse à un stress induit par la chaleur (Wang et al., 2011b). De plus, des tRF-5D et des tRF-3D ont aussi été caractérisés chez *O. sativa* chez qui 82% des fragments correspondent à des tRF-5D Ala AGC de 20, 21 et 23 nucléotides (Chen et al., 2011). D'autres expériences de séquençage haut débit effectuées sur l'orge ont aussi montré que les tRFs étaient produits en réponse à un stress de carence en phosphate et que le tRFs le plus abondant correspondait au tRNA<sup>His(GTG)</sup> plastidial (Hackenberg et al., 2013).

Chez *A. thaliana* les tRFs ont pour la première fois été décrits en 2009 par Hsieh et al. (Hsieh et al., 2009). Cette étude montre sur des racines et des seedlings que les tRFs sont plus abondants lors de carence en phosphate. De plus, la majeure partie des fragments obtenus dans les racines (80%) correspondent à un tRF-5D de 19nt appartenant au tRNA<sup>Gly</sup> TCC (Hsieh et al., 2009). D'autres expériences menées par Thompson et collaborateurs sur des seedlings d'*A. thaliana* ont mis en évidence une accumulation du clivage des tRNA<sup>Arg</sup> (TCT), Asp(GTC), Gly(GCC), Glu(TCC), His(GTG), Lys (CTT) et Met(CAT) cytosoliques dans la boucle de l'anticodon, lorsque les plantes étaient soumises à un stress oxydatif. Il semble de plus, que les tRFs dérivant du tRNA<sup>Glu(TCC)</sup> soient plus abondants chez les fleurs, ce qui suggère une différence dans la biogenèse des tRFs selon les tissus (Thompson et al., 2008).

Organisme	Références	Tissus	types de tRFs	tRNA majoritaire (pili y à lieu)
<i>Cucurbita maxima</i>	Zhang et al., 2009	phloème	-	-
<i>Brassica rapa</i>	Wang et al., 2011b	seedlings	tRF-5A	-
<i>Oryza sativa</i>	Chen et al., 2011	seedlings	tRF-5D et tRF-3T	Ala-AGC
<i>Hordeum vulgare</i>	Hackenberg et al., 2013	seedlings	tRF-5D et tRF-3T	His-GTG
<i>Arabidopsis thaliana</i>	Hsieh et al., 2009	racines, seedlings	tRF-5D	Gly-TCC
	Thompson et al., 2008	seedlings	tRF-5A	Arg-CCT, Trp-CCA et His-GTG
	Nowacka et al., 2013	feuilles	tRF-5A, tRF-3A, tRF-5D et tRF-3T	-
	Loss-Morah et al., 2013	fleurs, feuilles racines, seedlings	tRF-5D et tRF-3T	Ala-AGC, Arg-CCT, Arg-TCG et Gly-TCC

Figure 15 : Liste des différentes plantes chez qui les tRFs ont été décrits.

Cependant tous les tRFs ne sont pas produits au cours de stress, c'est notamment ce qu'a démontré une équipe polonaise en identifiant une quarantaine de fragments dont 75% sont des tRF-3 tandis que 25% sont des tRF-5 chez la plante non stressée (Nowacka et al., 2013).

Enfin, alors que je débutais ma thèse, des hypothèses quant aux rôles que pouvaient occuper les tRFs ont commencé à émerger. L'un des rôles les plus récurrents, et qu'on retrouve chez quasiment tous les eucaryotes chez qui les tRFs ont été identifiés, concerne la capacité qu'ils auraient à inhiber la transcription (*i.e* PTGS : Post-transcriptional Gene Silencing). Cette hypothèse est, notamment, étayée chez *A. thaliana* par des résultats produits par Loss-Morais et collaborateur qui ont montré que des tRFs de 24nt étaient co-immunoprécipités avec AGO 1 et 4 (Loss-Morais et al., 2013) ce qui suggère une fonction des tRFs dans la régulation de l'expression des gènes chez *A. thaliana*.

C'est donc dans un contexte où les données disponibles chez *A. thaliana* sont très parcellaires, que nous avons choisi d'analyser différentes banques de petits ARN préparées par différentes équipes au sein de l'institut ou publiées par le passé (Wang et al., 2011a). La population en tRFs de chaque banque a été obtenue en croisant la population totale en petits ARN avec celle d'une banque de données établie dans l'équipe et comportant l'ensemble des séquences des tRNA d'*A. thaliana* correctement annotés (figure 16) (Cognat et al., 2013). Nous avons ainsi pu produire une analyse inédite et bien plus exhaustive que celles effectuées par le passé. Notre but étant d'adresser des questions telles que les raisons du déséquilibre constaté entre tRF-5 et tRF-3 (ces derniers sont très minoritaires), où encore le devenir des tRFs d'organites.

Cette étude fait donc l'objet d'un article, actuellement en cours de préparation, dont je suis le co-premier auteur avec Valérie Cognat bio-informaticienne à l'IBMP. Voici la version préliminaire :



## Genome-wide analysis of Arabidopsis tRNA-derived RNA fragments

Geoffrey Morelle<sup>1\*</sup>, Valérie Cognat<sup>1\*</sup>, Cyrille Migel<sup>1</sup>, Stéphanie Lalonde<sup>1</sup>, Jean Molinar, Ian Small<sup>2</sup>, Anne-Marie Duchêne<sup>1</sup>, Laurence Martchal-Bromand<sup>1</sup>

<sup>1</sup>Institut de Biologie Moléculaire des Plantes, UPR 2357 CNRS, associated with Strasbourg University, 12 rue du Général Zeller 67084 Strasbourg cedex, France

<sup>2</sup>Australian Research Council Centre of Excellence in Plant Energy Biology, The University of Western Australia, Crawley, WA6009, Australia

\*Contributed equally

\*To whom correspondence should be addressed. Tel. +333 88 41 72 40. Fax. +333 87 45 51 90. Email: [morelle@unistra.fr](mailto:morelle@unistra.fr)

### ABSTRACT

In the expanding repertoire of small noncoding RNAs (ncRNA), tRNA-derived RNA fragments (tRFs) have been identified in all domains of life. Their existence in plants has been already proven but no detailed analysis has been performed so far. Here, short tRFs of 19-26 nucleotides were retrieved from various *Arabidopsis thaliana* small ncRNA libraries. Very specific tRFs populations were found in various tissues, in plants submitted to abiotic stress and in immuno-precipitated ARGONAUTE 1 (AGO1) fractions. Depending on the tRNA, either tRFs-5 (due to a cleavage in the D region) or tRFs-3 (via a cleavage in the T region) are in general specifically observed. Surprisingly, many hot spots of tRNA cleavages have been identified in double-stranded RNA stems. Interestingly, up to 25% of the tRFs originate from organellar (mainly plastidial) tRNAs thanks to very specific cleavage sites. We provide evidence that these tRFs are not created within chloroplasts or mitochondria but rather they are generated outside the organelles by cytosolic endonuclease activities. Some tRFs-5 (but no tRF-3) deriving from nuclear-encoded but also from plastidial-encoded tRNAs are strongly enriched in AGO1 immuno-precipitates. These observations may suggest that some organellar tRFs play important regulatory functions within the cell.

## INTRODUCTION

Small noncoding RNAs (snRNAs) implicated in the regulation of gene expression have been identified in evolutionary divergent organisms (e.g. (1)). Small-interfering RNAs (siRNAs), microRNAs (miRNAs) and piwi-interacting RNAs (piRNAs) are the most well known. Many more classes of snRNAs are now emerging. Among them, long- (30-32 nt, tRF-A) and short (15-28 nt, tRF-5D and tRF-3T) tRNA-derived RNA fragments (tRFs) (**supplemental Figure S1A**) have been found in eukaryotic organisms belonging to all domains of life such as yeast, protozoans, plants and metazoans (for reviews see (2-5)).

Two classes of short tRFs deriving from mature tRNAs have been described. They originate either from the 5' extremity of mature tRNAs after cleavage in the D region (tRF-5D) or from the 3' end of mature tRNAs after cleavage in the T region (tRF-3T). The general nomenclature recently proposed will be used here (5). Although most tRFs described so far originate from nucleus-encoded tRNAs, a few tRFs deriving from mitochondrion- or chloroplast-encoded tRNAs have been identified.

It is suspected that short tRFs are likely not just tRNA degradation products implicated for example in the recycling of key nutrients such as phosphate or nitrogen upon starvation. Indeed, a few of them were recently shown implicated in important regulatory and biological processes. For instance, some tRFs are implicated in the inhibition of protein synthesis in Archaeobacteria or in human cells (6,7). Other tRFs were found associated with Argonaute (AGO) complexes in *Schizosaccharomyces pombe*, in human or in plants, suggesting their implication in the regulation of gene expression (8-11). In the protozoan *Tetrahymena thermophila*, the interaction of tRFs-3T with Twi12, a piwi Argonaute protein, is essential for the nuclear RNA decay pathway (12). Finally, a tRF-3T involved in the priming of a human viral reverse transcriptase may be important for viral infection (13).

Although evidence that cutting tRNAs into pieces to generate tRFs with potential important biological functions is increasing, detailed data on tRF population in various organisms and tissues or under various environmental conditions are still scarce as well as their potential functions. This is particularly true in plants. The first report on the existence of plant tRFs is from 2009 where tRFs were found in the phloem of pumpkin (14). The same year, Hsieh and collaborators showed the overexpression of very specific tRFs (e.g. tRF-5D from tRNA<sup>Asp</sup>(GTC) or tRNA<sup>Gly</sup>(TCC)) upon phosphate starvation in *Arabidopsis* roots (15). Genome-wide analysis of snRNAs from rice callus allowed the identification of several tRFs-5 and tRFs-3 (16). In the chinese cabbage *Brassica rapa*, tRFs originating from

plastidial tRNAs were shown to be regulated by heat stress (17). In barley, differential expression of sncRNAs including tRFs originating from nucleus- or chloroplast-encoded tRNAs was also observed upon phosphate deficiency (18). Finally in 2013, several *Arabidopsis* tRFs were predominantly found to be associated with various AGO proteins or to be induced by various stresses (11) but no detailed analysis was performed in particular concerning organellar tRFs.

The goal of the present work was to provide a detailed analysis of the tRFs population present in various *A. thaliana* sncRNA libraries. A special focus was given to plastidial tRFs as they represent up to 24% of the total tRFs population. Interestingly, a few of them were found specifically associated with AGO1. The problem of the existence of mitochondrial tRFs was also addressed. Here, we now provide evidence that organellar tRFs are not generated within chloroplasts and mitochondria but rather in the cytosol, likely as the tRFs deriving from nucleus-encoded tRNAs. Altogether, our work allows us to propose a very specific set of tRFs to be used as a starting point to answer many other questions regarding their biogenesis and their functions in higher plants.

## MATERIALS AND METHODS

### Small ncRNA libraries

Several deep sequencing libraries were retrieved from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>). In addition, four home-made libraries corresponding to *A. thaliana* plants treated or not with UV-C were used for deep sncRNA sequencing. For that purpose, *A. thaliana* plants, ecotype Columbia (Col-0), were first germinated *in vitro* on solid GM medium (MS salts (Duchefa), 1% sucrose, 0.8% Agar-agar ultrapure (Merck), pH 5.8) in a culture chamber under a 16-h-light (21°C) and 8-h-dark (19°C) photoperiod for 10 days. Afterward seedlings were transferred in soil (1 plant per pot) and put in a growth chamber (21°C/19°C; 70% humidity) for 2 weeks. For treated samples, plants (12 per replicate) were irradiated with UV-C (254 nm, 3000 Jm<sup>-2</sup>) using a Stratalinker (Stratagene). Thirty minutes upon UV-C exposure leaves number 3 and 4 were harvested and rapidly frozen into liquid nitrogen. Leaves of un-irradiated control plants were harvested prior UV-C exposure. SncRNAs were prepared using Trizol (Sigma), following manufacturer's instructions, and sent to the Fasterris company (Plan-les-Ouates, Switzerland) for library preparation and deep sequencing.

The list of small RNA libraries analyzed in this study is presented in **supplemental Table S1**.

### **Bioinformatics analysis**

As the size of the sncRNAs present in the different libraries was heterogeneous, we mainly focused our analysis on tRFs with a size between 19 and 26 nt (i.e. mainly tRF-5D and tRF-3T), a size range common to all libraries. As a first step towards an analysis of longer tRFs (i.e. tRF-5A), one novel sncRNA library comprising RNA fragments up to 40 nt in length was also filtered for the presence of tRFs. In order to retrieve the tRFs population of each deep-sequenced sncRNA library, the sequence reads obtained were filtered through an *A. thaliana* tRNA database comprising the full set of mature tRNAs with the CCA 3' extremity from the three compartments where a genetic information is present (nucleus, chloroplast and mitochondrion). This database was extracted from the global PlantRNA database (19). Reads corresponding to tRF sequences were aligned on full-length tRNA sequences using PatMaN (20). In order to take into account sequence bias due to the presence of modified nucleotides, the best matches were kept. In addition, *A. thaliana* mitochondria contain 6 chloroplast-like tRNAs (tRNA<sup>Asp</sup>, tRNA<sup>Asn</sup>, tRNA<sup>Pro</sup>, tRNA<sup>Leu</sup>, tRNA<sup>Trp</sup>(OGA) and tRNA<sup>Trp</sup>) and their sequences are usually identical or mostly identical to true chloroplastic tRNAs (19). It was therefore impossible to differentiate between true chloroplast-like mitochondrial tRFs and chloroplastic tRFs. Assuming that the abundance of tRFs originating from chloroplast-specific tRNAs is very high as compared to those deriving from mitochondrial-specific tRNAs in leaves, tRFs whose sequences are identical between chloroplastic and mitochondrial tRNAs were considered as originating from chloroplastic tRNAs.

### **Preparation of *A. thaliana* total, cytosolic, mitochondrial and chloroplast small RNA fractions**

Mitochondria and chloroplasts were isolated from *A. thaliana* seedlings and from young leaves according to (21) and to (22) respectively. The cytosolic fraction corresponds to the supernatant obtained after the first low speed centrifugation when preparing chloroplasts. Total, cytosolic, mitochondrial and chloroplast small RNA fractions were prepared as described in (23).

### **Preparation of a crude *A. thaliana* protein extract**

Leaves (2g) of *A. thaliana* (Col-0) were frozen in liquid nitrogen and ground in a mortar until a fine powder is obtained. Two ml of protein extraction buffer (Tris-HCl pH 7.5 50 mM, MgCl<sub>2</sub> 10 mM, glycerol 10%, EDTA 1 mM, β-mercaptoethanol 1 mM, PMSF 50 μM,

protease inhibitor cocktail 4% (Pierce) were added to the powder. After quick thawing, the mixture was centrifuged at 10 000g and the supernatant used as a crude protein enzymatic extract stored at -80°C.

#### ***In vitro* synthesis of radiolabeled RNA transcripts**

The constructs containing *A. thaliana* cytosolic tRNA<sup>Asp</sup> gene and plastidial tRNA<sup>Thr</sup> sequences were obtained previously (24,25). The tRNA gene sequences were directly fused to a T7 RNA polymerase promoter at the 5' terminus and included a BstNI site at the 3' terminus. These constructs were used as substrate to synthesize *in vitro* radiolabeled transcript with T7 polymerase using Ribomax<sup>TM</sup> large scale transcription kit (Promega, Madison WI) as described in (26).

#### ***In vitro* cleavage assay of tRNA transcripts**

One µg of tRNA transcript was incubated at room temperature in water in the presence or not of 16 µl of crude *A. thaliana* protein extract (total volume: 160 µl). To perform kinetic experiments, at each time point (between 0 and 30 min), 30 µl of the mixture was rapidly taken and added to 170 µl of RNA extraction buffer (Tris-HCl pH 7.5 10mM, MgCl<sub>2</sub> 10mM, SDS 1%) and 200 µl of water-saturated phenol to extract RNAs. After centrifugation (10min, 2300 g), each aqueous phase was submitted to ethanol precipitation.

#### **Northern blot analysis**

RNAs were separated on 15% (w/v) polyacrylamide gels, electrotransferred onto Hybond-N<sup>+</sup> nylon membranes (Amersham) and hybridized to oligonucleotide probes at 42°C in perfectHyb<sup>TM</sup> Plus (Sigma-Aldrich). Two washes (10 min) were performed at 42°C in 2xSSC followed by one wash (10 min) at 42°C in 2xSSC, 0.1% SDS. The oligonucleotides used as probes are the following: cytosolic (nuclear-encoded) tRNA<sup>Asp</sup>(AGC), 5' ACCATCTGAGCTACATCCCC 3' ; organellar tRNA<sup>Thr</sup>(GUG), 5' TCCACTTGGCTACATCCCC 3'.

## **RESULTS**

### **One fourth of plant tRFs originates from organellar tRNAs in leaves**

As a first step towards a detailed analysis of plant tRFs, two miRNA libraries prepared from *A. thaliana* leaves were analyzed first. We focused our work on short (19 to 26 nt) tRFs. The two libraries (named L1 and L2 **supplemental Table S1**) gave highly reproducible data (see

for example the error bars shown on **Figure 1**). About 2.8% of mRNAs correspond to tRFs. While three fourth of them derive from nucleus-encoded tRNAs (they will be called nuclear tRFs), around 22% of them originate from chloroplastic tRNAs (i.e. chloroplastic tRFs) and small RNA fragments deriving from mitochondrial-specific tRNAs represent less than 1% (**Figure 1A**). Most tRFs belong to the two major classes of tRFs deriving from mature tRNAs described so far. About 50% of nuclear and chloroplastic tRFs belong to the tRF-5D class, meaning their sequence starts at the 5' extremity of mature tRNAs. The tRF-3T population (i.e. RNA fragments ending at the CCA extremity of mature tRNAs) represents about 30% and 25% of the total nuclear or plastidial tRF populations respectively (**Figure 1B**). The remaining tRFs population belongs to the "other" category and mainly represents RNA fragments localized all along tRNA molecules.

As shown on **Figure 1C**, some tRFs appear strongly enriched in leaves. Among them we find nuclear tRF-5D (Ala-AGC, Arg-TGC, Arg-TCG, Asp-GTC and Val-CAC) or plastidial tRF-5D (Gly-GCC, His-GTG, Gln-TTG and Val-TAC). For these tRNAs, mostly no tRF-3T are found. By contrast, an enrichment in tRF-3T was observed for nuclear (Gln-CTG, Ile-AAT, Trp-CCA) or for chloroplastic (Asn-GTT). In a few cases, both types of tRFs generated from the same tRNA sequence are stable *in planta*. As examples, we can cite nuclear (Arg-ACG, Gln-TTC) or plastidial (Leu-TAG, Phe-GAA or Tyr-GTA). In chloroplasts, tRNA abundance is correlated with codon usage (27) and in the cytosol, on one hand, a very strong correlation exists between codon usage and the number of tRNA genes in plants (28) and on the other hand, a strong correlation between tRNA abundance and codon usage has been shown in the green alga *Chlamydomonas reinhardtii* (29). Thus, the amount of stable tRF-5 or tRF-3 is not correlated with the abundance of the corresponding tRNA *in vivo*. This is also in agreement with the absence of correlation between codon usage and tRF abundance (**Supplemental Figure S1**). As a whole, this shows that the pool of stable tRFs is not correlated with tRNA abundance and strongly confirms that they are not just random degradation products.

#### **Specific features of Arabidopsis tRFs**

As mentioned above, around 80% of *Arabidopsis* tRFs are not RNA fragments randomly localized all along the tRNA molecules (**Figure 2, supplemental Figure S2**). However, they do not share a very specific size (**Figure 1D**). Both nuclear and chloroplastic tRF-5D of 19 and 20 nt are the most abundant but they are not the only ones. Concerning tRF-3T, there is also no real size specificity although plastidial tRF-3T of 24 nt appears to be twice more abundant than the others. The tRF-5D originating from ten nucleus-encoded tRNAs and from

13 chloroplastic-encoded tRNAs represent 67% and 75% of their respective total tRF-5D population. Similarly, 48% and 71% of the tRF-3T population originate from only 6 nuclear and 5 plastidial tRNAs respectively. The sequences of these most abundant tRFs are presented on **Figure 2A**. Considering the fact that they all derive from mature tRNAs and thus already contain a few invariant or semi-invariant nucleotides, tRFs do not appear to share extensive sequence homology. Only two plastidial tRFs out of the 28 tRF sequences presented begin with a U, a specific feature of siRNAs and most of them have a G at their 5' extremity. Strikingly however, hot spots of localization for the 3' extremity of tRF-5D and for the 5' extremity of tRF-3T exist (**Figure 2B**). These hot spots appear to be similar between nucleus-encoded and chloroplastic tRNAs. Concerning tRF-5D, the endonucleolytic cleavage of tRNAs occurs mainly either in the D-loop but also more surprisingly in the D-stem. Concerning tRF-3T, they are generated through cleavages occurring mainly in the T-stem and much less in the T-loop. These data suggest that the *Arabidopsis*'s endonuclease(s) responsible for tRNA cleavages do not only cut single-stranded RNA regions. For each tRNA isoacceptor, there is no strong specific site of cleavage. Rather, generally, there is a zone of cleavage thus generating a complex population of tRFs from the same tRNA sequence. This is illustrated for a few tRNAs in **Figure 2C**. These tRFs are present in leaves under normal condition of growth. We then wondered whether the population of tRFs is tissue specific and whether this population fluctuates under stress conditions.

#### **The root tRF population quantitatively differs from those of other tissues**

To investigate whether the population of *Arabidopsis* tRFs fluctuates depending on the tissue or developmental stage considered, four miRNA public libraries (R1, L3, S1, F1, **supplemental Table S1**) from roots, leaves, seedlings and flowers and already mentioned in (11) were analyzed into details. As shown on **Figure 3** and **supplemental Figure S3**, some fluctuations in the tRF-5D populations were observed when comparing leaves, flowers and seedlings. However more notable differences were found when comparing them with the root RNA library. Several tRF-5D originating from nucleus-encoded tRNAs are either down- (e.g. Ala-AGC) or up-regulated (e.g. Arg-TCG, Cys-GCA, Gly-TCC). No significant fluctuation was observed when comparing the tRF-3T. Strikingly, fluctuations are also seen when analyzing the plastidial tRFs. In particular, tRF-5D from His-GTG and tRF-3T from Arg-ACG are strongly enriched in roots. As the chloroplastic tRNA population is expected to be low in *Arabidopsis* roots as compared for example to green tissues, the reason of such enrichment in very specific plastidial tRFs is intriguing.

#### Variation of tRFs population in plants submitted to environmental stresses

Long tRFs, i.e. tRFs of about 30-35 nt generated after cleavage in the region of the anticodon (mainly tRF-5A) were often shown to be induced in various organisms by a large panoply of stresses (e.g. (30,31)). However less data exist concerning the population of stress-induced tRF-5D and tRF-3T. In order to get some insights on this question in plants, we focused our analysis on *Arabidopsis* plants submitted to various abiotic stresses.

First, we compared miRNA libraries from young *Arabidopsis* plants irradiated (L11, L12, **supplemental Table S1**) or not (L1, L2, see above) with UV-C, already known to trigger a stress response in plants (32). As shown on **Figure 4** and **supplemental Figure S4**, mostly no variations are observed for tRF-5D (except for nuclear tRF-5 (Gly-GCC and Val-AAC) which are present in significantly higher amounts in UV-stress plants). By contrast, several specific tRF-3T of nuclear but also of chloroplastic origin, are up-regulated in UV-irradiated plants. This is for instance the case for nuclear tRF-3T (Arg-CTG, His-AAT, Trp-CCA) or for plastidial tRF-3T (Arg-TCT, Phe-GAA). All tRNAs are not submitted to the same increase of endonucleolytic cleavage, in particular at the level of the T-region, thus suggesting that it is not just the consequence of a general tRNA degradation pathway. Another possibility is that the increased amount of some tRFs is due to an increase in their stability, may be because they are engaged in complex and protected against degradation.

Second, we also benefit from publicly available miRNA deep sequencing libraries (see also (11)) to analyze the tRFs profiles of *Arabidopsis* shoots submitted to drought, cold and salt stresses (D1, C1, S1, **supplemental information Table S1**) as compared to control plants (Co) grown in the same conditions. Considering nuclear tRFs, the expression patterns observed under these different environmental conditions are not significantly different to the patterns observed for control plants (**supplemental Figure S5**). When looking at chloroplastic tRFs, similar profiles are also obtained with the exception of tRF-5D originating from tRNA<sup>Arg</sup>(ACG) which is significantly up-regulated under drought condition. The reason of this increase is presently unknown. As a whole, it seems that, in contrast to long tRFs-A, the response of *Arabidopsis* upon abiotic stress in term of production/degradation of short tRFs is rather low and only the amount of some specific tRFs seem to be affected in response to UV-C.

#### **Nuclear but also plastidial tRFs are associated with AGO1**

Among the potential roles of tRFs, their implication in the regulation of gene expression via the classical RNA silencing pathways has already been put forward. As for several organisms, the existence of plant tRFs associated with AGO proteins has been observed (11) but not detailed analysis provided. We focused our work on AGO1 and analyzed the population of tRFs associated with this protein in roots and flowers (Figure 5, supplemental Figure S6). Public libraries from roots and flowers (R1, R-AGO1, F1 and F-AGO1) were used here. The first remarkable point is that the set of tRFs highly associated with AGO1 is limited and corresponds to RNA fragments deriving not only from nucleus-encoded tRNAs but also from chloroplastic-encoded tRNAs. The second important point to be noted is that while in the total snRNA fractions, tRF-3T are well represented, in particular in flowers, the tRF population present in the RNA fractions associated with AGO1 in roots or flowers is mainly exclusively constituted by tRF-5D. This is the case for nuclear tRFs but also for chloroplastic tRFs. Among them, four nuclear tRFs-5D are highly enriched both in roots and flowers (Arg-ACG, Arg-CCT, Arg-TCG, Gly-TCC), (Figure 5C), although they are not the most abundant in total roots or flowers RNA fractions. By contrast, the tRF-5D (Ala-AGC) found in high amount in flowers or leaves is never found associated with AGO1. In the same vein, only two chloroplastic tRFs are highly enriched in AGO1-immunoprecipitated RNA fractions of roots or flowers. They correspond to tRF-5D (Arg-ACG and His-GTG) and represent about 90% of the chloroplastic tRFs population associated with AGO1 (Figure 5C). The tRFs associated with AGO1 are 19 or 20 nt long but their sequences do not share extensive sequence homology. The identification of AGO1-associated tRFs originating from chloroplastic tRNAs led us to address the question where these tRFs are generated within the plant cell.

#### **Biogenesis of organellar tRFs is occurring outside the organelles**

To our knowledge, plant AGO1 has never been found within chloroplasts or mitochondria and is mainly localized in the cytosol and the nucleus. Thus, either the chloroplastic tRFs found associated with AGO1 are generated within the organelles and can they escape their cellular compartments to reach AGO1 or tRFs are generated outside the organelles and in that case, it is the tRNAs that need to escape the organelles. To discriminate between these two possibilities, RNA fractions extracted from total, cytosolic and highly purified chloroplasts and mitochondria were prepared (Figure 6A) and Northern experiments were performed with a radiolabeled oligonucleotide complementary to the 5' extremity of the chloroplastic tRNA<sup>Arg</sup>. As shown on Figure 6B, neither tRF-5D, nor tRF-5A can be detected in the

chloroplastic RNA fraction. By contrast, both classes of tRFs are present in the cytosolic or total mcRNA fractions. As mitochondria contain a chloroplast-like tRNA<sup>Trn</sup>, we also wondered whether these tRFs could be generated within mitochondria, but, as for chloroplasts, the RNA fragments were not detected. Our data show that chloroplastic tRNAs need to escape the organelles to be cleaved outside and generate tRFs. A similar situation may be also hypothesized for potential mitochondrial tRFs. We also wondered whether endoribonuclease activities able to generate tRFs exist in an *Arabidopsis* fraction. To answer that question, we set up an *in vitro* cleavage assay. For this purpose *in vitro* synthesized transcripts corresponding to either nucleus-encoded tRNA<sup>Asp</sup>(AGC) or to organellar tRNA<sup>Trn</sup>(GTG) were incubated with a total crude *Arabidopsis* protein extract. RNAs were extracted, fractionated on polyacrylamide gels and analyzed by Northern blots (**Figure 6C**). For each tRNA transcript, tRF-5D and tRF-5A were generated in the presence of the protein extract while, in its absence, only mechanical tRNA cleavages are detected. As a whole, our data suggest that both nucleus-encoded and chloroplastic-encoded tRNAs are likely cleaved by very similar endoribonucleases within the plant cell.

#### **Are tRFs generated from mitochondrial tRNAs?**

In a plant cell, there are three compartments with tRNAs: the cytosol, the chloroplast and the mitochondrion. Above, we discussed into details about tRFs originating from either nucleus-encoded or from chloroplast-encoded tRNAs. The situation in mitochondria is more complex (37). First, plant mitochondria contain few "native" tRNAs expressed from true mitochondrial tRNA genes. In addition, i) they also possess "chloroplast-like" tRNAs expressed from chloroplastic genes inserted into the mitochondrial genome, tRNAs that are mostly identical to true chloroplastic tRNAs and ii) they also import numerous nucleus-encoded tRNAs to compensate the lack of mitochondrial tRNA genes. Consequently, studying tRFs originating from tRNAs localized into mitochondria is problematic. The only tRFs we can be confident with concern those deriving from true mitochondrial tRNAs. In *Arabidopsis* they correspond to only 11 amino acids. Another difficulty resides in the low number of reads corresponding to "native" mitochondrial tRNAs in tncRNA deep sequencing libraries. When we compared the different tncRNA libraries prepared from various *Arabidopsis* tissues, it is in flowers that the percentage of "native" mitochondrial tRFs was the highest and we first choose the F1 library to look for potential mitochondrial tRFs. While for the majority of mitochondrial tRNAs, the low number of reads were localized randomly along tRNA molecules (see for example tRFs deriving from tRNA<sup>Trn</sup>(GCT) on **Figures 7A and 7B**) suggesting that these

RNA fragments rather corresponds to true degradation products, more specific and recognizable IRF patterns were identified for others. Indeed, IRF-5D and, to a less extent, IRF-3T from mitochondrial tRNA<sup>Gln</sup>(TTG) and tRNA<sup>Lys</sup>(TTT) were retrieved in this library but also in the RNA libraries from the other *Arabidopsis* tissues (Figures 7A and 7B). Furthermore, these IRFs appear very similar in size and in the position of the cleavage site on the corresponding tRNA molecule to what we already observed for nuclear and chloroplastic IRFs (Figures 2 and 7C). Altogether, these data strongly suggest that, as for chloroplasts, very specific IRFs are originating from mitochondrial tRNAs.

#### DISCUSSION

Specific tRNA endonucleolytic cleavages at the origin of nuclear IRFs have been reported in various organisms, such as protozoans, fungi or mammals. In *Arabidopsis*, only very scarce data on nuclear IRFs are reported (11,15,34). To provide a helpful support in deciphering the potential role(s) of some of these IRFs, it was first essential to provide a strong set of data to the scientific community. In plants, tRNAs are expressed from three different genomes, the nuclear, plastidial and mitochondrial ones. Do IRFs originate from these three populations of tRNAs? Here, by analyzing in depth the population of IRFs-5 and IRFs-3 found in various *Arabidopsis* tissues or organs, we first identified a major pool of nuclear IRFs-5D and IRFs-3T. Among the most abundant IRFs-5D observed are the four IRFs (Ala-AGC, Arg-TCG, Arg-CCT and Gly-TCC) already described in (11) but also a few others (Figure 4). Strikingly, the population of IRFs-3T usually originates from other tRNA species (e.g. Glu-CTG, Ile-AAT and Trp-CCA). In addition, the case where the abundance of IRFs-5 and IRFs-3 is generated from the same tRNA species is rare (e.g. Glu-TTC). Whether this is the result of differential cleavages of each tRNA isoacceptor either in the D- and/or T- region and/or whether it is due to differential stabilities of IRFs *in planta* is presently unknown. The role played by nucleotide modifications in protecting or facilitating tRNA cleavage is presently unknown in plants but has already been described in bacteria or eukaryotes (35,36). Cleavage sites on mature tRNAs are mainly restricted to very specific zones (Figure 2B) and the question whether the same endonuclease(s) is responsible of the biogenesis of IRF-5 and IRF-3 remains to be addressed in plants and more generally in various organisms. Indeed, Dicer was shown to be at the origin of some IRF-5D or IRF-3T in human (37,38) but other still unknown endonucleases are likely required (5). Here, the T-stem double-stranded RNA region represents a hot spot of cleavage, thus plant Dicer would also be good candidate. On the 5' side, cleavage rather occurs in the D-loop, thus likely implying another type of

endonucleases. So far yeast Rny1, a RNase T2, was shown to cleave yeast tRNAs in the anticodon loop (39), whether ribonuclease of the plant RNase T2 family (40) are able to cleave mature tRNAs in the D-loop is a question to be addressed in the future.

In addition to nuclear tRFs, 25% of the tRF population corresponds to chloroplastic tRFs-5D and tRFs-3T. This is in agreement with the presence of numerous tRFs deriving from plastidial tRNAs in *Brassica rapa* (17). These tRFs are due to cleavage in positions very similar to those found for nuclear tRFs. Furthermore, they are not generated within the organelle thus implying that chloroplastic tRNAs need to interact with the endonuclease responsible for their cleavage outside the organelle. Import of nucleus-encoded tRNAs into organelles is now largely documented (e.g. (41)) but the other way, i.e. export of organellar tRNAs to the cytosol has never been demonstrated so far. There are potentially two other ways to explain the presence of chloroplastic tRNAs outside the organelles. First, stromules (for stroma-filled tubules) are present at the surface of chloroplasts (for a review, see (42)). It is believed that these highly dynamic structures can allow the exchange of macromolecules between different compartments and this could be a way to transport tRNAs from the chloroplasts to the cytosol or to another compartment of the plant cell. The second way to deliver chloroplastic tRNAs outside the organelle is chlorophagy (for a review see (43)). For example, during senescence or upon stress, the delivery of chloroplasts to vacuoles has been reported. Also, chloroplastic proteins can be degraded via the ubiquitin-proteasome system (44) and it is easy to speculate that the cleavage of chloroplastic RNAs can occur via unknown non-organelle RNA degradation pathway.

Many tRFs were reported to be induced by a large variety of stresses in many eukaryotes. In *Arabidopsis* upon phosphate starvation, a group of 19 nt long tRFs-5D were shown to be expressed at high level (45). Upon oxidative stress, there is also up-regulation of some *Arabidopsis* tRFs-5A (30). By contrast, we show here that upon UV-C irradiation, only few tRFs-5D show a higher level of expression as compared to control plants, and mostly no differences were observed when plants were submitted to drought, cold or salt stresses. It is rather some tRFs-3T that are present at higher frequency. Altogether, it appears difficult to draw a general scheme on the fluctuation of tRF populations in response to various abiotic stresses. Do the fluctuations observed only reflect degradation of tRNAs to recycle nucleotide and phosphate under stress conditions or do some tRFs participate in regulation of very specific gene expression in order to respond to adverse environmental conditions will need to be further clarified.

Our analysis not only confirmed on one hand the existence of short nuclear tRFs (only tRFs-5D) but also of chloroplastic tRFs (mainly two major tRFs-5D) associated with AGO1 and on the other hand the complete absence of association with AGO1 of tRFs abundantly found in total snoRNA fractions. This strongly suggests that, indeed, the tRFs found associated with AGO1 can play important role in the regulation of gene expression via the RNA silencing pathway and in the case of chloroplastic tRFs can be key elements for retrograde signaling regulation. Target genes for tRF-AGO1 complex can be found but it is essential now to validate the candidates. Importantly, this also shows that very likely not all tRNAs play similar functions within the plant cell. As an example, tRFs-5D originating from nuclear or chloroplastic tRNAs<sup>524</sup> are abundant and their expression can fluctuate but they are not associated with AGO1. However they both start with a series of 4 G residues, a motif present at the extremity of human tRF-5A (Ala) and found essential for the interaction with the translational silencer YB-1 to inhibit translation (46). Whether these tRFs, and intriguingly the plastidial ones, have similar function in plants remains to be established.

As a whole, several nuclear but also plastidial tRFs have now been very well characterized here. They provide a very powerful support for further studies dealing with i) the identification of the endonucleases responsible for tRNA slicing and biogenesis of tRFs and ii) the functions to be attributed to these tRFs in plants.

#### **SUPPLEMENTARY INFORMATION**

Supplementary data are available at NAR online.

#### **FUNDING**

This work has been supported by Centre National de la Recherche Scientifique (CNRS) in association with the University of Strasbourg, by the Agence Nationale pour la Recherche (ANR) (grant ANR-09-BLAN-0240-01). This work has been also published under the framework of the LABEX (ANR-11-LABX-0057\_MITOCROSS) and benefits from a funding from the state managed by the French National Research Agency as part of the Investments for the future program. G.M. has been supported by the "région Alsace" and by the Australian Research Council. C.M. by the LabEx consortium « MitoCross » (ANR-11-LABX-0057\_MITOCROSS).

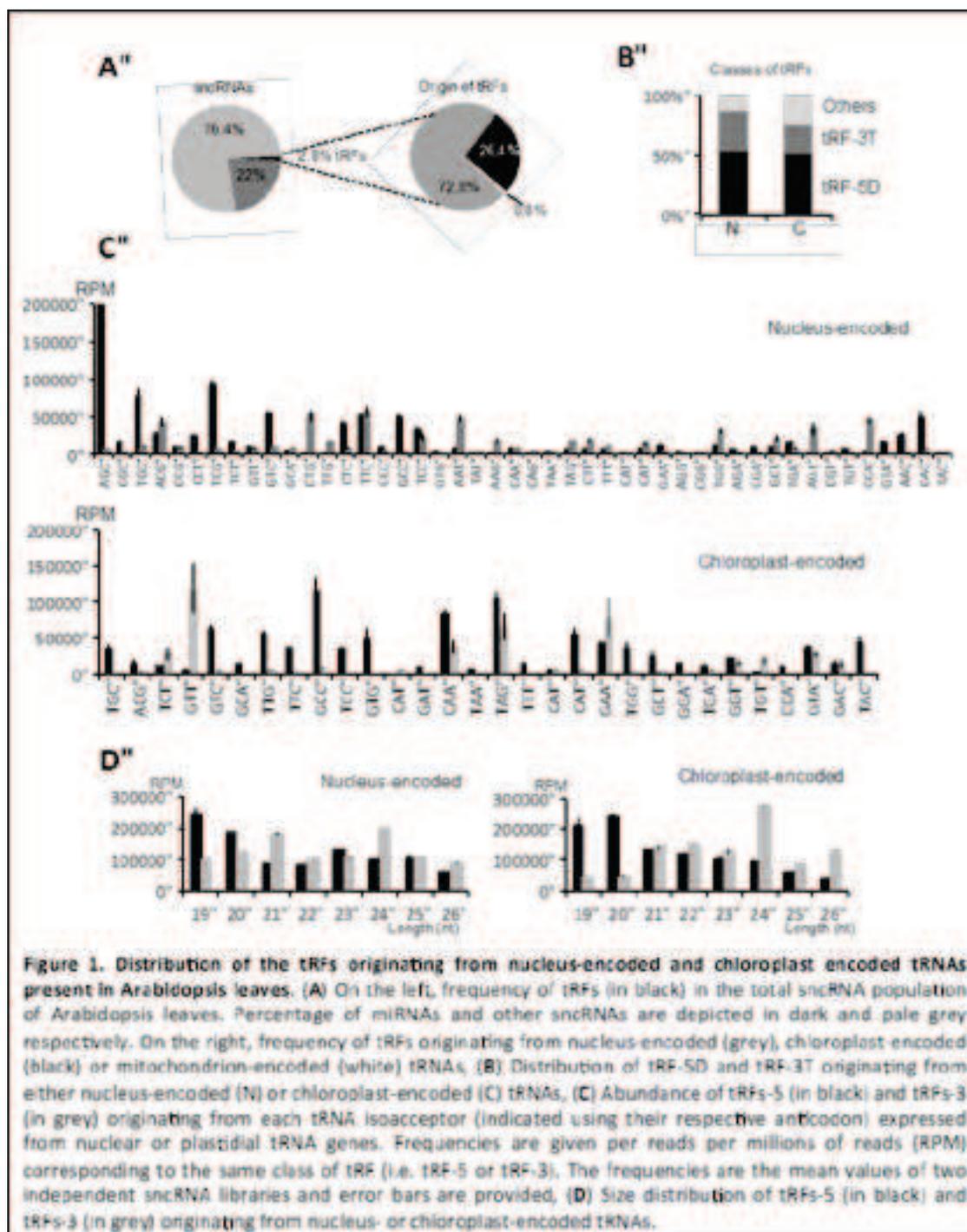
#### **REFERENCES**

1. Cech, T.R. and Steitz, J.A. (2014) The noncoding RNA revolution—tossing old rules to forge new ones. *Cell*, **157**, 77-94.
2. Sobala, A. and Huvagner, G. (2011) Transfer RNA-derived fragments: origins, processing, and functions. *RNA*, **2**, 853-862.
3. Gebetsberger, J. and Polacek, N. (2013) Slicing tRNAs to boost functional ncRNA diversity. *RNA Biol*, **10**, 1-9.
4. Raina, M. and Ibbá, M. (2014) tRNAs as regulators of biological processes. *Frontiers in Genet*, **5**, 1-14.
5. Megel, C., Morelle, G., Lalonde, S., Duchéne, A.M., Small, I. and Marichal-Drouard, L. (2015) Surveillance and Cleavage of Eukaryotic tRNAs. *Int. J. Mol. Sci.*, **16**, 1873-1893.
6. Gebetsberger, J., Zywicki, M., Kunzi, A. and Polacek, N. (2012) tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in *Haloflex volcanii*. *Archaea* **2012**, 260908.
7. Ivanov, P., Emara, Mohamed M., Villen, J., Ouyi, Steven P. and Anderson, P. (2011) Angiogenin-Induced tRNA Fragments Inhibit Translation Initiation. *Mol. Cell*, **43**, 613-623.
8. Hausviecker, D., Huang, Y., Lau, A., Parameswaran, P., Fire, A.Z. and Kay, M.A. (2010) Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA*, **16**, 673-695.
9. Burroughs, A.M., Ando, Y., de Hoon, M.J., Tomaru, Y., Suzuki, H., Hayashizaki, Y. and Daub, C.O. (2011) Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biol*, **8**, 156-177.
10. Kumar, P., Anaya, J., Mufsoni, S.B. and Datta, A. (2014) Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol*, **12**, 78.
11. Loss-Morris, G., Waterhouse, P.M. and Margis, R. (2013) Description of plant tRNA-derived RNA fragments (TRFs) associated with argonaute and identification of their putative targets. *Biologydirect*, **8**, 6.
12. Couvillion, M.T., Brunova, G., Parlom, E., Speed, T.P. and Collins, K. (2012) A Tetrahymena Piv1 Bound to Mature tRNA 3' Fragments Activates the Exonuclease Xrn2 for RNA Processing in the Nucleus. *Mol. Cell*, **48**, 509-520.

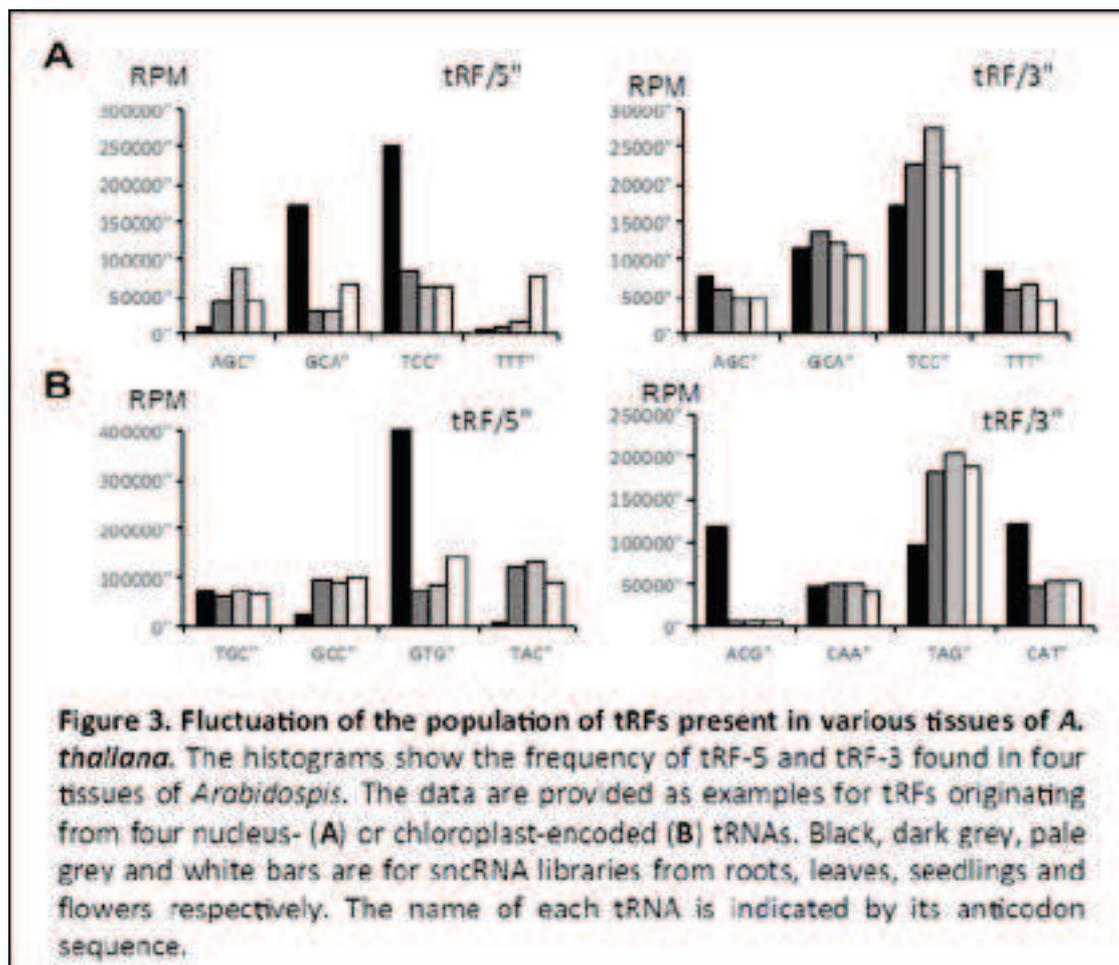
13. Ruggero, K., Guffanti, A., Corradin, A., Sharma, V.K., De Bellis, G., Corti, G., Grassi, A., Zanovello, P., Bronte, V., Ciminale, V. *et al.* (2014) Small noncoding RNAs in cells transformed by human T-cell leukemia virus type 1: a role for a tRNA fragment as a primer for reverse transcriptase. *J Virol*, **88**, 3612-3622.
14. Zhang, S., Sun, L. and Kragler, F. (2009) The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation. *Plant Phys*, **150**, 374-387.
15. Hsieh, L.C., Lin, S.J., Shih, A.C., Chen, J.W., Lin, W.Y., Tseng, C.Y., Li, W.H. and Chio, T.J. (2009) Uncovering small RNA-mediated responses to phosphate deficiency in Arabidopsis by deep sequencing. *Plant Phys*, **151**, 2120-2132.
16. Chen, C.J., Ju, Q., Zhang, Y.C., Qu, L.H., Chen, Y.Q. and Guetherot, D. (2011) Genome-wide discovery and analysis of microRNAs and other small RNAs from rice embryogenic callus. *RNA Biol*, **8**, 538-547.
17. Wang, L., Yu, X., Wang, H., Lu, Y.Z., de Ruiter, M., Prins, M. and He, Y.K. (2011) A novel class of heat-responsive small RNAs derived from the chloroplast genome of Chinese cabbage (*Brassica rapa*). *BMC genomics*, **12**, 289.
18. Hackenberg, M., Huang, P.J., Huang, C.Y., Shi, B.J., Gustafson, P. and Langridge, P. (2013) A comprehensive expression profile of microRNAs and other classes of non-coding small RNAs in barley under phosphorus-deficient and -sufficient conditions. *DNA Res*, **20**, 109-125.
19. Cognat, V., Dawlak, G., Duchêne, A.M., Daujat, M., Gigant, A., Salinas, T., Michaud, M., Gutmann, B., Giegé, P., Gebert, A. *et al.* (2012) PlantRNA, a database for tRNAs of photosynthetic eukaryotes. *Nucleic Acids Res*, **41**, 273-279.
20. Prüfer, K., Steuzel, U., Dammann, M., Green, R.E., Lachmann, M. and Kelso, J. (2008) PatMan: rapid alignment of short sequences to large databases. *Bioinformatics*, **24**, 1530-1531.
21. Kubiszewski-Jakubiak, S., Mezel, C., Ubrig, E., Salinas, T., Duchêne, A.M. and Maréchal-Drouard, L. (2015) In vitro RNA uptake studies in plant mitochondria. *Methods Mol Biol*, **in press**.
22. Pujol, C., Bailly, M., Keru, D., Maréchal-Drouard, L., Becker, H. and Duchêne, A.M. (2008) Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNA<sup>Gln</sup> synthesis in plants. *Proc Natl Acad Sci*, **105**, 6481-6485.
23. Maréchal-Drouard, L., Small, I., Wei, J.H. and Dietrich, A. (1995) Transfer RNA import into plant mitochondria. *Methods in enzymology*, **260**, 310-327.

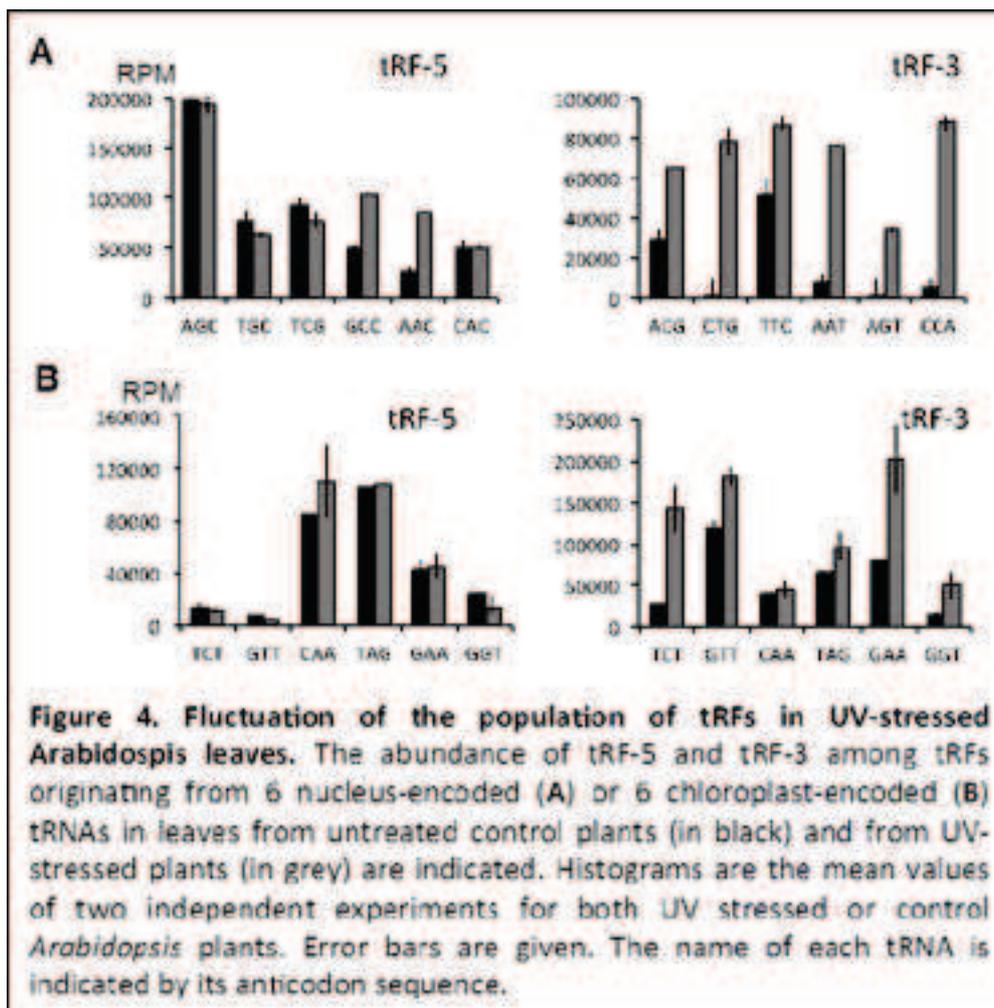
24. Carneiro, V.T., Dietrich, A., Maréchal-Drouard, I., Casset, A., Pelletier, G. and Small, I. (1994) Characterization of some major identity elements in plant alanine and phenylalanine transfer RNAs. *Plant Mol Biol.* **26**, 1841-1853.
25. Placido, A., Sieber, T., Gobert, A., Galerani, R., Guge, P. and Maréchal-Drouard, I. (2010) Plant mitochondria use two pathways for the biogenesis of tRNAs. *Nucleic Acids Res.* **38**, 7711-7717.
26. Placido, A., Gagliardi, D., Gallerani, R., Guenenberger, J.M. and Maréchal-Drouard, I. (2005) Fate of a leech-inherited tRNA precursor expressed in potato mitochondria. *J Biol. Chem.* **280**, 33571-33579.
27. Pribitzer, H., Gallemant, P., Weil, J.H. and Pillay, D.T.N. (1987) Adjustment of the tRNA population to the codon usage in chloroplasts. *Nucleic Acids Res.* **15**, 1377-1386.
28. Michaud, M., Cognat, V., Duchêne, A.M. and Maréchal-Drouard, I. (2011) A global picture of tRNA genes in plant genomes. *Plant J.* **66**, 80-93.
29. Cognat, V., Deragon, J.M., Vinogradova, E., Salinas, T., Remacle, C. and Maréchal-Drouard, I. (2008) On the evolution and expression of *Chlamydomonas reinhardtii* nucleus encoded transfer RNA genes. *Genetics* **179**, 113-123.
30. Thompson, D.M., Lu, C., Green, P.J. and Parker, R. (2008) tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *PNAS* **105**, 2095-2103.
31. Thompson, D.M. and Parker, R. (2009) Stressing Out over tRNA Cleavage. *Cell* **138**, 215-219.
32. Yao, Y., Danna, C.H., Zemp, F.J., Titov, V., Gillet, O.N., Przybylski, R., Anselmetti, F.M. and Kovachuk, I. (2011) UV-C-irradiated *Arabidopsis* and tobacco emit volatiles that trigger genome instability in neighboring plants. *Plant Cell* **23**, 3842-3852.
33. Maréchal-Drouard, I., Weil, J.H. and Diemsch, A. (1991) Transfer RNAs and Transfer RNA Genes in Plants. *Annals New York Acad. Sci.* **644**, 13-32.
34. Nowacka, M., Strojczyk, P.M., Jackowski, P., Eojka-Osimska, A., Szymanski, M. and Ingierowicz, M. (2013) Identification of stable, high copy number, medium-sized RNA degradation intermediates that accumulate in plants under non-stress conditions. *Plant Mol. Biol.* **83**, 191-204.
35. Tomita, K., Ojawa, T., Uozumi, T., Watanabe, K. and Masaki, H. (2000) A cytosolic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops. *Proc. Natl. Acad. Sci.* **97**, 8278-8283.

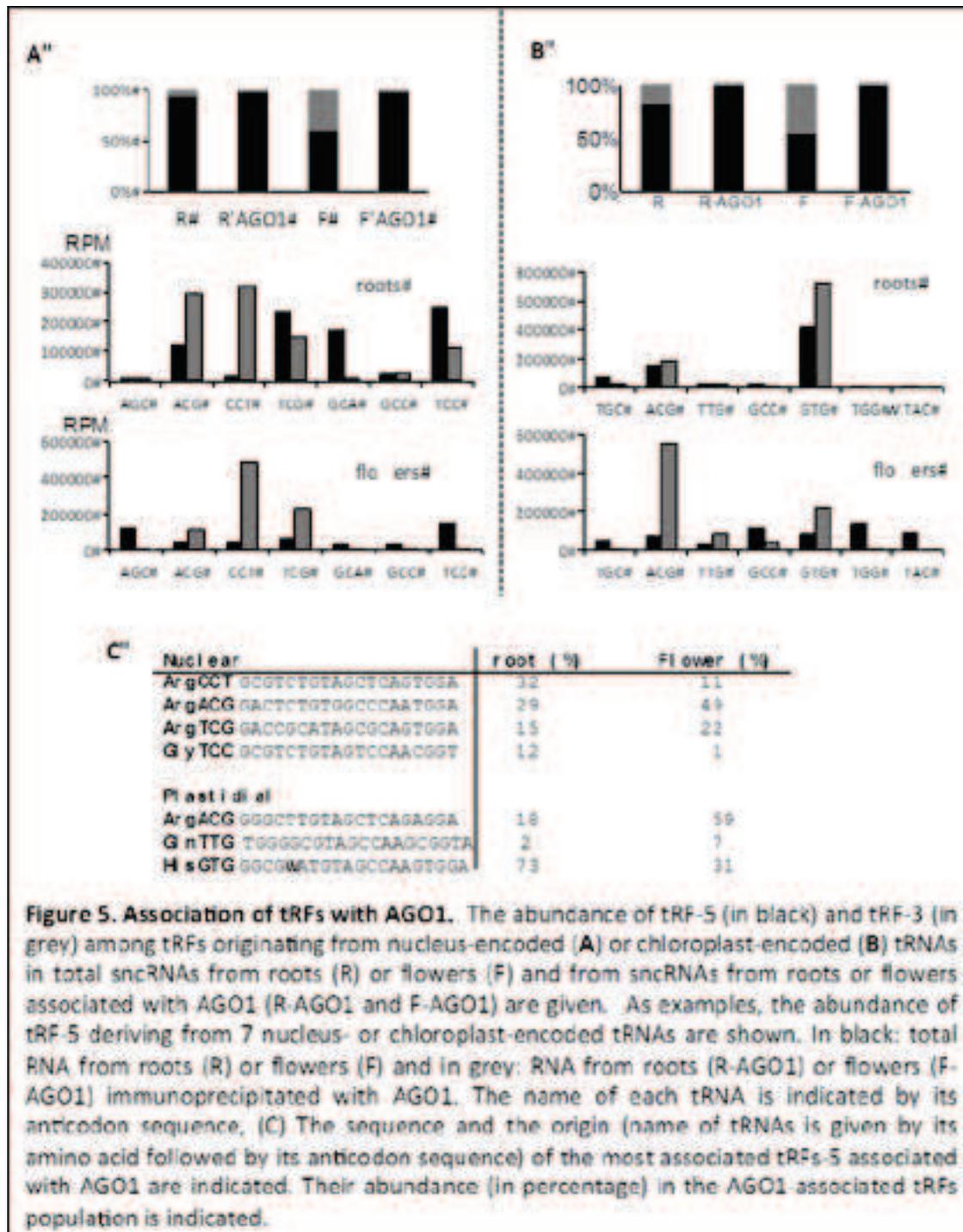
36. Tuorto, F., Liebers, R., Musch, T., Schaefer, M., Hofmann, S., Kellner, S., Frye, M., Helm, M., Staecklin, G. and Lyko, F. (2012) RNA cytosine methylation by Dnmt2 and Nsun2 promotes tRNA stability and protein synthesis. *Nature Struct. Mol. Biol.*, **19**, 900-905.
37. Cole, C., Sobala, A., Lu, C., Thatcher, S.R., Bowman, A., Brown, J.W.S., Green, P.J., Barton, G.J. and Hurvagner, G. (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA*, **15**, 2147-2160.
38. Maute, R.L., Schneider, C., Sumazin, P., Helmes, A., Califano, A., Basso, K. and Dalla-Favera, R. (2013) tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc. Natl. Acad. Sci.*, **110**, 1404-1409.
39. Thompson, D.M. and Parker, R. (2009) The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stress in *Saccharomyces cerevisiae*. *J. Cell. Biol.*, **185**, 43-50.
40. MacIntosh, G.C., Hillwig, M.S., Meyer, A. and Flagel, L. (2010) RNase T2 genes from rice and the evolution of secretory ribonucleases in plants. *Molecular Genet. Genomics*, **283**, 381-396.
41. Huot, J.L., Enkler, L., Mège, C., Karim, L., Laporte, D., Becker, H.D., Duchêne, A.M., Sissler, M. and Maréchal-Drouard, L. (2014) Idiosyncrasies in decoding mitochondrial genomes. *Biochimie*, **100**, 95-106.
42. Natesan, S.K., Sullivan, J.A. and Gray, J.C. (2005) Stromules: a characteristic cell-specific feature of plastid morphology. *J. Exp. Bot.*, **56**, 787-797.
43. Michaeli, S. and Galili, G. (2014) Degradation of organelles or specific organelle components via selective autophagy in plant cells. *Int. J. Mol. Cell.*, **16**, 7624-7638.
44. Ling, Q., Huang, W., Baldwin, A. and Jarvis, P. (2012) Chloroplast biogenesis is regulated by direct action of the ubiquitin-proteasome system. *Science*, **338**, 655-659.
45. Hsieh, L.C., Lin, S.J., Kuo, H.F. and Chiou, T.J. (2010) Abundance of tRNA-derived small RNAs in phosphate-starved *Arabidopsis* roots. *Plant Signal Behav.*, **5**, 537-539.
46. 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. *Proc. Natl. Acad. Sci.*, **111**, 18201-18206.

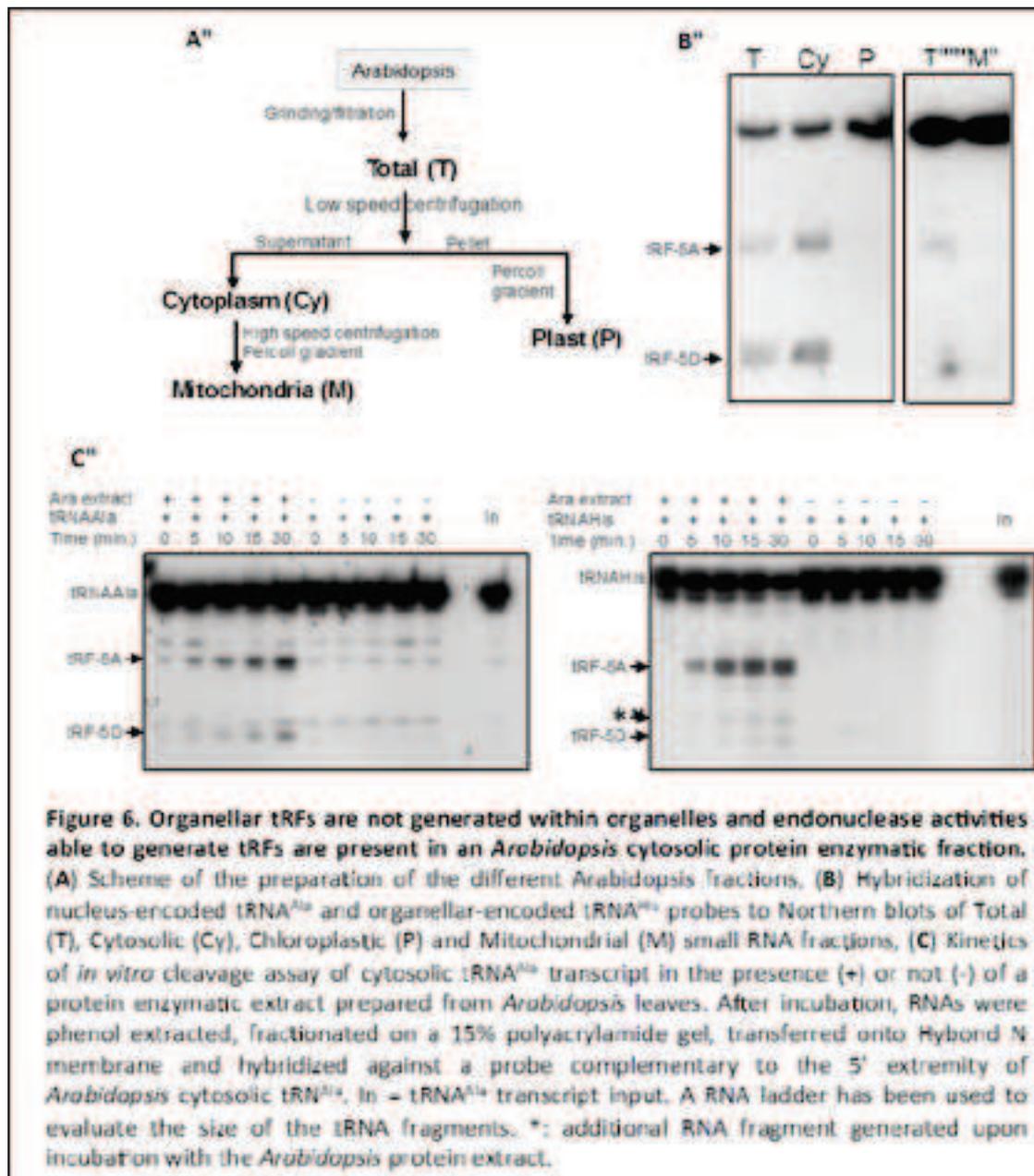


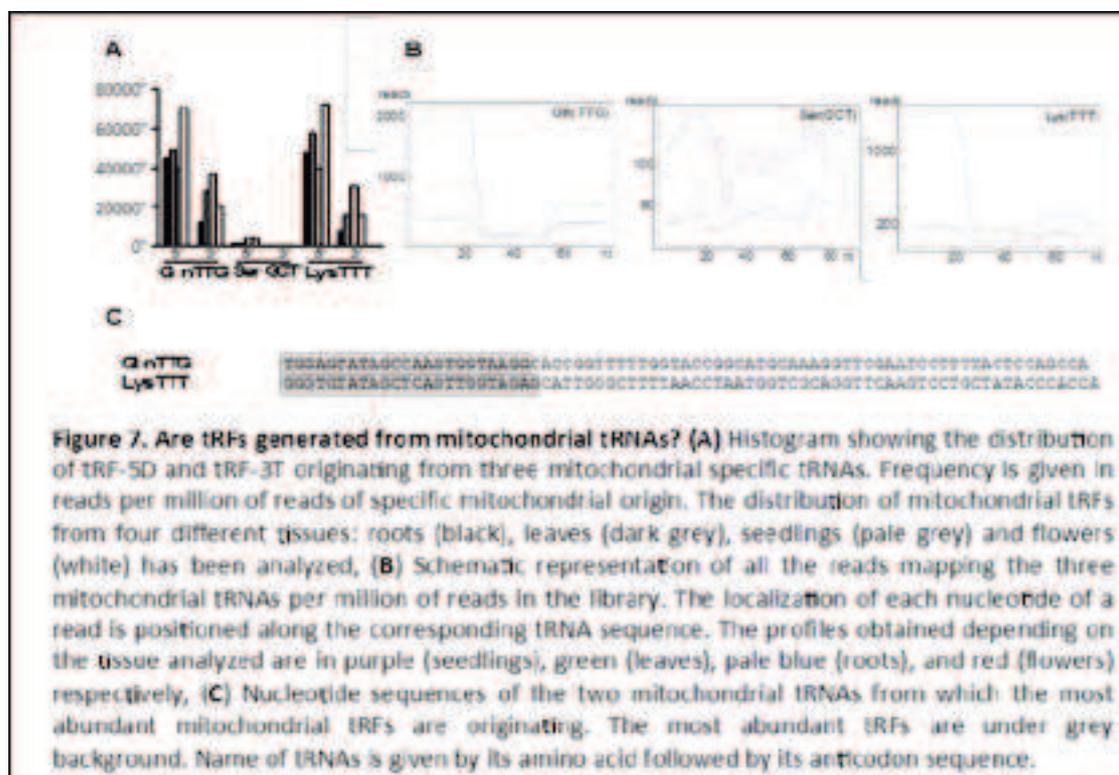








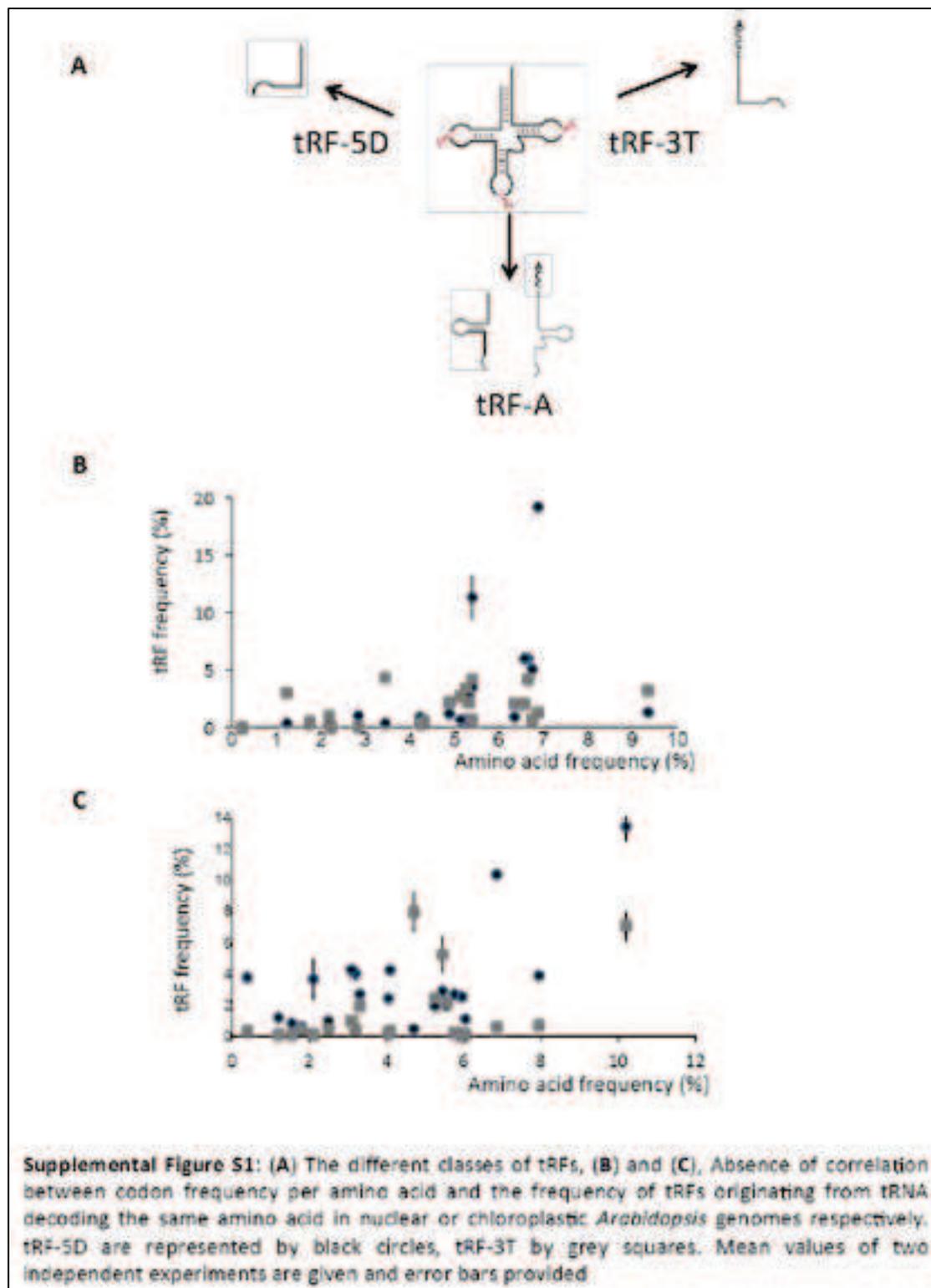


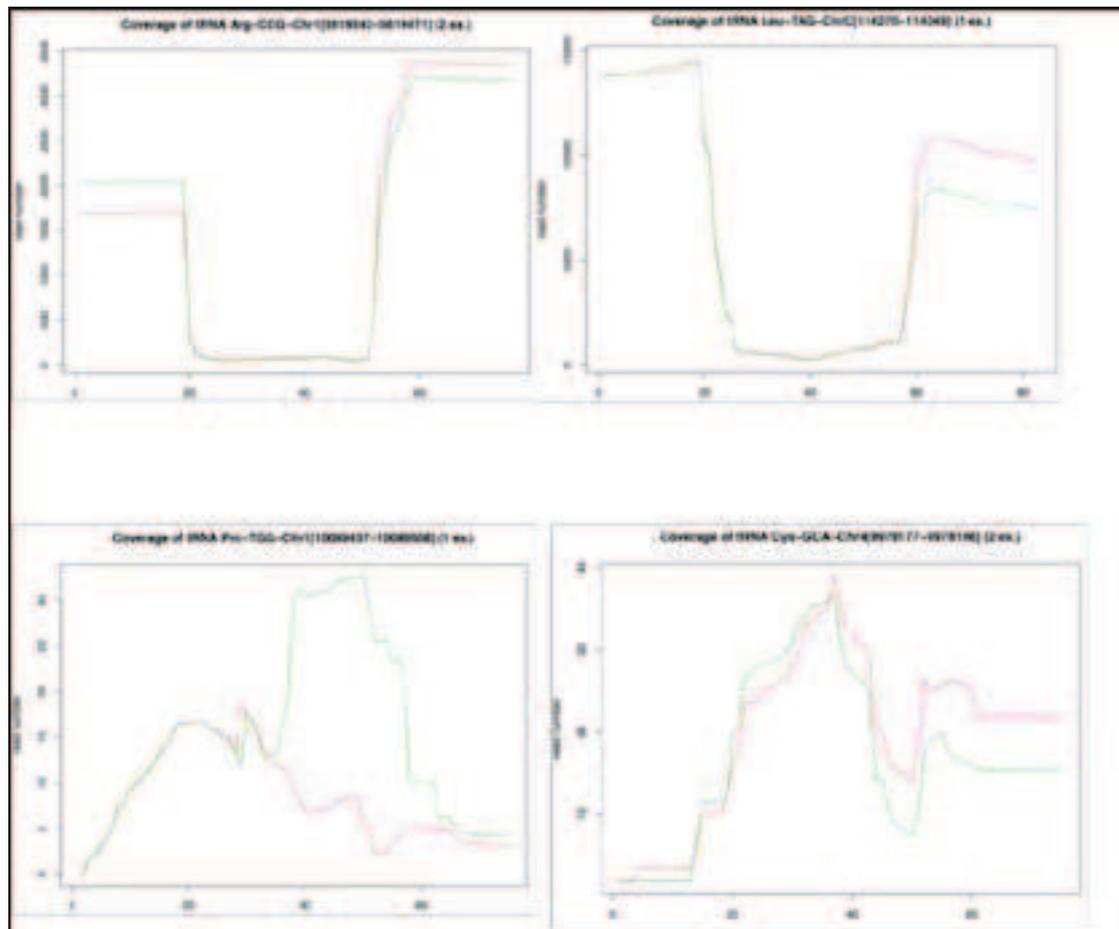


Description	Name	Tissue/organ	GEO Accessions	Related references
Col-0 leaves	L1	Leaves	-	This paper
Col-0 leaves	L2	Leaves	-	This paper
Col-0 roots	R1	Roots	GSM707680	(1,2)
Col-0 leaves	L3	Leaves	GSM707679	(1,2)
Col-0 seedlings	S1	Seedlings	GSM707681	(1,2)
Col-0 flowers	F1	Flowers	GSM707678	(1,2)
UV-C irradiated col-0 leaves	LU1	Leaves	-	This paper
UV-C irradiated col-0 leaves	LU2	Leaves	-	This paper
Control Col g1 RD29:LUC shoots	Co	Shoots	GSM647184	(1,2)
Drought stressed Col g1 RD29:LUC shoots	D1	Shoots	GSM647185	(1,2)
Cold stressed Col g1 RD29:LUC shoots	C1	Shoots	GSM647186	(1,2)
Salt stressed Col g1 RD29:LUC shoots	Sa1	Shoots	GSM647187	(1,2)
AGO1 associated sRNA in Col-0 roots	R-AGO1	Roots	GSM707691	(1,2)
AGO1 associated sRNA in Col-0 flowers	F-AGO1	Flowers	GSM707690	(1,2)

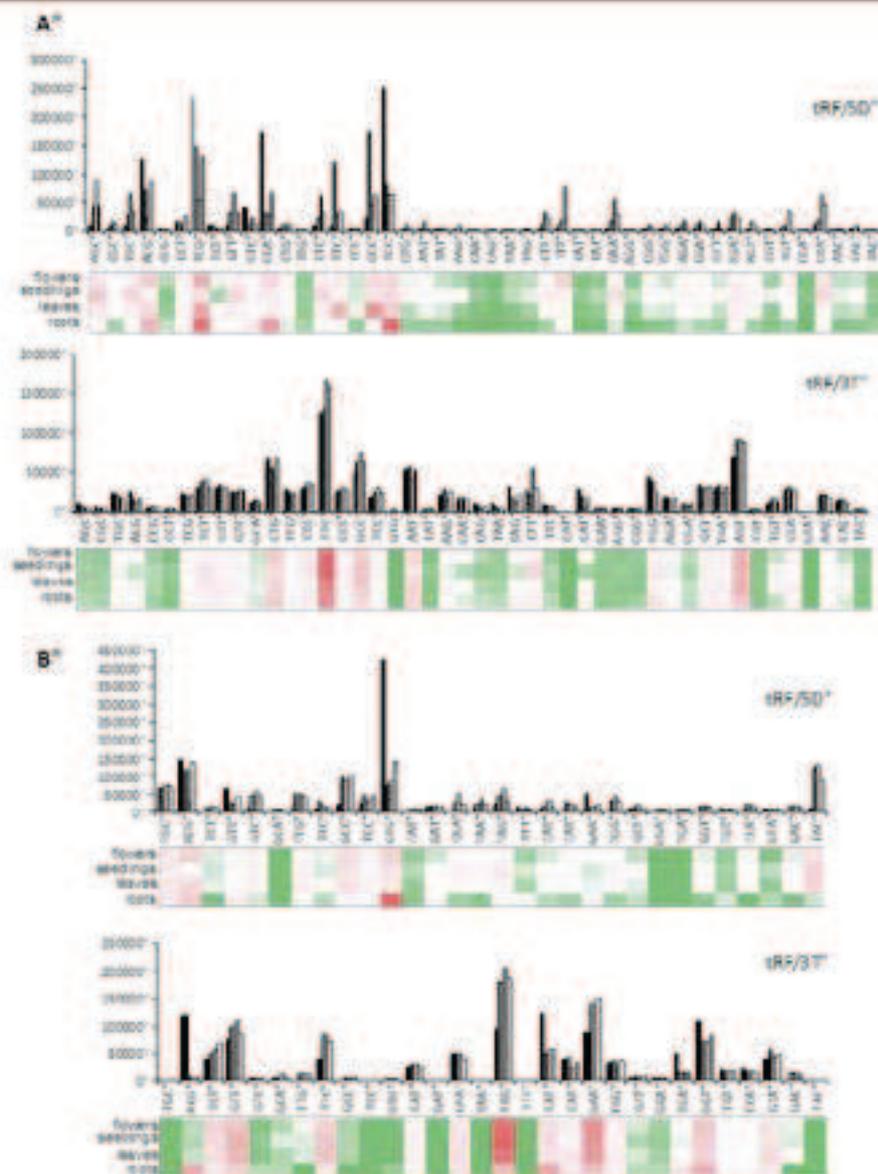
**Supplemental Table S1.** Description of the sncRNA deep sequencing libraries analyzed in this study.

1. Wang, H., Zhang, X., Liu, J., Kiba, T., Woo, J., Ojo, T., Hafner, M., Tuschl, T., Chua, N.H. and Wang, X.J. (2011) Deep sequencing of small RNAs specifically associated with Arabidopsis AGO1 and AGO4 uncovers new AGO functions. *Plant J.*, **67**, 292-304.
2. Loss-Morais, G., Waterhouse, P.M. and Margis, R. (2013) Description of plant tRNA-derived RNA fragments (tRFs) associated with argonaute and identification of their putative targets. *Biology direct*, **8**, 6.

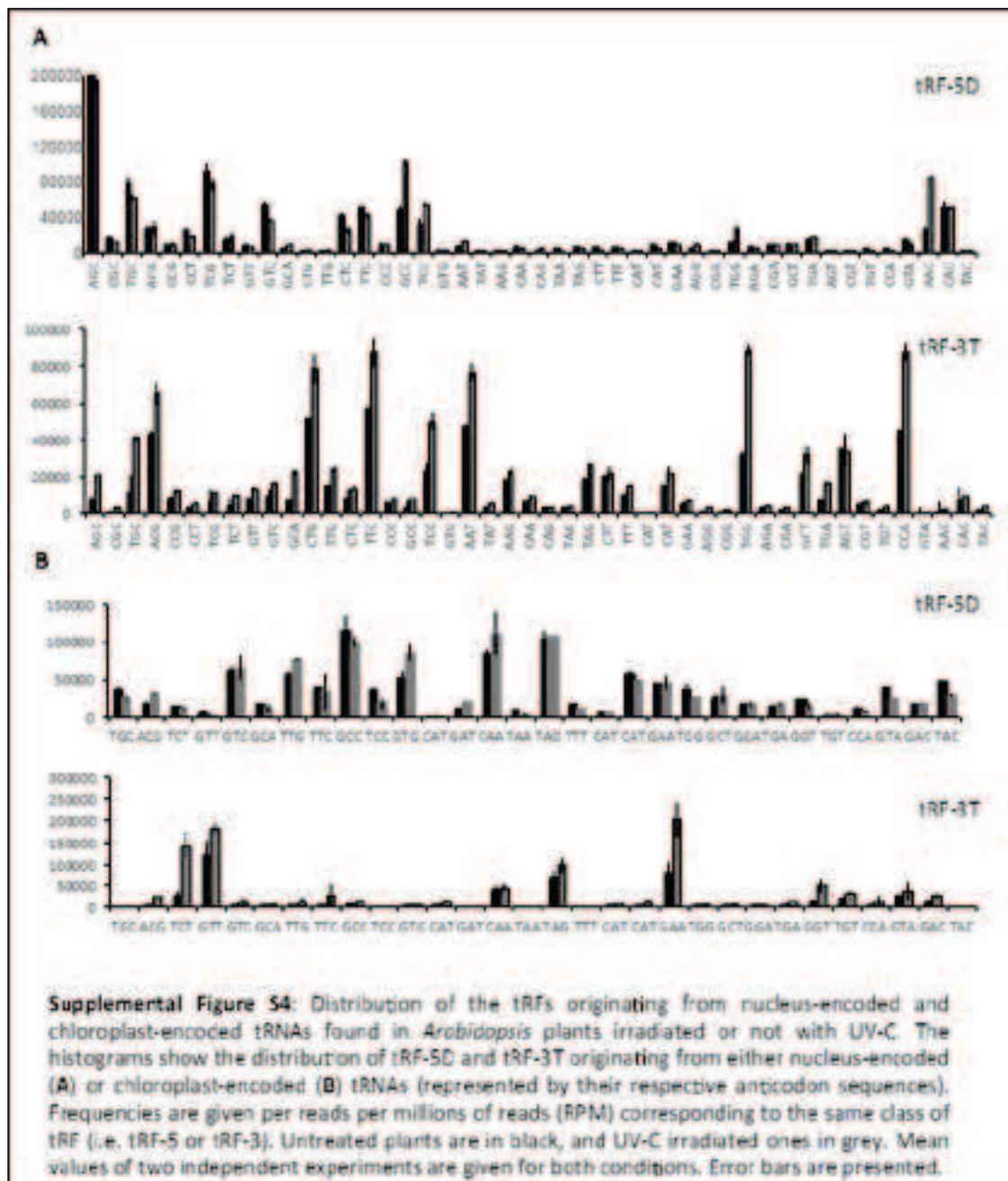




**Supplemental Figure S2:** Schematic representation of all the reads mapping specific tRNAs. The localization of each nucleotide of a read is positioned along the corresponding tRNA sequence. Four examples are given. Two for tRNAs where numerous tRF-5D and tRF-3T were observed, two for tRNAs where very few reads are present in the RNA libraries. The results of two independent sncRNA libraries are given in green and red respectively.



**Supplemental Figure S3:** Distribution of the tRFs originating from nucleus-encoded and chloroplast-encoded tRNAs found in various *Arabidopsis* tissues. The results are presented using two forms, histograms and color code. The histograms shows the distribution of tRF-50 and tRF-3T originating from either nucleus-encoded (**A**) or chloroplast-encoded (**B**) tRNAs (represented by their respective anticodon sequences). Frequencies are given per reads per millions of reads (RPM) corresponding to the same class of tRF (i.e. tRF-50 or tRF-3T). The tissues analyzed are roots (black), leaves (dark grey), seedlings (pale grey) and flowers (white). The color code is the following: red (for the most abundant tRFs), white (average level) to green (the less abundant tRFs)



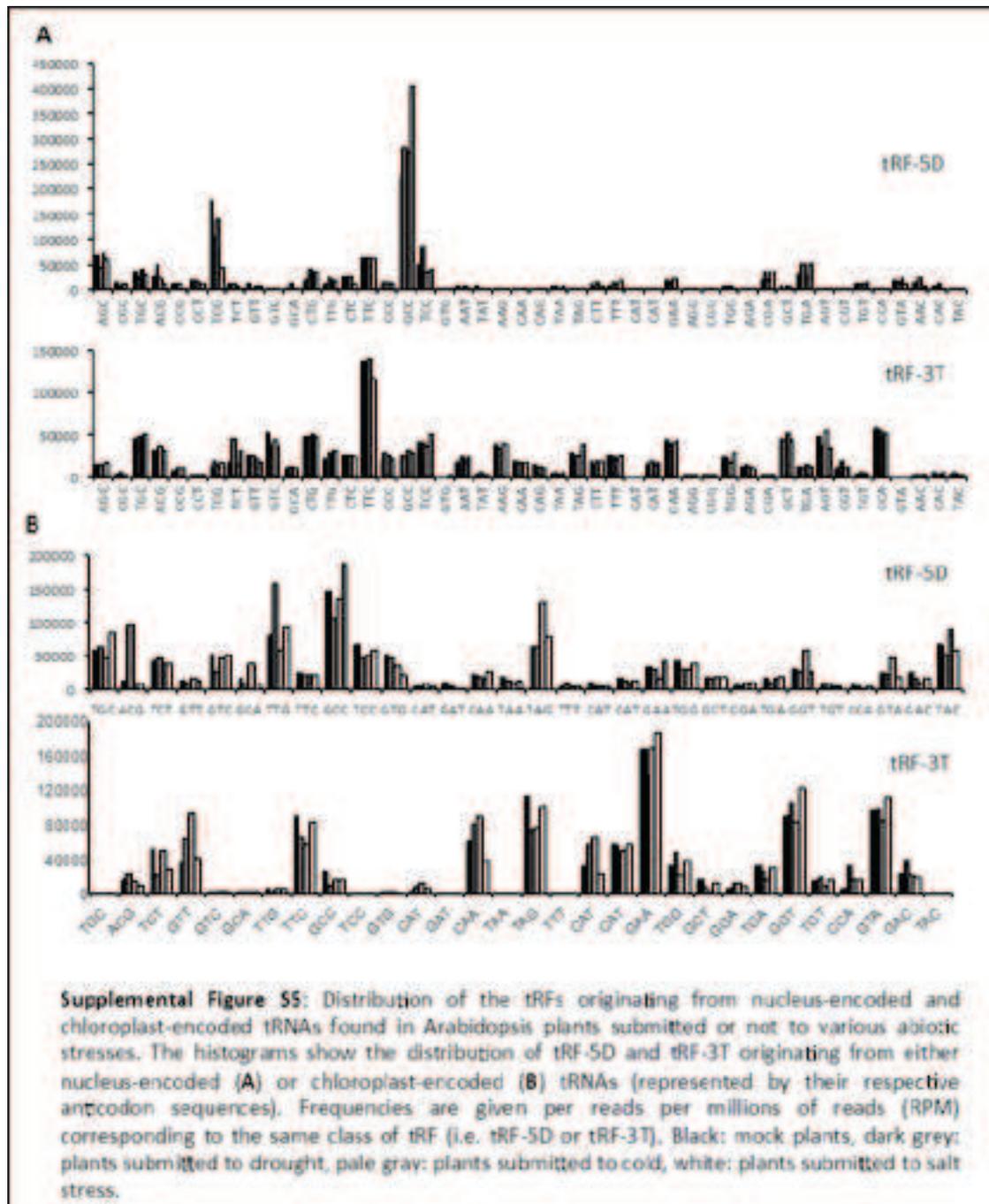






Figure 16 : Schéma synthétique de la création des banques de tRFs

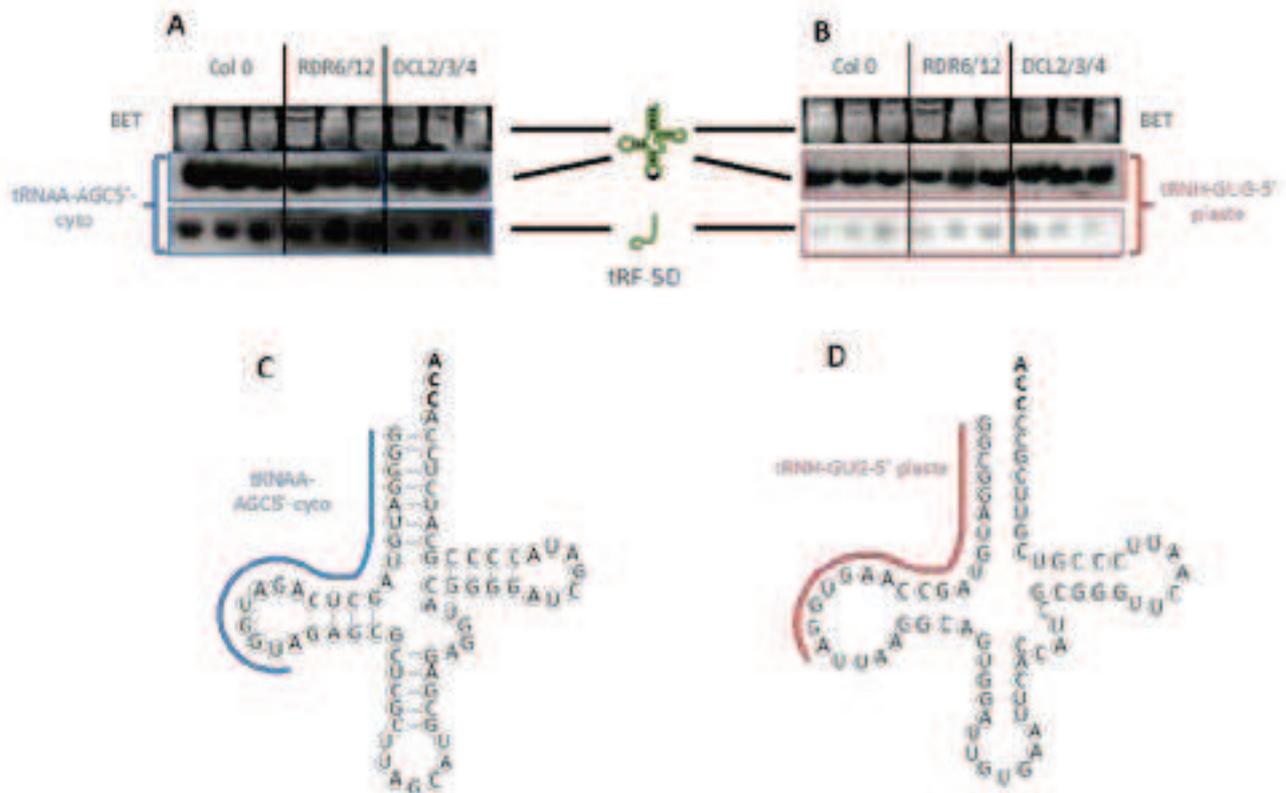
## 2 - Biogenèse des tRFs

Comme décrit dans la revue présentée dans l'introduction, il semble peu probable que les deux voies de dégradation des tRNA caractérisées chez *S. cerevisiae* soient responsables de la biogenèse des tRFs chez les eucaryotes et plus spécifiquement chez les plantes. Par contre chez certains organismes, des endonucléases ont pu être associées à la genèse des tRFs. Très brièvement, pour rappel, chez l'homme, une RNase de type A l'angiogénine et la RNase III DICER peuvent cliver des tRNA dans la boucle de l'anticodon et dans la boucle D respectivement (Cole et al., 2009; Fu et al., 2009). Chez *S. cerevisiae* il a été montré qu'une RNase de type T2, RNY1, clive les tRNA dans la boucle de l'anticodon (Luhtala and Parker, 2012). De même, chez *T. thermophila*, 3 endonucléases de type T2, Rnt2A, B et C, sont capables de cliver les tRNA dans l'anticodon.

En l'absence d'information dans le monde végétal nous avons, dans un premier temps, recherché si des endonucléases homologues à ces différentes enzymes, c'est-à-dire, à DICER, l'angiogénine et Rnt2, pouvaient être responsables de la genèse des tRFs chez *A. thaliana*. Nous avons donc, d'une part, étudié des mutants KO (Knock Out) pour les protéines DICER et d'autres part recherché des homologues aux RNases T2 et à l'angiogénine afin d'étudier leurs activités sur les tRNA *in vitro*.

### 2.1 – Etude de lignées mutantes *dcl* et *rdr* d'*A. thaliana*

La protéine DICER humaine est une RNase de type III intervenant dans l'interférence ARN. Elle possède 4 homologues chez *A. thaliana* nommés DCL (Dicer-like) et numérotés de 1 à 4 (DCL1-4), (Gascioli et al., 2005). DCL1 produit principalement des petits ARN de 21 nt alors que les produits de DCL2, DCL3 et DCL4 sont respectivement de 22, 24 et 21 nt. DCL1 intervient dans la biogenèse des miRNA, alors que les trois autres enzymes sont impliquées dans d'autres voies de petits ARN (cf. Introduction – 2.2.3). Les activités de ces enzymes sont partiellement redondantes. La lignée KO *dcl1* d'*Arabidopsis* est létale et uniquement des mutants *dcl1* hypomorphes sont disponibles. Cependant, le



**Figure 17** : Les enzymes de la voie de l'interférence ARN ne semblent pas impliquées dans la genèse des tRFs chez les plantes : A et B, coloration au BET d'un gel de polyacrylamide 15% dans lequel ont été déposés 10µg de petits ARN de fleurs issues de 3 plantes indépendantes pour chacune des 3 lignées testées. La première lignée est sauvage (col0), la deuxième lignée est KO pour les protéines RDR6 et 12, la troisième est KO pour DCL2: 3 et 4. A, northern blot à l'aide de la sonde tRNA<sup>Ala</sup>-AGC-5'-cyto spécifique du tRNA<sup>Ala</sup>(AGC) cytosolique. B, même membrane qu'en A mais hybridation à l'aide de la sonde tRNA<sup>His</sup>-GUG-5' plaste spécifique du tRNA<sup>His</sup>(GUG) plastidial. C et D, structure en feuille de trèfle des tRNA<sup>Ala</sup>(AGC) cytosolique et tRNA<sup>His</sup>(GUG) plastidial avec en bleu et en rouge, l'emplacement où s'hybride les sondes spécifiques aux extrémités 5' de ces tRNA.

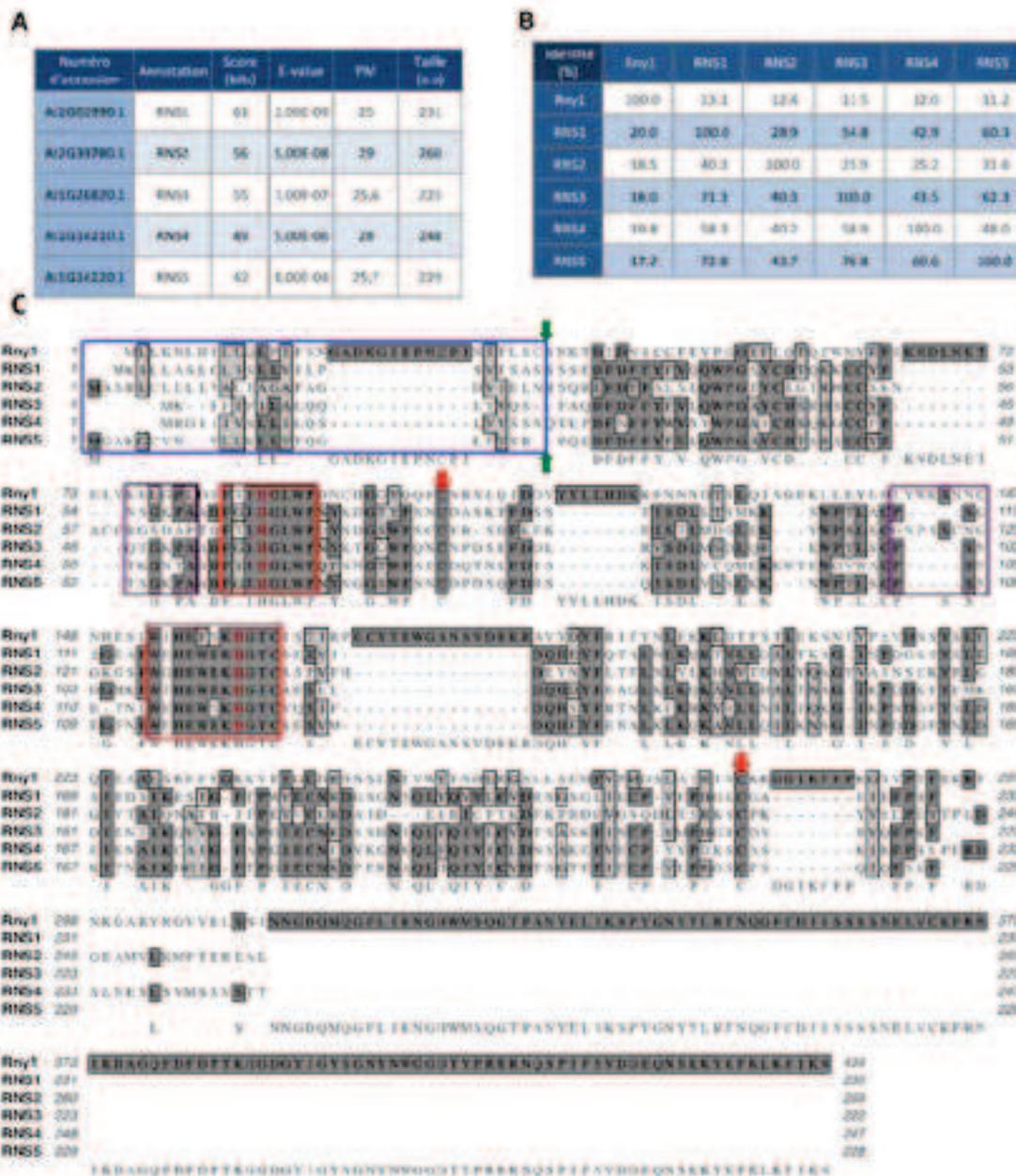
triple mutant KO *dcl2 dcl3 dcl4* (*dcl2/3/4*) est viable (Xie et al., 2004). Aucun clivage de tRNA par une de ces enzymes n'a été rapporté à ce jour.

La lignée triple KO *dcl2/3/4* est disponible à l'institut et nous a été fournie par Patrice Dunoyer (IBMP). Dans un premier temps, nous avons donc étudié ce triple mutant afin de savoir si ces enzymes étaient responsables du clivage des tRNA en particulier au niveau des régions D ou T. Pour cela nous avons analysé par northern blot trois extraits indépendants d'ARN provenant de fleurs de la lignée triple KO à l'aide de deux sondes : l'une spécifique de l'extrémité 5' du tRNA<sup>Ala(AGC)</sup> cytosolique, l'autre spécifique de l'extrémité 5' du tRNA<sup>His(GUG)</sup> chloroplastique d'*A. thaliana*. Comme le montre les figures 17A et 17B, il n'existe aucune différence significative entre la lignée sauvage *col0* et la lignée mutante *dcl2/3/4* quant à la quantité de petits ARN d'environ 20 nt dérivant soit du tRNA<sup>Ala(AGC)</sup>, soit du tRNA<sup>His(GUG)</sup>.

Ces résultats suggèrent que les protéines DICER ne sont pas à l'origine du clivage des tRNA chez *A. thaliana*, nous avons décidé de ne pas étendre notre étude aux lignées hypomorphes *dcl1*.

Toutefois, nous nous sommes intéressés à une autre lignée mutante pour le gène *rdr6* d'*A. thaliana* (lignée KO *rdr6-12*) qui code pour la RNA polymérase RNA dépendante RDR6 responsable du processus d'amplification des siRNA chez les plantes (cf. Introduction - 2.2.3). Comme le montre la figure 6 de l'introduction cette protéine occupe une fonction clé dans la synthèse des différents siRNA et si son absence provoque une diminution en tRFs cela nous encouragerait à poursuivre dans l'étude des enzymes de cette voie de biogenèse. Cette lignée KO *rdr6-12* nous a également été fournie par P. Dunoyer à partir d'une étude de Peragine et al. (Peragine et al., 2004). Comme ci-dessus, trois extraits indépendants d'ARN de fleurs ont été analysés par northern blot avec les mêmes sondes et aucune différence quant à la quantité de tRFs dérivant du tRNA<sup>Ala(AGC)</sup> ou du tRNA<sup>His(GUG)</sup> n'a été observée (figure 17).

Seuls les tRFs de deux tRNA, l'un d'origine cytosolique, l'autre d'origine plastidial ont été analysés au cours de cette étude. Toutefois les résultats obtenus suggèrent fortement que les enzymes impliquées dans la voie de l'interférence ARN chez les plantes supérieures ne sont pas responsables du clivage des tRNA pour générer des tRF, du moins ceux dérivant d'un clivage dans la région D des



**Figure 18 : Recherche d'homologues de RNY1 de levure chez *A. thaliana* :** **A**, liste des 5 premières séquences possédant le plus d'homologie avec RNY1 (Q02933.1) obtenues à l'aide du logiciel BlastP disponible sur TAIR (<http://www.arabidopsis.org/Blast/index.jsp>). La colonne "Score" indique l'homologie entre chaque protéine et RNY1 (plus le "score" est élevé plus l'homologie est forte). La colonne E-value indique les différences entre chaque protéine et RNY1 (plus la "e-value" est faible plus l'homologie est forte"). La colonne PM indique le poids moléculaire (en kDa) des protéines découvertes. **B**, taux d'homologie (en %) des différentes RNS entre elles ou avec RNY1; **C**, alignement des RNS avec RNY1 réalisé à l'aide du programme ClustalW dans le logiciel MacVector®. Le peptide signal est encadré en vert, le site de clivage de ce signal est symbolisé par une double flèche verte, les sites actifs de la protéine sont encadrés en rouge, les 2 résidus histidines essentiels à l'activité sont en rouge vif. Les 2 cystéines essentielles à la structure tridimensionnelle sont indiquées par 2 flèches rouges. La séquence approximative des sites B1 et B2 sont affichées en violet.

tRNA. Notre recherche de protéines candidates s'est alors focalisée sur l'existence potentielle d'homologues de l'angiogénine ou de RNase de type T2.

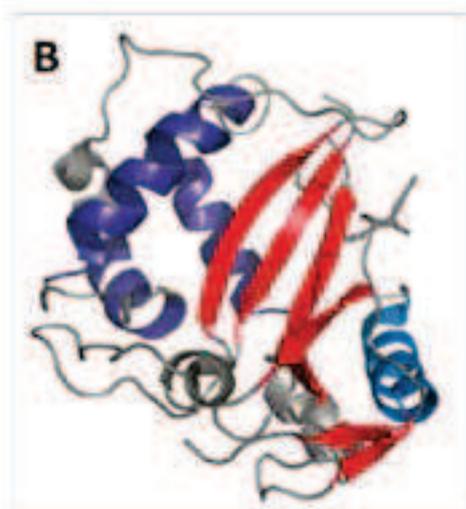
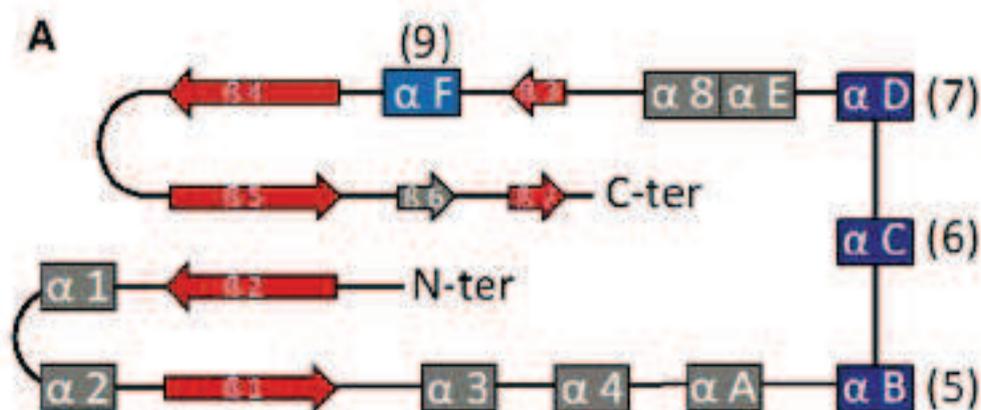
## 2.2 – Identification *in silico* d'homologues de rny1 de *S. cerevisiae* chez *A. thaliana*

Dans un deuxième temps, nous avons utilisé une approche *in silico* pour rechercher des homologues soit de l'angiogénine humaine soit de Rny1 de *S. cerevisiae*. La recherche à l'aide du programme d'alignement Blast (disponible sur TAIR : <http://www.arabidopsis.org/Blast/index.jsp>) à partir de la séquence protéique de l'angiogénine, la RNase de type A capable de cliver les tRNA au niveau de la boucle de l'anticodon chez les mammifères (Yamasaki et al., 2009), n'a abouti à l'identification d'aucun gène codant pour une protéine homologue sur le génome nucléaire d'*Arabidopsis*. Il est à noter qu'un résultat similaire a été obtenu chez *T. thermophila* (Andersen and Collins, 2011).

À l'inverse, l'utilisation de la séquence protéique de Rny1, la RNase de type T2 clivant également les tRNA au niveau de l'anticodon chez *S. cerevisiae* (Thompson and Parker, 2009a), a permis d'identifier 5 protéines candidates chez *A. thaliana* (figure 18A). Ces cinq protéines ont été nommées RNS1-5. Au niveau de leur séquence primaire, ces RNS ne possèdent qu'environ 20% d'identité de séquence avec RNY1 (figure 18B). Elles possèdent par contre une structure tridimensionnelle très proche (figure 19A). Alors que RNY1 a une taille de 50 kDa, les RNS d'*A. thaliana* ont des tailles qui varient entre 25 et 29 kDa. Cette différence de taille est liée entre autre à la présence d'une grande extension en C-terminal chez RNY1 de levure. Toutes ces protéines possèdent un peptide signal d'adressage au réticulum endoplasmique en N-terminal (figure 18C).

## 2.3 – Caractéristiques des RNases T2

Les RNases T2 ont été identifiées pour la première fois chez le champignon *Aspergillus oryzae* (Noguchi 1924). Depuis lors, il a été montré qu'elles étaient présentes chez pratiquement tous les groupes d'organismes vivants parmi lesquels les virus, les champignons, les plantes et les animaux

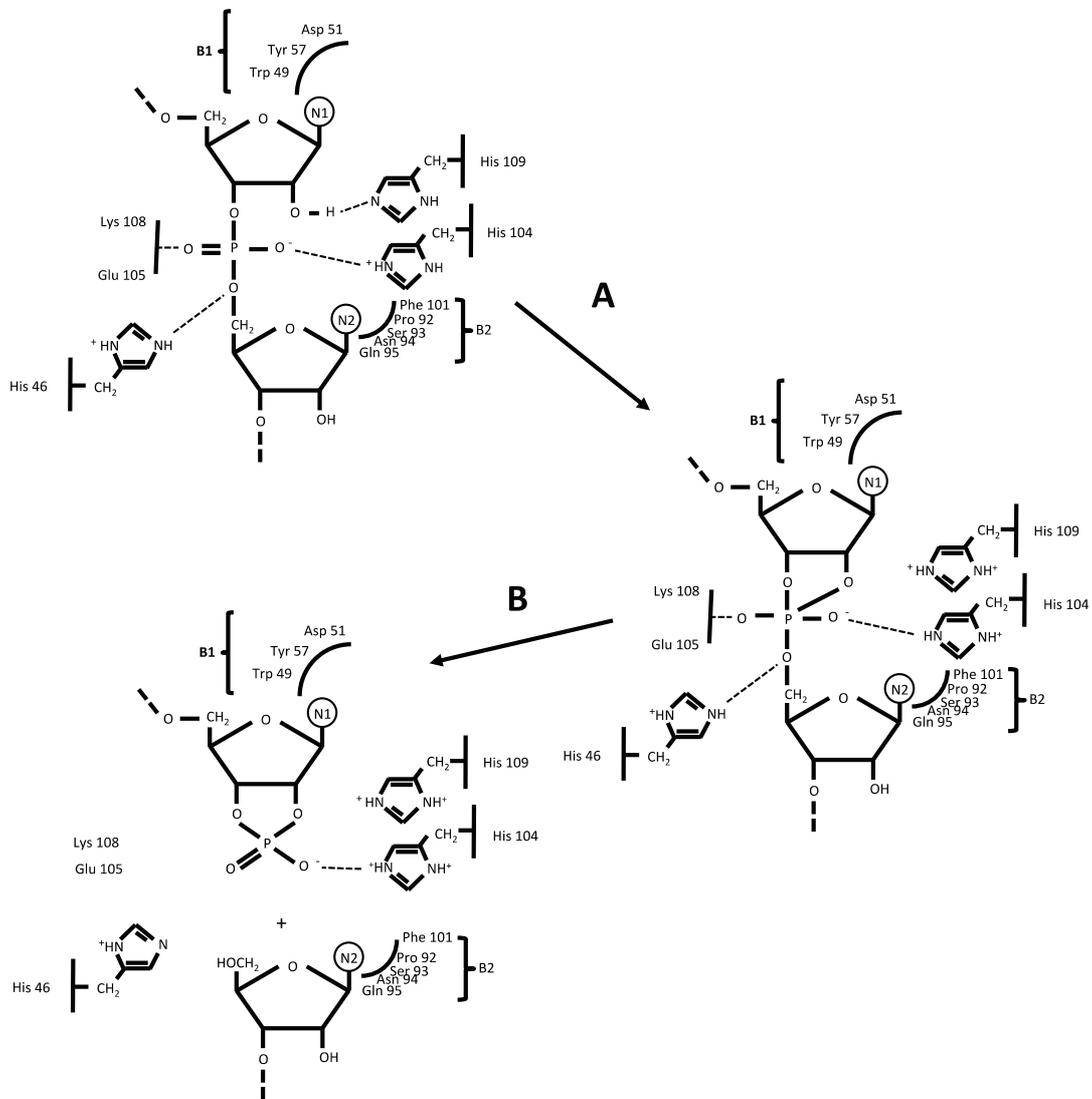


**Figure 19 : Schéma de la structure globale des RNase T2 :** **A**, diagramme topologique montrant les différents éléments de structures présent chez les RNase T2. Les éléments conservés chez toutes les RNase T2 sont symbolisés en **rouge** (feuillets  $\beta$ , flèches) et en **violet** (hélices  $\alpha$ , rectangles). Les éléments conservés partiellement sont symbolisés en **bleu**, et les éléments non conservés sont affichés en **gris** (figure d'après Macintosh et al. 2011). **B**, schémas de la structure 3D de la RNase LE de tomate (*Lycopersicon esculentum*, PDB ID : 1DIX) (Tanaka et al., 2000). Le code couleur utilisé est le même qu'en A (figure issue de Macintosh et al. 2011).

(Hillwig et al., 2011; Irie, 1999; MacIntosh et al., 2010) mais pas les archéobactéries (Condon and Putzer, 2002). Cette famille de protéines est particulièrement conservée chez les eucaryotes (Garcia-Silva et al., 2010; MacIntosh, 2011). Alors que certaines sont sécrétées, d'autres sont localisées dans la cellule au niveau de la vacuole ou du lysosome (*e.g* RNY1 de *S. cerevisiae*, RNase T2 Humaine et de *Danio rerio*) (Campomenosi et al., 2006; Haud et al., 2011; Thompson and Parker, 2009).

Ces endonucléases ont un poids moléculaire qui varie entre 20 et 50kDa (Deshpande and Shankar, 2002). Elles portent généralement des modifications post-traductionnelles comme des N et/ou O glycosylations (Oxley et al., 1996; Parry et al., 1998), de plus ces RNases sont inductibles par des blessures (Kawano et al., 2002, 2006). La structure tridimensionnelle de RNase T2 de différents organismes dont plusieurs plantes, comme *Nicotiana glauca*, *Pyrus pyrifolia* (Ida et al., 2001; Matsuura et al., 2001), *Lycopersicon esculentum* (RNase LE) (Tanaka et al., 2000), *Momordica charantia* (RNase MC1) (Nakagawa et al., 1999) et *Nicotiana glutinosa* (RNase NW et NT) a été obtenue par cristallographie. Une structure consensus a ainsi pu être établie pour les RNases T2 de bactéries, champignons et plantes (figure 19A) (MacIntosh, 2011). Celle-ci se décompose en deux séries de deux feuillets  $\beta$  anti-parallèles (1, 2, et 4, 5) suivi de deux petits feuillets  $\beta$  (3 et 7) antiparallèles ainsi que 3 hélices  $\alpha$  très conservées (B, C et D) (figure 19A). Il existe une hélice  $\alpha$  moins bien préservée ( $\alpha$  F, ou 9 selon les nomenclatures), et d'autres absolument pas ( $\alpha$  1, 2, 3, 4, A, 8 et E). Ces RNases T2 possèdent plusieurs résidus cystéine (en nombre variable) formant des ponts disulfures à divers endroits de la protéine (figure 18C). Ceux-ci la maintiennent en conformation active. Deux de ces ponts disulfures sont très conservés chez toutes les RNases T2 (Irie, 1999) tandis que deux autres le sont moins chez les plantes et les animaux (MacIntosh, 2011).

Les RNases T2 catalysent le clivage d'ARN simple brin *via* la formation d'un intermédiaire 2'-3' phosphate cyclique à l'extrémité 3' (figure 20). Elles clivent les ARN simple brin en libérant une extrémité 5'OH et 3'P. Leur pH optimum se situe entre 4 et 5,5 (Deshpande and Shankar, 2002; Irie, 1999; Luhtala and Parker, 2012). Le mécanisme d'action des RNases T2 a majoritairement été étudié chez les champignons. On estime cependant que



**Figure 20 : Première étape du clivage des ARN par les RNase T2: A**, Lors de la première étape l'His 46 agit comme un acide et l'His 109 comme une base. L'His 109 va ainsi capturer l'atome d'hydrogène en position 2' du ribose en amont du site de clivage. **B**, Ceci permettra l'attaque acide de l'histidine 46 sur le groupement phosphate libérant ainsi une extrémité 3'-P cyclique et une autre 5'-OH. La deuxième étape de cette catalyse (non représentée) consistera en la conversion de l'extrémité 3'-P cyclique en 3'-P (figure d'après Irie et al. 1999).

celui-ci est similaire chez toutes les espèces dans la mesure où les sites actifs de ces enzymes sont très conservés. Deux résidus histidine sont très importants pour l'activité catalytique de ces RNases, les His 46 et 109 (selon la numérotation de la RNase Rh, une des RNase T2 les plus étudiées à ce jour appartenant au champignon *Rhizopus niveus*) (Kawata et al., 1990; MacIntosh, 2011; Parry et al., 1997). Ces 2 histidines sont présentes dans 2 sites actifs (CAS: "Conserved catalytic Sites"), CAS I et CAS II dont les séquences sont respectivement les suivantes: F/WTL/IHGLWP et FWXHEWXKHGTC (Irie, 1999). Ces deux sites catalytiques ainsi que les résidus histidines sont présents chez les RNS d'*A. thaliana* (figure 18C). Le processus de clivage se découpe en deux étapes. La première consiste au clivage de la liaison phosphodiester *via* la formation transitoire d'une extrémité 3'-P cyclique qui sera convertie en 3'-P lors de la seconde étape (figure 20, seule la première étape est représentée). Tout d'abord, l'His 109 va agir comme une base afin de capter l'atome d'hydrogène se situant dans le groupement 2'-OH. Ceci générera un intermédiaire 2'-3' phosphate cyclique (figure 20A). Cependant, l'histidine 46, dont le pK<sub>a</sub> est plus élevé que celui de l'His 109, va agir comme un acide sur l'extrémité 5'-O du nucléotide en aval du site de clivage pour lui donner un proton et ainsi cliver l'ARN (figure 20B) (Kawata et al., 1990; Ohgi et al., 1992). Puis, au cours de la deuxième étape (non représentée sur le schéma), le rôle des 2 histidines est inversé, l'histidine 46 agit donc comme une base et l'histidine 109 comme un acide. L'histidine non-protonée va alors activer une molécule d'eau et former un groupement OH<sup>-</sup> qui va attaquer la liaison en 2' du phosphate cyclique tandis qu'un proton sera donné par l'histidine 109 qui agira comme un acide. Ceci établira alors une extrémité 5'-OH et 3'-P de part et d'autre du site de clivage. Deux autres sites actifs B1 et B2 ont été identifiés. Chacun d'eux correspond à une poche hydrophobe, dans laquelle va se fixer chacune des extrémités de la liaison phosphodiester clivée. Ces deux poches participent ainsi à l'interaction entre le substrat et la RNase T2. Chez les plantes, les résidus tryptophane et tyrosine du site B1 sont très conservés (Kawano et al., 2002, 2006; Tanaka et al., 2000), et ils sont à l'origine de l'interaction entre l'ARN et la protéine (MacIntosh, 2011).

Les RNases T2 n'ont pas de spécificité de clivage. Toutefois une préférence pour certains nucléotides a pu être observée en fonction de

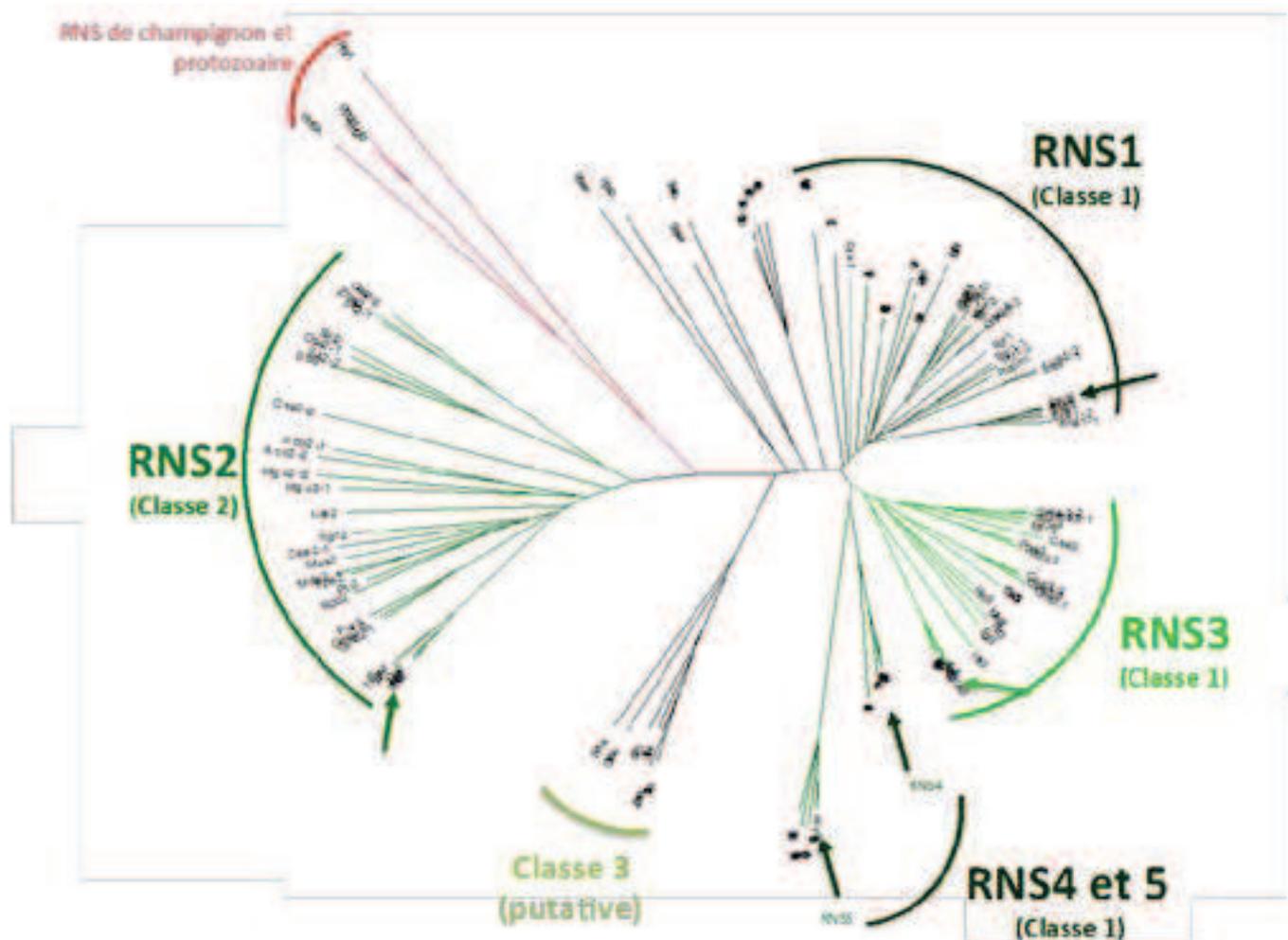
Embranchements	Organisme	Nombre de gènes de RNase T2
Chlorophyta	<i>Chlamydomonas reinhardtii</i>	1
	<i>Volvox carteri</i>	1
	<i>Ostreococcus tauri</i>	1
Marchantiophyta	<i>Marchantia polymorpha</i>	1
Pinophyta	<i>Picea glauca</i>	3
Magnoliophyta	<i>Arabidopsis thaliana</i>	5
	<i>Populus trichocarpa</i>	5
	<i>Solanum lycopersicum</i>	3
	<i>Nicotiana glutinosa</i>	3
	<i>Glycine max</i>	11
	<i>Antirrhinum hispanicum</i>	1
	<i>Calystegia sepium</i>	1
	<i>Oryza sativa</i>	8
	<i>Sorghum bicolor</i>	6
	<i>Zea mays</i>	2
	<i>Luffa aegyptiaca</i>	2
	<i>Momordia charantia</i>	1
	<i>Cicer arietinum</i>	1
	<i>Prunus dulcis</i>	2
	<i>Zinnia violacea</i>	2
	<i>Drosera adelae</i>	1
	<i>Nicotiana alata</i>	1
	<i>Fagopyrum esculentum</i>	1
<i>Triticum aestivum</i>	4	
<i>Hordeum vulgare</i>	4	
Lycopodiophyta	<i>Selaginella moellendorffii</i>	2
Pteridophyta	<i>Adiantum capillus-veneris</i>	1
Bryophyta	<i>Physcomitrella patens</i>	3

Figure 21 : Nombre de gènes de RNase T2 chez diverses plantes et algues.

l'endonucléase considérée (pour une revue : Deshpande and Shankar, 2002). Ainsi, par exemple, la RNase Rh clive préférentiellement après une adénine. Les RNases LE et NW, quant à elle, coupent après une guanine, tandis que les RNases MC1 et CL1 le font après une uridine (MacIntosh, 2011). D'autres en revanche n'ont montré aucune préférence (*e.g.* la RNase NT) (MacIntosh, 2011). Il est intéressant de noter que la RNase LE de tomate, dont la position 51 est occupée par un résidu asparagine, clive préférentiellement après une guanine, or la mutation de cet asparagine pour un résidu aspartique entraîne un changement dans la préférence de clivage pour une uridine (MacIntosh, 2011).

L'analyse phylogénétique des RNases T2 de plantes et d'animaux a montré que le nombre de gènes codants pour ces protéines est plus élevé chez les plantes (MacIntosh et al., 2010). Alors que la plupart des animaux n'ont conservé qu'un gène codant pour une RNase T2, on dénombre par exemple huit gènes codants pour une RNase T2 chez le riz (Hillwig et al., 2011; MacIntosh, 2011) et 5 chez *Arabidopsis* (voir aussi tableau 21). Chez les plantes, on distingue 3 classes de RNase T2 (figure 22) (Igc and Kohn, 2001; MacIntosh et al., 2010; Roalson and McCubbin, 2003; Steinbachs and Holsinger, 2002). Les RNases T2 de classe I sont spécifiques de certains tissus et leur production est généralement régulée par des stress (Zhang et al., 2000). Les gènes de classe III sont dérivés de gènes de classe I. Ils jouent généralement un rôle dans la floraison (MacIntosh et al., 2010; Roalson and McCubbin, 2003; Steinbachs and Holsinger, 2002). Contrairement aux RNases T2 des classes I et III, les endonucléases de classe II semblent avoir conservé des caractéristiques ancestrales. Il n'existe qu'un gène de classe II par génome végétal, et ce dernier est exprimé constitutivement suggérant un rôle essentiel pour la cellule (Köthke and Köck, 2011; MacIntosh et al., 2010). Chez *A. thaliana* RNS 1, 3, 4 et 5 sont des RNases T2 de classe I, RNS 2, quant à elle, appartient à la classe II (Nowak et al., 2011).

Ces données font écho à l'étude phylogénétique menée récemment par C. Megel doctorant dans l'équipe qui met plus en évidence les différents groupes de RNS qui ne correspondent pas forcément aux grandes classes connues de RNases T2. L'arbre phylogénétique dégagé de cette étude (figure 22) permet de distinguer 4 grands groupes de RNS (RNS1, RNS2, RNS3 et RNS4/5). Ces groupes sont relativement distincts les uns des autres, cependant RNS2 est plus éloigné



**Abbreviations :**

<i>Aco</i>	<i>Aquilegia coerulea</i>	<i>Gma</i>	<i>Glycine max</i>	<i>Ptr</i>	<i>Populus trichocarpa</i>
<i>Aly</i>	<i>Arabidopsis lyrata</i>	<i>Lus</i>	<i>Linum usitatissimum</i>	<i>Pvu</i>	<i>Phaseolus vulgaris</i>
<i>Bdl</i>	<i>Brachypodium distachyon</i>	<i>Mdo</i>	<i>Malus domestica</i>	<i>Rco</i>	<i>Ricinus Communis</i>
<i>Bra</i>	<i>Brassica rapa</i>	<i>Mes</i>	<i>Manihot esculenta</i>	<i>Tha</i>	<i>Thellungiella halophila</i>
<i>Ccl</i>	<i>Citrus clementina</i>	<i>Mgu</i>	<i>Mimulus guttatus</i>	<i>Sbi</i>	<i>Sorghum bicolor</i>
<i>Cpa</i>	<i>Carica papaya</i>	<i>Mtr</i>	<i>Medicago truncatula</i>	<i>Sit</i>	<i>Setaria italica</i>
<i>Cru</i>	<i>Capsella rubella</i>	<i>Osa</i>	<i>Oryza sativa</i>	<i>Sma</i>	<i>Selaginella moellendorffii</i>
<i>Csa</i>	<i>Cucumis sativus</i>	<i>Ppa</i>	<i>Physcomitrella patens</i>	<i>Vvi</i>	<i>Vitis vinifera</i>
<i>Csi</i>	<i>Citrus sinensis</i>	<i>Ppe</i>	<i>Prunus persica</i>	<i>Zma</i>	<i>Zeamays</i>
<i>Egr</i>	<i>Eucalyptus grandis</i>				

**Figure 22 :** Arbre phylogénétique des RNase T2 provenant de plantes ainsi que d'un champignon (*S. cerevisiae*) et d'un protozoaire (*T. thermophila*): Les RNS d'*A. thaliana* sont indiquées par une flèche. Les différents groupes auxquels chaque RNS appartient sont indiquées en vert (RNS1, RNS2, RNS3 et RNS4 et 5). Les RNases T2 connus pour cliver des tRNA chez d'autres organismes que les plantes sont affichées en rouges : *S. cerevisiae* (RNY1) (champignon) et *T. thermophila* (protozoaire) (Rnt2A, Rnt2B et Rnt2H) (Arbre créé par C. Megel).

phylogénétiquement des 3 autres groupes. Les RNases T2 de *S. cerevisiae* et *T. thermophila* sont par contre phylogénétiquement plus proche du groupe RNS2. Il est difficile de déterminer à quel groupe appartiennent les RNS de bryophyte (e.g *Physcomitrella patens*) ou de lycophyte (e.g *Selaginella moellendorffii*).

## 2.4 - Données bibliographiques des RNS d'*A. thaliana*

Les données de la littérature concernant la famille des RNS d'*A. thaliana* restent encore très parcellaires. Trois d'entre elles ont été partiellement caractérisées, RNS 1, 2 et 3, il y a maintenant une vingtaine d'années. Les recherches de l'équipe de Pamela J. Green ont ainsi permis de montrer que RNS1 était excrétée dans le milieu de culture d'une lignée de cellules d'*Arabidopsis thaliana* T87-C33 *in vitro*, tandis que RNS2 était séquestrée dans la cellule (Bariola et al., 1999). RNS2 est localisée dans les vacuoles ou le réticulum endoplasmique ainsi que les vésicules qui en dérivent (Hillwig et al., 2011). L'étude de banques de cDNA a permis de déterminer que toutes ces protéines étaient ubiquitaires (Taylor and Green, 1991; Taylor et al., 1993). RNS2 semble toutefois la plus exprimée (Taylor and Green, 1991). Il y existe cependant une prévalence de l'une ou l'autre dans certains tissus, RNS1 est ainsi abondante dans les fleurs (Bariola et al., 1994; Taylor et al., 1993), tandis que RNS3 est plutôt exprimée dans les tiges, les racines et les fleurs. Toutefois, l'expression de ces RNS n'a pas à l'époque été analysée sur tous les organes de la plante. Ainsi, les données de microarrays disponibles sur Genevestigator (<https://genevestigator.com/gv/plant.jsp>) ou BAR (<http://bar.utoronto.ca/welcome.htm>) montrent que RNS1 est particulièrement présente dans les fleurs, les siliques et les graines matures. RNS2 est quant à elle exprimée à faible niveau dans tout l'organisme et en grande quantité dans le pollen tandis que RNS3 est très abondante dans les pétales, le pollen et, dans une moindre mesure, dans les graines. Enfin RNS 4 et 5 sont, semble t'il, produites dans les racines et l'hypocotyle (figure 23).

Il a aussi été montré que ces 3 protéines étaient surexprimées lors de la sénescence cellulaire et lors d'un stress nutritionnel lié à une carence en phosphate (Bariola et al., 1994, 1999; Taylor et al., 1993). L'hypothèse est que

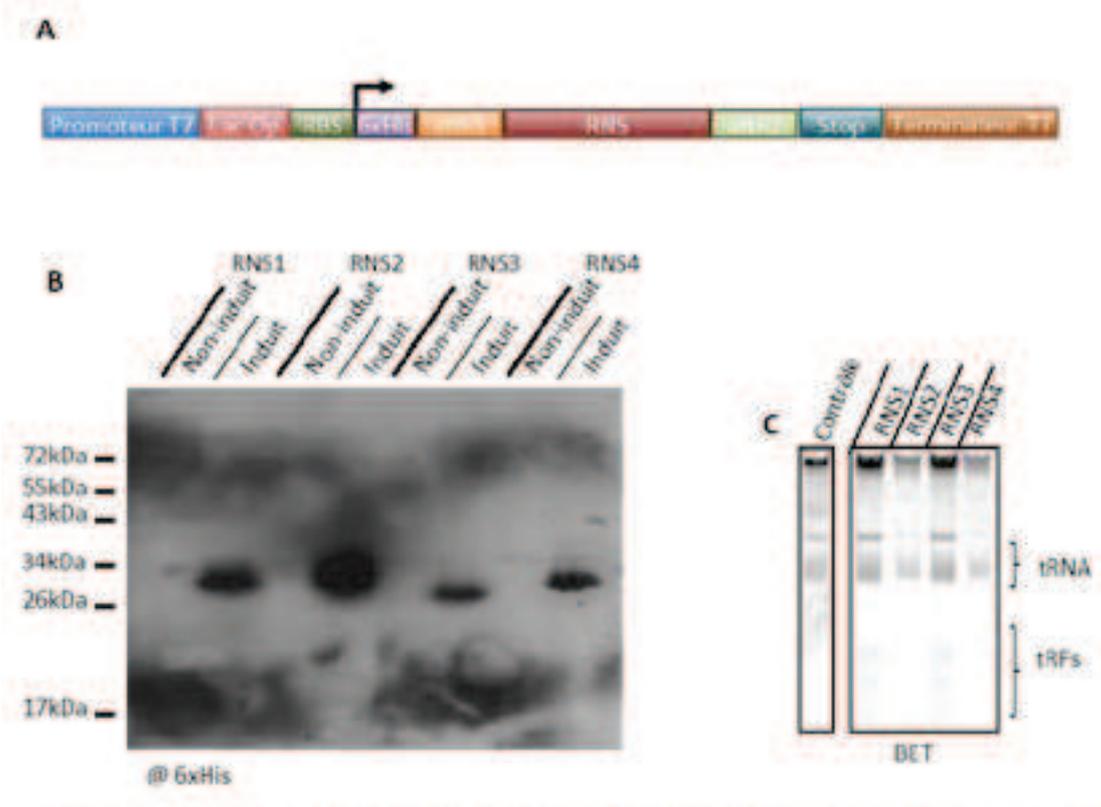
	Racines	Hypocotyle	Feuilles	Bourgeons	Fleurs	Pétales	Pollen	Siliques	Jeunes graines	Graines matures
RNS1	-	-	-	+	++	-	-	++	++	+++
RNS2	+	+	+	+	+	+	+++	+	+	+
RNS3	-	-	-	-	-	+++	++	-	+	++
RNS4	+++	+++	-	-	-	-	-	-	+	-
RNS5	+++	+++	-	-	-	-	-	-	-	-

**Figure 23** : Indications sur l'expression des RNS dans différents tissus selon *Arabidopsis* EFP Browser. Données d'après le site <http://bar.utoronto.ca>. Les tissus n'exprimant que peu ou pas les RNS dans les différents tissus sont symbolisés à l'aide de "-". Les tissus exprimant les RNS sont symbolisés à l'aide de "+", tandis que ceux chez qui les RNases sont abondantes ou très abondantes portent les symboles "++" ou "+++".

ces endonucléases dégradent les ARN stables de la cellule (rRNA, tRNA) afin de relarguer du phosphate pour compenser la carence. D'après les travaux de Hillwig et collaborateurs, il semble que les rRNA soient l'un des substrats de RNS2 (Hillwig et al., 2011). Par ailleurs il a été montré que l'inhibition de RNS1 entraînait une augmentation de la production d'anthocyanine. Or, il est bien connu que la carence en phosphate entraîne une augmentation en anthocyanine chez *Arabidopsis* (Trull et al., 1997). Il a ainsi été proposé que RNS1 et 2 étaient partie prenante du système de secours utilisé lors de ce stress, en servant notamment au recyclage du phosphate à partir d'ARN. Ces taux d'anthocyanine élevés sont ainsi le symptôme de plantes incapables de recycler le phosphate car elles n'ont plus les enzymes requises pour l'extraire des ARN (Bariola et al., 1999). Il a aussi été montré que RNS1 est reliée à la signalisation des blessures via une voie indépendante de l'acide jasmonique (Hillwig et al., 2008).

## 2.5 – Clonage des cDNA codant pour les RNS

Afin de déterminer si parmi les RNS d'*A. thaliana* identifiés par l'approche *in silico* une ou plusieurs de ces protéines sont capables de cliver les tRNA de plantes, nous avons choisi l'approche suivante : cloner les cDNA correspondants, exprimer les protéines chez *E. coli* et tester leur capacité à cliver les transcrits. À partir d'amorces spécifiques (voir Matériel et méthodes – 1.2.1) pour chacune des RNS d'*A. thaliana*, les cDNA codant pour les RNS à l'exception de RNS5, ont été amplifiés par PCR (résultats non montrés). Ceci peut s'expliquer par le fait que le gène de RNS5 (At1g14220) ne possède pas d'EST dans les banques de données d'*A. thaliana* et qu'il succède directement au gène codant pour RNS4 (At1g14210) dans le génome. Ceci suggère que RNS4 et RNS5 sont des paralogues issus d'un événement de duplication. RNS5 pourrait correspondre à un pseudogène, toutefois d'après les données de microarrays disponibles sur BAR, RNS5 serait exprimée préférentiellement dans les racines. En absence d'information supplémentaire et du fait du clonage infructueux du cDNA codant pour RNS5, nous avons pour l'instant concentré nos efforts sur les quatre autres RNS.



**Figure 24 : Production des RNS en système bactérien :** **A**, schéma de la construction permettant la surexpression des RNS dans le vecteur pHGWA sous contrôle du promoteur du phage T7. **B**, les bactéries transformées à l'aide des vecteurs contenant les constructions présentées en A ont ensuite été induites ou non à l'IPTG, puis lysées à l'aide de tampon laemmli pour être chargés sur un gel de polyacrylamide 12%. Le gel est ensuite transféré sur membrane, puis les protéines sont révélées à l'aide d'anticorps dirigé contre le tag 6xHis. **C**, séparation sur gel de polyacrylamide 15% des petits ARN issus de bactéries transformées, ou non (contrôle) avec les RNS puis induites. Coloration au bromure d'éthidium. Les bandes correspondant aux tRNA et aux tRFs sont indiqués.

## 2.6 – Production des RNS chez *E. coli*

Les RNS1-4 ont été clonées par la technologie Gateway dans le vecteur d'expression bactérien pHGWA (figure 24A). Ces plasmides recombinants sont maintenus dans la souche bactérienne TOP10F qui ne permet pas la production de protéines recombinantes, évitant ainsi le problème de leur toxicité. Les plasmides recombinants ont été vérifiés par séquençage puis utilisés pour transformer des bactéries BL21, permettant la production de protéines recombinantes sous contrôle du promoteur T7 après induction à l'IPTG. Après 3 heures d'induction, les protéines provenant des cultures bactériennes ont été fractionnées sur gel de polyacrylamide et après coloration au bleu de Coomassie, seule une très faible bande supplémentaire, à la taille attendue (environ 25 kDa) a été observée pour la construction RNS2 (non montré). Ces mêmes extraits protéiques ont ensuite été analysés par western blot à l'aide d'un anticorps dirigé contre le tag Histidine présent à l'extrémité N-terminal de chacune des RNS produites. Comme le montre la figure 24B, après induction à l'IPTG, les quatre RNS ont pu être détectées. Nous avons donc tenté de les purifier en condition native sur colonne d'affinité au nickel, mais nos résultats ont été infructueux du fait, soit de la quantité insuffisante de RNS exprimée, soit de leur agrégation rendant inaccessible le tag Histidine.

À cause du très faible rendement d'expression, la question de la toxicité des RNS pour les bactéries se posait. En effet, cette toxicité pouvait provenir de plusieurs facteurs non exclusifs. En premier lieu, l'activité endonucléasique de ces RNS sur les ARN bactériens pouvait entraîner leur dégradation et en conséquence l'arrêt de la traduction. De fait, comme le montre la figure 24C, l'expression des RNS, même en faible quantité, entraîne une dégradation importante des tRNA dans les extraits bactériens correspondants. De plus, il faut se rappeler que les RNS possèdent en N-terminal un peptide signal susceptible de s'ancrer dans les membranes bactériennes, pouvant provoquer ainsi une déstabilisation membranaire toxique pour les bactéries. Comme nous le verrons plus loin, des souches bactériennes exprimant des RNS mutantes dépourvues d'activité catalytique sont également toxiques pour les bactéries, suggérant effectivement que le peptide signal soit toxique. Ceci a été confirmé par le fait que des bactéries transformées avec des constructions RNS délétées de leur

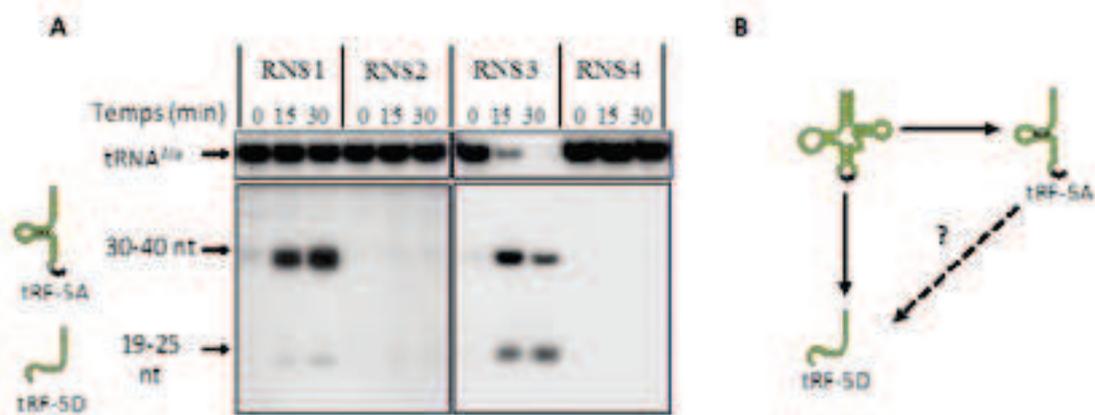
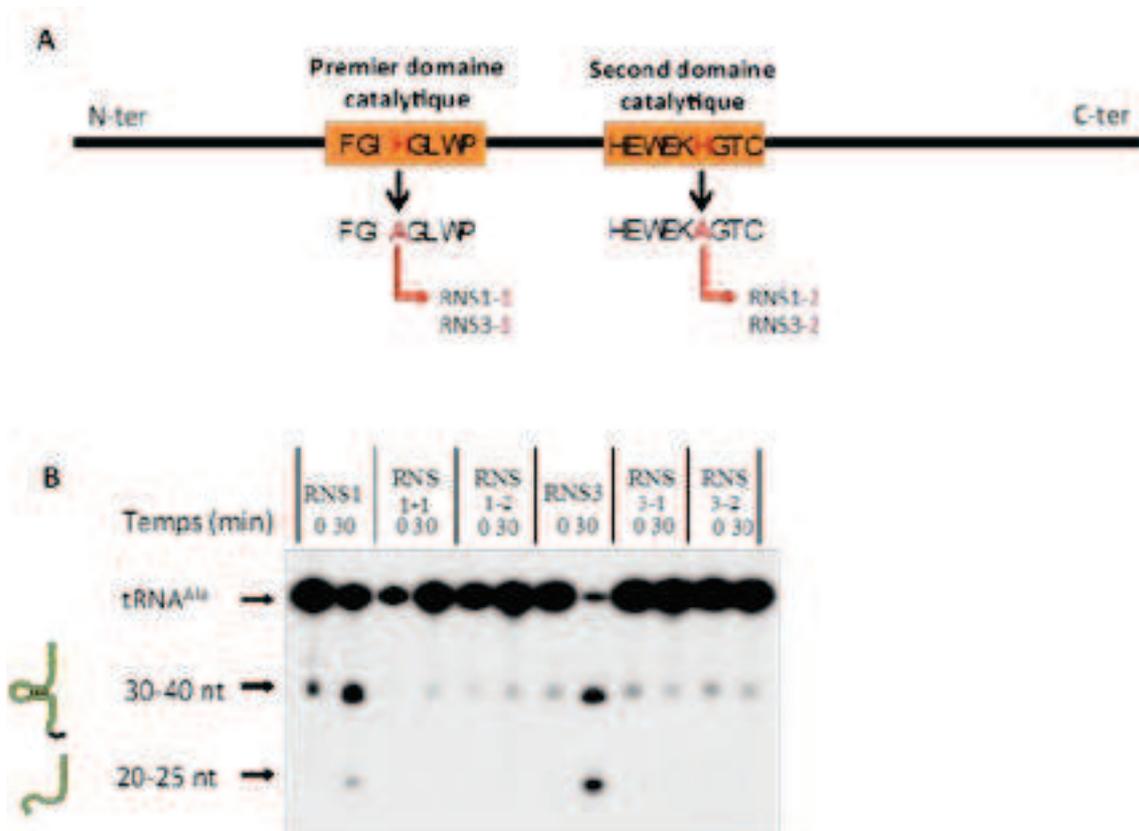


Figure 25 : Test de l'activité des RNS en présence de tRNA synthétisés *in vitro* : **A**, northern blot de tRNA<sup>UGC</sup> transcrits *in vitro* et incubé 0, 15 et 30 minutes avec un extrait de protéine brut issus de bactéries transformées avec RNS 1, 2, 3 et 4. **B**, représentation des fragments générés par le clivage des tRNA grâce aux RNS1 et 3. En pointillé est représenté la possibilité de clivage des TRF-SA pouvant donner des TRF-SD.

peptide signal exprimaient très fortement les endonucléases mais celles-ci avaient tendance à former des corps d'inclusions. Ce travail a été réalisé à Strasbourg durant la période où j'étais en Australie. Il a ainsi été possible de purifier les RNS1 à 4 sur colonne de nickel et en condition dénaturante afin de produire les quantités de protéines requise pour la préparation d'anticorps spécifique contre les RNS (société Covalab).

## 2.7 – Activité de clivage des RNS recombinantes sur des tRNA transcrits *in vitro*

L'expression bactérienne des RNS de plantes ne permettant pas d'envisager leur purification en conditions natives, une solution alternative a été mise en place pour tester leur activité de clivage. Pour ce faire, nous nous sommes demandé s'il était possible de faire des tests de clivage de tRNA avec des extraits bactériens bruts provenant des souches d'*E. coli* BL21 contenant les plasmides recombinants et exprimant, même faiblement, les RNS. Pour réaliser ces tests de clivage, nous avons choisi comme substrat le transcrit du tRNA<sup>Ala</sup> cytosolique de plante. Ce transcrit a été choisi car les tRFs d'une vingtaine de nucléotides correspondant à ce tRNA ont été trouvés en grande quantité dans les extraits d'*A. thaliana* et le clone permettant sa synthèse *in vitro* était disponible au laboratoire (voir Matériel et Méthodes - 1.2.1).



**Figure 26** : Test de l'activité des mutants catalytiques des RNS en présence de tRNA transcrits *in vitro* : **A**, schéma des mutants catalytiques avec en rouge les histidines mutées en alanine dans les deux domaines catalytiques. **B**, tRNA<sup>Ala</sup> synthétisé *in vitro* et incubé en présence d'extraits bruts bactériens surexprimant RNS1 et RNS3 ainsi que chacun de leurs 2 mutants pendant les temps indiqués. Après incubation et extraction phénolique, les échantillons ont été fractionnés sur gel de polyacrylamide 15% dénaturant. Le gel a été transféré sur membrane et hybridé à la sonde tRNA<sup>Ala</sup>-AGC-5'-cyto complémentaire à l'extrémité 5' du tRNA<sup>Ala</sup>.

Des expériences de clivage ont ainsi été faites et, comme le montre la figure 25A, les extraits protéiques bactériens exprimant, les protéines RNS2 et 4 ne sont pas capables de cliver le tRNA<sup>Ala</sup>. RNS2 connu jusqu'à présent pour son rôle dans le clivage et la dégradation des rRNA (Hillwig et al., 2011) ne semble donc pas capable de cliver les tRNA, il en va de même pour RNS4. En revanche, les extraits bactériens exprimant RNS1 et RNS3 génèrent des tRF-5A et des tRF-5D mais dans des proportions différentes (Figure 25A). En effet, RNS1 clive majoritairement le tRNA<sup>Ala</sup> au niveau de la boucle de l'anticodon et très faiblement au niveau de la région D, alors que RNS3 coupe très efficacement tant dans la boucle de l'anticodon que dans la boucle D. De plus, aucun clivage n'a été observé en l'absence d'extrait protéique bactérien n'exprimant pas la protéine. Des résultats similaires ont été obtenus avec un tRNA<sup>His</sup> plastidial transcrit *in vitro* (donnée non-montrée).

En résumé, concernant RNS2 et RNS4, aucune activité de clivage n'est visible, malgré plusieurs essais, mais nous ne pouvons exclure une perte de cette activité lors de leur expression en système bactérien. Par contre, nos résultats montrent que RNS1 et RNS3 sont les deux premières endonucléases de type T2 pouvant cliver des tRNA au niveau de la région D. Elles semblent avoir phylogénétiquement divergé des RNases T2 capables de cliver les tRNA uniquement au niveau de l'anticodon chez *S. cerevisiae* et *T. thermophila*, enzymes qui semblent plus proches de RNS2 d'après l'arbre phylogénétique (figure 22). Par ailleurs, nous ne savons pas actuellement si les tRF-5D sont générés directement à partir du transcrit mature et/ou s'ils sont générés après clivage dans la boucle de l'anticodon (figure 25B). Des expériences complémentaires à partir de transcrits de tRNA tronqués permettront de répondre à cette question. Enfin les clivages obtenus *in vitro* au niveau de la région de l'anticodon ou de la boucle D corréleront également avec les données *in vivo* provenant de l'analyse de banques de petits ARN et des expériences de Northern blots (présentées en première partie des résultats). L'ensemble de ces

Gènes	Nombre d'accèsion	Logiciels						Consensus
		AtaBoost	BaCella	Flart-mPoc	Protostar v1.03	WolF PSORT	Yolo	
RR51	ADG02990.1	Extracellulaire	Extracellulaire, ER, Golgi	Extracellulaire	Extracellulaire, ER, Golgi	Extracellulaire	Extracellulaire	Extracellulaire
RR52	ADG10790.1	Extracellulaire	Extracellulaire, ER, Golgi	Extracellulaire	Extracellulaire, ER, Golgi	Extracellulaire	Extracellulaire	Extracellulaire
RR53	AR029820.1	Extracellulaire	Extracellulaire, ER, Golgi	Extracellulaire	Extracellulaire, ER, Golgi	Cytosole	Extracellulaire	Extracellulaire
RR54	ADG14210.1	Extracellulaire	Extracellulaire, ER, Golgi	Extracellulaire	Extracellulaire, ER, Golgi	Vacuole	Extracellulaire	Extracellulaire
RR55	AR014220.1	Extracellulaire	Extracellulaire, ER, Golgi	Extracellulaire	Extracellulaire, ER, Golgi	Membr.	Extracellulaire	Extracellulaire

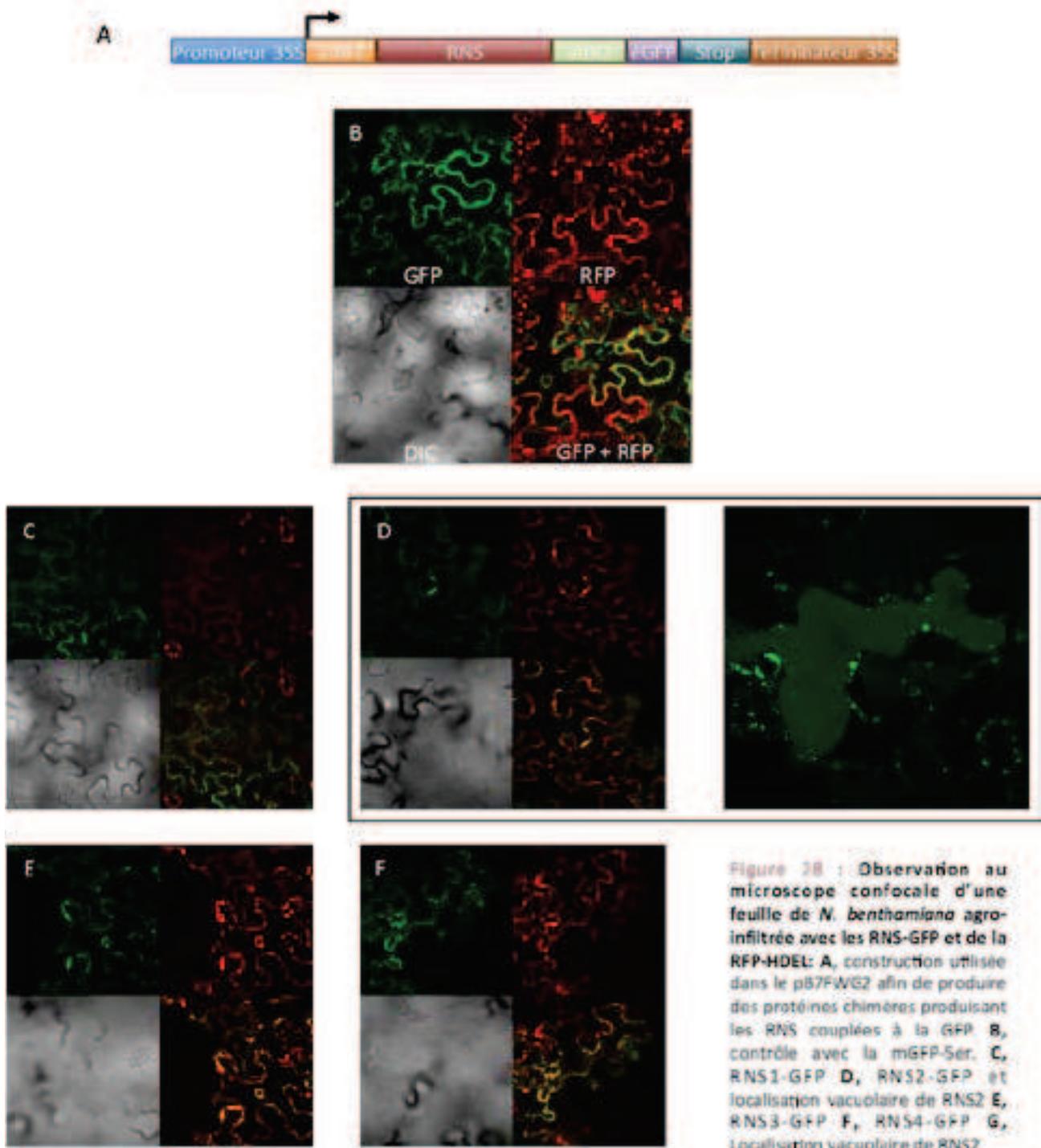
Figure 27 : Prédiction de localisation des RNS à l'aide de divers logiciels. Le consensus de la localisation a été déterminé par le logiciel SUBA 3 (<http://suba.plantenergy.uwa.edu.au>).

résultats suggère fortement que RNS1 et RNS3 sont les endonucléases responsables du clivage des tRNA chez *A. thaliana*.

## 2.8 – Mutagenèse dirigée des RNS et activité de clivage des protéines recombinantes mutantes

Afin de valider les résultats présentés ci-dessus et ainsi exclure la possibilité que les activités endonucléasiques observées ne proviennent d'enzymes bactériennes, les sites catalytiques de RNS1 et RNS3 ont été mutés (figure 26A). Pour ce faire, les résidus histidine contenus dans les deux régions catalytiques ont été remplacés par des résidus alanine. Ces résidus se trouvent respectivement pour RNS1 en position 65 et 123, et en position 57 et 115 pour RNS3. Quatre constructions ont été réalisées : RNS1-1, RNS1-2, RNS3-1 et RNS3-2 (figure 26A). Comme précédemment, les extraits bruts bactériens contenant les protéines recombinantes mutées ont alors été testés pour leur capacité à cliver le transcrit de tRNA<sup>Ala</sup>. Comme le montre la figure 26B, alors que les extraits bactériens produisant les endonucléases RNS1 ou RNS3 sauvages permettent le clivage du transcrit de tRNA au niveau de l'anticodon et de la région D, aucune activité n'est détectée à partir des extraits bactériens exprimant les formes mutées des RNS1 ou RNS3 (figure 26B), démontrant ainsi que l'activité endonucléasique provient vraisemblablement de l'expression des RNS1 ou RNS3. Il faut noter cependant la présence, dans l'ensemble des extraits, d'une faible quantité de tRF-5A dans les différents échantillons analysés. Cette présence de tRF-5A semble être le reflet d'une coupure mécanique lors de l'extraction, la région de l'anticodon étant la région du tRNA la plus accessible au clivage.

Ces résultats confortent les précédents, à savoir que les RNS1 et RNS3 d'*A. thaliana* sont capables de générer tant des tRF-5A que des tRF-5D et que les résidus histidine des deux régions catalytiques sont des résidus primordiaux pour l'activité de clivage.



## 2.9 – Localisation *in vivo* des RNS

Afin de comprendre à la fois la biogenèse et les fonctions des tRF, il est important de déterminer leur localisation *in planta* ainsi que celle des RNases qui les génèrent. Peu de données sur la localisation des tRFs existent. Si les tRFs ont été localisés dans des granules de stress dans le cytoplasme chez le protozoaire *T. cruzi* (Reifur et al., 2012), aucune autre donnée n'est disponible quant à leur localisation, notamment chez les plantes. Pourtant, connaître la localisation des tRFs est un élément clé dans la compréhension de leurs potentielles fonctions. En raison de la paroi végétale il est par conséquent très difficile d'accéder à la cellule afin de marquer les tRFs en utilisant la technique de FISH. Nos efforts se sont donc portés dans un premier temps à la localisation des RNS.

Comme pour les tRFs peu de données sont disponibles concernant la localisation des RNS. Les seules données dont nous disposons nous sont fournies par des sites de prédiction de localisation (figure K). Le consensus de localisation est donné par le logiciel disponible sur le site (<http://suba.plantenergy.uwa.edu.au>) (figure K, colonne Consensus). Ainsi, la majorité de ces programmes donne un peptide signal en N-terminal, ainsi qu'un signal d'excrétion en C-terminal vers la vacuole. Ceci suggère que ces protéines sont vouées au système sécrétoire (figure 27, colonne logiciel). Cette prédiction des RNS dans ce compartiment est en corrélation avec les résultats obtenus pour Rny1. En effet, chez *S. cerevisiae* la suppression de Rny1 entraîne des modifications de morphologie de la vacuole, suggérant la localisation vacuolaire de Rny1 (Luhtala and Parker, 2012).

Afin de déterminer la localisation intracellulaire des RNS, nous les avons fusionné à la GFP. La GFP a été placée du côté C-terminal (figure 28A) afin de ne pas masquer le peptide signal. Ces constructions ont alors été exprimées de façon transitoire dans des feuilles de *N. benthamiana*. Ces mêmes feuilles ont également été co-transformées avec un second plasmide codant pour la RFP fusionnée à un signal HDEL de localisation au réticulum endoplasmique (RE). La construction mGFP5-RE a été utilisée comme contrôle positif de localisation au RE.

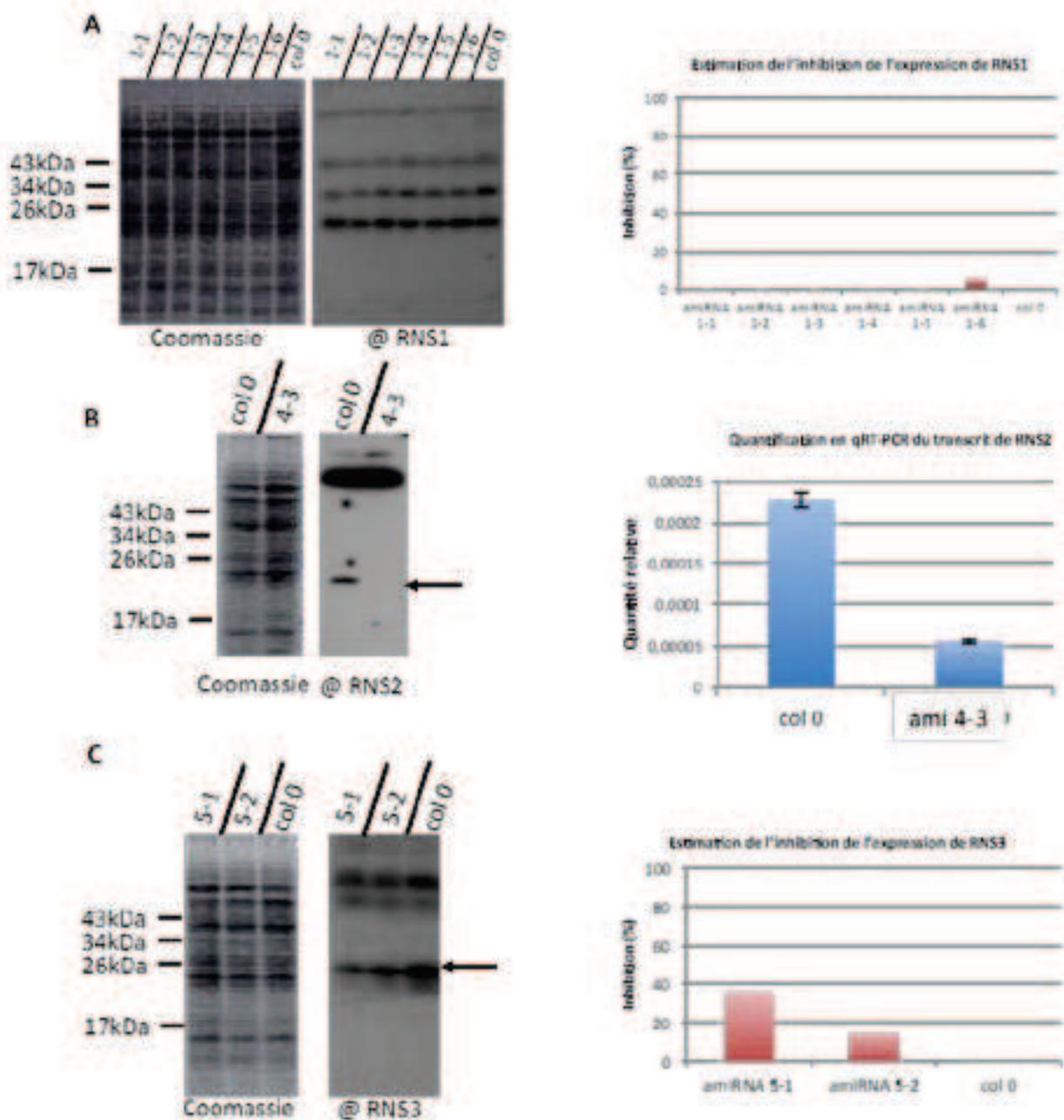


Figure 29 : Analyse des lignées "Knock-Down" (KD) pour RNS 1, 2 et 3 en western blot, ou pour RNS2 en qRT-PCR : **A**, western blot et estimation de l'inhibition de production de RNS1. Ce western blot est effectué à l'aide d'extrait de protéines de fleurs d'*A. thaliana* issus des lignées transgéniques ami 1-1 à ami 1-6, inhibant la production de RNS1, puis révéillé à l'aide d'un anticorps spécifique pour RNS1. **B**, même test que précédemment avec la lignée ami4-3 inhibant la production de RNS2 à l'aide d'un anticorps spécifique de RNS2. Puis cette lignée a été testée en qRT-PCR à l'aide d'amorces spécifiques pour RNS2 sur des ARN extraits de feuilles d'*A. thaliana*. **C**, même test qu'en A mais pour des plantes inhibant la production de RNS3 (ami 5-1 et 5-3), et estimation de l'inhibition de production de RNS3.

Deux jours après co-infiltration des deux constructions, les feuilles sont observées au microscope confocal (figure 28 A, B, C, D, E, F). Il y est alors clairement visible que la mGFP5-RE ainsi que les RNS-GFP co-localisent exactement avec la RFP-HDEL, prouvant que toutes ces protéines se trouvent dans le RE. Très occasionnellement, nous avons aussi pu observer que RNS2 était localisée dans la vacuole (figure 28D). Bien que l'état physiologique de la cellule indique que celle-ci subit de gros stress probablement consécutifs à l'infiltration d'agrobactéries, ou de la trop forte expression des protéines (*e.g* RE rétracté et concentré en de petits points). Ceci suppose malgré tout, une localisation vacuolaire de RNS2 durant des stress.

## 2.10 – Mutants KO ou lignées inhibant la production des RNS d'*A. thaliana*

Afin d'apporter la preuve *in vivo* que les RNS sont bien les enzymes responsables du clivage des tRNA chez *A. thaliana*, nous avons cherché à obtenir sur le site NASC des lignées d'insertions pour chacune de ces RNS. Pour RNS1 et 3 nous disposons des lignées d'insertion FLAG\_566A08 et FLAG\_164A04 dont le T-DNA est inséré dans un exon. Nous disposons aussi de 2 autres lignées potentiellement KO pour RNS4 (SALK\_055304) et RNS5 (SALK\_053974) ainsi que d'une dernière lignée d'insertion KO pour RNS2 : SALK\_069588. Les lignées KO pour RNS1 et 3 ne sont pas homozygotes, tandis que toutes les autres le sont. Cependant, seule la lignée SALK\_069588 a été testée (Hillwig et al., 2011). Il nous faudra donc préparer les lignées homozygotes pour RNS1 et 3 tout en s'attendant à ce que ces lignées soient létales, ou a croissance retardée au vue des fonctions primordiales que semblent occuper les RNS. C'est pourquoi nous avons préparé plusieurs constructions de amiRNA (artificial miRNA) permettant la sous-expression des RNS (cf. matériel et méthode – 2.1.5). Huit constructions ont été préparées, 2 par RNS, numérotées de 1 à 8 : ami 1 et 2 pour RNS1, 3 et 4 pour RNS2, 5 et 6 pour RNS3 et enfin 7 et 8 pour RNS4. Après vérification, celles-ci ont été insérées dans des vecteurs binaires d'expressions *in planta* et nous avons effectué des transformations stables par "floral-dip". Puis nous avons sélectionné et génotypé 3 fois successivement les différentes lignées de plantes afin d'en faire des homozygotes. À l'heure d'aujourd'hui ces plantes sont en cours



d'analyse, soit par western blot afin de détecter les protéines matures, soit par qRT-PCR afin de détecter les transcrits de RNS. Nous avons pu confirmer l'effet, par western blot et qRT-PCR que de trois lignées (figure 29), une sous-exprimant RNS2 (en western blot et qRT-PCR) quasiment totalement (75%) et 2 sous exprimant RNS3 (47,5% environ, western blot seulement, figure 29). Si la lignée RNS2 est la seule à présenter un phénotype moléculaire c'est probablement parce que les tests en qRT-PCR n'ont été effectués que sur des cDNA préparés à partir d'ARN totaux de feuilles, or RNS2 y est plus exprimée que les autres RNS (figure 23). En ce qui concerne les western blots de protéines extraites de fleurs sur d'autres lignées d'amiRNA (RNS1 et 4), ceux-ci n'ont pas donné de résultats interprétables (données non montrées) car les RNS étaient indétectables tant dans le contrôle, que dans les plantes transgéniques. Ceci est probablement dû au niveau d'expression faible, et donc indétectable en western blot, des RNS durant des conditions physiologique normale. C'est pourquoi afin de connaître le réel potentiel de ces lignées nous avons pour but de tester les différentes lignées en qRT-PCR, qui est plus sensible que le western blot, à l'aide de cDNA préparés à partir de fleurs.

## 3 – Les tRFs et le séquençage haut débit

### 3.1 – RNase T2, séquençage haut débit

La technique de séquençage Illumina, basée sur la méthode de Sanger (*cf.* Matériel et méthodes – 2.2.7), prévoit le clonage des fragments d'ARN entre 2 oligonucléotides, ou adaptateurs, afin de les séquencer. Cependant, cette technique requiert que chacun des fragments possèdent une extrémité 5'-P et 3'-OH afin d'y cloner les adaptateurs. Or comme précisé en 2.3 les RNases T2 sont des endonucléases qui clivent les ARN simples brins en libérant une extrémité 5'-OH et une autre 3'-P. Les fragments issus de ce clivage ne sont donc pas adaptés à la technique de séquençage Illumina.



### 3.1.1 – Rétablissement des extrémités 5' et 3' normales

La présence de ces extrémités inhabituelles de part et d'autre des tRFs nécessite d'adapter le protocole de clonage utilisé dans les kits pour séquençage Illumina. L'option la plus simple que nous avons choisi est de rétablir les extrémités 5'-P et 3'-OH de part et d'autre des tRFs. Deux activités sont nécessaires, une activité kinase afin de phosphoryler l'extrémité 5'-OH, et une activité phosphatase qui permet la déphosphorylation de l'extrémité 3'-P. Pour limiter les pertes de matériel entre chaque étape, ainsi que les dégradations inhérentes à la fragilité des ARN, nous avons cherché à effectuer ces 2 étapes en même temps. C'est pourquoi nous avons choisi d'utiliser la T4 polynucléotides kinase. Car cette enzyme, connue pour sa faculté de phosphoryler l'extrémité 5'-OH d'un ARN, est aussi capable d'en déphosphoryler l'extrémité 3'-P (Cameron and Uhlenbeck, 1977). Ainsi, après traitement, une fois les tRFs clonés entre les adaptateurs, nous les avons séquencés.

Deux banques de petits ARN de feuilles d'*A. thaliana* Col-0 ont été créées : une banque d'ARN non-traitée à la T4 polynucléotide kinase, ainsi qu'une autre qui l'est. Le séquençage de ces 2 banques a permis d'identifier 308 914 et 3 435 136 reads apparentés aux tRFs dans les banques non-traitée et traitée. Après normalisation par le nombre de reads totaux de chaque banque (2 862 356 pour la banque non-traitée et 7 560 686 pour la banque traitée) nous obtenons 107 922 rpm (reads par millions) pour la banque non-traitée et 454 341 rpm pour la banque traitée. Ce qui suggère un enrichissement en tRFs dans cette banque traitée lorsqu'on la compare à la banque non-traitée.

### 3.1.2 – Comparaison de la banque traitée avec la T4 polynucléotides kinase avec la banque non traitée

#### 3.1.2.1 – Comparaison de l'origine des tRFs avec, ou sans traitement

L'observation de la provenance de l'ensemble des tRFs (figures 30, 31 et 32) montre qu'avec ou sans traitement à la T4 polynucléotides kinase, les mêmes tRFs sont présents dans chacune de ces banques, bien que leurs quantités

## tRFs mitochondriaux

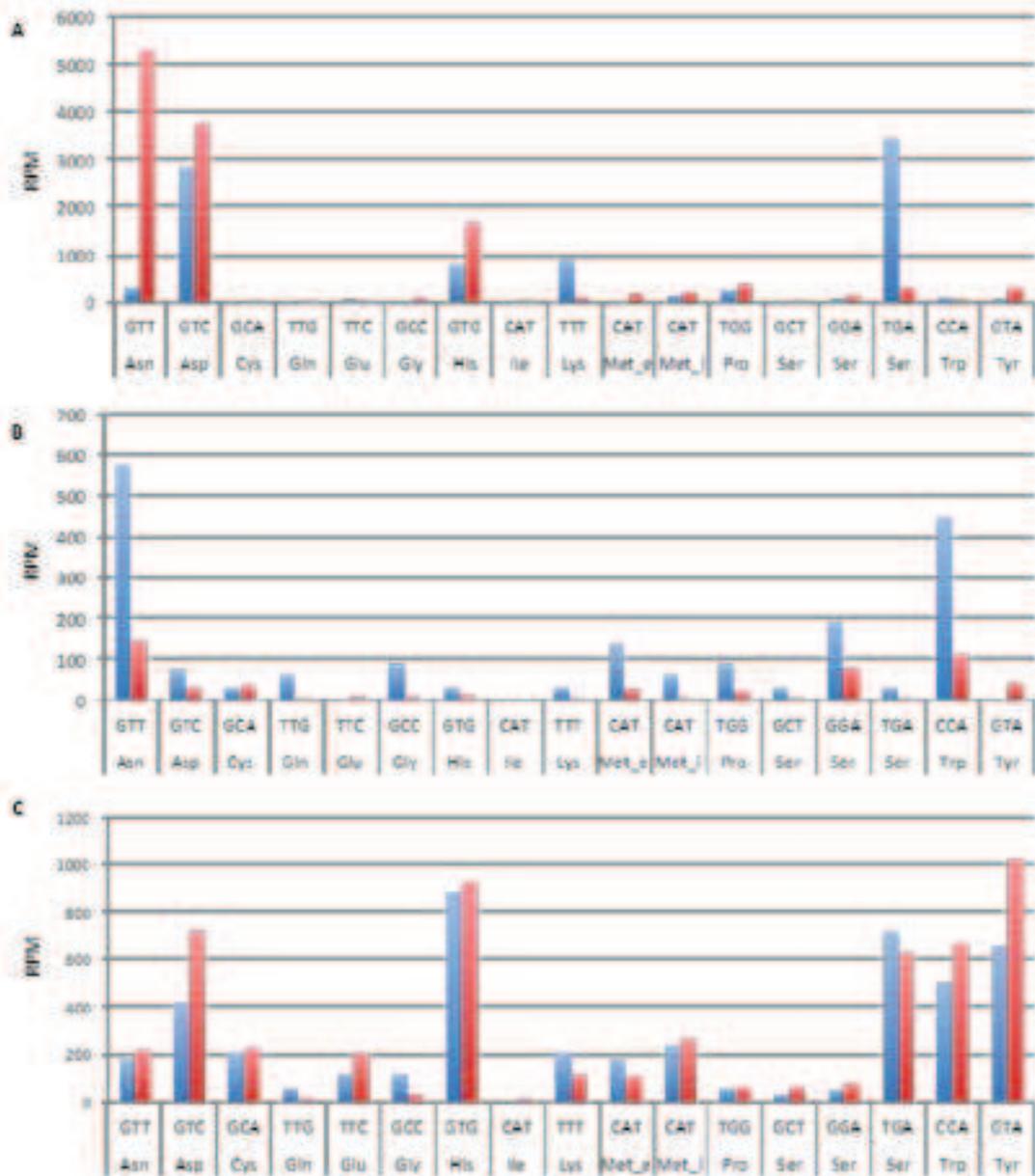


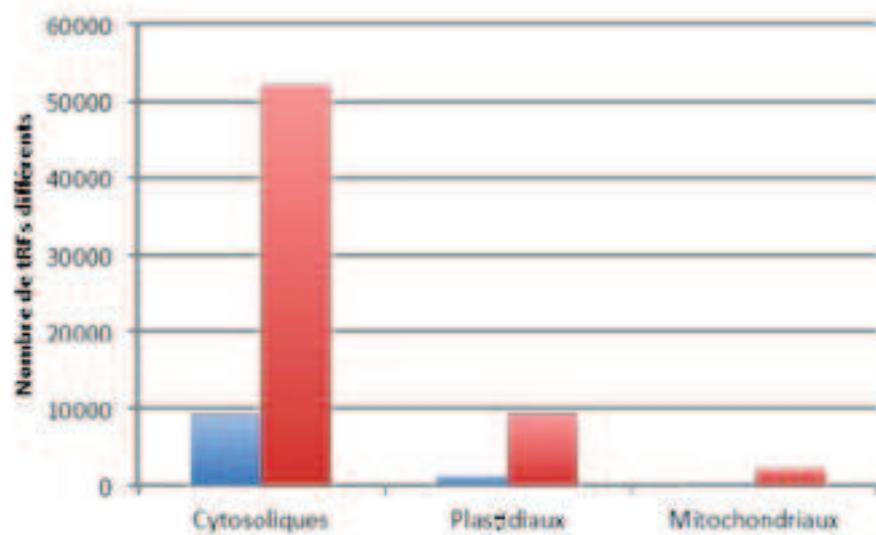
Figure 32 : Quantité en reads par million (rpm) en tRFs mitochondriaux dans les banques non traitées (bleu) et traitées (rouge) selon les différents isodécodons : Les quantités sont affichées pour chaque tRFs en fonction des isodécodons dont ils proviennent. A, tRF-5 D, T et A; B, tRF-3 D, T et A; C, tRF n'appartenant à aucune des classes précédentes.

varient. On remarque aussi qu'avec le traitement, la quantité en tRF-5 augmente que ce soit pour les tRF-5 cytosoliques (figure 30A), plastidiaux (figure 31A), ou mitochondriaux (figure 32A). À l'inverse, la quantité en tRF-3 nucléaire, plastidiale ou mitochondriale diminue avec le traitement (figure 30B, 31B et 32B). Enfin concernant les autres tRFs, aucune dynamique n'est clairement identifiable bien que dans le cytosol le traitement puisse multiplier jusqu'à 9 fois leur quantité (figure 30C, tRF-Gly TCC). Dans les plastides, la quantité du tRF-Leu surpasse de beaucoup les autres (figure 31C, 300000 rpm) et il semble qu'elle diminue drastiquement avec le traitement (de 300000 à 13200 rpm). Enfin les autres tRFs originaires de la mitochondrie restent stable peu importe le traitement (figure 32C).

Notre hypothèse de travail suppose que la régénération d'extrémité 5'-P et 3'-OH à l'extrémité de ces fragments augmente la quantité de tous les types de tRFs qui peuvent donc ainsi être clonés et séquencés. Or ce n'est pas ce qui est observé, car les tRF-3 diminuent après traitement des banques tandis que les autres tRFs sont parfois plus abondants (cytosol, figure 30C), d'autres fois moins abondants (chloroplaste figure 31C), ou restent stable (mitochondrie, figure 32C).

Une première hypothèse pouvant expliquer cette incohérence concerne le nombre limité de fragment qu'il est possible de cloner sur une même puce. En effet, puisque la quantité en tRF-5 et en tRFs cytosoliques de tous types croît d'environ 25 à 93% (figure 30A et 31C, 30A et 32A) dans les banques traitées, il est possible que cette abondance de fragments et d'autres non apparentés aux tRNA mais clivés par les RNS (*i.e* rRNA) viennent réduire l'espace disponible sur la puce aux autres ARN qui n'ont pas été autant amplifiés durant les étapes de préparation des banques. Ceci est illustré par l'augmentation de la diversité en tRFs dans les banques traitées qui y est jusqu'à 8 fois supérieure que dans les banques contrôle (figure 33). Ainsi le traitement permettrait une augmentation de la diversité en différentes espèces de tRFs dans les banques.

Une autre hypothèse, pouvant expliquer cette diminution en tRF-3 après traitement, porte sur le caractère réversible de la polynucléotides kinase qui possède une activité phosphorylase et kinase. Ainsi, si on suppose que les extrémités 5'-OH issues du clivage par les RNS sont converties en 5'-P par la



**Figure 33 : Nombre de tRFs différents identifiés appartenant aux banques non traitées (bleu) et traitées (rouge) ; Les quantités sont affichées pour chaque compartiment.**

cellule, en amont de notre action, alors le traitement à l'aide de la kinase peut potentiellement endommager ces extrémités en ôtant de nouveau le phosphate et donc en empêchant le clonage de ces tRF-3.

Si une telle hypothèse est à prendre en considération, plusieurs points restent obscurs. Tout d'abord, pourquoi la kinase n'aurait d'effet délétère que sur les extrémités 5' des tRF-3 et pas sur l'extrémité 5' des tRF-5 ? Ensuite, aucune information n'est disponible quant aux possibles moyens que la cellule utiliserait pour convertir l'extrémité 5'-OH des tRF-3 en 5'-P. On peut donc supposer l'existence de kinase ayant cette activité spécifique, ou d'exonucléases qui clivent le premier nucléotide du fragment en laissant une extrémité 5'-P. Cette seconde supposition prend plus de sens si l'on rappelle que la population en tRFs augmente lors de carences en phosphate (Hsieh et al., 2009) tandis que les RNS1 et 3 sont surexprimées (Bariola et al., 1994). Ainsi, si le but des tRFs est en partie d'être recyclés pour en faire du phosphate, il est possible que cette extrémité 5'-P soit le signe avant coureur de la future dégradation du tRFs. De plus, s'il s'avère que de tels exonucléases existent, le retour d'extrémités "classiques" pourrait expliquer comment les tRFs ont été clonés dans d'autres banques par le passé ainsi que ceux clonés dans notre banque contrôle.

### 3.1.2.2 – Comparaison des positions de clivage des tRFs avec, ou sans traitement

Si l'on compare l'aspect général des positions auxquelles les tRFs débutent (figure 34A, 35A et 36A) ou les positions auxquelles ils se terminent (figure 34B, 35B, 36B) on observe que celle-ci sont quasiment similaires que les fragments soient traités avec la T4 polynucléotides kinase, ou non.

De plus, à l'exception des positions "start" 7 et 8 (figure 35A) et "stop" 21 plastidiales (figure 35B), ainsi que la position "stop" 23 mitochondriale (figure 36B) qui diminuent fortement après le traitement, la majeure partie des positions semble augmenter dans des proportions très diverses allant de 4 à 98%. Ces augmentations sont peut-être le reflet des effets du traitement T4 polynucléotides kinase.

D'autres part, lorsque l'on compare les positions auxquelles les tRFs débutent avec, ou sans traitement, une très grande majorité d'entre

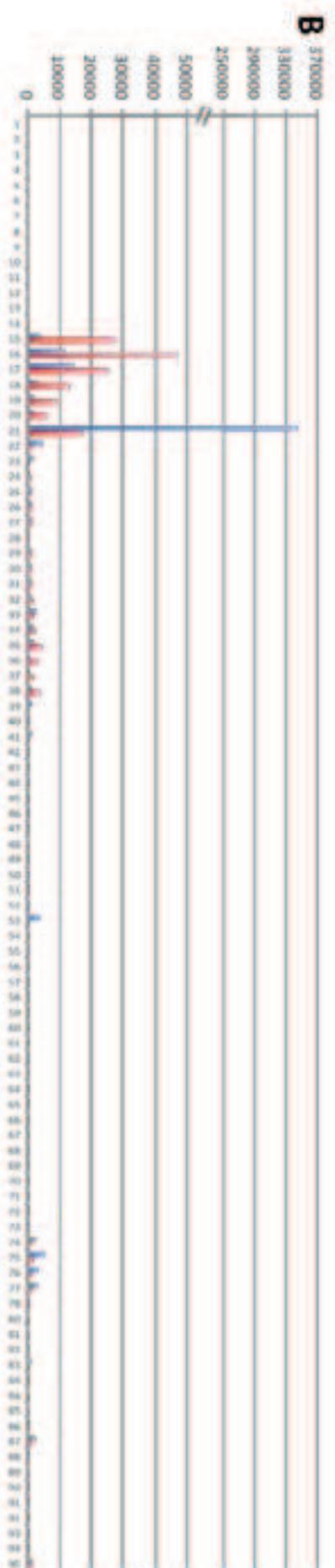
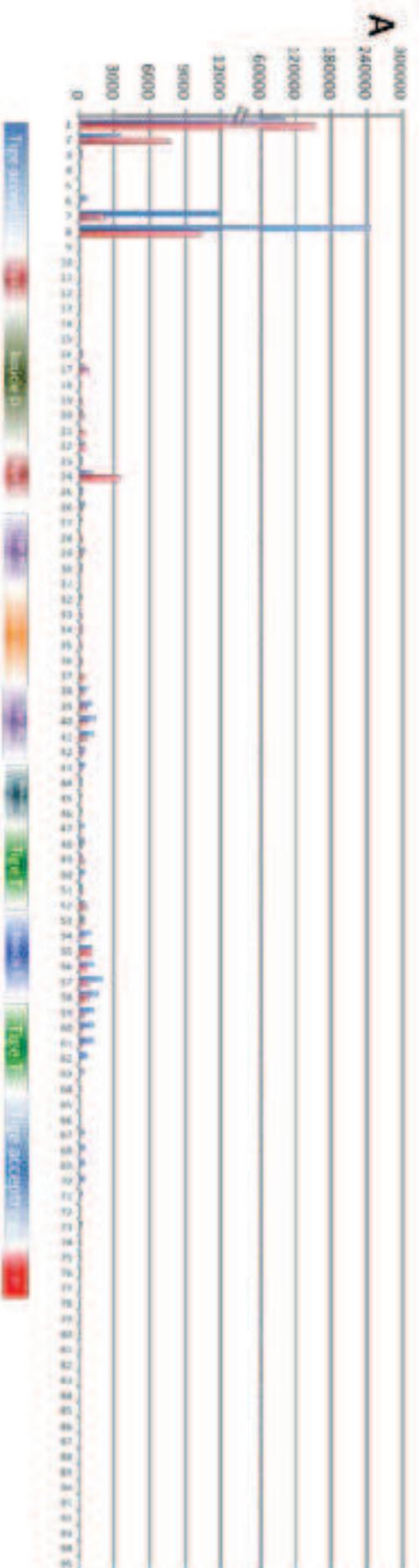


eux débutent en position 1, ou dans les 3 nucléotides en aval (figure 34A, 35A et 36A). D'une façon similaire, pour les clivages internes de la séquence des tRNA il n'existe pas de site de clivage franc et net sur les figures 34, 35 et 36. Ainsi la totalité des boucles D et de l'anticodon sont les cibles privilégiées des RNS. Cependant lorsque l'on s'attarde sur les principaux sites de clivages des tRFs majoritaires, (figures 37 et 38) on remarque qu'avant traitement il n'y a que très peu de sites de clivages, et bien plus après. Ces profils de clivage s'expliquent de diverses façons. Tout d'abord, pour les figures 34, 35 et 36 il faut rappeler que celles-ci représentent une moyenne, une somme de tRNA de tailles diverses dont la structure peut varier, et donc dont les positions de clivages sont légèrement différentes. Ceci peut expliquer l'aspect en pyramide des histogrammes au niveau des boucles D et de l'anticodon. Ensuite, bien qu'il ait été montré, ou qu'il soit suspecté, que certains tRFs ont un rôle dans la cellule, ils ont aussi été impliqués dans des processus de carences en phosphate (Hsieh et al., 2009). Or l'une des sources en phosphate dans la cellule sont les ARN, et il se pourraient donc que les tRFs soient des fragments dont le devenir seraient d'être recyclés afin d'en extraire le phosphate. Ce processus de dégradation peut ainsi faire appel à des exonucléases qui débutent la dégradation des tRFs. Le produit de dégradation de ces tRFs tronqués de seulement quelques nucléotides à leurs extrémités pourrait expliquer les raisons de ce profil de clivage particulier.

Concernant les figures 37 et 38, il est possible que le traitement endommage les fragments ou les tRNA en apportant de nouveaux sites de clivage indépendant des RNS. Nous sommes donc dans l'incapacité d'affirmer que ces sites sont dus à des fragments qui n'étaient pas clonés avant le traitement ou à des fragments générés par le traitement.

### 3.1.2.3 – Comparaison des tailles des différents tRFs avec, ou sans traitement

Lorsque l'on observe le profil des graphiques présentés en figures 39, 40 et 41 on remarque 2 tailles principales de tRFs, les tRFs de 15 à 26 nt qui correspondent à des clivages dans la boucle D et des fragments allant de 34 à 39 nucléotides ce qui correspond à des sites de clivage dans la boucle de l'anticodon. Ces résultats sont d'ailleurs confirmés par l'observation des positions "start" et

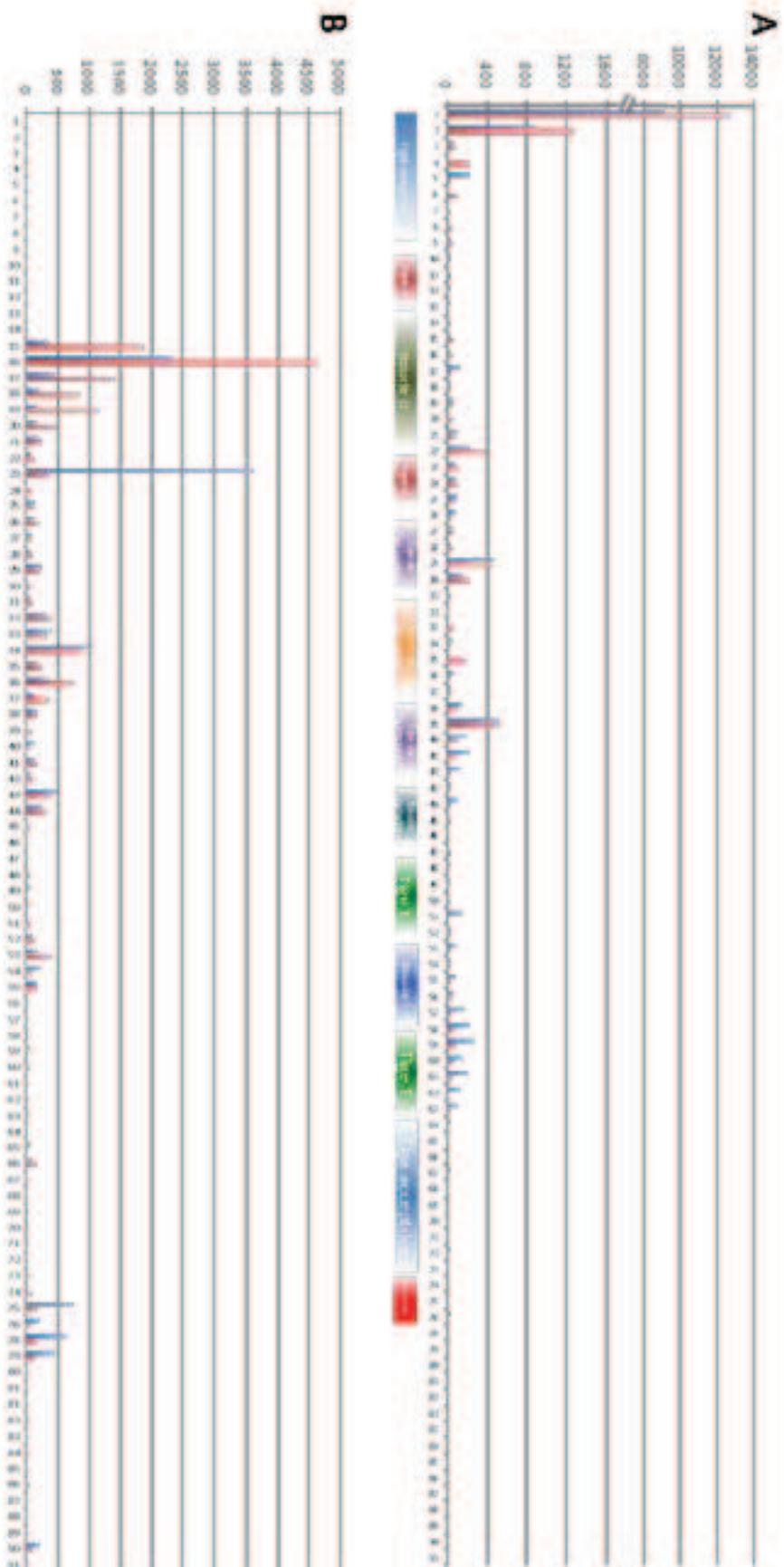


**Figure 35 : Quantité en reads par million (rpm) en tRFs plastidiaux dans les banques non traitées (bleu) et traitées (rouge) selon leurs extrémités :** Les quantités sont affichées en fonction de : **A**, leur première position (start); **B**, leur dernière position (stop). Entre les 2 histogrammes est représenté la structure générique d'un tRNA avec ses différents domaines. Cette structure générique est d'environ 75 nucléotides de long. Les reads se trouvant au delà de la position 77 correspondent souvent à des tRF-3 dérivant de tRNA plus grands à cause d'un domaine variable plus long (ici 95 nucléotides).

“stop“ des différents tRFs (figures 34, 35 et 36). Là encore la diversité des tailles de tRFs est la même quelle que soit la banque observée. Cependant on dénote un enrichissement en tRF-5 d'environ 35 à 85% dans les banques traitées qui est accompagné d'une diminution de quelques tailles tRFs qui étaient majoritaire avant le traitement (*e.g* tRFs plastidiaux de 21 nucléotides de long et mitochondriaux de 23nt de long (figure 40A et 41A). Concernant les tRF-3 la dynamique est très nettement inversée par rapport aux tRF-5, et tous diminuent après traitement (figure 39B, 40B et 41B). Là encore les causes d'un tel changement sont difficiles à expliquer et sont probablement les mêmes que celles énoncées plus haut.

### 3.1.3 – Clivage des tRFs dans l'anticodon dans la banque non traitée

Ces banques ont aussi pour avantage de couvrir tous les tRFs d'une taille inférieures à 50 nucléotides. Par conséquent, nous avons pu observer les fragments consécutifs d'un clivage dans la boucle de l'anticodon (figure 42). Ces tRF-5A et tRF-3A sont relativement nombreux et ils représentent en tout 17% de tous les tRFs présents dans la banque d'ARN non traités. Parmi ces tRFs 84% de ces fragments sont des tRF-5A cytosoliques, 1,3% de ces fragments sont des tRF-3A cytosoliques, 12% sont des tRF-5A chloroplastiques et enfin 2,7% sont des tRF-3A chloroplastiques. Si l'on s'intéresse maintenant au profil de clivage dans la tige-boucle de l'anticodon des principaux tRF-5 (figure 42) on remarque que les tRF-5A cytosoliques sont toujours le produit d'un clivage des tRNA dans la boucle de l'anticodon, en revanche les tRF-5A chloroplastiques sont souvent issus de clivages dans la tige de l'anticodon.



**Figure 36 : Quantité en reads par million (rpm) en tRFs mitochondriaux dans les banques non traitées (bleu) et traitées (rouge) selon leurs extrémités :** Les quantités sont affichées en fonction de : **A**, leur première position (Start); **B**, leur dernière position (Stop). Entre les 2 histogrammes est représentée la structure générique d'un tRNA avec ses différents domaines. Cette structure générique est d'environ 75 nucléotides de long. Les reads se trouvant au delà de la position 79 correspondent souvent à des tRF-3 dérivant de tRNA plus grands à cause d'un domaine variable plus long (ici 91 nucléotides).



Figure 37 : Représentation des tRFs cytosoliques majoritaires dans les banques non traitées et traitées : **A** et **B**. Chaque tableau liste au moins les 30 premiers pourcents en tRFs des compartiments cytosoliques et nucléaires pour les banques non-traitée et traitée. En noir, figure la séquence du tRNA source et en rouge est représentée la séquence du tRFs. \* : tRFs couvrant la même séquence mais non identiques entre eux, à cause d'un mismatch **C** et **D**, représentation des sites de clivage pour les tRFs présentés en **A** et **B** dans la banque non traitée **C** ou traitée **D**.

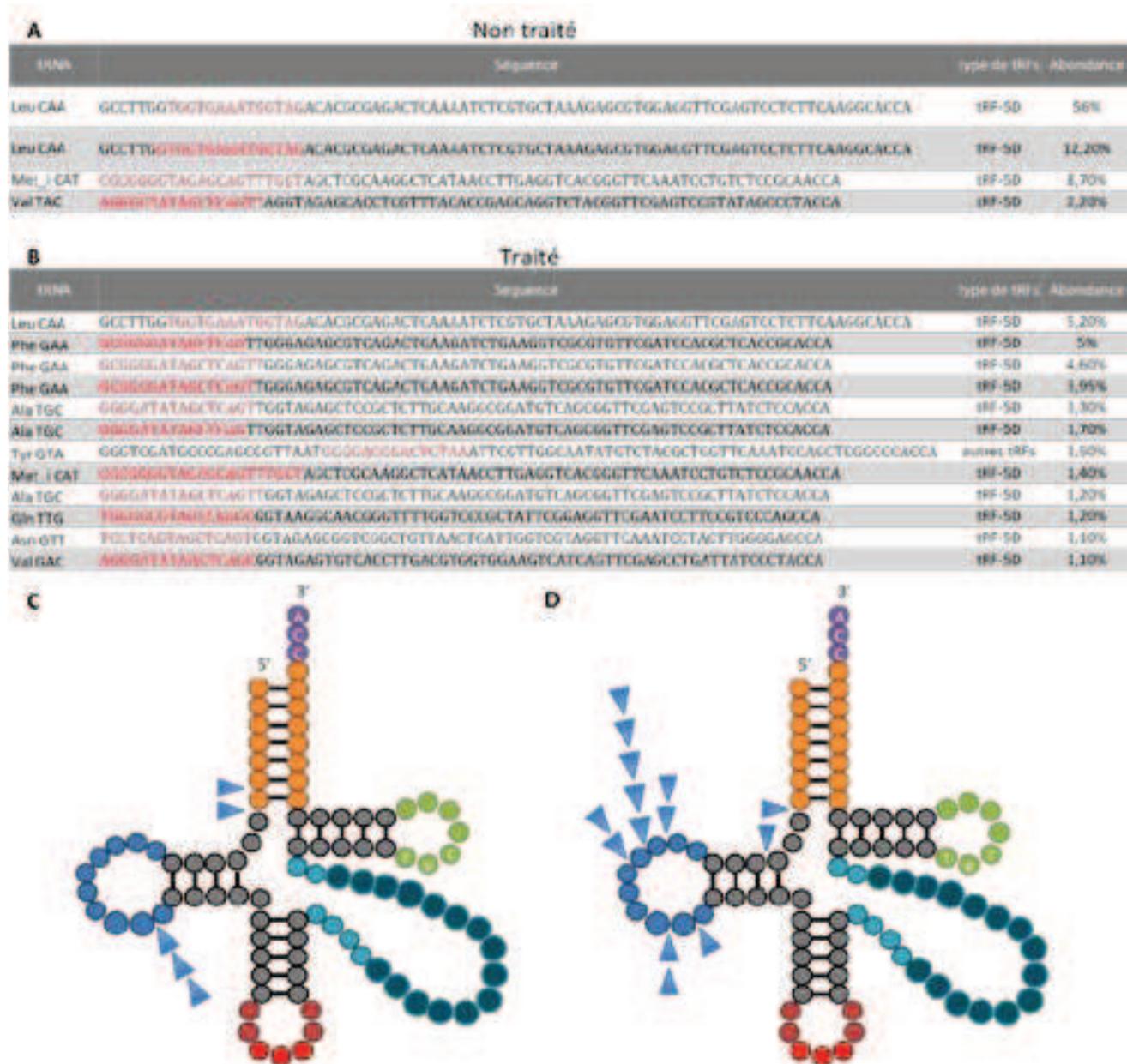
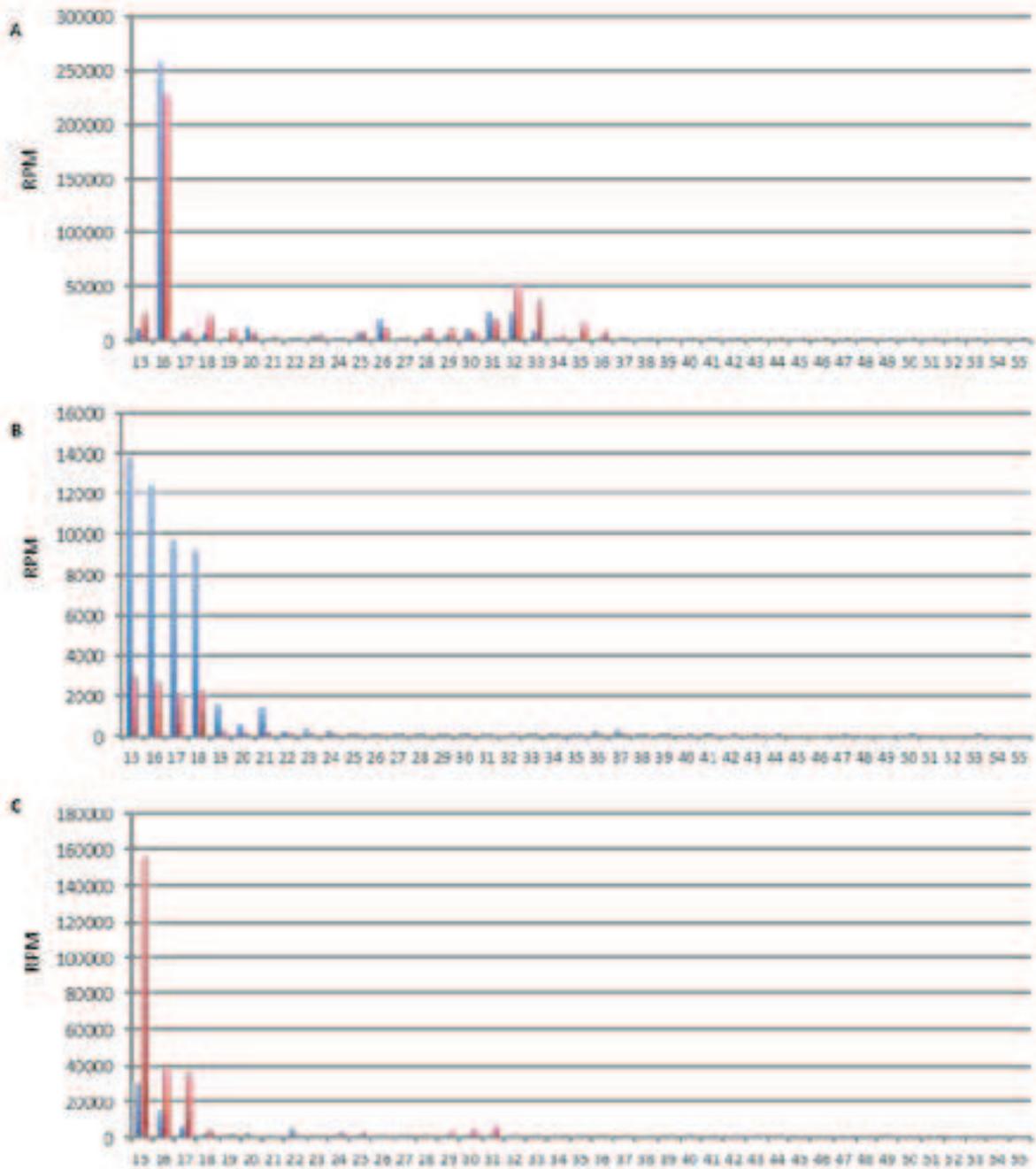


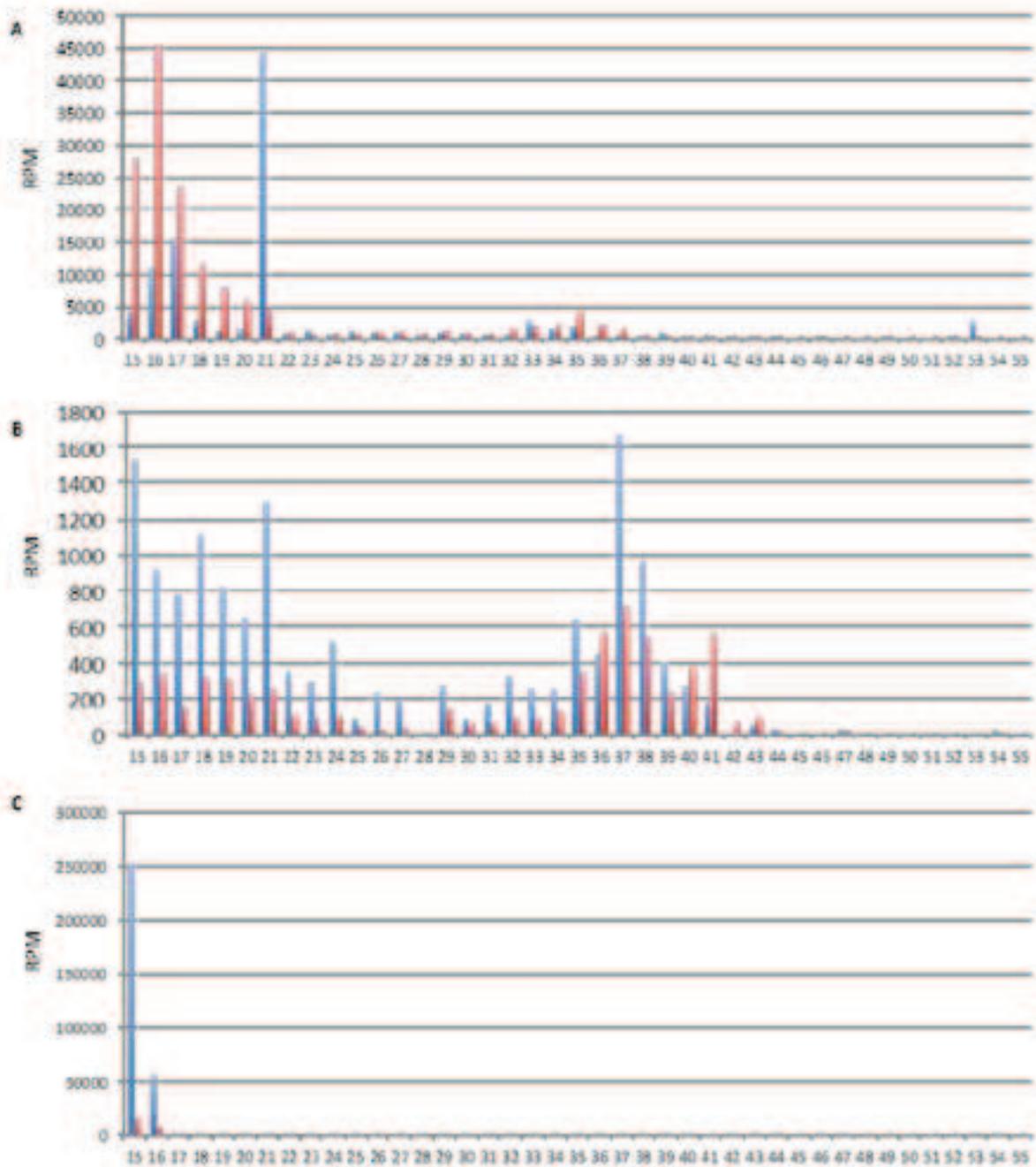
Figure 38 : Représentation des tRFs cytosoliques majoritaires dans les banques non traitées et traitées : A et B, Chaque tableau liste au moins les 30 premiers pourcents en tRFs des compartiments cytosoliques et nucléaires pour les banques non-traitée et traitée. En noir, figure la séquence de tRNA source et en rouge est représentée la séquence du tRFs. C et D, représentation des sites de clivage pour les tRFs présentés en A et B dans la banque non traitée C ou traitée D.

## tRFs cytosoliques



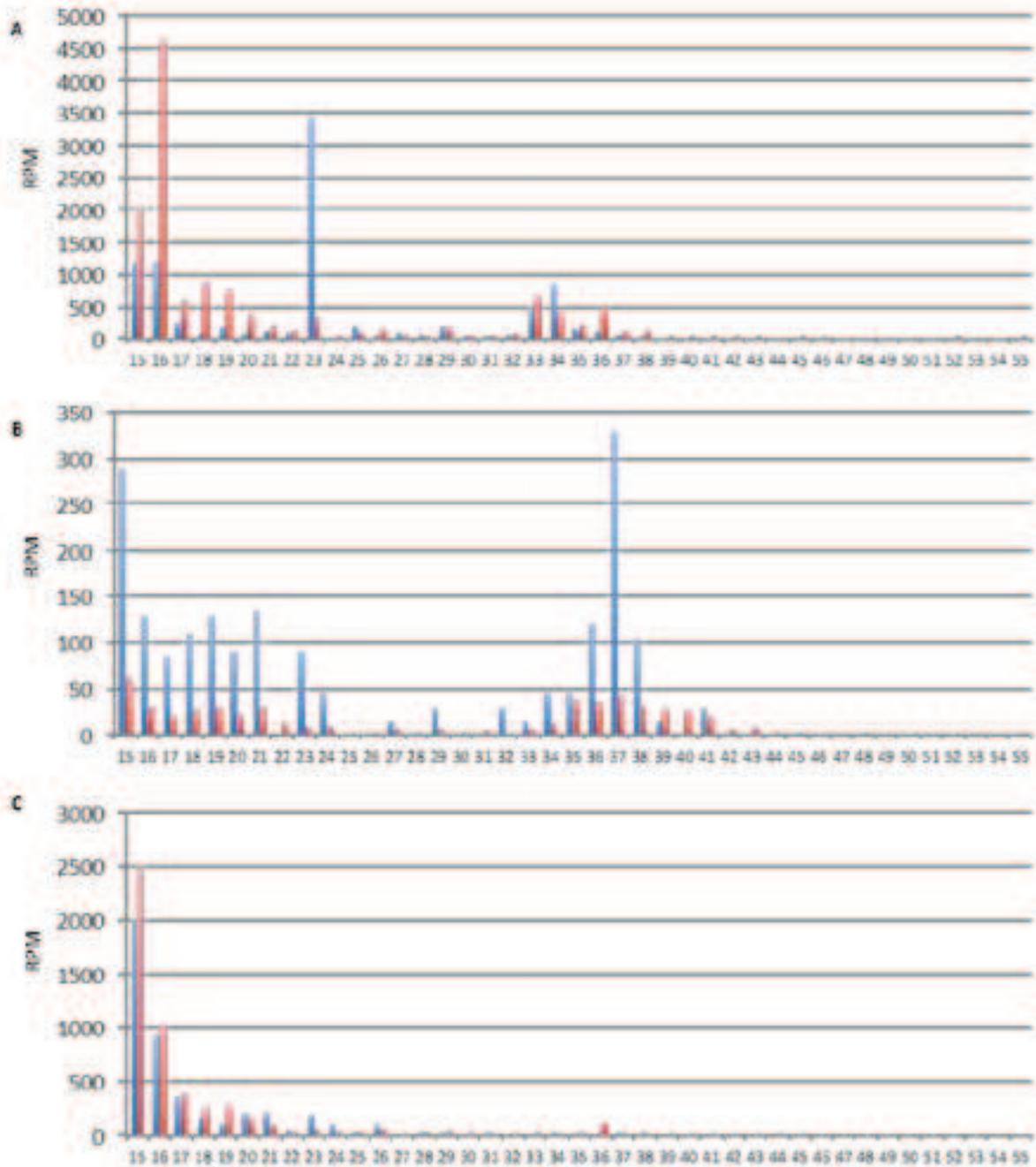
**Figure 39** : Quantité en reads par million (rpm) en tRFs cytosoliques de tailles différentes dans les banques non traitées (bleu) et traitées (rouge) : Les quantités sont affichées pour chaque tailles de tRFs. A, tRF-5 D, T et A; B, tRF-3 D, T et A; C, tRF n'appartenant à aucune des classes précédentes.

## tRFs plastidiaux



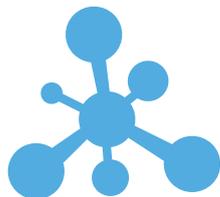
**Figure 40 :** Quantité en reads par million (rpm) en tRFs plastidiaux de tailles différentes dans les banques non traitées (bleu) et traitées (rouge) : Les quantités sont affichées pour chaque tailles de tRFs. A, tRF-5 D, T et A; B, tRF-3 D, T et A C, tRF n'appartenant à aucune des classes précédentes.

## tRFs mitochondriaux



**Figure 41 :** Quantité en reads par million (rpm) en tRFs mitochondriaux de tailles différentes dans les banques non traitées (bleu) et traitées (rouge) : Les quantités sont affichées pour chaque tailles de tRFs. A, tRF-5 D, T et A; B, tRF-3 D, T et A C, tRF n'appartenant à aucune des classes précédentes.





---

# *DISCUSSION*



# Discussion et perspectives

## 1 – Conclusion générale : récapitulatif des résultats

### 1.1 – Analyse des banques de tRFs

L'analyse des banques de petits ARN chez *A. thaliana* a permis l'identification d'une population en tRFs ne représentant qu'une faible proportion des sncRNA totaux (quelques % des petits ARN se situant entre 19 et 26nt). En effet, bien que les tRFs mitochondriaux soient très minoritaires (<1%), il a été possible d'identifier des tRFs issus de tRNA des trois compartiments possédant un génome, notamment le chloroplaste dont 25% des tRFs sont originaires dans les feuilles d'Arabidopsis. Au cours de ces analyses trois types de tRFs ont été caractérisés, les tRF-5D, les tRF-3T et enfin les autres tRFs (*i.e.* tRFs n'appartenant à aucune de ces deux classes). La catégorie « autre » est minoritaire. Elle contient souvent des tRFs distribués tout au long des séquences de tRNA. Elle n'a pas été étudiée en détail et nous ne pouvons exclure que certains de ces tRFs ne soient pas plus fortement représentés et ne soient pas impliqués dans des fonctions biologiques importantes. Il est d'ailleurs à remarquer que peu de tRFs (même si certains sont trouvés, données non montrées) sont issus des grandes boucles variables que possèdent certains tRNA et qui *a priori* pourraient être facilement accessibles au clivage mécanique. Pour ce qui est des deux autres classes de tRFs, les tRFs-5D et tRFs-3T, leur distribution varie grandement d'un compartiment à l'autre. Toutefois, les tRF-5D y sont toujours majoritaires. Il est aussi intéressant de remarquer que la grande majorité des tRFs, qu'ils soient issus de l'une ou l'autre des extrémités ou de l'un ou l'autre des compartiments, dérivent d'un jeu réduit de tRNA. Ainsi, 11 et 13 séquences de tRFs représentent respectivement 67 et 75% des tRF-5D cytosoliques et plastidiaux, tandis que 6 et 5 autres séquences représentent 48 et 71% des tRF-3D cytosoliques et plastidiaux.

De plus, il semble que la fréquence de certains tRFs-5D (même s'ils sont peu nombreux) varie en fonction des tissus ou organes concernés. Ainsi, par exemple, la quantité de tRF-5D Ala-AGC cytosolique diminue très fortement dans



les racines. D'autres tRF-5D cytosoliques (*e.g.* Arg-TCG, Cys-GCA, Gly-TCC) voient eux aussi leur quantité augmentée dans les racines lorsqu'on les compare aux autres banques. De façon extrêmement intéressante, il en est de même pour le tRF-5D His-GTG chloroplastique, alors que la quantité de chloroplastes dans les racines est très fortement réduite par rapport à un extrait de feuilles. En revanche, très peu de fluctuations existent quant aux tRFs-3T. La seule exception semble concerner le tRF-3T Arg-AGG chloroplastique qui augmente dans les racines, en dépit de la quasi-absence de chloroplastes dans ce tissu.

La quantité en tRFs ne varie pas que selon les tissus dans lesquels ils se trouvent, il semble en effet, que certains tRFs soient induits au cours de stress abiotiques, en particulier lors d'un stress UV. Toutefois les variations semblent très peu marqués, en particulier lorsque les plantes sont soumises à des stress sécheresse, froid ou encore salin. D'après les données de la littérature, les stress liés à des carences en nutriments, comme la carence en phosphate semble entraîner des variations plus importantes quant aux populations en tRFs. Il est aussi apparu au cours de l'analyse de banques de petits ARN co-immunoprécipités avec AGO1 qu'une petite population de tRFs-5D de 19 à 20nt était immunoprécipitée avec AGO1. Outre la *quasi* absence de tRF-3T dans ces banques nous avons aussi remarqué que les tRFs-5D immunoprécipités étaient redondants malgré les différents tissus analysés. Parmi les tRFs les plus abondants on dénombre les tRFs-5D cytosoliques ArgACG, ArgCCT, ArgTCG et GlyTCC. Des tRFs-5D chloroplastiques sont aussi enrichis dans ces banques de racines et de feuilles : ArgACG et HisGTG, ils représentent à eux deux 90% de tous les fragments chloroplastiques séquencés dans ces banques.

Enfin, des northern blots effectués à l'aide d'extrait des différentes fractions de cellule d'*A. thaliana* (*i.e.*, mitochondriale/chloroplastique, cytosolique et totale) ont révélé que les tRFs d'organites n'étaient générés ni dans la mitochondrie ni dans le chloroplaste.

## 1.2 – Biogenèse des tRFs

Le second objectif de ma thèse était d'identifier les enzymes responsables du clivage des tRNA chez *A. thaliana*. Notre première approche visant à vérifier



l'activité des protéines DICER (DCL 2, 3 et 4) d'*Arabidopsis* s'est révélée infructueuse. De plus, l'analyse des plantes KO pour RDR6 ne montre aucun changement dans la quantité en tRFs. Ceci nous suggère que beaucoup des voies impliquées dans la biogenèse des siRNA ne sont pas impliquées dans celles des tRFs.

Nous nous sommes donc orientés vers la recherche d'homologues à l'angiogénine humaine, ou à RNY1 de levure chez *A. thaliana*. Seule RNY1, connue pour cliver les tRNA dans l'anticodon, a permis l'identification de 5 homologues chez *A. thaliana*. Ces enzymes, nommées RNS, sont des RNases de type T2. Elles clivent donc les ARN simples brins en libérant une extrémité 5'-OH et une autre 3'-P. Après clonage des RNS en système recombinant, afin de les exprimer chez *E. coli*, nous avons testé l'activité de chacune d'entre elles sur les tRNA<sup>Ala</sup> cytosoliques et tRNA<sup>His</sup> chloroplastiques transcrits *in vitro*. Dans ces conditions, seules les protéines RNS1 et 3 recombinantes ont montré une activité de clivage, non seulement dans la boucle de l'anticodon, mais aussi dans la région D. Ces résultats ont été confirmés par l'analyse des RNS1 et 3 dont les sites catalytiques ont été mutés et chez qui aucune activité n'a été détectée. Ces 2 protéines sont donc les premières RNases à avoir été caractérisées comme étant capables de cliver les tRNA dans la boucle D en plus de la boucle de l'anticodon. Il est pour l'instant difficile de savoir si l'absence d'activité de clivage des RN2 et RNS4 sont une réalité biologique où s'il s'agit d'un problème de mauvaise conformation des protéines produites en système bactérien.

Des transformations transitoires de feuilles de *N. benthamiana* avec des constructions codantes pour chacune des RNS fusionnées à la GFP ont permis la localisation de ces enzymes dans le réticulum endoplasmique.

Bien qu'il ne fasse aucun doute de leur activité *in vitro*, nous manquons d'éléments pour affirmer que ces enzymes occupent ce rôle *in planta*. Nous avons donc cherché à produire des lignées KO pour chacune des RNS. Or, seules les lignées d'insertion homozygotes pour RNS2 et 3 sont disponibles, et il n'existe que des hétérozygotes pour les autres RNS. Devant le peu de lignées KO disponibles, nous avons aussi cherché à produire des lignées hypomorphes pour chacune des RNS par une approche de RNA silencing. Celles-ci ont été obtenues à



la fin de ma thèse, elles n'ont pas encore fait l'objet d'analyses détaillées. Tant pour la lignée KO que pour la lignée knock down, il semble très difficile d'abolir ou de diminuer l'expression de RNS1. À ce stade, il est difficile de savoir s'il s'agit un problème technique ou si cette difficulté a une signification biologique.

### 1.3 – Clonage des tRFs, activité des RNS

Après avoir pris conscience que les RNS clivaient les tRNA de façon non conventionnelle, nous avons tenté d'adapter le protocole illumina en traitant au préalable les ARN à la T4 polynucléotides kinase dans le but de produire des banques plus fiables offrant une meilleure perspective sur les tRFs et leur population dans la cellule.

L'analyse des 2 banques (traitées et non traitées) a permis d'identifier que l'action de la kinase en amont du séquençage semble avoir fonctionné. Ainsi, la quantité en tRFs tant en reads qu'en diversité augmente après le traitement. Cette augmentation en tRFs est notamment due à l'amélioration de la qualité du clonage des fragments générés par les RNS. De plus, au cours de cette analyse nous avons pu observer une augmentation en autres tRFs dans la banque traitée. Les autres tRFs sont des fragments très peu caractérisés dans la littérature, issus d'un clivage entre les boucles et ne comportant donc aucune extrémité.

Cependant, il s'avère que ces résultats ne sont pas aussi positifs qu'il n'y paraît. Ainsi, même si la quantité en tRFs d'un point de vue général semble avoir augmenté, ce n'est pas le cas en détail. En effet, tandis que la quantité en tRF-5 suit la tendance générale, il semble que la majorité des tRF-3 diminue après traitement. Nous avons dès lors supposé que le traitement enzymatique avait au moins partiellement endommagé les tRNA causant donc la diminution constatée en tRF-3 ainsi qu'une partie de l'augmentation en autres tRFs.

## 2 – Discussion

### 2.1 - Les tRFs

De nombreux points sont encore à éclaircir concernant l'origine des tRFs et les rôles que ceux-ci occupent chez *A. thaliana*. L'un des premiers points de



discussion porte sur l'aspect général des banques et les raisons du déséquilibre entre tRFs-5D et tRFs-3T. En effet, il semble que quel que soit l'origine des fragments (cytosoliques ou plastidiaux), les stress (UV), ou les traitements (immunoprécipitations) les tRFs-3T sont moins représentés que les tRFs-5D.

La moins probable de ces explications réside dans la façon dont les banques ont été créées. En effet, celles-ci ne s'étendent que sur des tailles allant de 19 à 26 nt. Or, en raison de l'ajout post-transcriptionnel du CCA sur les tRNA, l'extrémité 3' est plus grande de 4 nucléotides que l'extrémité 5'. Ce qui, après clivage, peut emmener des tRF-3D à dépasser les 26nt et donc à ne pas être séquencés. Cependant, si l'on s'intéresse au profil de clivage des tRNA (figure 2 – article en 1.1) il semble que la majorité des sites de coupures des tRF-3T se situe en aval de la position 48 et les clivages à cette position donnent lieu à des tRF-3T de 26 nucléotides. Par conséquent, la plupart de ces fragments sont pris en compte dans nos banques, et seule une petite fraction d'entre eux n'a pas été analysée. Il semble donc peu crédible que les quelques reads consécutifs de ce biais dans la création des banques soit responsable de l'écart constaté entre les 2 classes de tRFs.

Une autre explication à ce déséquilibre entre tRF-5D et 3T concerne peut-être la structure des tRNA, elle-même. En effet, comme montré dans la figure 9C, les boucles D et T sont très impliquées dans la structure tertiaire des tRNA. Il est donc possible que cette absence de tRFs-3T soit due à un clivage différentiel des tRNA favorisant les coupures dans la boucle D, elles-mêmes facilitées par la structure des tRNA chez qui cette boucle est rendue plus accessible que l'autre.

Une troisième explication est apportée par la banque d'ARN non traitée, créée sur Perth. En effet, dans cette banque, les tailles étudiées s'étalent de 16 à 50 nucléotides. Par conséquent, nous avons été en mesure d'étudier les clivages dans la boucle de l'anticodon. Or, pour ce type de clivage aussi, nous pouvons faire le même type d'observation : les tRFs-5 sont très abondants et les tRFs-3 ne le sont pas. Pourtant, par définition, un clivage dans la boucle de l'anticodon génère automatiquement 2 moitiés de tRNA, et donc un tRF-5 et un tRF-3. Cependant, ces derniers sont, sous représentés dans notre banque, nous pouvons supposer que, pour des raisons encore inconnues à ce jour, les tRFs-3A sont dégradés par la cellule peu de temps après leur clivage. De plus, si les tRF-



3A sont détruits, il est tout aussi probable que les tRFs-3T subissent le même processus expliquant alors le déséquilibre observé dans le premier jeu de banques étudiées. Ceci nous permet donc de sérieusement émettre le postulat que certains tRFs occupent une fonction dans la cellule, et que ces tRFs sont sciemment choisis et conservés par la cellule. Cette hypothèse pourrait en outre expliquer pourquoi sur les 46 tRNA isoaccepteurs connus dans le génome d'*A. thaliana* seulement 11 donnent principalement lieu à des tRF5-D et 6 à des tRF-3T (13 et 5 dans le génome plastidial).

Les variations en tRFs observées dans les différents tissus ou au cours des stress UV abondent aussi dans le sens d'un rôle des tRFs dans la cellule. Ils suggèrent que les besoins, soit en tRFs, soit en tRNA, de la cellule évoluent selon ses conditions de vies. Or, bien que l'on sache peu de choses sur le taux d'expression des tRNA en fonction des tissus, il est probable que leur abondance suive les besoins de la traduction, et donc de la cellule. De la même manière, il n'a jamais été démontré pour les tRFs que la cellule conservait certains de ces fragments dans le but de les utiliser. Cependant, il semble que certaines données de séquençages obtenues pour les banques de petits ARN co-immunoprécipitées avec AGO1 suggèrent qu'aux moins quelques tRFs sont conservés, et qu'ils occupent une place bien précise dans la régulation de l'expression génique. En effet, malgré le fait que ces banques aient été produites par différents expérimentateurs, sur plusieurs tissus ou organes, et à des moments différents, on remarque que ce sont les mêmes tRFs qui y sont enrichis (voir article en préparation et données non montrées) (*e.g* tRFs-5D ArgTCG et GlyTCC). AGO1 étant une protéine de la voie de l'interférence ARN, il est donc tout à fait concevable que ces tRFs co-immunoprécipités avec cette protéine occupent un rôle dans la régulation de l'expression de certains gènes. La recherche de cibles potentielles devra donc être menée.

L'avant dernier point qu'il est nécessaire d'aborder concerne les localisations des tRFs observées en northern blots. En effet, il semble que ceux-ci ne soient pas produits dans les organites. Plusieurs hypothèses sont alors possibles. Soit les tRFs sont générés dans le cytosol, soit dans des compartiments intracellulaires. Cette dernière hypothèse est actuellement mis en avant grâce à

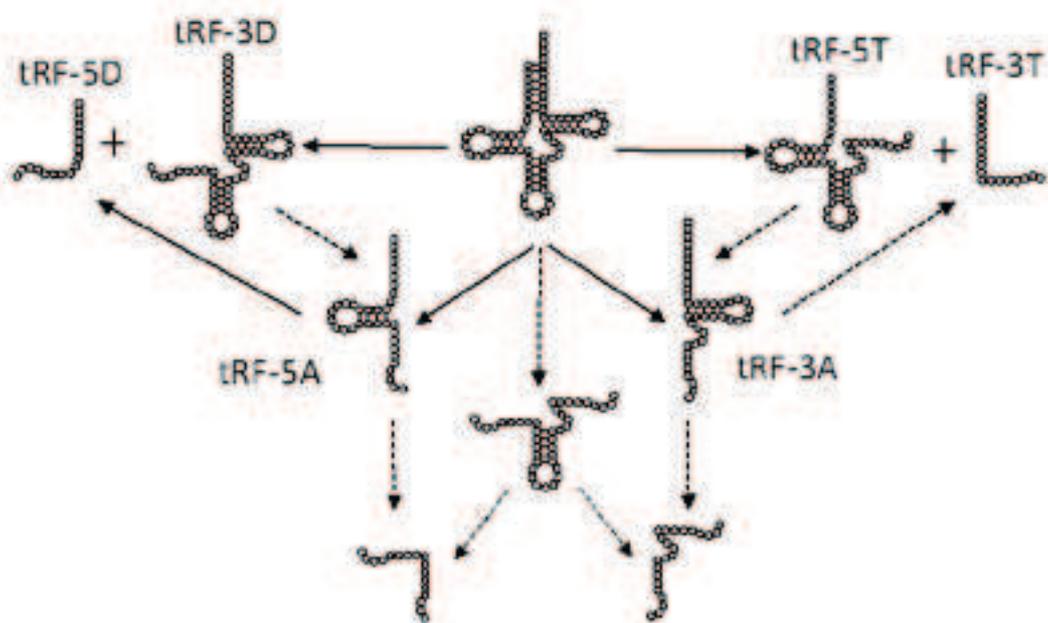


des données obtenues chez *T. cruzi* chez qui les tRFs ont été identifiés, en FISH, dans de petites vésicules de stress (Garcia-Silva et al., 2010).

## 2.2 – Biogenèse des tRFs et localisation des RNS

De nombreuses questions restent en suspens concernant les RNS. La première de ces questions porte sur les différences entre les plantes, les champignons et les protozoaires d'une part, et l'homme d'autre part. En effet, chez ce dernier le clivage des tRNA n'est pas assuré par des RNases T2 mais par une RNase A (Angiogénine, homme, Yamasaki et al., 2009), ou une RNase III (DICER, homme, Haussecker et al., 2010), des RNases Z (ELAC2, human, Lee et al., 2009). Les raisons expliquant ces différences sont à ce jour inconnues, d'autant que le génome humain, comme la majorité des eucaryotes, code pour au moins une RNase T2 (MacIntosh, 2011). Cependant, l'enzyme humaine (*i.e* RNASE T2, impliquée dans des processus de tumorigenèse) n'a pas encore été testée sur des tRNA il est donc possible qu'elle clive les tRNA. Quoiqu'il en soit, la multiplicité d'enzymes de différentes familles, clivant les tRNA chez l'homme suppose plusieurs voies de synthèse des tRFs. Or, chez *A. thaliana*, bien que des données de microarray (BAR) confirmées par des western blots effectués dans notre équipe (données non montrées, expériences effectuées par C. Megel) situent les RNS dans divers tissus très différents, chacune d'entre elles appartient à un groupe d'enzyme similaire dont certaines répondent aux mêmes stress (Bariola et al., 1994, 1999; Hillwig et al., 2011). Cependant, les fragments produits par chacune des enzymes, que ce soit les RNS ou les enzymes humaines, ne sont pas très différents les uns des autres. Ce qui tend à montrer que les différences constatées avec l'homme sont dues à des divergences d'évolutions et que malgré des activités différentes, ces enzymes occupent la même fonction chez ces 2 organismes.

Ceci soulève aussi la question du rôle des RNS dans la cellule et des raisons qui les amènent à être exprimées dans des tissus aussi variés. On peut raisonnablement penser que, comme les RNases T2 de *N. alata*, RNS1, 2 et 3 sont exprimées dans le pollen, les fleurs et les siliques afin d'empêcher l'autofécondation des plantes par leur propre pollen (Pour une revue MacIntosh, 2011). On suppose d'ailleurs que ces mécanismes d'auto-incompatibilité

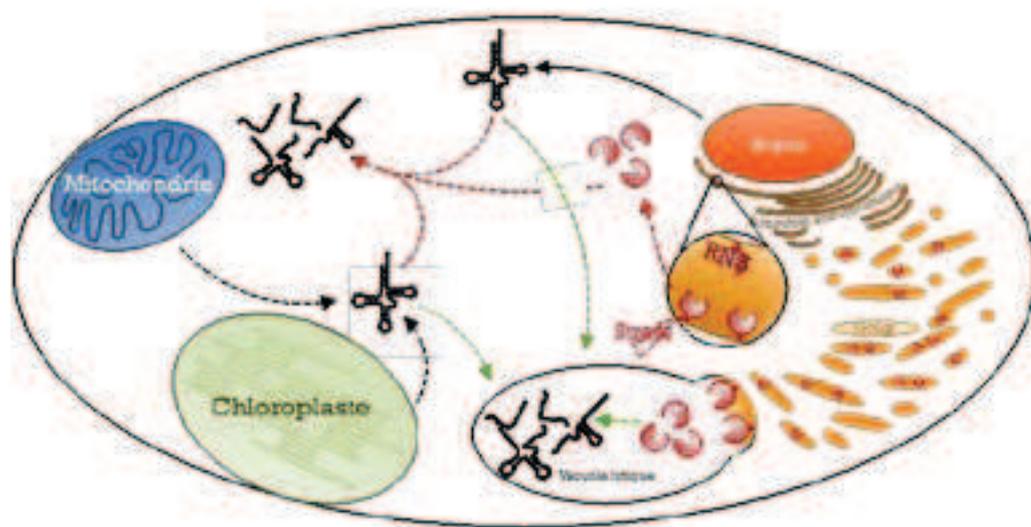


**Figure 43 : Schéma de la production des tRFs à partir d'un tRNA.** Les flèches simples représentent un clivage qui a déjà été caractérisé *in vitro* par l'étude des RNS ou *in silico* par l'étude des banques. Les flèches en pointillés symbolisent des clivages qui n'ont pas encore été observés bien que les fragments qui en découlent aient été observés en de rare occasion.

trouvent leurs origines dans le recrutement d'une ribonucléase impliquée dans des mécanismes de défenses contre les pathogènes (Lee et al. 1992; Hiscock et al. 1996; Nasrallah 2005). Cependant, les RNases T2 n'ont pas totalement disparu du processus de défense contre les pathogènes. Il a ainsi été montré que les RNases NE et Nk1 de *Nicotiana tabacum* ainsi que la RNase NW de *Nicotiana glutinosa* sont surexprimées durant des stress d'infections dû à des virus, et des champignons (Galiana et al., 1997; Kurata et al., 2002; Ohno and Ehara, 2005). D'autres RNases T2 sont induites en réponse à des blessures notamment, RNS1 d'*A. thaliana*, les RNases NW et Nk1 de *N. tabacum* ou encore la RNase LE de tomate (*Lycopersicon esculentum*) (Pour une revue MacIntosh, 2011).

L'activité des RNases T2 sur les tRNA est donc nouvelle, mais pas incompatible avec les rôles qui leurs ont été attribués par le passé car les cibles de ces RNS n'étaient alors pas connus. Il est ainsi tout à fait possible que les tRFs participent aux processus énoncés plus haut *via* des mécanismes pour lesquels ils ont été caractérisés par le passé (*i.e.*, inhibition de la traduction, PTGS) (Ivanov et al., 2011; Kumar et al., 2014). Ceci n'explique cependant pas, pourquoi RNS2 et 4 n'ont pas d'activité sur les tRNA. On suppose que *E. coli* n'est tout simplement pas adapté à la production de ces enzymes comportant des ponts disulfures. Il serait sûrement souhaitable d'utiliser des systèmes de production eucaryotique (*S. cerevisiae*, *Pichia pastoris* ou encore cellules d'insectes) pour produire ces protéines afin de caractériser leurs activités. Un autre point de questionnement concerne la reconnaissance des sites de clivages par les RNS (figure 43). Ainsi bien que nous ayons réussi à identifier que les RNases T2 d'*Arabidopsis* clivaient les tRNA dans les boucles de l'anticodon et D, il semble que les résultats obtenus notamment avec RNS3 sur le tRNA<sup>Ala</sup> transcrit *in vitro* (figure 25) montre que les tRF5-A sont clivés par cette RNS, suggérant que chaque fragment est à la fois produit et substrat. Afin de confirmer cette hypothèse nous pourrions incuber chacune des RNS avec des moitiés de tRNA transcrits *in vitro* et observer le profil obtenu en northern blot.

De plus, des imprécisions demeurent concernant la localisation des RNS. Bien qu'ayant été caractérisées dans le réticulum endoplasmique, il est probable que ceci ne soit pas leur localisation définitive où elles exercent leur activité endonucléasique. Il est cependant impossible que des enzymes ayant une activité



**Figure 44 : Schéma des localisations possibles des RNS et du lieu de clivage des tRNA.** La flèche simple représente le changement de compartiment connu des tRNA étant exporté du noyau vers le cytosol. Les flèches en pointillés représentent des actions pour lesquelles aucune preuve n'a encore été apportée. Les flèches rouges font références à la voie dans laquelle les tRNA sont clivés dans le cytosol. Les flèches vertes symbolisent la voie dans laquelle les tRNA sont clivés dans une vacuole de lyse.

aussi délétère sur les tRNA soient libres dans le cytosol. Ces enzymes et leurs activités sont donc très probablement contenues dans un compartiment à l'intérieur de la cellule et actuellement nous pouvons émettre deux hypothèses quant au compartiment occupé par les RNS (figure 44).

La première hypothèse suppose que les RNS sont séquestrées au niveau du réticulum endoplasmique dans l'attente du clivage de leur peptide signal par une enzyme produite en réponse à un stress. Une fois clivées les RNS seraient relarguées dans le cytosol pour cliver tRNA et rRNA. Cette théorie a l'avantage d'apporter un élément de réponse quant à la façon dont la cellule pourrait répondre aux différents stress, et notamment l'inhibition de la traduction qui est une réponse rapide des plantes soumises à un stress (Muñoz and Castellano, 2012). Cependant cette théorie ne permet pas d'expliquer comment les tRNA plastidiaux et mitochondriaux sont clivés car à l'heure d'aujourd'hui leur export vers le cytosol n'a pas été démontré. Deux hypothèses sont actuellement à l'étude. Les stromules ("stroma filled tubules") sont des structures présentes à la surface des chloroplastes et ils pourraient permettre l'export de tRNA vers le cytosol (Natesan et al., 2005). Ou un processus nommé chlorophagie dont le but est de recycler les chloroplastes vieillissants dans des vacuoles de lyses se situant dans le cytosol (Michaeli and Galili, 2014).

Il se pourrait aussi que les RNS ne soient pas séquestrées au niveau du réticulum endoplasmique mais plutôt que celles-ci suivent la voie de sécrétion à travers le golgi. Les enzymes termineraient leur maturation dans une vacuole lytique où les tRNA seraient également présents. Les tRNA mitochondriaux et plastidiaux quant à eux pourraient être clivés au cours d'un processus de recyclage des organites (*i.e* autophagie) dans ces vacuoles de lyse (Michaeli and Galili, 2014), ce qui expliquerait la présence de tRFs d'organites dans le cytosol (cf. hypothèse plus haut). Cette théorie corrobore les données bibliographiques expliquant que les RNAses T2 préfèrent un pH acide en raison de leurs ponts disulfures. Or, ce pH se trouve notamment dans de telles vacuoles. De plus, cette hypothèse fait écho à deux études séparées montrant d'une part l'augmentation en RNS dans des plantes carencées en phosphate, et d'autre part une augmentation des tRFs (Bariola et al., 1994, 1999; Megel et al., 2015). De telles vacuoles de lyses constituent donc le compartiment idéal pour la cellule afin de



dégrader ses ARN pour en extraire le phosphate. Cette hypothèse n'explique pas comment les tRFs, échappent à leur complète dégradation dans cette vacuole, pour jouer d'autres potentiels rôles dans les processus énoncés plus haut.

## 2.3 – T4 polynucléotides kinase et tRFs

Nous savons que l'analyse des banques traitées à l'aide de la T4 polynucléotides kinase est grandement compliquée par l'éventualité que cette enzyme puisse endommager les tRNA et les tRFs en générant de nouveaux fragments n'ayant aucune signification biologique. Cependant, même si l'on suppose qu'une partie de ces fragments sont des artefacts, l'analyse de ces banques a permis pour la première fois d'identifier des tRFs internes aux tRNA, ce que nous appelons "autres tRFs". La simple preuve de l'existence de tels fragments soulève deux hypothèses.

Si ces fragments ne sont pas en temps normal conservés et que le traitement est le seul responsable de leur présence dans les banques, cela signifie que si ces tRFs sont éventuellement générés *in vivo*, ils sont très vite dégradés, car la seule façon de les observer est de les générer artificiellement. Ceci viendrait étayer l'hypothèse qu'il existe des tRFs qui sont conservés parce qu'ils sont impliqués dans certains processus et stabilisés dans des complexes.

À l'inverse, si ces fragments sont bien le produit du clivage des RNS et que le traitement permet leur clonage, cela apporte la preuve que l'ensemble des fragments est équitablement conservé et que les différentes variations observées jusqu'alors ne proviendraient pas d'une variation de la quantité en tRFs mais d'une variation de la quantité en tRNA. Il serait alors concevable que l'ensemble des tRFs ne soit qu'un sous-produit du clivage des tRNA en attente de dégradation.

Plusieurs démarches permettraient de répondre à cette problématique. La création d'une banque de données libre de tous biais mais pas produite à l'aide d'un traitement T4 polynucléotides kinase permettrait de différencier les fragments clonés, générés *in vivo* des fragments générés par le traitement. Nous pourrions aussi tester en northern blot, sur un transcrit de tRNA<sup>Ala</sup> l'activité de la



kinase sur les tRNA et voir si elle génère des fragments, et ainsi confirmer l'une des 2 hypothèses.

### 3 – Perspectives

Durant les trois ans de mon doctorat nous avons réussi à montrer que *A. thaliana* possède une population en tRFs complexe qui fluctue en fonction de divers paramètres développementaux ou environnementaux. D'autre part, nous avons apporté la preuve que des RNAses d'*A. thaliana* étaient capables de cliver les tRNA. Cependant, à l'heure actuelle nous n'avons pas encore apporté la preuve *in vivo* que ces enzymes sont responsables de la biogenèse des tRFs. C'est pourquoi il faut continuer à étudier les diverses lignées KO et hypomorphes, présentées précédemment, afin de prouver définitivement que les RNS sont responsables de la biogenèse des tRFs chez *A. thaliana*. De plus, nous ne sommes toujours pas certains que RNS2 et 4 n'ont pas d'activité sur les tRNA car celles-ci ont été exprimées en système bactérien. Des expériences sont donc actuellement conduites par Stéphanie Lalande, doctorante dans notre équipe, elles visent à produire les RNS chez *S. cerevisiae* afin de se rapprocher au plus près des conditions de productions des RNS *in planta* de manière à tester une nouvelle fois leurs activités. De plus, la production des RNS en système eucaryote nous permettra peut-être d'obtenir de grandes quantités de ces protéines afin d'en obtenir la structure par cristallographie, et pourquoi pas le mode d'interaction en co-cristallisant les RNS mutées (les RNS sauvages cliveraient leurs substrats) avec un tRNA.

À moyen terme, il serait aussi intéressant de s'attarder sur les rôles que peuvent occuper les tRFs dans la cellule notamment en clonant et surexprimant chez *A. thaliana* les tRFs récurrents découverts dans les différentes banques de petits ARN immunoprécipités avec AGO1. Un autre élément clé nécessaire à la compréhension du rôle des tRFs chez *A. thaliana* porte sur leur localisation subcellulaire qui est à ce jour inconnue. Plusieurs techniques ont été tentées au sein du laboratoire, notamment le FISH (Fluorescence *in situ* Hybridization), mais sans succès car cette démarche est compliquée par la présence de la paroi cellulaire.



Enfin à très long terme, il nous faudrait identifier les voies de signalisations qui déclenchent le clivage des tRFs par les RNS, et sur les interactants participant à ces voies de signalisations, s'ils existent.



# Bibliographie

Adhya, S., Sarkar, P., Valenzuela, D., and Maitra, U. (1979). Termination of transcription by *Escherichia coli* RNA polymerase: influence of secondary structure of RNA transcripts on rho-independent and rho-dependent termination. *Proc. Natl. Acad. Sci. U.S.A.* 76, 1613–1617.

Akama, K., Junker, V., and Beier, H. (2000). Identification of two catalytic subunits of tRNA splicing endonuclease from *Arabidopsis thaliana*. *Gene* 257, 177–185.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2002). *Molecular biology of the cell* (New York: Garland Science).

Allen, E., and Howell, M.D. (2010). miRNAs in the biogenesis of trans-acting siRNAs in higher plants. *Semin. Cell Dev. Biol.* 21, 798–804.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell* 31, 785–799.

Arimbasseri, A.G., Rijal, K., and Maraia, R.J. (2013). Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription*. 4, 6.

Arnold, J.J., Smidansky, E.D., Moustafa, I.M., and Cameron, C.E. (2012). Human mitochondrial RNA polymerase: structure-function, mechanism and inhibition. *Biochim. Biophys. Acta* 1819, 948–960.

Ashby, J., Boutant, E., Seemanpillai, M., Sambade, A., Ritzenthaler, C., and Heinlein, M. (2006). Tobacco Mosaic Virus Movement Protein Functions as a Structural Microtubule-Associated Protein. *J. Virol.* 80, 8329–8344.

Bachellerie, J.P., Cavaillé, J., and Hüttenhofer, A. (2002). The expanding snoRNA world. *Biochimie* 84, 775–790.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science* 234, 179–186.



Bariola, P.A., Howard, C.J., Taylor, C.B., Verburg, M.T., Jaglan, V.D., and Green, P.J. (1994). The Arabidopsis ribonuclease gene RNS1 is tightly controlled in response to phosphate limitation. *Plant J.* *6*, 673–685.

Bariola, P.A., MacIntosh, G.C., and Green, P.J. (1999). Regulation of S-like ribonuclease levels in Arabidopsis. Antisense inhibition of RNS1 or RNS2 elevates anthocyanin accumulation. *Plant Physiol.* *119*, 331–342.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* *136*, 215–233.

Bear, D.G., and Peabody, D.S. (1988). The E. coli Rho protein: an ATPase that terminates transcription. *Trends Biochem. Sci.* *13*, 343–347.

Björk, P., and Wieslander, L. (2014). Mechanisms of mRNA export. *Semin. Cell Dev. Biol.* *32*, 47–54.

Björk, G.R., Jacobsson, K., Nilsson, K., Johansson, M.J., Byström, A.S., and Persson, O.P. (2001). A primordial tRNA modification required for the evolution of life? *EMBO J.* *20*, 231–239.

Bratkovič, T., and Rogelj, B. (2014). The many faces of small nucleolar RNAs. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* *1839*, 438–443.

Braun, K.A., and Young, E.T. (2014). Coupling mRNA Synthesis and Decay. *Mol. Cell. Biol.* *34*, 4078–4087.

Von Braun, S.S., Sabetti, A., Hanic-Joyce, P.J., Gu, J., Schleiff, E., and Joyce, P.B.M. (2007). Dual targeting of the tRNA nucleotidyltransferase in plants: not just the signal. *J. Exp. Bot.* *58*, 4083–4093.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* *322*, 1387–1392.

Brown, G.G., Colas des Francs-Small, C., and Ostersetzer-Biran, O. (2014). Group II intron splicing factors in plant mitochondria. *Front Plant Sci* *5*, 35.

Budde, B.S., Namavar, Y., Barth, P.G., Poll-The, B.T., Nürnberg, G., Becker, C., van Ruissen, F., Weterman, M.A.J., Fluiter, K., te Beek, E.T., et al. (2008). tRNA splicing endonuclease mutations cause pontocerebellar hypoplasia. *Nat. Genet.* *40*, 1113–1118.



- Bulmer, M. (1987). Coevolution of codon usage and transfer RNA abundance. *Nature* 325, 728–730.
- Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. *Mol Cell* 36, 541–546.
- Burger, G., Gray, M.W., Forget, L., and Lang, B.F. (2013). Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol Evol* 5, 418–438.
- Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10, 1957–1966.
- Cameron, V., and Uhlenbeck, O.C. (1977). 3'-Phosphatase activity in T4 polynucleotide kinase. *Biochemistry* 16, 5120–5126.
- Campomenosi, P., Salis, S., Lindqvist, C., Mariani, D., Nordström, T., Acquati, F., and Taramelli, R. (2006). Characterization of RNASET2, the first human member of the Rh/T2/S family of glycoproteins. *Arch. Biochem. Biophys.* 449, 17–26.
- Canino, G., Bocian, E., Barbezier, N., Echeverría, M., Forner, J., Binder, S., and Marchfelder, A. (2009). Arabidopsis Encodes Four tRNase Z Enzymes. *Plant Physiol.* 150, 1494–1502.
- Carneiro, V.T., Dietrich, A., Maréchal-Drouard, L., Cosset, A., Pelletier, G., and Small, I. (1994). Characterization of some major identity elements in plant alanine and phenylalanine transfer RNAs. *Plant Mol. Biol.* 26, 1843–1853.
- Cashel, M., and Gallant, J. (1969). Two Compounds implicated in the Function of the RC Gene of *Escherichia coli*. *Nature* 221, 838–841.
- Chan, P.P., and Lowe, T.M. (2009). GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucl. Acids Res.* 37, D93–D97.
- Chen, C.-J., liu, Q., Zhang, Y.-C., Qu, L.-H., Chen, Y.-Q., and Gautheret, D. (2011). Genome-wide discovery and analysis of microRNAs and other small RNAs from rice embryogenic callus. *RNA Biology* 8, 538–547.
- Chen, P., Jäger, G., and Zheng, B. (2010). Transfer RNA modifications and genes for modifying enzymes in *Arabidopsis thaliana*. *BMC Plant Biology* 10, 201.



Cheung, A.C.M., and Cramer, P. (2012). A Movie of RNA Polymerase II Transcription. *Cell* *149*, 1431–1437.

Cognat, V., Deragon, J.-M., Vinogradova, E., Salinas, T., Remacle, C., and Maréchal-Drouard, L. (2008). On the Evolution and Expression of *Chlamydomonas reinhardtii* Nucleus-Encoded Transfer RNA Genes. *Genetics* *179*, 113–123.

Cognat, V., Pawlak, G., Duchêne, A.-M., Daujat, M., Gigant, A., Salinas, T., Michaud, M., Gutmann, B., Philippe Giegé, Gobert, A., et al. (2013). PlantRNA, a database for tRNAs of photosynthetic eukaryotes. *Nucleic Acids Res.* *41*, D273–D279.

Cole, C., Sobala, A., Lu, C., Thatcher, S.R., Bowman, A., Brown, J.W.S., Green, P.J., Barton, G.J., and Hutvagner, G. (2009). Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* *15*, 2147–2160.

Condon, C., and Putzer, H. (2002). The phylogenetic distribution of bacterial ribonucleases. *Nucleic Acids Res.* *30*, 5339–5346.

Cooley, L., Appel, B., and Söll, D. (1982). Post-transcriptional nucleotide addition is responsible for the formation of the 5' terminus of histidine tRNA. *Proc. Natl. Acad. Sci. U.S.A.* *79*, 6475–6479.

Copeland, C.S., Mann, V.H., Morales, M.E., Kalinna, B.H., and Brindley, P.J. (2005). The Sinbad retrotransposon from the genome of the human blood fluke, *Schistosoma mansoni*, and the distribution of related Pao-like elements. *BMC Evol. Biol.* *5*, 20.

Couvillion, M.T., Bounova, G., Purdom, E., Speed, T.P., and Collins, K. (2012). A *Tetrahymena* Piwi bound to mature tRNA 3' fragments activates the exonuclease Xrn2 for RNA processing in the nucleus. *Mol. Cell* *48*, 509–520.

Davanloo, P., Sprinzl, M., Watanabe, K., Albani, M., and Kersten, H. (1979). Role of ribothymidine in the thermal stability of transfer RNA as monitored by proton magnetic resonance. *Nucleic Acids Res.* *6*, 1571–1581.

Delage, L., Duchêne, A.-M., Zaepfel, M., and Maréchal-Drouard, L. (2003). The anticodon and the D-domain sequences are essential determinants for plant cytosolic tRNA(Val) import into mitochondria. *Plant J.* *34*, 623–633.

Delannoy, E., Fujii, S., Francs-Small, C.C. des, Brundrett, M., and Small, I. (2011). Rampant Gene Loss in the Underground Orchid *Rhizanthella gardneri* Highlights Evolutionary Constraints on Plastid Genomes. *Mol Biol Evol* *28*, 2077–2086.



- Deshpande, R.A., and Shankar, V. (2002). Ribonucleases from T2 family. *Crit. Rev. Microbiol.* 28, 79–122.
- Deutscher, M.P. (1982). 7 tRNA Nucleotidyltransferase. In *The Enzymes*, Paul D. Boyer, ed. (Academic Press), pp. 183–215.
- Dieci, G., Yukawa, Y., Alzapiedi, M., Guffanti, E., Ferrari, R., Sugiura, M., and Ottonello, S. (2006). Distinct modes of TATA box utilization by the RNA polymerase III transcription machineries from budding yeast and higher plants. *Gene* 379, 12–25.
- Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The expanding RNA polymerase III transcriptome. *Trends Genet.* 23, 614–622.
- Dietrich, A., Small, I., Cosset, A., Weil, J.H., and Maréchal-Drouard, L. (1996). Editing and import: strategies for providing plant mitochondria with a complete set of functional transfer RNAs. *Biochimie* 78, 518–529.
- Dombroski, B.A., Feng, Q., Mathias, S.L., Sassaman, D.M., Scott, A.F., Kazazian, H.H., and Boeke, J.D. (1994). An in vivo assay for the reverse transcriptase of human retrotransposon L1 in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 14, 4485–4492.
- Doucet, A.J., Hulme, A.E., Sahinovic, E., Kulpa, D.A., Moldovan, J.B., Kopera, H.C., Athanikar, J.N., Hasnaoui, M., Bucheton, A., Moran, J.V., et al. (2010). Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet.* 7, 6.
- Duchêne, A.-M., and Maréchal-Drouard, L. (2001). The Chloroplast-Derived trnW and trnM-e Genes Are Not Expressed in Arabidopsis Mitochondria. *Biochemical and Biophysical Research Communications* 285, 1213–1216.
- Duchêne, A.-M., Pujol, C., and Maréchal-Drouard, L. (2009). Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. *Curr Genet* 55, 1–18.
- Dunoyer, P., Brosnan, C.A., Schott, G., Wang, Y., Jay, F., Alioua, A., Himber, C., and Voinnet, O. (2010). An endogenous, systemic RNAi pathway in plants. *The EMBO Journal* 29, 1699–1712.
- Duret, L. (2000). tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. *Trends in Genetics* 16, 287–289.



El Yacoubi, B., Lyons, B., Cruz, Y., Reddy, R., Nordin, B., Agnelli, F., Williamson, J.R., Schimmel, P., Swairjo, M.A., and de Crécy-Lagard, V. (2009). The universal YrdC/Sua5 family is required for the formation of threonylcarbamoyladenine in tRNA. *Nucleic Acids Res.* *37*, 2894–2909.

Englert, M., Latz, A., Becker, D., Gimple, O., Beier, H., and Akama, K. (2007). Plant pre-tRNA splicing enzymes are targeted to multiple cellular compartments. *Biochimie* *89*, 1351–1365.

Evguenieva-Hackenberg, E., and Klug, G. (2011). New aspects of RNA processing in prokaryotes. *Current Opinion in Microbiology* *14*, 587–592.

Feng, L., Sheppard, K., Namgoong, S., Ambrogelly, A., Polycarpo, C., Randau, L., Tumbula-Hansen, D., and Söll, D. (2004). Aminoacyl-tRNA synthesis by pre-translational amino acid modification. *RNA Biol* *1*, 16–20.

Fey, J., and Maréchal-Drouard, L. (1999). Expression of the two chloroplast-like tRNA(Asn) genes in potato mitochondria: mapping of transcription initiation sites present in the trnN1-trnY-nad2 cluster and upstream of trnN2. *Curr. Genet.* *36*, 49–54.

Fey, J., Weil, J.H., Tomita, K., Cosset, A., Dietrich, A., Small, I., and Maréchal-Drouard, L. (2002). Role of editing in plant mitochondrial transfer RNAs. *Gene* *286*, 21–24.

Frechin, M., Duchêne, A.-M., and Becker, H.D. (2009). Translating organellar glutamine codons: a case by case scenario? *RNA Biol* *6*, 31–34.

Fu, H., Feng, J., Liu, Q., Sun, F., Tie, Y., Zhu, J., Xing, R., Sun, Z., and Zheng, X. (2009). Stress induces tRNA cleavage by angiogenin in mammalian cells. *FEBS Lett.* *583*, 437–442.

Galej, W.P., Nguyen, T.H.D., Newman, A.J., and Nagai, K. (2014). Structural studies of the spliceosome: zooming into the heart of the machine. *Curr Opin Struct Biol* *25*, 57–66.

Galiana, E., Bonnet, P., Conrod, S., Keller, H., Panabières, F., Ponchet, M., Poupet, A., and Ricci, P. (1997). RNase activity prevents the growth of a fungal pathogen in tobacco leaves and increases upon induction of systemic acquired resistance with elicitor. *Plant Physiol.* *115*, 1557–1567.

Galli, G., Hofstetter, H., and Birnstiel, M.L. (1981). Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. *Nature* *294*, 626–631.

Ganot, P., Caizergues-Ferrer, M., and Kiss, T. (1997). The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.* *11*, 941–956.



Garcia-Silva, M.R., Frugier, M., Tosar, J.P., Correa-Dominguez, A., Ronalte-Alves, L., Parodi-Talice, A., Rovira, C., Robello, C., Goldenberg, S., and Cayota, A. (2010). A population of tRNA-derived small RNAs is actively produced in *Trypanosoma cruzi* and recruited to specific cytoplasmic granules. *Mol. Biochem. Parasitol.* *171*, 64–73.

Gascioli, V., Mallory, A.C., Bartel, D.P., and Vaucheret, H. (2005). Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr. Biol.* *15*, 1494–1500.

Gerber, A.P., and Keller, W. (1999). An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* *286*, 1146–1149.

Ghosh, A., and Lima, C.D. (2010). Enzymology of RNA cap synthesis. *Wiley Interdisciplinary Reviews - RNA* *1*, 152-172.

Giegé, R., Sissler, M., and Florentz, C. (1998). Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* *26*, 5017–5035.

Giegé, R., Jühling, F., Pütz, J., Stadler, P., Sauter, C., and Florentz, C. (2012). Structure of transfer RNAs: similarity and variability. *Wiley Interdiscip Rev RNA* *3*, 37–61.

Giuliodori, S., Percudani, R., Braglia, P., Ferrari, R., Guffanti, E., Ottonello, S., and Dieci, G. (2003). A composite upstream sequence motif potentiates tRNA gene transcription in yeast. *J. Mol. Biol.* *333*, 1–20.

Gonda, D.K., Bachmair, A., Wüning, I., Tobias, J.W., Lane, W.S., and Varshavsky, A. (1989). Universality and structure of the N-end rule. *J. Biol. Chem.* *264*, 16700–16712.

Graciet, E., Hu, R.-G., Piatkov, K., Rhee, J.H., Schwarz, E.M., and Varshavsky, A. (2006). Aminoacyl-transferases and the N-end rule pathway of prokaryotic/eukaryotic specificity in a human pathogen. *Proc. Natl. Acad. Sci. U.S.A.* *103*, 3078–3083.

Green, N.J., Grundy, F.J., and Henkin, T.M. (2010). The T box mechanism: tRNA as a regulatory molecule. *FEBS Lett.* *584*, 318–324.

Grosjean, H., de Crécy-Lagard, V., and Marck, C. (2010). Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett.* *584*, 252–264.



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

Hackenberg, M., Huang, P.-J., Huang, C.-Y., Shi, B.-J., Gustafson, P., and Langridge, P. (2013). A Comprehensive Expression Profile of MicroRNAs and Other Classes of Non-Coding Small RNAs in Barley Under Phosphorous-Deficient and -Sufficient Conditions. *DNA Res* *20*, 109–125.

Hafez, M., Burger, G., Steinberg, S.V., and Lang, F. (2013). A second eukaryotic group with mitochondrion-encoded tmRNA. *RNA Biol* *10*, 1117–1124.

Hamada, M., Huang, Y., Lowe, T.M., and Maraia, R.J. (2001). Widespread use of TATA elements in the core promoters for RNA polymerases III, II, and I in fission yeast. *Mol. Cell. Biol.* *21*, 6870–6881.

Hanaoka, M., Kanamaru, K., Takahashi, H., and Tanaka, K. (2003). Molecular genetic analysis of chloroplast gene promoters dependent on SIG2, a nucleus-encoded sigma factor for the plastid-encoded RNA polymerase, in Arabidopsis thaliana. *Nucleic Acids Res.* *31*, 7090–7098.

Haseltine, W.A., and Block, R. (1973). Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* *70*, 1564–1568.

Haud, N., Kara, F., Diekmann, S., Henneke, M., Willer, J.R., Hillwig, M.S., Gregg, R.G., Macintosh, G.C., Gärtner, J., Alia, A., et al. (2011). RNaset2 mutant zebrafish model familial cystic leukoencephalopathy and reveal a role for RNase T2 in degrading ribosomal RNA. *Proc. Natl. Acad. Sci. U.S.A.* *108*, 1099–1103.

Haussecker, D., Huang, Y., Lau, A., Parameswaran, P., Fire, A.Z., and Kay, M.A. (2010). Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* *16*, 673–695.

Helm, M., Florentz, C., Chomyn, A., and Attardi, G. (1999). Search for differences in post-transcriptional modification patterns of mitochondrial DNA-encoded wild-type and mutant human tRNA<sup>Lys</sup> and tRNA<sup>Leu</sup>(UUR). *Nucleic Acids Res.* *27*, 756–763.

Hillwig, M.S., Lebrasseur, N.D., Green, P.J., and Macintosh, G.C. (2008). Impact of transcriptional, ABA-dependent, and ABA-independent pathways on wounding regulation of RNS1 expression. *Mol. Genet. Genomics* *280*, 249–261.

Hillwig, M.S., Contento, A.L., Meyer, A., Ebany, D., Bassham, D.C., and Macintosh, G.C. (2011). 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*, 1093–1098.



- Hofstetter, H., Kressman, A., and Birnstiel, M.L. (1981). A split promoter for a eucaryotic tRNA gene. *Cell* 24, 573–585.
- Hopper, A.K., and Shaheen, H.H. (2008). A decade of surprises for tRNA nuclear-cytoplasmic dynamics. *Trends Cell Biol.* 18, 98–104.
- Howard, B.H., and de Crombrughe, B. (1976). ATPase activity required for termination of transcription by the *Escherichia coli* protein factor rho. *J. Biol. Chem.* 251, 2520–2524.
- Hsieh, L.-C., Lin, S.-I., Shih, A.C.-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 Physiol.* 151, 2120–2132.
- 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 Signal Behav* 5, 537–539.
- Huh, W.-K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O’Shea, E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686–691.
- Ibba, M., and Söll, D. (2000). Aminoacyl-tRNA synthesis. *Annual Review of Biochemistry* 69, 617–650.
- Ida, K., Norioka, S., Yamamoto, M., Kumasaka, T., Yamashita, E., Newbigin, E., Clarke, A.E., Sakiyama, F., and Sato, M. (2001). The 1.55 Å resolution structure of *Nicotiana glauca* S(F11)-RNase associated with gametophytic self-incompatibility. *J. Mol. Biol.* 314, 103–112.
- Igic, B., and Kohn, J.R. (2001). Evolutionary relationships among self-incompatibility RNases. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13167–13171.
- Irie, M. (1999). Structure-function relationships of acid ribonucleases: lysosomal, vacuolar, and periplasmic enzymes. *Pharmacol. Ther.* 81, 77–89.
- Ivanov, P., Emara, M.M., Villen, J., Gygi, S.P., and Anderson, P. (2011). Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol. Cell* 43, 613–623.
- Jacob, Y., Seif, E., Paquet, P.-O., and Lang, B.F. (2004). Loss of the mRNA-like region in mitochondrial tmRNAs of jakobids. *RNA* 10, 605–614.



Jones-Rhoades, M.W., Bartel, D.P., and Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. *Annual Review of Plant Biology* 57, 19–53.

Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-Wide Profiling and Analysis of Arabidopsis siRNAs. *PLoS Biol* 5, e57.

Kawamura, K. (2012). Drawbacks of the ancient RNA-based life-like system under primitive earth conditions. *Biochimie* 94, 1441–1450.

Kawano, S., Kakuta, Y., and Kimura, M. (2002). Guanine binding site of the *Nicotiana glutinosa* ribonuclease NW revealed by X-ray crystallography. *Biochemistry* 41, 15195–15202.

Kawano, S., Kakuta, Y., Nakashima, T., and Kimura, M. (2006). Crystal structures of the *Nicotiana glutinosa* ribonuclease NT in complex with nucleoside monophosphates. *J. Biochem.* 140, 375–381.

Kawata, Y., Sakiyama, F., Hayashi, F., and Kyogoku, Y. (1990). Identification of two essential histidine residues of ribonuclease T2 from *Aspergillus oryzae*. *Eur. J. Biochem.* 187, 255–262.

Keiler, K.C., Shapiro, L., and Williams, K.P. (2000). tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: A two-piece tmRNA functions in *Caulobacter*. *Proc. Natl. Acad. Sci. U.S.A.* 97, 7778–7783.

Kilchert, C., and Vasiljeva, L. (2013). mRNA quality control goes transcriptional. *Biochem. Soc. Trans.* 41, 1666–1672.

Kim, M., and Buratowski, S. (2013). Transcription Termination by RNA Polymerase II. In *Posttranscriptional Gene Regulation*, J. Wu, ed. (Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA), pp. 19–40.

Köthke, S., and Köck, M. (2011). The *Solanum lycopersicum* RNaseLER is a class II enzyme of the RNase T2 family and shows preferential expression in guard cells. *J. Plant Physiol.* 168, 840–847.

Kruszka, K., Barneche, F., Guyot, R., Ailhas, J., Meneau, I., Schiffer, S., Marchfelder, A., and Echeverría, M. (2003). Plant dicistronic tRNA-snoRNA genes: a new mode of expression of the small nucleolar RNAs processed by RNase Z. *EMBO J.* 22, 621–632.

Kühn, K., Weihe, A., and Börner, T. (2005). Multiple promoters are a common feature of mitochondrial genes in *Arabidopsis*. *Nucleic Acids Res.* 33, 337–346.



Kurata, N., Kariu, T., Kawano, S., and Kimura, M. (2002). Molecular Cloning of cDNAs Encoding Ribonuclease-related Proteins in *Nicotiana glutinosa* Leaves, as Induced in Response to Wounding or to TMV-infection. *Bioscience, Biotechnology, and Biochemistry* 66, 391–397.

Layer, G., Reichelt, J., Jahn, D., and Heinz, D.W. (2010). Structure and function of enzymes in heme biosynthesis. *Protein Sci.* 19, 1137–1161.

Leal-Klevezas, D.S., Martínez-Soriano, J.P., and Nazar, R.N. (2000). Cotranscription of 5S rRNA-tRNA(Arg)(ACG) from *Brassica napus* chloroplasts and processing of their intergenic spacer. *Gene* 253, 303–311.

Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* 23, 4051–4060.

Lee, Y.S., Shibata, Y., Malhotra, A., and Dutta, A. (2009). A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev.* 23, 2639–2649.

Lewis, B.P., Shih, I.-hung., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. *Cell* 115, 787–798.

Li, W., Zhang, P., Fellers, J.P., Friebe, B., and Gill, B.S. (2004). Sequence composition, organization, and evolution of the core Triticeae genome. *Plant J.* 40, 500–511.

Longevialle, A.F. de, Small, I.D., and Lurin, C. (2010). Nuclearly Encoded Splicing Factors Implicated in RNA Splicing in Higher Plant Organelles. *Mol. Plant* 3, 691–705.

Loss-Morais, G., Waterhouse, P.M., and Margis, R. (2013). Description of plant tRNA-derived RNA fragments (tRFs) associated with argonaute and identification of their putative targets. *Biol. Direct* 8, 6.

Luhtala, N., and Parker, R. (2012). Structure-Function Analysis of Rny1 in tRNA Cleavage and Growth Inhibition. *PLoS ONE* 7,

Lund, E., and Dahlberg, J.E. (2006). Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb. Symp. Quant. Biol.* 71, 59–66.

Machnicka, M.A., Milanowska, K., Oglou, O.O., Purta, E., Kurkowska, M., Olchowik, A., Januszewski, W., Kalinowski, S., Dunin-Horkawicz, S., Rother, K.M., et al. (2013). MODOMICS: a database of RNA modification pathways—2013 update. *Nucl. Acids Res.* 41, 262–267.



MacIntosh, G.C. (2011). RNase T2 Family: Enzymatic Properties, Functional Diversity, and Evolution of Ancient Ribonucleases. In *Ribonucleases*, A.W. Nicholson, ed. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 89–114.

MacIntosh, G.C., Hillwig, M.S., Meyer, A., and Flagel, L. (2010). RNase T2 genes from rice and the evolution of secretory ribonucleases in plants. *Mol. Genet. Genomics* 283, 381–396.

Magnusson, L.U., Farewell, A., and Nyström, T. (2005). ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol.* 13, 236–242.

Maréchal-Drouard, L., Weil, J.H., and Dietrich, A. (1993). Transfer RNAs and Transfer RNA Genes in Plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 44, 13–32.

Marquet, R., Isel, C., Ehresmann, C., and Ehresmann, B. (1995). tRNAs as primer of reverse transcriptases. *Biochimie* 77, 113–124.

Martínez de Alba, A.E., Elvira-Matelot, E., and Vaucheret, H. (2013). Gene silencing in plants: A diversity of pathways. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1829, 1300–1308.

Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* 15, 108–121.

Matera, A.G., Terns, R.M., and Terns, M.P. (2007). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* 8, 209–220.

Matsuura, T., Sakai, H., Unno, M., Ida, K., Sato, M., Sakiyama, F., and Norioka, S. (2001). Crystal structure at 1.5-Å resolution of *Pyrus pyrifolia* pistil ribonuclease responsible for gametophytic self-incompatibility. *J. Biol. Chem.* 276, 45261–45269.

McCarthy, E.M., and McDonald, J.F. (2004). Long terminal repeat retrotransposons of *Mus musculus*. *Genome Biol.* 5, R14.

Megel, C., Morelle, G., Lalande, S., Duchêne, A.-M., Small, I., and Maréchal-Drouard, L. (2015). Surveillance and Cleavage of Eukaryotic tRNAs. *Int J Mol Sci* 16, 1873–1893.

Mei, Y., Yong, J., Liu, H., Shi, Y., Meinkoth, J., Dreyfuss, G., and Yang, X. (2010). tRNA binds to cytochrome c and inhibits caspase activation. *Mol. Cell* 37, 668–678.



- Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol* *19*, 560–567.
- Michaeli, S., and Galili, G. (2014). Degradation of Organelles or Specific Organelle Components via Selective Autophagy in Plant Cells. *International Journal of Molecular Sciences* *15*, 7624–7638.
- Michaud, M., Cognat, V., Duchêne, A.-M., and Maréchal-Drouard, L. (2011). A global picture of tRNA genes in plant genomes. *Plant J.* *66*, 80–93.
- Mills, D.R., Peterson, R.L., and Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. U.S.A.* *58*, 217–224.
- Mogk, A., Schmidt, R., and Bukau, B. (2007). The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol.* *17*, 165–172.
- Morris, K.V., and Mattick, J.S. (2014). The rise of regulatory RNA. *Nat Rev Genet* *15*, 423–437.
- Muñoz, A., and Castellano, M.M. (2012). Regulation of Translation Initiation under Abiotic Stress Conditions in Plants: Is It a Conserved or Not so Conserved Process among Eukaryotes? *International Journal of Genomics* *2012*.
- Munroe, D., and Jacobson, A. (1990). Tales of poly(A): a review. *Gene* *91*, 151–158.
- Muramatsu, T., Nishikawa, K., Nemoto, F., Kuchino, Y., Nishimura, S., Miyazawa, T., and Yokoyama, S. (1988). Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature* *336*, 179–181.
- Murchison, E.P., and Hannon, G.J. (2004). miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr. Opin. Cell Biol.* *16*, 223–229.
- Natesan, S.K.A., Sullivan, J.A., and Gray, J.C. (2005). Stromules: a characteristic cell-specific feature of plastid morphology. *J. Exp. Bot.* *56*, 787–797.
- Nakagawa, A., Tanaka, I., Sakai, R., Nakashima, T., Funatsu, G., and Kimura, M. (1999). Crystal structure of a ribonuclease from the seeds of bitter melon (*Momordica charantia*) at 1.75 Å resolution. *Biochim. Biophys. Acta* *1433*, 253–260.



- Németh, A., Perez-Fernandez, J., Merkl, P., Hamperl, S., Gerber, J., Griesenbeck, J., and Tschochner, H. (2013). RNA polymerase I termination: Where is the end? *Biochim. Biophys. Acta* *1829*, 306–317.
- Novoa, E.M., and Ribas de Pouplana, L. (2012). Speeding with control: codon usage, tRNAs, and ribosomes. *Trends in Genetics* *28*, 574–581.
- Novoa, E.M., Pavon-Eternod, M., Pan, T., and Ribas de Pouplana, L. (2012). A role for tRNA modifications in genome structure and codon usage. *Cell* *149*, 202–213.
- Nowacka, M., Strozycki, P.M., Jackowiak, P., Hojka-Osinska, A., Szymanski, M., and Figlerowicz, M. (2013). Identification of stable, high copy number, medium-sized RNA degradation intermediates that accumulate in plants under non-stress conditions. *Plant Mol. Biol.* *83*, 191–204.
- Nowak, M.D., Davis, A.P., Anthony, F., and Yoder, A.D. (2011). Expression and Trans-Specific Polymorphism of Self-Incompatibility RNases in *Coffea* (Rubiaceae). *PLoS One* *6*.
- O'Donnell, K.A., and Boeke, J.D. (2007). Mighty Piwis defend the germline against genome intruders. *Cell* *129*, 37–44.
- Ohgi, K., Horiuchi, H., Watanabe, H., Iwama, M., Takagi, M., and Irie, M. (1992). Evidence that three histidine residues of a base non-specific and adenylic acid preferential ribonuclease from *Rhizopus niveus* are involved in the catalytic function. *J. Biochem.* *112*, 132–138.
- Ohno, H., and Ehara, Y. (2005). Expression of Ribonuclease Gene in Mechanically Injured or Virus-Inoculated *Nicotiana tabacum* Leaves. *Tohoku Journal of Agricultural Res.*
- Ohshima, K., and Okada, N. (2005). SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. *Cytogenet. Genome Res.* *110*, 475–490.
- Orellana, O., Cooley, L., and Söll, D. (1986). The additional guanylate at the 5' terminus of *Escherichia coli* tRNA<sup>His</sup> is the result of unusual processing by RNase P. *Mol. Cell. Biol.* *6*, 525–529.
- Oxley, D., Munro, S.L., Craik, D.J., and Bacic, A. (1996). Structure of N-glycans on the S3- and S6-allele stylar self-incompatibility ribonucleases of *Nicotiana glauca*. *Glycobiology* *6*, 611–618.
- Park, M.Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., and Poethig, R.S. (2005). Nuclear processing and export of microRNAs in *Arabidopsis*. *PNAS* *102*, 3691–3696.



- Parker, R. (2012). RNA degradation in *Saccharomyces cerevisiae*. *Genetics* *191*, 671–702.
- Parry, S., Newbiggin, E., Currie, G., Bacic, A., and Oxley, D. (1997). Identification of active-site histidine residues of a self-incompatibility ribonuclease from a wild tomato. *Plant Physiol.* *115*, 1421–1429.
- Parry, S., Newbiggin, E., Craik, D., Nakamura, K.T., Bacic, A., and Oxley, D. (1998). Structural analysis and molecular model of a self-incompatibility RNase from wild tomato. *Plant Physiol.* *116*, 463–469.
- Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. *Genes Dev.* *18*, 2368–2379.
- Pfizinger, H., Guillemaut, P., Weil, J.-H., and Pillay, D.T.N. (1987). Adjustment of the tRNA population to the codon usage in chloroplasts. *Nucl. Acids Res.* *15*, 1377–1386.
- Phizicky, E.M., and Hopper, A.K. (2010). tRNA biology charges to the front. *Genes Dev.* *24*, 1832–1860.
- Piekna-Przybylska, D., DiChiacchio, L., Mathews, D.H., and Bambara, R.A. (2010). A sequence similar to tRNA 3 Lys gene is embedded in HIV-1 U3-R and promotes minus-strand transfer. *Nat. Struct. Mol. Biol.* *17*, 83–89.
- Piekna-Przybylska, D., Dykes, C., Demeter, L.M., and Bambara, R.A. (2011). Sequences in the U3 region of human immunodeficiency virus 1 improve efficiency of minus strand transfer in infected cells. *Virology* *410*, 368–374.
- Placido, A., Sieber, F., Gobert, A., Gallerani, R., Giegé, P., and Maréchal-Drouard, L. (2010). Plant mitochondria use two pathways for the biogenesis of tRNA<sup>His</sup>. *Nucleic Acids Res.* *38*, 7711–7717.
- Pontier, D., Picart, C., Roudier, F., Garcia, D., Lahmy, S., Azevedo, J., Alart, E., Laudié, M., Karlowski, W.M., Cooke, R., et al. (2012). NERD, a Plant-Specific GW
- Protein, Defines an Additional RNAi-Dependent Chromatin-Based Pathway in *Arabidopsis*. *Molecular Cell* *48*, 121–132.
- Potuschak, T., Stary, S., Schlögelhofer, P., Becker, F., Nejnskaia, V., and Bachmair, A. (1998). PRT1 of *Arabidopsis thaliana* encodes a component of the plant N-end rule pathway. *Proc. Natl. Acad. Sci. U.S.A.* *95*, 7904–7908.



Pribnow, D. (1975). Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proc. Natl. Acad. Sci. U.S.A.* *72*, 784–788.

Rai, R., and Kashina, A. (2005). Identification of mammalian arginyltransferases that modify a specific subset of protein substrates. *Proc. Natl. Acad. Sci. U.S.A.* *102*, 10123–10128.

Raina, M., and Ibba, M. (2014). tRNAs as regulators of biological processes. *Front. Genet.* *5*, 171.  
Reifur, L., Garcia-Silva, M.R., Poubel, S.B., Alves, L.R., Arauco, P., Buiar, D.K., Goldenberg, S., Cayota, A., and Dallagiovanna, B. (2012). Distinct subcellular localization of tRNA-derived fragments in the infective metacyclic forms of *Trypanosoma cruzi*. *Mem. Inst. Oswaldo Cruz* *107*, 816–819.

Roalson, E.H., and McCubbin, A.G. (2003). S-RNases and sexual incompatibility: structure, functions, and evolutionary perspectives. *Mol. Phylogenet. Evol.* *29*, 490–506.

Ross, W., Vrentas, C.E., Sanchez-Vazquez, P., Gaal, T., and Gourse, R.L. (2013). The magic spot: a ppGpp binding site on *E. coli* RNA polymerase responsible for regulation of transcription initiation. *Mol. Cell* *50*, 420–429.

Roy, B., and Jacobson, A. (2013). The intimate relationships of mRNA decay and translation. *Trends Genet.* *29*, 691–699.

Salinas, T., Duchêne, A.-M., and Maréchal-Drouard, L. (2008). Recent advances in tRNA mitochondrial import. *Trends Biochem. Sci.* *33*, 320–329.

Samarsky, D.A., Fournier, M.J., Singer, R.H., and Bertrand, E. (1998). The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *EMBO J* *17*, 3747–3757.

Sanmiguel, P., and Bennetzen, J.L. (1998). Evidence that a Recent Increase in Maize Genome Size was Caused by the Massive Amplification of Intergene Retrotransposons. *Ann Bot* *82*, 37–44.

Sarkies, P., and Miska, E.A. (2014). Small RNAs break out: the molecular cell biology of mobile small RNAs. *Nat Rev Mol Cell Biol* *15*, 525–535.

Schramm, L., and Hernandez, N. (2002). Recruitment of RNA polymerase III to its target promoters. *Genes Dev.* *16*, 2593–2620.

Schulman, A.H. (2013). Retrotransposon replication in plants. *Current Opinion in Virology* *3*, 604–614.



- Schümperli, D., and Pillai, R.S. (2004). The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *CMLS, Cell. Mol. Life Sci.* *61*, 2560–2570.
- Schwab, R., Ossowski, S., Riester, M., Warthmann, N., and Weigel, D. (2006). Highly specific gene silencing by artificial microRNAs in Arabidopsis. *Plant Cell* *18*, 1121–1133.
- Sengupta, R., Vainauskas, S., Yarian, C., Sochacka, E., Malkiewicz, A., Guenther, R.H., Koshlap, K.M., and Agris, P.F. (2000). Modified constructs of the tRNA TPsiC domain to probe substrate conformational requirements of m(1)A(58) and m(5)U(54) tRNA methyltransferases. *Nucleic Acids Res.* *28*, 1374–1380.
- Seto, A.G., Kingston, R.E., and Lau, N.C. (2007). The coming of age for Piwi proteins. *Mol. Cell* *26*, 603–609.
- Shandilya, J., and Roberts, S.G.E. (2012). The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* *1819*, 391–400.
- Shiina, T., Tsunoyama, Y., Nakahira, Y., and Khan, M.S. (2005). Plastid RNA polymerases, promoters, and transcription regulators in higher plants. *Int. Rev. Cytol.* *244*, 1–68.
- Shrader, T.E., Tobias, J.W., and Varshavsky, A. (1993). The N-end rule in Escherichia coli: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene *aat*. *J. Bacteriol.* *175*, 4364–4374.
- Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* *12*, 246–258.
- Spieth, J., and Lawson, D. (2006). Overview of gene structure. *WormBook* 1–10.
- Stange, N., Beier, D., and Beier, H. (1991). Expression of nuclear tRNA(Tyr) genes from Arabidopsis thaliana in HeLa cell and wheat germ extracts. *Plant Mol. Biol.* *16*, 865–875.
- Stark, B.C., Kole, R., Bowman, E.J., and Altman, S. (1978). Ribonuclease P: an enzyme with an essential RNA component. *Proc Natl Acad Sci U S A* *75*, 3717–3721.
- Steinbachs, J.E., and Holsinger, K.E. (2002). S-RNase-mediated gametophytic self-incompatibility is ancestral in eudicots. *Mol. Biol. Evol.* *19*, 825–829.



- Sy, J., and Lipmann, F. (1973). Identification of the synthesis of guanosine tetraphosphate (MS I) as insertion of a pyrophosphoryl group into the 3'-position in guanosine 5'-diphosphate. *Proc. Natl. Acad. Sci. U.S.A.* *70*, 306–309.
- Tanaka, N., Arai, J., Inokuchi, N., Koyama, T., Ohgi, K., Irie, M., and Nakamura, K.T. (2000). Crystal structure of a plant ribonuclease, RNase LE. *J. Mol. Biol.* *298*, 859–873.
- Tasaki, T., Sriram, S.M., Park, K.S., and Kwon, Y.T. (2012). The N-End Rule Pathway. *Annual Review of Biochemistry* *81*, 261–289.
- Taylor, C.B., and Green, P.J. (1991). Genes with Homology to Fungal and S-Gene RNases Are Expressed in *Arabidopsis thaliana*. *Plant Physiol.* *96*, 980–984.
- Taylor, C.B., Bariola, P.A., delCardayré, S.B., Raines, R.T., and Green, P.J. (1993). RNS2: a senescence-associated RNase of *Arabidopsis* that diverged from the S-RNases before speciation. *Proc. Natl. Acad. Sci. U.S.A.* *90*, 5118–5122.
- Theologis, A., Ecker, J.R., Palm, C.J., Federspiel, N.A., Kaul, S., White, O., Alonso, J., Altfafi, H., Araujo, R., Bowman, C.L., et al. (2000). Sequence and analysis of chromosome 1 of the plant *Arabidopsis thaliana*. *Nature* *408*, 816–820.
- Thompson, D.M., and Parker, R. (2009). The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stress in *Saccharomyces cerevisiae*. *J. Cell Biol.* *185*, 43–50.
- Thompson, D.M., Lu, C., Green, P.J., and Parker, R. (2008). tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA* *14*, 2095–2103.
- Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991). The N-end rule in bacteria. *Science* *254*, 1374–1377.
- Trotta, C.R., Miao, F., Arn, E.A., Stevens, S.W., Ho, C.K., Rauhut, R., and Abelson, J.N. (1997). The yeast tRNA splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tRNA endonucleases. *Cell* *89*, 849–858.
- Trull, M.C., Guiltinan, M.J., Lynch, J.P., and Deikman, J. (1997). The responses of wild-type and ABA mutant *Arabidopsis thaliana* plants to phosphorus starvation. *Plant, Cell & Environment* *20*, 85–92.
- Ulmasov, B., and Folk, W. (1995). Analysis of the role of 5' and 3' flanking sequence elements upon in vivo expression of the plant tRNATrp genes. *Plant Cell* *7*, 1723–1734.



Ulmasov, B., Topin, A., Chen, Z., He, S.H., and Folk, W.R. (1998). Identity elements and aminoacylation of plant tRNA<sup>Trp</sup>. *Nucleic Acids Res.* *26*, 5139–5141.

Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gascioli, V., Mallory, A.C., Hilbert, J.-L., Bartel, D.P., and Crété, P. (2004). Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. *Mol. Cell* *16*, 69–79.

Vogel, A., Schilling, O., Späth, B., and Marchfelder, A. (2005). The tRNase Z family of proteins: physiological functions, substrate specificity and structural properties. *Biol. Chem.* *386*, 1253–1264.

Walters, R., and Parker, R. (2014). Quality control: Is there quality control of localized mRNAs? *J. Cell Biol.* *204*, 863–868.

Wang, C., Sobral, B.W., and Williams, K.P. (2007). Loss of a Universal tRNA Feature. *J. Bacteriol.* *189*, 1954–1962.

Wang, H., Zhang, X., Liu, J., Kiba, T., Woo, J., Ojo, T., Hafner, M., Tuschl, T., Chua, N.-H., and Wang, X.-J. (2011a). Deep sequencing of small RNAs specifically associated with Arabidopsis AGO1 and AGO4 uncovers new AGO functions. *Plant J.* *67*, 292–304.

Wang, L., Yu, X., Wang, H., Lu, Y.-Z., de Ruiter, M., Prins, M., and He, Y.-K. (2011b). A novel class of heat-responsive small RNAs derived from the chloroplast genome of Chinese cabbage (*Brassica rapa*). *BMC Genomics* *12*, 289.

Wang, X.-J., Reyes, J.L., Chua, N.-H., and Gaasterland, T. (2004). Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets. *Genome Biol.* *5*, R65.

Weber, F., Dietrich, A., Weil, J.H., and Maréchal-Drouard, L. (1990). A potato mitochondrial isoleucine tRNA is coded for by a mitochondrial gene possessing a methionine anticodon. *Nucleic Acids Res.* *18*, 5027–5030.

Wende, S., Platzer, E.G., Jühling, F., Pütz, J., Florentz, C., Stadler, P.F., and Mörl, M. (2014). Biological evidence for the world's smallest tRNAs. *Biochimie* *100*, 151–158.

Wenke, T., Holtgräwe, D., Horn, A.V., Weisshaar, B., and Schmidt, T. (2009). An abundant and heavily truncated non-LTR retrotransposon (LINE) family in *Beta vulgaris*. *Plant Mol Biol* *71*, 585–597.



Wenke, T., Döbel, T., Sørensen, T.R., Junghans, H., Weisshaar, B., and Schmidt, T. (2011). Targeted Identification of Short Interspersed Nuclear Element Families Shows Their Widespread Existence and Extreme Heterogeneity in Plant Genomes. *Plant Cell* 23, 3117–3128.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., et al. (2007). A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982.

Wilusz, J.E., Whipple, J.M., Phizicky, E.M., and Sharp, P.A. (2011). tRNAs marked with CCACCA are targeted for degradation. *Science* 334, 817–821.

Wolfe, K.H., Morden, C.W., and Palmer, J.D. (1992). Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci U S A* 89, 10648–10652.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). Genetic and Functional Diversification of Small RNA Pathways in Plants. *PLoS Biol* 2,.

Xie, Z., Allen, E., Wilken, A., and Carrington, J.C. (2005). DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12984–12989.

Xue, S., Calvin, K., and Li, H. (2006). RNA recognition and cleavage by a splicing endonuclease. *Science* 312, 906–910.

Yamasaki, S., Ivanov, P., Hu, G.-F., and Anderson, P. (2009). Angiogenin cleaves tRNA and promotes stress-induced translational repression. *J. Cell Biol.* 185, 35–42.

Yoshihisa, T., Ohshima, C., Yunoki-Esaki, K., and Endo, T. (2007). Cytoplasmic splicing of tRNA in *Saccharomyces cerevisiae*. *Genes to Cells* 12, 285–297.

Yu, Q.-B., Huang, C., and Yang, Z.-N. (2014). Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants. *Front Plant Sci* 5, 316.

Yukawa, Y., Sugita, M., Choisne, N., Small, I., and Sugiura, M. (2000). 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 J.* 22, 439–447.

Zhai, L., and Xiang, S. (2014). mRNA quality control at the 5' end. *J Zhejiang Univ Sci B* 15, 438–443.



Zhang, J., Dyer, K.D., and Rosenberg, H.F. (2000). Evolution of the rodent eosinophil-associated RNase gene family by rapid gene sorting and positive selection. *Proc. Natl. Acad. Sci. U.S.A.* *97*, 4701–4706.

Zhang, S., Sun, L., and Kragler, F. (2009). The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation. *Plant Physiol.* *150*, 378–387.

Zhang, Y.-T., Jiang, F., Tian, Z.-X., Huo, Y.-X., Sun, Y.-C., and Wang, Y.-P. (2014). CRP-Cyclic AMP Dependent Inhibition of the Xylene-Responsive  $\sigma$ <sup>54</sup>-Promoter Pu in *Escherichia coli*. *PLoS ONE* *9*.

Zhao, K., Liu, M., and Burgess, R.R. (2005). The global transcriptional response of *Escherichia coli* to induced sigma 32 protein involves sigma 32 regulon activation followed by inactivation and degradation of sigma 32 in vivo. *J. Biol. Chem.* *280*, 17758–17768.

Zhou, Y.N., Kusukawa, N., Erickson, J.W., Gross, C.A., and Yura, T. (1988). Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor sigma 32. *J. Bacteriol.* *170*, 3640–3649.





## Geoffrey MORELLE

### **Les ARN de transfert, une nouvelle source de petits ARN non-codants chez *A. thaliana*.**

Au cours de ces 10 dernières années une nouvelle classe de petits ARN non-codants nommés “tRNA-derived fragments” (tRFs) a été caractérisée. Tandis que le rôle des tRNA concernant la synthèse des protéines est bien connu, les raisons pour lesquels des fragments de tRNA s’accumulent dans la cellule restent inconnus. Actuellement, peu d’informations sont disponibles sur leurs biogenèses et leurs rôles biologiques, mais les preuves montrant leurs importances dans la régulation de l’expression des gènes augmentent régulièrement. Cependant chez les plantes, bien que l’existence des tRFs ait été montrée, peu de données sont disponibles concernant leurs natures ou leurs biogenèse.

À l’aide d’expérience de “deep-sequencing” et de northern blot nous avons identifié chez *A. thaliana*, nous avons confirmé l’existence d’une grande population en tRFs d’origine variée.

À la suite de ces observations, trois questions sont établies. Tout d’abord, quelles sont les enzymes responsables de la biogenèse des tRFs. Ensuite, où les tRFs sont générés. Enfin, est-ce que les tRFs sont des sous-produits de la dégradation des tRFs où ont-t’ils une fonction biologique ?

**Mots-clés :** tRNA, tRFs, deep-sequencing, petits ARN non-codants, endonucléase, RNS

During the last decade, a new class of small non-coding RNAs called tRNA-derived RNA fragments (tRFs) has emerged. Whilst full-length tRNAs have a well-defined role in protein synthesis, the reason(s) why stable tRNA-derived fragments remains are largely enigmatic. Indeed, the number of tRFs has rapidly increased in various evolutionary divergent eukaryotic cells, such as yeast or humans. To date, only substantial information on their biogenesis and on their biological roles is known but their importance in the regulation of gene expression and in cell life is continuously and rapidly expanding. In plants, the existence of tRFs has also been reported but only few data are available.

Using deep-sequencing and computational analysis on various small RNA libraries from *Arabidopsis thaliana* and northern blot experiments, we confirmed the existence of large but specific population of tRFs.

Following these observations, three major questions are now addressed. First what are the enzymes responsible for tRF biogenesis, second where are tRFs generated and third, are tRFs merely degradation by-products or do they have biological functions?

**Mots-clés :** tRNA, tRFs, deep-sequencing, small non-coding RNA, endonuclease, RNS