

THÈSE

Présentée et soutenue publiquement en vue de l'obtention du grade de

Docteur ès Sciences de la vie et de la santé

Filière doctorale: Régulation de l'expression des génomes

Spécialité: Biologie moléculaire

par

Franck VAZQUEZ

Analyse moléculaire de mutants affectés dans les contrôles épigénétiques post-transcriptionnels chez *Arabidopsis thaliana*.

Le 21 Octobre 2004 devant la commission d'examen composée de MM:

Vincent COLOT	Directeur de Recherche, CNRS	Président du Jury
Martin CRESPI	Directeur de Recherche, CNRS	Rapporteur
Eric MEYER	Directeur de Recherche, ENS	Rapporteur
Theo HENDRIKS	Professeur, USTL	Examineur
Jean-Louis HILBERT	Professeur, USTL	Directeur de thèse
Patrice CRÉTÉ	MCF, Univ. Méditerranée	Encadrant de thèse

TABLE DES MATIERES

TABLE DES ILLUSTRATIONS	4
Résumé.....	6
Résumé en anglais.....	7
Abréviations	8
Lexique	9
I. REVUE BIBLIOGRAPHIQUE	10
Introduction	12
1. Inactivations épigénétiques impliquant des transgènes	21
1.1. Mise en évidence des siRNAs et mécanismes d'action	22
1.1.1. Découverte des siRNAs impliqués dans les processus de PTGS	23
1.1.2. Bases moléculaires des différents processus de PTGS	24
1.1.3. Rôle de la PTGS	30
1.1.4. Bases moléculaires de la TGS induite par des siRNAs	32
1.2. L'ARN interférence.....	33
1.2.1. L'ARN interférence: un nouvel outil de laboratoire	34
1.2.2. L'ARN interférence: Un nouvel espoir pour la santé humaine?	35
2. Les micro-ARNs et les siRNAs : Des petits ARN non codants endogènes possédant des fonctions régulatrices.....	37
2.1. Les micro-ARNs: des ARNnc d'une importance majeure.....	37
2.1.1. Qu'es-ce qu'un micro-ARN?	38
2.1.2. Biosynthèse.....	41
2.1.3. Mécanisme d'action des micro-ARNs.....	48
2.1.3.1. Déterminants du mécanisme d'action.....	48
2.1.3.2. Les complexes miRNPs	51
2.1.3.2.A. Caractéristiques communes des siRNAs et des micro-ARNs.....	51
2.1.3.2.B. Composants identifiés du complexe siRNP	53
2.1.3.2.C. Formation du complexe RISC.....	55
2.1.3.2.D. Activité "slicer" du complexe RISC	57
2.1.4. Implication des micro-ARNs dans différents processus cellulaires et physiologiques.....	61
2.1.5. Régulation de l'expression des gènes de micro-ARNs?.....	68
2.2. La méthylation induite par des siRNAs endogènes.....	70
2.3. Conclusion.....	72

II. RESULTATS & DISCUSSION	74
PREMIÈRE PARTIE: Contribution à l'identification de protéines impliquées dans la voie de régulation des gènes par les micro-ARNs chez <i>Arabidopsis thaliana</i>	75
1.1. Introduction	75
1.2. Contribution à la mise en évidence de l'implication des protéines HEN1 et HYL1 dans la voie de biosynthèse des micro-ARNs et de la protéine AGO1 dans les étapes effectrices de régulation des gènes par les micro-ARNs.	76
1.2.1. Quantification des ARNm <i>uidA</i> par la technique de PCR quantitative en temps réel	77
1.2.2. Caractérisation, par la technique modifiée de 5' RLM (RNA ligase-mediated)-RACE, de sites de clivage sur des ARNm cibles potentielles de micro-ARNs.	79
1.2.3. Quantification des ARNm cibles de micro-ARNs par PCR quantitative en temps réel	85
DEUXIÈME PARTIE: ARTICLE: Endogenous trans-acting siRNAs regulate the accumulation of <i>Arabidopsis</i> mRNAs.	92
1.1. Introduction	92
1.2. ARTICLE	97
1.3. Conclusion	117
1.4. Discussion générale et Perspectives	119
III. ANNEXES	123
ARTICLE annexe N°1:	124
ARTICLE annexe N°2	134
ARTICLE annexe N°3	143
IV. RÉFÉRENCES BIBLIOGRAPHIQUES	155

TABLE DES ILLUSTRATIONS

Figures:

Figure 1: Cycle de vie de <i>T. thermophila</i>	15
Figure 2: Elimination des IES lors de la différenciation du macronoyau chez <i>T. thermophila</i>	17
Figure 3: Mécanisme d'action des ARNnc chez les procaryotes.....	19
Figure 4: Mécanisme d'inactivation épigénétique post-transcriptionnelle chez les plantes.	25
Figure 5: Mécanisme d'inactivation épigénétique transcriptionnelle guidée par des siRNAs endogènes chez les plantes	31
Figure 6: Structures de différents précurseurs de micro-ARNs.	39
Figure 7: Mécanisme de biosynthèse des micro-ARNs chez les animaux.	42
Figure 8: Mécanisme de biosynthèse des micro-ARNs chez les animaux.	44
Figure 9: Mécanisme de biosynthèse des micro-ARNs chez <i>A. thaliana</i>	46
Figure 10: Mécanismes d'action des micro-ARNs chez les animaux.....	49
Figure 11: Entrée des siRNAs dans le complexe siRNP appelé RISC chez la drosophile.....	54
Figure 12: Entrée des micro-ARNs dans le complexe miRNP appelé RISC chez la drosophile.	56
Figure 13: Modélisation du mécanisme d'action de l'activité "slicer" d'AGO2 chez l'homme.....	58
Figure 14 : GUS Activity and <i>GUS</i> mRNA Accumulation in Wild-Type and mutant plants.	78
Figure 15: Principe de la technique de 5' RLM-RACE et position des sites de clivage sur le transcrit <i>SCL6-III</i> guidé par le micro-ARN miR171.	80
Figure 16 : Diagrammatic Representation of At1g27370/ <i>SPL10</i> , At1g06580/ <i>PPR</i> , and At1g48410/ <i>AGO1</i> mRNA Cleavage Sites Determined by Modified 5' RLM-RACE.....	82
Figure 17 : <i>miR171</i> miRNA and <i>SCL</i> mRNA Accumulation in Wild Type and Mutant Plants.....	84
Figure 18 : miRNA Target mRNA Accumulation in <i>hen1</i> , <i>hyl1</i> , and <i>dcl1</i> mutants	86
Figure 19 : <i>ago1</i> , <i>hen1</i> , and <i>hyl1</i> mutants have increased steady-state levels of miRNA targets.	88
Figure 20 : Silent mutations in the <i>miR168</i> complementary site of the <i>AGO1</i> mRNA induce developmental defects.	90
Figure 21: Profil de cDNA- AFLP.....	94
Figure 22: Mécanismes de biosynthèse et d'action des petits ARNnc endogènes régulateurs de l'expression des gènes au niveau post-transcriptionnel chez <i>A. thaliana</i>	118

Tableaux:

Tableau 1: Comparaison de la taille des génomes d'organismes modèles et de la proportion de région non codante de ces génomés.....	11
Tableau 2: Classification des ARN selon leur fonction et exemples d'ARN non codants.	13
Tableau 3: Protéines impliquées dans les processus d'inactivation épigénétique post-transcriptionnelle chez différentes espèces.....	28
Tableau 4: micro-ARNs impliqués dans différents processus cellulaires et physiologiques chez les animaux.	60
Tableau 5: micro-ARNs impliqués dans différents processus cellulaires et physiologiques chez les plantes.	64
Tableau 6: Cibles de micro-ARNs dont le clivage a été validé par des expériences de 5'RACE.....	67
Tableau 7: Gènes correspondants aux séquences obtenues pour les couples d'amorces de cDNA AFLP présentant une bande plus intense chez les mutants <i>sgs2/rdr6</i> et <i>sgs3</i> comparé à la plante sauvage.....	95

REMERCIEMENTS

Je tiens à remercier en premier lieu les Professeurs Jacques Vasseur et Jean-Louis Hilbert pour leur accueil au sein du laboratoire de Physiologie de la Différenciation végétale, ainsi que pour les conditions dans lesquelles ces travaux ont pu être réalisés.

Un immense merci à Patrice Créte, pour tout. D'abord, pour m'avoir transmis sans compter, tout au long du DEA et de la thèse. Egalement, pour avoir transformé notre relation de travail en amitié sincère.

Encore un grand merci à Hervé Vaucheret, pour l'accueil au sein de son équipe lors de mes venues à l'INRA de Versailles, pour la collaboration étroite que Patrice et moi avons eu avec lui au cours de ma thèse. Merci également pour sa grande générosité.

Je remercie également les membres de son équipe pour leur accueil et leur sympathie. Un merci particulier à Taline Elmayan pour son soutien.

Un remerciement spécial aux amitiés qui ont émergées au fil des rencontres: David Windels, Sylvain Legrand, Philippe Hance, Chris Lepers, p'tite Sophie Alvarez, Clarita Vanappelghem, Eugénie Fournier, quelques membres du Journal club: Mélissanne DeWispelaere, Frédérique Potet, Sophie Germann, Valérie Gaudin et Hélène Lucas.

Un grand merci à ma famille pour le soutien qu'ils m'ont apporté et plus particulièrement à mon amoureuse pour avoir oublié de regarder sa montre pendant trois ans...

Merci ...

Franck

A handwritten signature in black ink, appearing to be 'F. U. 2006', written in a cursive style.

Résumé

L'inactivation épigénétique post-transcriptionnelle correspond à la dégradation séquence-spécifique d'ARNm, guidée par des petits ARN non codants complémentaires, appelés siRNAs (pour small interfering RNAs), associés à un complexe protéique nommé RISC (pour RNA-Induced Silencing Complex). Parmi les inducteurs de la PTGS, on trouve des transgènes sens (S-PTGS), des transgènes en répétition inverse (IR-PTGS) et les virus.

L'observation que tous les mutants d' *A. thaliana*, affectés dans la S-PTGS, présentaient des anomalies dans le développement foliaire, nous a conduit à faire l'hypothèse qu'il devait exister, chez la plante sauvage, une régulation de type PTGS de gènes assurant un développement foliaire normal.

Une comparaison des transcriptomes d'une plante sauvage et de deux mutants affectés dans la S-PTGS a permis l'identification d'une nouvelle classe de petits ARN endogènes non codants, correspondant à des siRNAs. Ces siRNAs sont produits par le gène At2g27400. Nous avons partiellement caractérisé la voie de régulation de l'expression des gènes par ces siRNAs endogènes. Comme pour les micro-ARNs, autres petits ARNs endogènes non codants, les protéines DCL1, HEN1 et HYL1 sont nécessaires à la production des siRNAs endogènes, et la protéine AGO1 est essentielle au processus de clivage des ARNm ciblés par ces siRNAs. De plus, la production des siRNAs endogènes requiert les activités d'une ARN polymérase ARN dépendante, RDR6, et d'une protéine de fonction inconnue, SGS3. RDR6 utiliserait donc l'ARN de At2g27400 pour produire un long ARN double brin, clivé par DCL1 en de multiples siRNAs d'environ 21 nucléotides. Nous avons montré que ces siRNAs sont capables de guider le clivage d'ARNm dont les protéines sont de fonction encore inconnue.

L'identification de cette nouvelle voie de régulation de l'expression des gènes par des siRNAs endogènes apporte une nouvelle dimension aux régulations post-transcriptionnelles chez les végétaux.

Résumé en anglais

Posttranscriptional gene silencing (PTGS) is a sequence specific mRNA degradation process guided by complementary non coding small RNAs termed siRNAs (small interfering RNAs) incorporated into a protein complex termed RISC (for RNA-Induced Silencing Complex). Among PTGS inducers are sense transgenes (S-PTGS), transgenes with inverted repeats of the same gene (IR-PTGS) and viruses.

Observations that all *A. thaliana* mutants impaired in S-PTGS exhibit leaf development abnormalities, has lead to the hypothesis that certain genes implied in leaf development are regulated by a PTGS-like phenomenon in wild-type plants.

Comparison of the transcriptomes of a wild-type plant and of two mutants impaired in S-PTGS allowed the identification of a new class of endogenous non-coding regulatory RNAs, corresponding to siRNAs. These siRNAs are produced from the At2g27400 gene. We partially characterized the pathway of gene regulation by these endogenous siRNAs. As for microRNAs, an other class of small endogenous non-coding RNAs, DCL1, HEN1 and HYL1 proteins are required for production of these endogenous siRNAs, and AGO1 is essential for cleavage of siRNAs-targeted mRNAs. In addition, the production of these siRNAs requires the activities of RDR6, an RNA-directed RNA polymerase and of SGS3, a protein with unknown function. RDR6 seems to use At2g27400 RNA as a template for the synthesis of a long double-stranded RNA which is subsequently cleaved by DCL1 in multiple siRNAs of about 21 nucleotides. It is shown that these siRNAs are able to direct the cleavage of complementary mRNAs which encode proteins with unknown functions. The identification of this silencing pathway adds yet another dimension to posttranscriptional mRNA regulation in plants.

Abréviations

Abréviations françaises:

ADN: Acide DésoxyRiboNucléique

ARN: Acide RiboNucléique

ARNab: ARN aberrant

ARNdb: ARN double brin

ARNnc: ARN non codant

ARNsb: ARN simple brin

Complexe RNP: Complexe RiboNucléoProtéique

miR: micro-ARN

Abréviations anglaises:

ESTs: Expressed Sequence Tags

cDNA AFLP: cDNA Amplified Fragment Length Polymorphism.

NLS: Nuclear Localization Sequence

ORF: Open Reading Frame

PCR: Polymerase Chain Reaction

Pre-miR: precursor for microRNAs

Pri-miR: primary precursor for microRNAs

PTGS: Post-Transcriptional Gene Silencing

RdDM: RNA directed DNA Methylation

RdRP: RNA dependent RNA Polymerase

RISC: RNA Induced Silencing Complex

RNAi: RNA-mediated interference

siRNAs: small interfering RNAs

stRNAs: small temporal RNAs

TGS: Transcriptional Gene Silencing

VIGS: Virus Induced Gene Silencing

Lexique

Cosuppression: Processus d'inactivation épigénétique post-transcriptionnelle induit par l'introduction de transgènes et conduisant à une dégradation concomitante des ARNm du transgène et du gène endogène homologue.

Épigénétique: Étude des changements dans l'expression des gènes, héritable en mitose et/ou en méiose, ne pouvant être expliqués par des changements dans leur séquence nucléotidique.

Micro-ARN: ARN non codant endogène de 21-22 nucléotides, généralement conservé entre les espèces, produit à partir d'un transcrit replié en une structure en épingle à cheveux et régulant des gènes en trans.

PTGS: Nom générique donné aux processus d'inactivation épigénétique post-transcriptionnelle conduisant à la dégradation des ARNm.

Quelling: Nom donné au processus de PTGS chez les Champignons filamenteux.

siRNAs: Petits ARN double brin produits par découpage successif, tous les 21 nucléotides, d'un long ARN double brin.

TGS: Nom générique donné aux processus d'inactivation épigénétique transcriptionnelle conduisant à la formation d'un état transcriptionnel inactif.

RdDM: Méthylation de l'ADN guidée par des siRNAs.

-

I. REVUE BIBLIOGRAPHIQUE

-

	Organisme	Taille du génome (Mb)	Nombre de gènes prédits en milliers	% non codant du génome ¹⁰
Procaryotes	<i>Escherichia coli</i> O157:H7 ¹	5,5	5,4	16
Eucaryotes	<i>Saccharomyces cerevisiae</i> ²	12,1	5,9	29,5
	<i>Caenorhabditis elegans</i> ³	97	18,5	73
	<i>Arabidopsis thaliana</i> ⁴	125	25,5	71,2
	<i>Drosophila melanogaster</i> ⁵	180	13,6	86,6
	<i>Oryza sativa ssp. Japonica</i> ⁶	420	50	ND
	<i>Zea mays</i> ⁷	2 500	50	ND
	<i>Homo sapiens</i> ^{8,9}	3 200	39,1	98

¹Hayashi *et al.* 2001; ²Goffeau *et al.* 1996; ³The *C. elegans* Sequencing Consortium 1998; ⁴Rouzé *et al.* 1999; ⁵Adams *et al.* 2000; ⁶Goff *et al.* 2002; ⁷Walbot et Petrov 2001; ⁸Lander *et al.* 2001; ⁹Venter *et al.* 2001; ¹⁰Szymanski *et al.* 2003

Tableau 1 : Comparaison de la taille des génomes d'organismes modèles et de la proportion de région non codante de ces génomes.

Introduction

Depuis les découvertes de l'ADN comme support de l'information génétique (Watson et Crick 1953), du décryptage du code génétique (Nirenberg et Matthaei 1961), et de la régulation de l'opéron lactose, archétype de la régulation génique (travaux pour lesquels Jacob, Monod et Lwoff ont partagé le prix Nobel en 1965) au milieu du 20^{ème} siècle, il a été largement admis que le terme de gène ou de transcrit était synonyme de production de protéines et qu'hormis les gènes d'ARN structuraux (et parfois catalytiques) tels que les ARN de transferts, les ARN ribosomiaux, ou les ARN présents dans les spliceosomes (tous dédiés à la production de protéines), les gènes étaient purement dédiés au codage de protéines. Cette notion dérive directement du dogme central de la génétique énonçant le flux de l'information génétique de l'ADN vers les protéines en passant par un intermédiaire: l'ARNm.

Grâce au séquençage de nombreux génomes procaryotes et eucaryotes et aux très nombreuses données d'expression des gènes, ce dogme a été remis en cause. Chez les Procaryotes, la majorité du génome est composée de séquences codantes juxtaposées les unes aux autres avec leurs éléments 5' et 3' de régulation en *cis*. Cependant, il a été mis en évidence récemment que ces génomes procaryotes, malgré leurs petites tailles et la faible proportion non codante du génome, contiennent plusieurs gènes permettant la production d'ARN ne codant pas une protéine: on parle d'ARN non codants (ARNnc). Ceux-ci ne représenteraient que 0,5% des gènes bactériens.

Les génomes des organismes eucaryotes sont, quant à eux, beaucoup plus complexes que ceux des Procaryotes (Tableau 1). Leurs tailles dépassent d'un facteur 20 à 500 celles des Procaryotes alors que le nombre de leurs séquences codantes augmenterait au maximum d'un facteur 10. Le génome de *Saccharomyces cerevisiae*, probablement le plus petit génome eucaryote, est 260 fois plus petit que celui de l'homme et il ne posséderait que 6 fois moins de gènes, contrairement à ce que l'on pensait avant l'annotation du génome humain (Liang *et al.* 2000). Cette observation est également valable chez les végétaux supérieurs dont la taille des génomes peut atteindre la taille de celui de l'homme, voire la dépasser (homme: 3.200Mb; blé: 16.000Mb).

ARN CODANTS	ARN NON CODANTS	
	sans fonction régulatrice	fonction régulatrice
<p>ARNm codent les protéines</p>	<p>ARNt Transfert des acides aminés durant la traduction</p> <p>ARNr Composants des ribosomes</p> <p>ARNsn Composants des spliceosomes</p> <p>ARNsno Modification des ARN (Ψ Urydilation)</p> <p>ARN de télomérase Composant des Télomérases</p> <p>ARN RNaseP Maturation des extrémités 5' des pré-ARNt</p> <p>ARN 4,5S Export de protéines chez les Procaryotes</p> <p>ARN 7SL Export de protéines chez les Eucaryotes</p> <p>ARNtm Trans-traduction</p> <p>ARNhY Particule Ro</p> <p>ARN RNase MRP Dégradation des ARN</p>	<p>Régulateurs transcriptionnels :</p> <p>XIST(16,5kb chez l'homme)¹ <i>Air</i> (100kb chez l'homme)² <i>roX1</i> et <i>roX2</i> chez la drosophile (3,7kb et 0,5kb)³ <i>Nespas</i> chez la souris (1,4kb)⁴ <i>Ube3A-ATS</i> chez l'homme (460kb)⁵ <i>HBA2 antisens</i> chez l'homme (1,5kb)⁶ <u>siRNAs chez les plantes</u> (25bp)⁷ <i>scanRNAs</i> chez les Ciliés (21-28bp)^{8, 21}</p> <p>Régulateurs post-transcriptionnels :</p> <p><u>Micro-ARNs</u> chez les Eucaryotes multicellulaires (21-23nt)⁹ <i>siRNAs Su(Ste)</i> chez la drosophile (25-27nt)²² <i>RyhB</i> chez <i>E. coli</i> (80nt)¹⁰</p> <p>Régulateurs traductionnels :</p> <p><i>MicF</i> chez <i>E. coli</i> (93nt)¹¹ <i>DsrA</i> chez <i>E. coli</i> (85nt)¹² <i>BC1</i> chez la souris (152nt)¹³</p> <p>Régulateurs de localisation cellulaire :</p> <p>Localisation des ARNm <i>Xlsirt</i> chez le xénope (80nt)¹⁴</p> <p>Localisation de protéines <i>hsr-Ω-c</i> (474nt chez la drosophile)¹⁵ <i>enod40</i> (700nt chez <i>Medicago Trunculata</i>)¹⁶</p> <p>Modulateurs de fonction protéique :</p> <p><i>smRNA NRSE</i> chez l'homme (21bp)¹⁷ <i>OxyS</i> (109nt chez <i>E. coli</i>)¹⁸ <i>Rsa1</i> (875nt chez l'homme)¹⁹ <i>ARN 7SK</i> (328nt chez l'homme)²⁰</p>

¹Plath *et al.* 2002; ²Sleutels *et al.* 2002; ³Andersen and Panning 2003; ⁴Ogawa and Lee 2002; ⁵Yamasaki *et al.* 2003; ⁶Tufarelli *et al.* 2003; ⁷Xie *et al.* 2004; ⁸Mochizuki and Gorovsky 2004; ⁹Bartel 2004; ¹⁰Masse *et al.* 2003; ¹¹Andersen and Delihis 1990; ¹²Majdalani *et al.* 1998; ¹³Wang *et al.* 2002; ¹⁴Kloc *et al.* 1993; ¹⁵Hogan *et al.* 1995; ¹⁶Sousa *et al.* 2001; ¹⁷Kuwabara *et al.* 2004; ¹⁸Altuvia *et al.* 1998; ¹⁹Lanz *et al.* 1999; ²⁰Murphy *et al.* 1984; ²¹Garnier *et al.* 2004; ²²Aravin *et al.* 2001.

Tableau 2 : Classification des ARN selon leur fonction et exemples d'ARN non codants.

On pensait encore récemment que ces séquences non codantes, ou séquences intergéniques, n'étaient présentes que pour espacer les gènes, structurer l'ADN ou encore que ces séquences faisaient partie des éléments de régulation des gènes. Cependant, la majorité des séquences non codantes contenues dans les génomes eucaryotes sont transcrites et conduisent donc à la production d'ARNnc. Il a été estimé que 97% du contenu en transcrit d'une cellule humaine consiste en de l'ARNnc (Mattick 2001). Cette estimation est basée sur l'observation que les introns (non codants) représentent 95% des transcrits primaires (Venter *et al.* 2001; Lander *et al.* 2001) et sur de nombreuses études suggérant qu'une grande proportion des transcrits, atteignant probablement 50% du transcriptome, correspond à des transcrits non codants (Wong *et al.* 2001). Le nombre de gènes répertoriés chez la levure, l'homme et les plantes n'est pas si différent. Cependant, la très forte augmentation du nombre de séquences non codantes présentes au sein de leurs génomes s'explique très aisément grâce aux données récentes sur les ARNnc. Il semble en effet exister deux niveaux de programmation génétique chez les organismes eucaryotes. Le premier est la programmation des composants fonctionnels nécessaires au fonctionnement de la cellule, représenté par les protéines. Le nombre des gènes varie peu d'un organisme à l'autre puisque ce niveau de programmation est conservé chez tous les organismes. Le second niveau de programmation est l'orchestration de l'expression de ces protéines, de leur assemblage ou de leur destruction au cours des processus de différenciation et de développement. Ce niveau de régulation de l'expression des protéines serait essentiellement réalisé par les ARNnc. Il est donc facile de ce point de vue de comprendre l'augmentation du nombre de ces séquences non codantes. En effet, il semble largement plus complexe d'orchestrer l'ensemble des processus développementaux nécessaires à la formation d'un organisme humain fonctionnel que ceux nécessaires à la vie d'une levure. L'orchestration de l'expression des protéines par les ARNnc peut être réalisée à différents niveaux (Tableau 2).

Ce n'est que très récemment que les ARN non codants régulateurs ont commencé à être étudiés de façon systématique. En effet, il a d'abord fallu quelques exemples sporadiques d'études présentant leur rôle dans la régulation de l'expression des gènes et les impliquant dans des processus

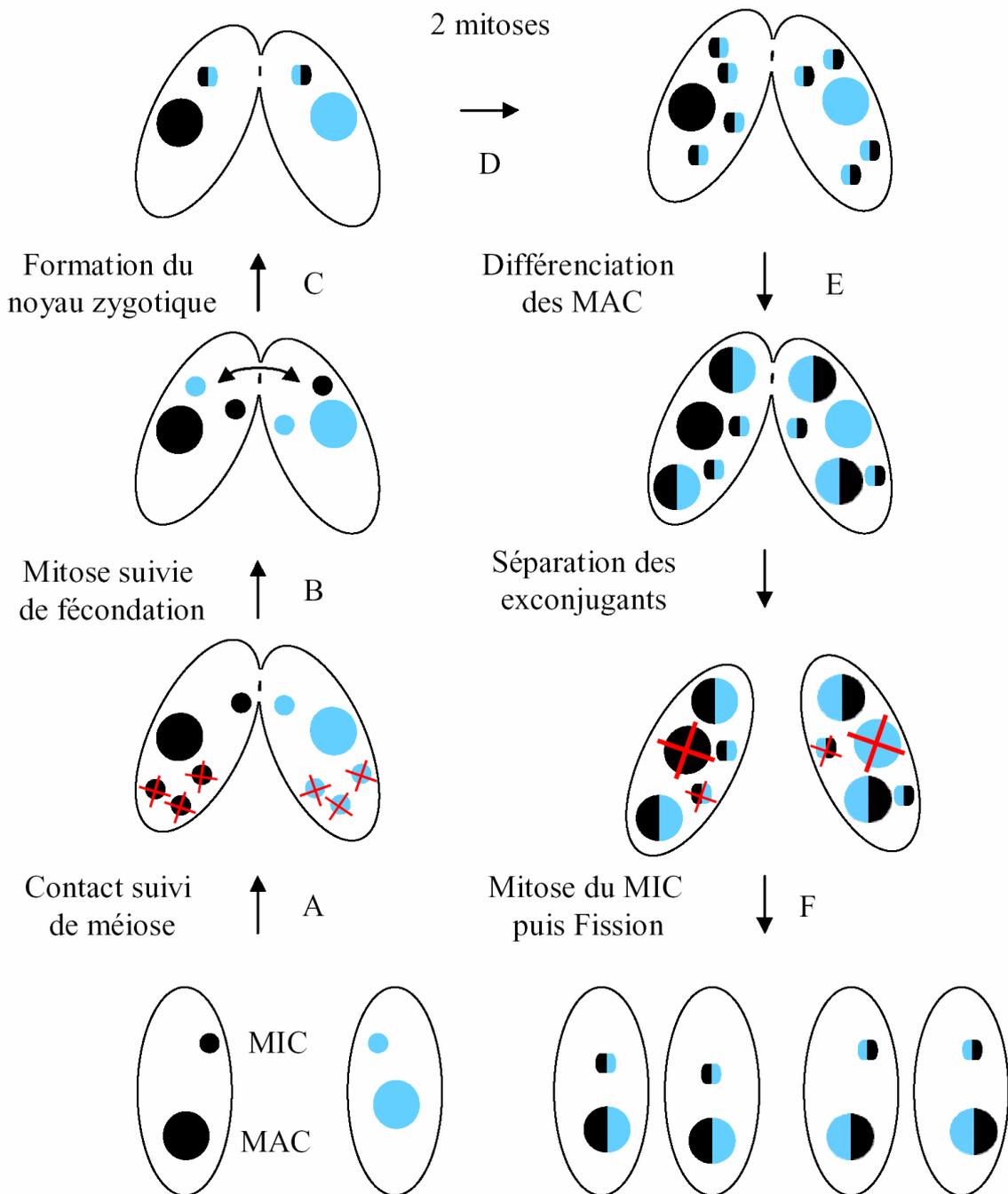


Figure 1 : Cycle de vie de *T. thermophila*.

Les Ciliés possèdent deux noyaux : le micronoyau (MIC) et le macronoyau (MAC). Ce dernier est le seul à être transcrit pendant la vie de la cellule. Après échanges des noyaux reproducteurs (A et B) et formation des noyaux zygotiques (C et D), deux noyaux néoformés sur les quatre obtenus par mitoses vont se différencier en macronoyau avec l'aide de l'ancien MAC encore présent €. Ce dernier dégénèrera après différenciation des nouveaux macronoyaux (F).

physiologiques et développementaux. Dès lors, de nombreux exemples de transcrits non codants possédant des fonctions régulatrices ont été décrits chez divers organismes (Tableau 2). Ces ARNnc ont des tailles très différentes allant actuellement de 21 nucléotides jusqu'à plusieurs milliers de nucléotides. Ces ARNnc régulent l'expression des gènes à différents niveaux, allant du réarrangement du génome, à la régulation de transcription, de la traduction ou de la localisation sub-cellulaire d'ARNm ou de protéines.

Ainsi, chez les protozoaires ciliés, et plus particulièrement chez *Tetrahymena thermophila*, un processus de réarrangement chromosomique a récemment été associé à des ARNnc.

T. thermophila, comme la plupart des Ciliés, possède deux noyaux différents d'un point de vue structure et fonction au sein d'une unique cellule. Son micronoyau (MIC), à fonction reproductrice, est diploïde, alors que son macronoyau (MAC) permettant l'expression des gènes est polyplöïde. Ces deux noyaux dérivent d'un même noyau zygotique initial, formé par le processus sexuel qu'est la conjugaison (Figure 1).

Avant la conjugaison et afin de transmettre un micronoyau haploïde à l'exconjuguant, le micronoyau subit la méiose (Figure 1A). Des quatre micronoyaux ainsi formés, un seul reste dans la cellule, les trois autres dégénèrent. Une copie du micronoyau haploïde obtenue par une mitose est transmise à la cellule conjuguant, et réciproquement (Figure 1B). Ces micronoyaux fusionnent pour former le noyau zygotique (Figure 1C) qui donnera naissance aux deux noyaux différents: micronoyau et macronoyau.

Pour cela, après séparation physique des exconjuguants, ce micronoyau zygotique diploïde subit 2 mitoses successives (Figure 1D). Ainsi, cette cellule en cours de différenciation possède 4 micronoyaux diploïdes identiques et un macronoyau en cours de dégénérescence. Les deux micronoyaux placés à l'avant de la cellule vont se différencier en deux nouveaux macronoyaux (Figure 1E), et un seul des deux micronoyaux restant va subsister. Après une mitose de ce nouveau micronoyau, chaque exconjuguant pourra donc se diviser pour donner deux cellules filles (karyonides) possédant chacune un micronoyau et un macronoyau (Figure 1F).

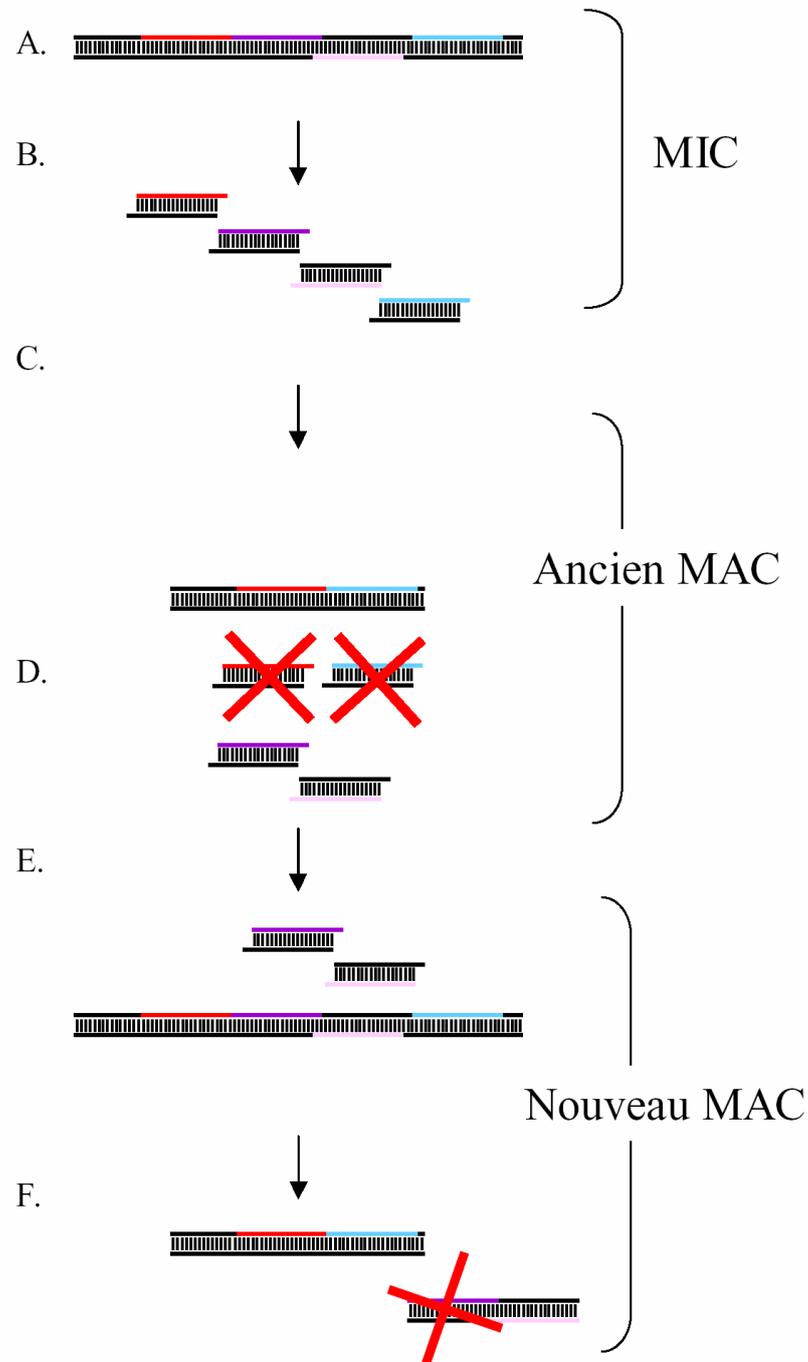


Figure 2 : Elimination des IES lors de la différenciation du macronoyau chez *T. thermophila*.

L'ensemble du génome est transcrit de façon bidirectionnelle à partir du micronoyau (MIC) et donne naissance à de longs ARNdb (A). Le découpage successif des ARNdb par une RNase III de type DICER produirait les scanRNAs (B). Les scanRNAs rejoindraient l'ancien MAC (C) et tous les scanRNAs possédant une homologie avec les séquences ADN du macronoyau seraient éliminés (D). Les scanRNAs restant correspondant aux séquences à éliminer, rejoignent le micronoyau en association avec la protéine TWI1 (E). Les séquences génomiques homologues aux scanRNAs seraient ainsi éliminées et formant le nouveau macronoyau (F).

Au cours de la différenciation du macronoyau, les chromosomes vont subir deux types de réarrangements: des délétions et une fragmentation des chromosomes. Les séquences déletées, nommées IES (Internally Eliminated Sequences) sont éliminées de façon reproductible. Le modèle actuel d'élimination des séquences IES repose sur la mise en évidence de petits ARN, nommés scanRNAs ou scnRNAs, d'environ 28 nucléotides (Mochizuki et Gorovsky 2004) (Figure 2).

Alors que pendant le cycle végétatif de *T. termophila*, tous les transcrits sont produits à partir du macronoyau, au cours de la conjugaison l'ensemble du génome serait transcrit de façon bidirectionnelle à partir du micronoyau (Figure 2A) (Chalker et Yao 2001). Les scanRNAs seraient produits par découpage successif des ARNdb ainsi formés, par une RNase III de type DICER (en français "coupeur") à identifier (Figure 2B). Après avoir rejoint l'ancien macronoyau (Figure 2C), la protéine TWI1, spécifiquement produite durant la conjugaison et capable d'interagir avec les scanRNAs, permettrait d'éliminer tous les scanRNAs possédant une homologie avec les séquences ADN du macronoyau (Figure 2D) (Mochizuki et Gorovsky 2004). Les scanRNAs restants, correspondants aux séquences à éliminer au sein du nouveau macronoyau, rejoignent le macronoyau en association avec la protéine TWI1 (Figure 2E) (Mochizuki *et al.* 2002; Yao *et al.* 2003; Mochizuki et Gorovsky 2004). Dans le nouveau macronoyau, les séquences génomiques homologues aux scanRNAs seraient ainsi éliminées (Figure 2F).

De la même façon, chez *Paramecium tetraurelia*, le processus de réarrangement chromosomique est lié à la présence de scanRNAs d'environ 21-22nt (Garnier *et al.* 2004). La différence de taille des scanRNAs décrits chez ces deux espèces pourrait refléter une différence de spécificité de clivage par la protéine DICER de chaque espèce (Garnier *et al.* 2004). Les processus de réarrangement décrits chez ces deux espèces semblent cependant très proches.

Les processus de compensation du dosage en gènes chez la drosophile et les mammifères sont également dépendants d'ARNnc (Tableau 2). Ces processus permettent d'équilibrer le dosage des gènes transcrits à partir du chromosome X chez les mâles et les femelles. Chez la drosophile, ce processus fait intervenir deux ARNnc, *roX1* et *roX2*. Ces ARNnc, intégrés dans un complexe

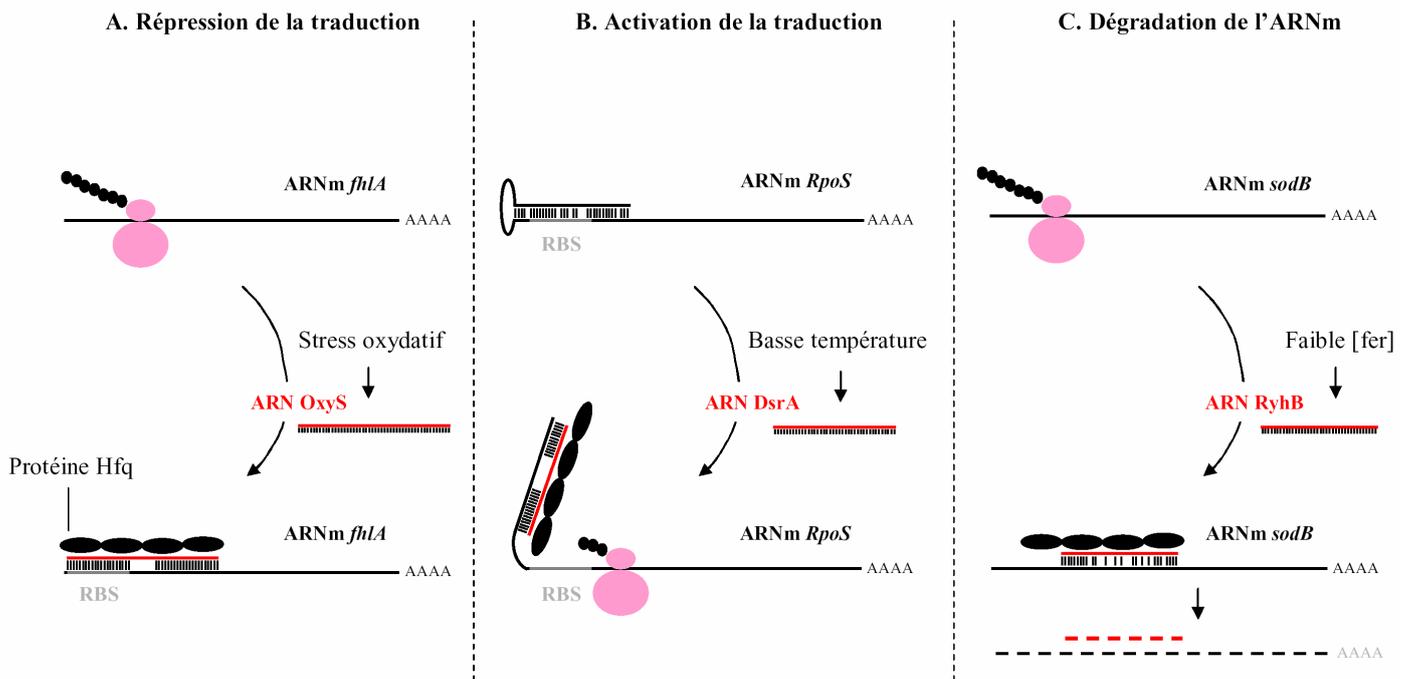


Figure 3 : Mécanisme d'action des ARNnc chez les procaryotes.

Chez les procaryotes, la majorité des ARNnc sont transcrits dans des conditions particulières et tous les ARNnc connus agissent par appariement avec l'ARNm complémentaire au sein d'un complexe protéique formé de la protéine Hfq qui va alors :

- A. masquer le site d'entrée du ribosome (RBS) et affecter la traduction de l'ARNm correspondant.
- B. démasquer le site d'entrée du ribosome par compétition d'appariement avec une zone complémentaire de l'ARNm.
- C. induire le recrutement d'une activité nucléolytique conduisant à la dégradation concomitante des deux brins du duplex formé.

nommé Male-Specific Lethal (MSL), interagissent avec le chromosome X au niveau de sites d'entrée localisés le long du chromosome X. La fixation de ce complexe ribonucléoprotéique MSL permet l'acétylation de la lysine 16 des histones H4 du chromosome. Ceci entraîne un relâchement de la chromatine et conduit, chez la drosophile mâle, à un doublement du taux de transcription de chacun des gènes portés par son chromosome X.

A l'inverse, chez les mammifères la compensation du dosage en gènes se produit chez les femelles et conduit à l'inactivation transcriptionnelle de l'un des deux chromosomes X (Plath *et al.* 2002). Cette inactivation de l'X est un processus aléatoire où les chromosomes paternels et maternels ont la même probabilité d'être inactivés. Tout comme chez la drosophile, un complexe ribonucléoprotéique se propageant le long du chromosome X jusqu'à le recouvrir entièrement est responsable de cette compensation. L'ARNnc impliqué dans ce processus est nommé XIST chez l'homme et *Xist* chez la souris (Plath *et al.* 2002).

Chez les procaryotes, les ARNnc régulateurs décrits (Tableau 2) sont principalement transcrits dans des conditions de cultures particulières (Storz *et al.* 2004). Suite à leur transcription, ces ARNnc, associés à un complexe protéique, peuvent s'apparier avec un ARNm complémentaire (Figure 3). Cet appariement conduit selon les cas à l'obstruction (Figure 3A) ou à la libération du site d'entrée des ribosomes (RBS) (Figure 3B), mais peut également conduire à la dégradation de l'ARNm (Figure 3C). Plusieurs études récentes ont permis de mettre en évidence l'existence d'au moins 40 ARNnc exprimés dans des conditions particulières (Argaman *et al.* 2001; Rivas *et al.* 2001; Wassarman *et al.* 2001).

Les ARNnc peuvent également agir en modulant le fonctionnement de protéines, comme dans le cas des "small modulatory RNA" (smRNA) dont la production est essentielle pour la différenciation des cellules neuronales (Kuwabara *et al.* 2004).

D'autres ARNnc sont capables de guider la localisation sub-cellulaire d'ARNm ou de protéines (Tableau 2). Par exemple, le transcrit du gène *enod40*, codant deux peptides très courts dont les fonctions sont inconnues, est impliqué dans la formation de nodules racinaires lors de symbioses avec certaines espèces de *rhizobium*. Durant le processus de nodulation, la protéine MtRBP1 (*Medicago truncatula* RNA Binding Protein1), normalement localisée dans le noyau, est exportée vers le cytoplasme en corrélation directe avec son interaction avec le transcrit *enod40* (Campalans *et al.* 2004).

Les quelques processus de régulation génique par des ARNnc endogènes évoqués ci-dessus ne sont qu'un échantillon non exhaustif de l'ensemble des régulations réalisées par les ARNnc identifiés à ce jour.

Les régulations géniques par des ARNnc régulateurs, tels que les micro-ARNs, sont actuellement largement étudiées car essentielles pour que les programmes développementaux des organismes eucaryotes soient réalisés de façon correcte. De plus, les micro-ARNs sont les premiers ARNnc phylogénétiquement conservés. D'autres ARNnc régulateurs, les petits ARN interférants ou "small interfering RNAs" (siRNAs) endogènes sont également très étudiés chez les plantes et sont impliqués dans le maintien de l'intégrité du génome et de l'état inactif des rétroéléments.

La mise en évidence des processus de régulations par les siRNAs et les micro-ARNs, n'a été possible que grâce à la compréhension des processus d'inactivation épigénétique transcriptionnelle (ou TGS pour Transcriptional Gene Silencing) et post-transcriptionnelle (ou PTGS pour Post-Transcriptional Gene Silencing) impliquant des transgènes.

Durant ma thèse je me suis intéressé à ces deux types d'ARNnc régulateurs, je vais donc présenter de façon plus approfondie ce que l'on connaît des siRNAs et des micro-ARNs.

1. Inactivations épigénétiques impliquant des transgènes

1.1. Mise en évidence des siRNAs et mécanismes d'action

Les premières études ayant révélé le processus d'inactivation épigénétique post-transcriptionnelle ont été réalisées chez les plantes au début des années 90. Simultanément, les équipes de Mol et de Jorgensen ont rapporté chez le pétunia, que l'introduction de copies supplémentaires du gène de la chalcone synthase A (*chsA*) (dont l'enzyme est impliquée dans la production de flavonoïdes responsables de la pigmentation des fleurs) conduisait à une dépigmentation totale ou partielle des fleurs (Napoli *et al.* 1990; van der Krol *et al.* 1990). Ces études montrèrent que cette extinction de la production de flavonoïdes due à l'inactivation du gène endogène et du transgène introduit, nommée co-suppression, se faisait sans changement du niveau de transcription du gène *chsA* endogène. La co-suppression correspondait donc à une dégradation spécifique et simultanée des ARNm produits par le gène endogène et par le transgène. Ils montrèrent que le phénotype observé co-ségrégait avec la présence du transgène, mais que la présence du transgène n'était pas suffisante pour déclencher la co-suppression (Napoli *et al.* 1990; van der Krol *et al.* 1990). L'épigénétique est définie comme "l'étude des changements dans l'expression des gènes, héritable en mitose et/ou en méiose, ne pouvant être expliqués par des changements dans la séquence nucléotidique" (Riggs *et al.* 1996), la co-suppression est donc également nommée inactivation épigénétique post-transcriptionnelle ou "Post-Transcriptional Gene Silencing" (PTGS). Des processus agissant de façon similaire, c'est-à-dire par dégradation spécifique d'un transcrit donné, ont ensuite été décrits chez de très nombreux organismes, tels que des protozoaires (*Trypanosoma brucei* (Ngo *et al.* 1998), *Plasmodium falciparum* (McRobert et McConkey 2002), *Toxoplasma gondii* (Al-Anouti et Ananvoranich 2002)), des invertébrés (*Caenorhabditis elegans* (Fire *et al.* 1991), *Drosophila melanogaster* (Pal-Bhadra *et al.* 1997)), des champignons (*Neurospora Crassa*, *Schyzosaccharomyces pombe*) (Cogoni *et al.* 1996), des algues (*Chlamydomonas reinhardtii*) (Wu-Scharf *et al.* 2000) ou encore des mammifères (*Mus musculus* (Elbashir *et al.* 2001a), *Homo sapiens sapiens* (Wianny et Zernicka-Goetz 2000)). Certains acteurs de ces processus sont conservés d'un point de vue phylogénétique, indiquant une origine ancestrale commune de tous ces processus

(Cogoni et Macino 1999a; Cogoni et Macino 2000; Dalmay *et al.* 2000b; Fagard *et al.* 2000; Mourrain *et al.* 2000; Smardon *et al.* 2000).

Ces processus de PTGS peuvent être déclenchés par différents types d'inducteurs, tels qu'un transgène en orientation antisens ou encore en répétition inversée (IR) du même gène (Fagard et Vaucheret 2000; Wesley *et al.* 2001). Ce dernier type de transgène induit le processus de PTGS de façon beaucoup plus efficace (Wesley *et al.* 2001). De nombreuses hypothèses concernant la base moléculaire de l'induction de ces processus post-transcriptionnels ont été formulées: production d'ARN aberrants tels que des ARN tronqués, non polyadénylés ou encore des ARN méthylés et l'existence d'une machinerie permettant de détecter une quantité d'ARN anormalement abondant (Meins et Kunz 1995; Meins 2000). Le mécanisme responsable de la spécificité de ces processus a suscité un très grand intérêt.

1.1.1. Découverte des siRNAs impliqués dans les processus de PTGS

Ce n'est qu'avec les travaux de Fire et col. (1998), puis de ceux d'Hamilton et Baulcombe (1999) que la nature de l'inducteur commun aux processus de PTGS a pu être clairement établie. Les premiers montrèrent chez *C. elegans* que l'injection de longs ARN purifiés simples brins (ARNsb) sens ou antisens ne produisait qu'une faible réduction de la quantité d'ARNm cible, alors qu'un mélange des deux brins appariés permettait une dégradation complète et spécifique de l'ARNm endogène homologue chez les individus injectés et leur descendance. Ils mirent ainsi en évidence la nature double brin de l'inducteur des processus d'inactivation épigénétique post-transcriptionnelle. Ces expériences ouvrirent également la voie à une nouvelle technique, nommée ARN interférence (RNAi), permettant la production rapide de mutants pour un gène donné. Comme nous le verrons dans le paragraphe 1.2, cette technique a pris un essor considérable chez de très nombreux organismes.

Les expériences clés ayant permis d'identifier le composant imposant la spécificité des processus de PTGS ont été réalisés par Hamilton et Baulcombe (1999). Chez plusieurs lignées transgéniques

inactivées, ils mirent en évidence la présence de petits ARN, certains complémentaires et d'autres homologues au transcrit cible. Ces petits ARN, dont la taille fût estimée à environ 25 nucléotides, étaient détectés dans les lignées ayant déclenché la PTGS mais étaient absents dans des lignées transgéniques témoins surexprimant le gène introduit. Les auteurs posèrent l'hypothèse que ceux-ci étaient les déterminants de la spécificité de PTGS. Ces petits ARN furent ensuite décrits dans des cultures de cellules de drosophile, après induction de RNAi par introduction de longues molécules d'ARN double brin de 500 nucléotides (Hammond *et al.* 2000), et lors d'expériences *in vitro* avec des extraits embryonnaires de drosophile (Zamore *et al.* 2000). Ainsi, la production de petits ARN de polarité sens et antisens d'environ 21-25 nucléotides, nommés siRNAs, était associée au processus de PTGS. Il fût ensuite montré que les siRNAs sont produits à partir de long ARNdb et possèdent un groupement phosphate en 5' et un groupement hydroxyle en 3' de chaque brin, caractéristiques de produits générés par l'action d'une RNase III (Elbashir *et al.* 2001b).

1.1.2. Bases moléculaires des différents processus de PTGS

La PTGS chez les plantes peut-être induite par différents inducteurs tels que les transgènes en orientation sens ou antisens en copie unique et les transgènes en répétitions inversées permettant la production directe d'ARNdb (Fagard et Vaucheret 2000). Chez les plantes, la PTGS peut également être induite par des virus (Angell et Baulcombe 1997; Dalmay *et al.* 2000a). Dans les cas où des virus recombinants possèdent des homologies au niveau de la séquence codante avec des gènes endogènes et/ou des transgènes, une dégradation coordonnée et spécifique des ARN du virus et du gène homologue est obtenue. Cette inactivation est appelée VIGS (Virus-Induced Gene Silencing). Dans le cas de transgènes possédant le promoteur de l'ARN 35S du virus de la mosaïque du chou-fleur (35S CaMV) et produisant des transcrits sens, la PTGS se met en place au cours du développement. L'expression du transgène est d'abord maximale en tout début de développement (juste après la germination), puis il y a une phase d'initiation, où localement une cellule ou un groupe de cellules vont déclencher la PTGS, suivie d'une seconde phase pendant laquelle il y a

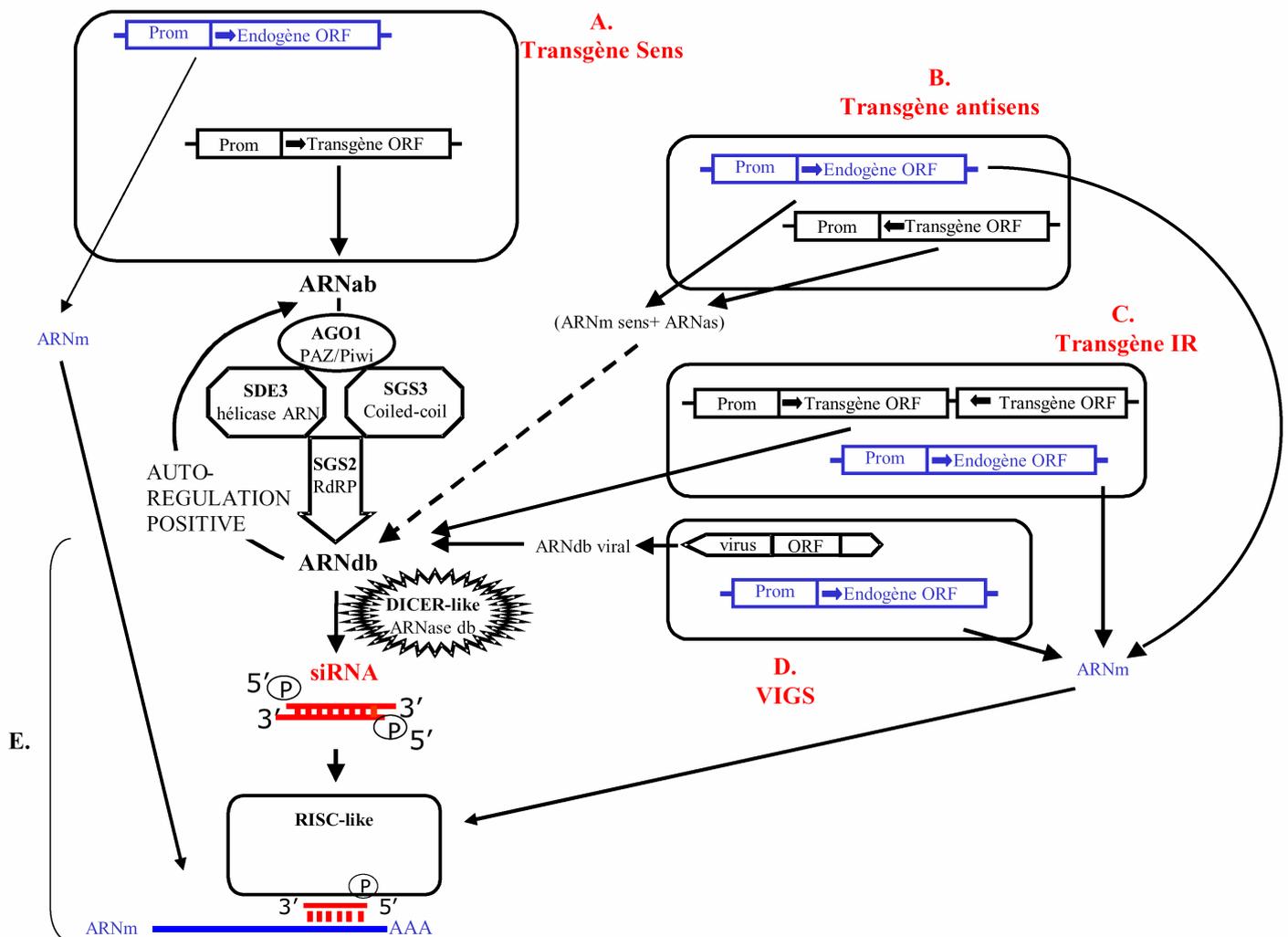


Figure 4 : Mécanisme d'inactivation épigénétique post-transcriptionnelle chez les plantes.

Les différents initiateurs du mécanisme de dégradation sont représentés et peuvent être groupés en 4 catégories selon qu'il y a production d'ARN sens aberrant (ARNab), d'ARN antisens (ARNas), d'ARN double brin (ARNdb) ou d'ARNdb viral.

(A.) Les transgènes sens produiraient un ARNab. Ces ARNab génèreraient des ARNdb grâce à l'intervention d'une ARN polymérase ARN dépendante (RdRP) (SGS2) après détection de ces ARNab par une hélicase ARN (SDE3), un facteur d'initiation de la traduction de la famille des protéines à domaines PAZ (AGO1) et une protéine à domaine coiled-coil (SGS3).

(B.) Dans le cas de la production d'ARN antisens, il est probable que les ARNm sens et antisens s'apparient pour former un ARNdb. La structure des transgènes en répétitions inversées(IR) donnerait également naissance à un ARNdb. Enfin, les virus en cours de répllication sont sous forme ARNdb.

Les ARNdb seraient ensuite dégradés par une ARNase db de type DICER (DICER-Like) dégradant les ARNdb à intervalle de 21-25 nucléotides et générant des petits ARN de 21-25 nucléotides (siRNAs). Ces siRNAs seraient alors pris en charge par un complexe de type RISC (RISC-like) permettant ainsi la dégradation de tous les ARNm homologues à l'ARNdb. Ce mécanisme serait auto-entretenu grâce à la détection des ARNdb considérés comme ARNab et permettant à la RdRP d'amplifier ces ARNdb.

propagation de l'inactivation aux cellules voisines, *via* les plasmodesmes, et à l'ensemble de la plante de façon systémique *via* le système vasculaire (Palauqui *et al.* 1997; Voinnet *et al.* 1998). Enfin, durant la troisième phase, celle de maintenance, tous les tissus possèdent un transgène inactivé ainsi que le gène endogène homologue jusqu'à la fin de la vie de la plante. On observe dans ce cas une méthylation de la séquence codante du transgène mais pas du gène endogène. Ce patron de déclenchement, reproduit à chaque génération, semblait montrer qu'il existait un signal inducteur de la PTGS capable de se propager. Ceci a été confirmé par des expériences de greffes chez *A. thaliana*, *Nicotiana benthamiana* et le tabac (Palauqui *et al.* 1997; Crété *et al.* 2001). Ce signal inducteur porte la spécificité de séquence nucléique nécessaire à la dégradation des ARNm du transgène et des gènes endogènes homologues. Il a donc été suggéré que ce signal serait un acide nucléique de petite taille de nature ribonucléique. Les siRNAs impliqués dans le ciblage de la dégradation de l'ARNm ne correspondraient pas au signal mobile inducteur de la PTGS chez les plantes (Mallory *et al.* 2001). La molécule responsable de ce signal reste donc à identifier.

Le modèle de PTGS chez les plantes, présenté dans la figure 4, regroupe l'ensemble des données obtenues par des études génétiques et biochimiques chez les plantes et les animaux. Les études génétiques réalisées chez les plantes ont permis de mettre en évidence qu'il existe différents inducteurs conduisant à la PTGS.

L'étape commune à tous ces inducteurs est la formation d'un long ARNdb, qui après découpage successif tous les 21 nucléotides par une enzyme de type DICER-Like conduit à la production des siRNAs (Figure 4E). Dans le cas de transgènes sens, la production de l'ARNdb requiert l'intervention de composants spécifiques non requis pour les autres inducteurs (Béclin *et al.* 2002) (Figure 4A). En effet, dans le cas de transgènes en orientation antisens, l'ARNdb serait formé par l'appariement entre l'ARNm endogène et l'ARN produit par le transgène (Figure 4B). Dans le cas d'une induction de PTGS par un transgène en répétition inversée, l'ARN produit se replie en un ARN partiellement double brin (Figure 4C). De même pour les virus qui possèdent leur propre ARN polymérase ARN dépendante ou "RNA directed RNA Polymerase" (RdRP) nécessaire à la

production d'intermédiaire de réplication ARNdb (Figure 4D) (Dalmay *et al.* 2000b; Béclin *et al.* 2002).

Certains composants spécifiques, nécessaires à la production d'ARNdb lors d'une induction de la PTGS par des transgènes sens, permettraient la reconnaissance d'ARN aberrants (ARNab) produits par ces transgènes. Ces ARNab pourraient correspondre à des ARN méthylés, des ARN tronqués ou non polyadénylés. L'un des composants spécifiquement requis pour ce type d'inducteur correspond à une RdRP. Cette RdRP cellulaire, nommée RDR6, SGS2 ou encore SDE1 (notée ici RDR6), semble remplir cette fonction de synthèse d'ARNdb en utilisant comme matrice les ARNab (Dalmay *et al.* 2000b; Mourrain *et al.* 2000). D'autres RdRPs cellulaires ont été identifiées chez d'autres organismes lors de cribles de mutants visant à identifier les composants nécessaires à la PTGS. Ainsi, les RdRPs EGO1 et RRF-1 sont requises, respectivement, pour le RNAi dans la lignée germinale et dans la lignée somatique chez *C. elegans* et QDE1 pour le Quelling (processus de PTGS chez les champignons filamenteux) chez *N. crassa* (Cogoni et Macino 1999a; Smardon *et al.* 2000; Sijen *et al.* 2001a). Cependant ces RdRPs sont nécessaires, même dans le cas d'initiation par des ARNdb ou par injection de siRNAs. Le rôle de ces RdRPs est donc moins clair que celle de la RdRP chez les plantes. Ces RdRPs pourraient augmenter la quantité de siRNAs présents ou encore permettre la stabilisation du complexe siRNA/ARNm cible (Sijen *et al.* 2001a). Aucun homologue de RdRPs ne semble présent dans les génomes de la drosophile et de l'homme.

Chez les plantes, d'autres facteurs essentiels à l'induction de la PTGS par des transgènes sens ont été identifiés, tels que la protéine SGS3, dont la fonction reste à élucider (Figure 4A) (Mourrain *et al.* 2000; Hamilton *et al.* 2002). Celle-ci possède un domaine coiled-coil probablement impliqué dans une interaction protéine-protéine et ne possède pas d'homologue chez d'autres organismes. Les protéines SDE3, une hélicase à ARN putative, et AGO1, un homologue de facteur d'initiation de la traduction, sont également requises pour initier la PTGS sens (Figure 4A) (Elmayan *et al.* 1998; Dalmay *et al.* 2001; Béclin *et al.* 2002; Morel *et al.* 2002). Ces protéines, RDR6, SGS3, SDE3 et AGO1, ne sont pas nécessaires dans le cas d'une initiation de la PTGS par des transgènes antisens (Figure 4B), des transgènes en répétition inversée (Figure 4C) ou par des virus (Figure 4D).

Organisme	Protéine	Domaine	Fonction putative	Référence
<i>A. thaliana</i>	RDR6	RdRP	Production ARNdb	1, 2, 6
	SDE3	RecQ hélicase	Déplie de l'ARN	3
	AGO1	famille PPD	ND	4, 5, 6
	WEX	RNase D	ND	7
	SGS3/SDE2	domaine coiled-coil	Cofacteur SGS2	1, 2, 6
<i>N. crassa</i>	QDE1	RdRP	Production ARNdb	8
	QDE3	RecQ hélicase	Déplie de l'ARN	10
	QDE2	famille PPD	ND	9
	DCL1	RNase III	Découpage des ARNdb en siRNAs	11
	DCL2	RNase III	Découpage des ARNdb en siRNAs	11
<i>C. elegans</i>	EGO1	RdRP	Production ARNdb (lignée germinale)	13
	RRF-1	RdRP	Production ARNdb (lignée somatique)	17
	SMG-2	upflp hélicase	Déplie de l'ARN	16
	MUT14	DEAD box RNA hélicase	Déplie de l'ARN	15
	PPW1	famille PPD	ND	19
	RDE1	famille PPD	ND	12
	MUT7	RNase D	ND	14
	DCR-1	RNase III	Découpage des ARNdb en siRNAs	18
	RDE4	Liaison ARNdb	Production des siRNAs	12
SID1	Protéine transmembranaire	Canal pour signal systémique	20	
<i>D. melanogaster</i>	AGO2	famille PPD	ND	21, 22
	TUDOR-SN	Nucléase	ND	23
	VIG	Liaison à l'ARN	ND	24
	dFXR	Liaison à l'ARN	ND	24
	DICER-1	RNase III	Découpage des ARNdb en siRNAs	25
	DICER-2	RNase III	ND	25
	R2D2	Liaison ARNdb	ND	31
<i>D. discoideum</i>	RRPA	RdRP	Production ARNdb	26
<i>C. reinhardtii</i>	MUT6	hélicase	Déplie de l'ARN	27
Mammifères	AGO2	famille PPD	activité Slicer: clivage de l'ARNm cible	28; 29; 30
	DICER	RNase III	Découpage des ARNdb en siRNAs	28

¹Mourrain *et al.* 2000; ²Dalmay *et al.* 2000; ³Dalmay *et al.* 2001; ⁴Fagard *et al.* 2000; ⁵Morel *et al.* 2002; ⁶Beclin *et al.* 2002; ⁷Glazov *et al.* 2003; ⁸Cogoni and Macino 1999a; ⁹Catalanotto *et al.* 2002; ¹⁰Cogoni and Macino 1999b; ¹¹Catalanotto *et al.* 2004; ¹²Grishok *et al.* 2000; ¹³Smardon *et al.* 2000; ¹⁴Ketting and Plasterk 2000; ¹⁵Tijsterman *et al.* 2002a; ¹⁶Domeier *et al.* 2000; ¹⁷Sijen *et al.* 2001; ¹⁸Knight and Bass 2001; ¹⁹Tijsterman *et al.* 2002b; ²⁰Winston *et al.* 2002; ²¹Hammond *et al.* 2001; ²²Okamura *et al.* 2004; ²³Caudy *et al.* 2003; ²⁴Caudy *et al.* 2002; ²⁵Lee *et al.* 2004; ²⁶Martens *et al.* 2002; ²⁷Wu-Scharf *et al.* 2000; ²⁸Doi *et al.* 2003; ²⁹Song *et al.* 2004; ³⁰Liu *et al.* 2004; ³¹Liu *et al.* 2003.

Tableau 3 : Protéines impliquées dans les processus d'inactivation épigénétique post-transcriptionnelle chez différentes espèces.

D'autres protéines impliquées dans les processus d'inactivation post-transcriptionnelle ont été identifiées chez d'autres organismes (Tableau 3).

Notamment, la protéine impliquée dans la production des siRNAs de transgènes à partir de longs ARNdb a pu être identifiée chez différentes espèces. Les travaux de Bernstein et col. (2001) permirent d'identifier l'enzyme responsable de la production des siRNAs. Lors d'une approche gène candidat, ils testèrent les membres de trois familles de RNase III existant chez la drosophile (DROSHA, HOMELESS et la protéine CG4792) pour leur capacité à produire des siRNAs à partir de longs ARNdb. Ils montrèrent que l'incubation de longues molécules d'ARNdb avec les deux premières protéines immuno-précipitées ne permettait pas la production de siRNAs, alors que la protéine CG4792 en était capable. De part sa capacité à générer des siRNAs à intervalle régulier de 21 nucléotides à partir de longs ARNdb, les auteurs nommèrent cette protéine DICER-1 (la traduction de DICER est "coupeur") (Bernstein *et al.* 2001). La deuxième protéine de la même famille de RNase III que DICER-1 fut nommée DICER-2. L'activité de DICER-1, chez la drosophile, nécessite *in vitro* la présence d'ATP, probablement pour que l'activité hélicase que possède cette protéine puisse fonctionner (Bernstein *et al.* 2001). Cette famille de protéines DICER est phylogénétiquement conservée chez les plantes, les animaux et les champignons filamenteux. De plus, la fonction de cette famille de protéines est conservée dans des processus d'inactivation post-transcriptionnelle chez l'homme (Bernstein *et al.* 2001; Doi *et al.* 2003), chez *C. elegans* (Knight et Bass 2001), chez les plantes (Park *et al.* 2002), ou le poisson zèbre (Wienholds *et al.* 2003). Les nombreuses études des processus d'inactivation épigénétique post-transcriptionnelle chez les plantes (PTGS/cosuppression), chez les animaux (RNAi) ou chez les champignons filamenteux (Quelling), ont ainsi convergé vers un modèle commun (Figure 4). Le maillon central de ce modèle est une molécule d'ARN double brin clivée en de multiples siRNAs par une enzyme DICER, semblant agir par dimérisation intramoléculaire de ses deux domaines RNase III (Zhang *et al.* 2004). La protéine impliquée dans la production de siRNAs à partir de longs ARNdb n'est actuellement pas identifiée chez les plantes. Chez *A. thaliana*, quatre membres de la famille Hélicase-RNaseIII, nommées DICER-Like, sont répertoriées. L'obtention de mutants d'insertion

pour chacun de ces gènes (ainsi que de chaque combinaison de double mutants) devrait permettre de dire si l'un de ces membres (ou plusieurs) est impliqué dans cette étape.

Les siRNAs, intégrés dans un complexe ribonucléoprotéique effecteur, nommé complexe RISC (RNA induced Silencing Complex), permettent alors de guider la reconnaissance d'un ARNm complémentaire et de produire le clivage de l'ARNm au milieu de la zone d'appariement (Figure 4E).

Récemment, de petits ARN endogènes biochimiquement indiscernables des siRNAs, et nommés micro-ARNs (parfois notés miRNAs), ont été mis en évidence. L'entrée des siRNAs dans le complexe effecteur RISC semblant être très proche de celle des micro-ARNs, cette partie ainsi que la composition du complexe effecteur sera donc traitée dans la section consacrée au mode d'action des micro-ARNs.

1.1.3. Rôle de la PTGS

Chez les plantes, la PTGS est rapidement apparue comme un mécanisme de résistance aux virus. En effet, les mutants *sgs2/sde1*, *sgs3/sde2*, *sde3* et *ago1* sont plus sensibles aux infections par certains virus que les plantes sauvages (Mourrain *et al.* 2000; Dalmay *et al.* 2001; Morel *et al.* 2002). De plus, les virus sont capables de lutter contre ce processus de PTGS grâce à la production de protéines virales pouvant bloquer le mécanisme de PTGS à différentes étapes (Anandalakshmi *et al.* 1998; Béclin *et al.* 1998; Voinnet *et al.* 1999; Voinnet *et al.* 2000; Voinnet 2001; Silhavy et Burgyan 2004). Ces données montrent clairement l'évolution commune du mécanisme de résistance aux virus et de riposte engagée par la production d'inhibiteurs viraux.

Chez *C. reinhardtii*, *D. melanogaster* et *C. elegans* la PTGS pourrait être impliquée dans le contrôle de certains éléments transposables (Ketting *et al.* 1999; Wu-Scharf *et al.* 2000; Jensen *et al.* 2002). En effet, chez les mutants *mut-2*, *mut-7*, *rde-2* et *rde-3* chez *C. elegans*, le transposon Tc1 présente une activité de transposition plus importante (Grishok *et al.* 2000).

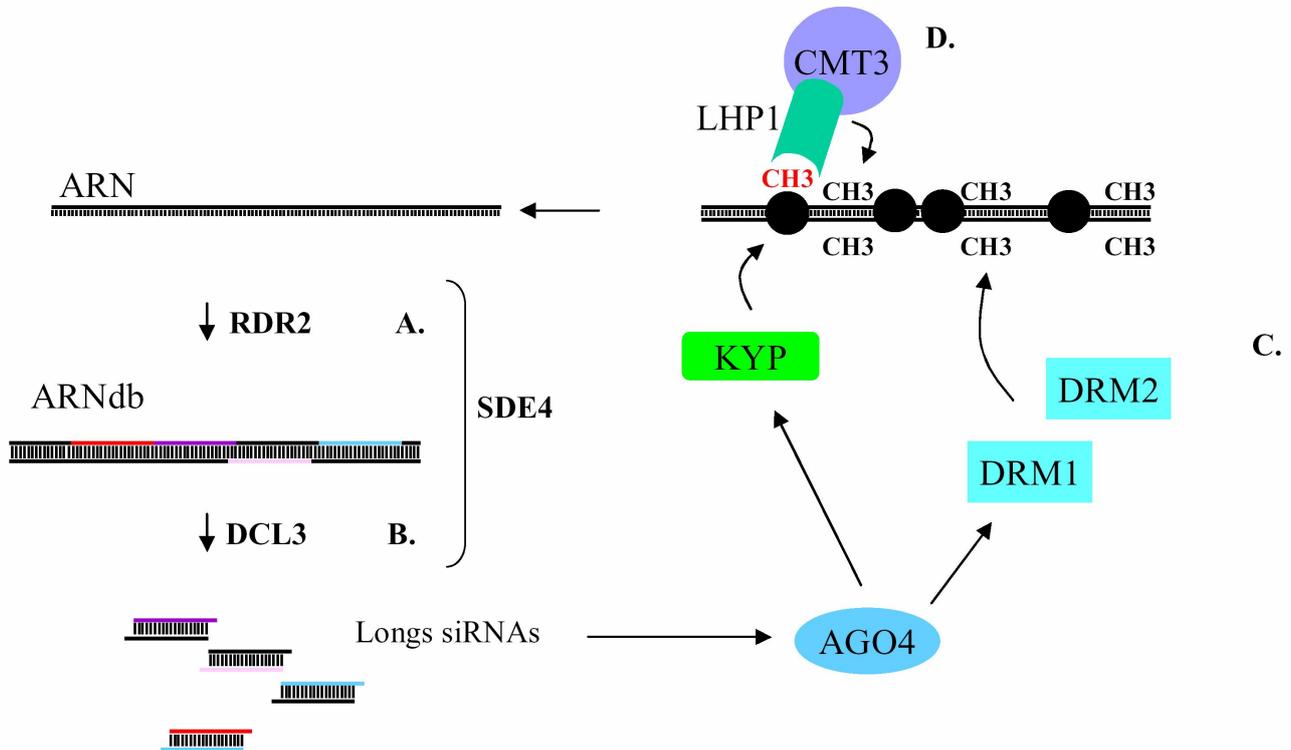


Figure 5 : Mécanisme d'inactivation épigénétique transcriptionnelle guidée par des siRNAs endogènes chez les plantes.

A certains loci (*FWA* ou *AtSN1*), la production des siRNAs serait réalisée par clivage successifs par DCL3 de longs ARNdb produits par une RdRP, RDR2 (A et B). Cette étape de production des siRNAs requiert la protéine SDE4, dont la fonction est inconnue. Ces siRNAs, en association avec la protéine AGO4, induiraient la mise en place de méthylations au niveau des cytosines de l'ADN par les protéines DRM1 et DRM2 et de méthylations au niveau des lysines 9 des histones H3 (CH3) par la protéine KRYPTONITE (KYP) (C). La protéine LHP1 reconnaîtrait l'état méthylé des lysines 9 des histones H3, ce qui permettrait la maintenance des méthylations de l'ADN aux sites CHG par CMT3 (D).

1.1.4. Bases moléculaires de la TGS induite par des siRNAs

L'inactivation épigénétique transcriptionnelle de transgènes semble également pouvoir être guidée par des siRNAs. Des résultats chez les végétaux montrent qu'il existe des liens d'un point de vue du mécanisme entre la TGS et la PTGS (Mette *et al.* 2000; Fagard et Vaucheret 2000). En effet, quand un transgène produit un ARNdb dont la séquence est homologue à un promoteur endogène, les siRNAs entraînent la méthylation des cytosines au niveau de la séquence du promoteur et la mise en place de méthylation au niveau des cytosines du transgène dans tous les contextes répertoriés: CG, CNG, ou asymétriques. Ce qui conduit à l'inactivation transcriptionnelle du gène correspondant. On parle alors de RNA-directed DNA methylation (RdDM) (Mette *et al.* 2000; Jones *et al.* 2001; Sijen *et al.* 2001b; Aufsatz *et al.* 2002). Grâce à l'utilisation d'inhibiteurs viraux, Hamilton et col. (2002) semblent avoir identifié la molécule responsable de la spécificité de la TGS induite par des ARNdb. En effet, ces auteurs ont mis en évidence deux classes de petits ARN double brin. Les premiers de 21-22 nucléotides, notés siRNAs courts, sont associés spécifiquement au clivage des ARN cibles après intégration au complexe RISC. La seconde classe, notée siRNAs longs, d'environ 25 nucléotides, est associée spécifiquement à la méthylation et à l'inactivation transcriptionnelle (Hamilton *et al.* 2002; Zilberman *et al.* 2003). Il ne semble pas que les siRNAs courts (21-22 nt) dérivent des siRNAs longs puisque la production de chacune des deux classes de siRNAs est dissociable par l'utilisation de différents inhibiteurs viraux (Hamilton *et al.* 2002).

Dans le cas du RdDM, il a été observé chez le mutant *sde4*, spécifiquement impliqué dans la production des siRNAs longs, une perte de méthylation au niveau des séquences de rétroéléments (Hamilton *et al.* 2002). Il semble donc que l'un des rôles du RdDM soit de contrôler l'activité transcriptionnelle et donc la transposition de ces rétroéléments. Des expériences utilisant le gène endogène *FWA*, naturellement inactivé au niveau transcriptionnel chez *A. thaliana*, ont permis de comprendre certaines étapes de la RdDM, et d'identifier plusieurs éléments contrôlant la synthèse des siRNAs longs et la mise en place de méthylation (Figure 5). La protéine RDR2 (l'une des 6

RdRPs codées par le génome d'*A. thaliana*), est requise pour la production d'un long ARNdb (Figure 5A) qui sera ensuite découpé en siRNAs longs par DCL3 (Figure 5B), l'une des quatre protéines hélicase-RNase III codées par le génome d'*A. thaliana*. Ces siRNAs, en association avec la protéine AGO4, induiraient la mise en place de méthylations au niveau des cytosines de l'ADN par les protéines DRM1 et DRM2 (*de novo* ADN méthyltransférases) et de méthylations au niveau des lysines 9 des histones H3 par la protéine KRYPTONITE (KYP, une histone méthyltransférase) (Figure 5C) (Cao et Jacobsen 2002; Cao *et al.* 2003; Zilberman *et al.* 2003; Xie *et al.* 2004; Zilberman *et al.* 2004).

Les protéines DDM1 (composant putatif d'un complexe de remodelage de la chromatine) et MET1 (aussi nommée DDM2, une ADN méthyltransférase) sont, comme la CHROMOMETHYLASE3 (CMT3, une ADN méthyltransférase), requises dans les étapes de maintenance de l'inactivation. Le maintien des méthylations de l'ADN aux sites CG est réalisé par MET1 et celle des sites CNG par CMT3 (Lindroth *et al.* 2001). Le maintien des méthylations de l'ADN par CMT3 (Figure 5D) serait réalisée grâce à une interaction avec la protéine LHP1 et à la reconnaissance de l'état méthylé des lysines 9 des histones H3 (Jackson *et al.* 2002).

Les acteurs impliqués dans la RdDM font partie de la même famille que les protéines impliquées dans la PTGS. Ce qui suggère ainsi l'existence d'un complexe RISC spécifique de la TGS au sein du noyau (Xie *et al.* 2004).

Très récemment, la RdDM a été mise en évidence chez les mammifères. Dans des cellules humaines, l'introduction de siRNAs ciblant les îlots CpG des promoteurs induit leur méthylation, conduit à des méthylations des lysines 9 des histones H3 et à l'inactivation transcriptionnelle du gène ciblé. Ce processus requiert les ADN méthyltransférases DNMT1 et DNMT3B (Kawasaki et Taira 2004).

1.2. L'ARN interférence

La mise en évidence et la compréhension des mécanismes d'inactivation épigénétique post-transcriptionnelle a fourni aux biologistes un nouvel outil permettant l'inactivation spécifique de gènes d'intérêts.

1.2.1. L'ARN interférence: un nouvel outil de laboratoire

La technique, nommée ARN interférence (RNAi), a rapidement été mise à contribution chez de nombreuses espèces. Elle a notamment permis de reproduire certains phénotypes observés chez des mutants nuls obtenus par d'autres techniques (KO ou mutagenèse), prouvant ainsi l'efficacité et la spécificité du ciblage de la technique (Bosher et Labouesse 2000).

Cette technique a rapidement été plébiscitée pour la production de mutants et le nombre de travaux ayant utilisé à ce jour l'ARN interférence afin de connaître la fonction d'un gène donné est considérable. En parallèle de ces études ciblées et de l'obtention des séquences complètes de plusieurs génomes, des projets ayant pour but d'identifier à grande échelle la fonction des gènes ont été développés. Par exemple, lors d'expériences visant à identifier des gènes impliqués dans le cycle cellulaire et le développement embryonnaire, des mutants ont été obtenus par RNAi pour l'ensemble des gènes des chromosomes I et III de *C. elegans* (Fraser *et al.* 2000; Gonczy *et al.* 2000). D'autres études ont également produit des mutants par RNAi pour 20.000 gènes de *C. elegans* permettant d'identifier des phénotypes pour plus de 1.700 d'entre eux (Kamath *et al.* 2003). Lors d'une autre étude, 192 mutants pour des gènes candidats ont été produits par RNAi et ont permis d'identifier des gènes impliqués spécifiquement dans la morphogenèse des chromosomes et l'organisation nucléaire au cours de la prophase de la lignée germinale de *C. elegans* (Colaiacovo *et al.* 2002).

Chez les plantes, l'inactivation des gènes est majoritairement réalisée par transformation, *via* *Agrobacterium tumefaciens*, avec un vecteur contenant une partie de la séquence du gène à inactiver en répétition inversée, qui donne ainsi naissance à un transcrit se repliant en une structure double brin (Chuang et Meyerowitz 2000; Wesley *et al.* 2001). D'autres techniques, telles que les techniques biolistiques ou utilisant un vecteur viral, sont parfois utilisées (Baulcombe 1999; Klahre

et al. 2002). L'identification de la fonction des gènes à grande échelle est en cours chez les plantes. Le projet AGRIKOLA, basé sur le système de recombinaison Gateway (invitrogen) a permis la construction de vecteurs avec le gène à inactiver en répétition inversée pour chacun des gènes d'*Arabidopsis* (www.agrikola.org).

L'ensemble des données apportées par ces études d'inactivation des gènes à grande échelle va avoir un impact considérable dans l'identification de leurs fonctions, notamment chez les espèces où les banques d'insertions ne sont pas disponibles. Cette technique peut également permettre d'inactiver en une seule expérience plusieurs membres, voire tous les membres d'une même famille de protéines.

Cette méthode pourrait également avoir un impact considérable en terme de santé humaine.

1.2.2. L'ARN interférence: Un nouvel espoir pour la santé humaine?

Au fur et à mesure des études montrant l'étendue de la liste des organismes chez lesquels le mécanisme de RNAi est fonctionnel, il est apparu que cette technique pouvait probablement être utilisée comme une nouvelle méthode thérapeutique chez l'homme.

Plusieurs études ont montré que le RNAi dans des cultures de cellules humaines est spécifique du gène ciblé (Chi *et al.* 2003; Jackson *et al.* 2003; Semizarov *et al.* 2003). L'utilisation de cette technique *in vivo* reste confrontée, comme les autres techniques de transferts d'acide nucléique, aux problèmes de ciblage cellulaire. Certaines études ont déjà permis d'apporter des résultats encourageants en terme de thérapie humaine. Il a par exemple été possible d'inactiver dans des cultures cellulaires le virus du HIV ou de l'hépatite C (Jacque *et al.* 2002; Yokota *et al.* 2003). Une étude a permis de montrer que la transfection préventive de siRNAs contre un poliovirus permettait une protection spécifique de ces cellules contre une infection intervenant quelques jours plus tard (Gitlin *et al.* 2002). Il a été montré que cette "vaccination" est séquence spécifique par une infection efficace d'un poliovirus variant d'un seul nucléotide au milieu de la zone d'appariement avec le siRNA.

Les exemples de RNAi pour la thérapie anti-cancéreuse sont également très encourageants. Il a été possible de bloquer l'expression d'oncogène dans divers types de cellules cancéreuses humaines en culture et ainsi d'induire spécifiquement leur apoptose (Wilda *et al.* 2002; Aoki *et al.* 2003).

Certaines études ouvrent la voie des essais à l'échelle de l'organisme. Chez la souris, les premiers essais ont permis de vérifier que la technique était fonctionnelle et spécifique *in vivo*. Des expériences de co-expression d'un gène rapporteur et de siRNAs contre ce gène ont permis de montrer son inactivation spécifique chez la souris (Lewis *et al.* 2002; McCaffrey *et al.* 2002). Les premières inactivations de gènes endogènes par RNAi réalisées chez la souris donnent des résultats intéressants. Il a pu être montré que l'injection intraveineuse de siRNAs contre un gène (Fas) impliqué dans l'apoptose des cellules du foie permettait de réduire de façon spécifique la quantité d'ARNm Fas. L'effet est maintenu pendant 10 jours après injection. De plus, alors que des souris traitées par injection d'un anticorps contre un antagoniste de Fas, meurent après 3 jours d'une hépatite fulminante, la co-injection de siRNAs anti-Fas permet à 82% des souris traitées d'être toujours en vie après 10 jours d'observation (Song *et al.* 2003a).

De nombreuses études ont montré que le RNAi est effectif chez plusieurs parasites humains. Ainsi *Trypanosoma brucei* (Ngo *et al.* 1998; Durand-Dubief *et al.* 2003), *Leishmania major* (Robinson et Beverley 2003), *Brugia malayi* (Aboobaker et Blaxter 2003), *Plasmodium falciparum* (Malhotra *et al.* 2002; McRobert et McConkey 2002) ou *Schistosoma mansoni* (Boyle *et al.* 2003; Skelly *et al.* 2003) sont des cibles potentielles de l'interférence *in vivo*.

Les résultats de Malhotra et col. (2002) montrent, dans un modèle murin de malaria, qu'il est possible de cibler un gène essentiel du parasite circulant *Plasmodium berghei*. Par des injections intraveineuses de siRNAs complémentaires du gène d'une cystéine protéase chez des souris infectées, il est observé chez le parasite un élargissement de la vacuole alimentaire et une accumulation d'hémoglobine caractéristique de l'effet observé lors d'utilisation d'inhibiteurs de cystéines protéases. Cet exemple indique donc clairement que le RNAi pourrait être une technique efficace dans la lutte anti-parasitaire.

2. Les micro-ARNs et les siRNAs : Des petits ARN non codants endogènes possédant des fonctions régulatrices

2.1. Les micro-ARNs: des ARNnc d'une importance majeure

En 1993, Lee et col. mettaient en évidence le premier micro-ARN chez *C. elegans*, mais ils étaient loin de se douter que leur découverte allait bouleverser la pensée concernant la régulation de l'expression des gènes. Ils montrèrent que le produit du transcrite *lin-4*, connu pour être impliqué dans la transition du premier stade larvaire vers le second (gène hétérochronique), n'était pas une protéine mais deux ARN non codant de 22nt et 61nt. Ils notèrent que le transcrite de 61nt était capable de se replier en une structure en épingle à cheveux et proposèrent que l'ARNnc de 61nt était le précurseur de l'ARNnc de 22nt (Lee *et al.* 1993). La même année, Wightman et col. (1993) mettaient en évidence la majeure partie des propriétés des micro-ARNs connues aujourd'hui. Ils notèrent en effet que le petit ARNnc de 22nt présentait plusieurs sites de complémentarité au sein de la région 3' non codante du transcrite *lin-14*. Cette région non codante ayant précédemment été rapportée comme essentielle à la régulation de *lin-14* par le produit du transcrite *lin-4* (Wightman *et al.* 1991). Ils purent donc montrer l'importance de l'appariement du petit ARN *lin-4* avec la région 3' non codante du transcrite *lin-14* (conservée entre *C. elegans* et *C. briggsae*), et que *lin-4* permettait ainsi une diminution de la quantité de protéine LIN-14 produite, sans variation de la quantité du transcrite *lin-14*. De plus, des expériences de fusion de la région 3' non codante en 3' d'un gène rapporteur montrèrent que cette région d'homologie est suffisante pour rendre compte de la régulation temporelle de l'expression du gène hétérochronique *lin-14*, et également l'importance de la régulation par le petit ARNnc *lin-4* dans la transition du premier au second stade larvaire. Il fut montré plus tard que l'inhibition de la traduction du transcrite *lin-14* par *lin-4* a lieu après la première étape d'initiation de la traduction (Olsen et Ambros 1999). Ce premier travail avait posé quasiment tous les fondements de la régulation de l'expression des gènes par les micro-ARNs, et ce, 10 ans

avant les premiers travaux ayant permis de comprendre les bases moléculaires de la production et de l'action des micro-ARNs. *lin-4* est donc actuellement reconnu comme l'archétype d'une très abondante famille de petits ARN: les micro-ARNs.

Suite au clonage des petits ARN endogènes chez de nombreuses espèces, la famille des micro-ARNs a littéralement explosée. On en recense actuellement une centaine chez *Arabidopsis* (Llave *et al.* 2002a; Park *et al.* 2002; Reinhart *et al.* 2002; Sunkar et Zhu 2004), autant chez le riz (Wang *et al.* 2004), la drosophile (Lai *et al.* 2003b) et *C. elegans* (Lagos-Quintana *et al.* 2001; Lau *et al.* 2001; Lee et Ambros 2001; Lim *et al.* 2003c; Ambros *et al.* 2003b), et plus de 200 chez la souris (Lagos-Quintana *et al.* 2002) et l'homme (Lagos-Quintana *et al.* 2001; Lagos-Quintana *et al.* 2003; Lim *et al.* 2003c; Mourelatos *et al.* 2002). Les micro-ARNs sont les premiers ARNnc régulateurs décrits dont le rôle au cours de différents processus cellulaires chez les Eucaryotes est aussi crucial. Le nombre des régulations des processus développementaux décrit à ce jour par les micro-ARNs n'est qu'une petite fraction des régulations prédites par les études informatiques (Rhoades *et al.* 2002; Enright *et al.* 2003; Grad *et al.* 2003; Lai *et al.* 2003a; Lewis *et al.* 2003; Bonnet *et al.* 2004; Griffiths-Jones 2004). Plusieurs exemples de régulations avérées seront approfondis à la fin du chapitre sur les micro-ARNs.

Le nombre de micro-ARNs identifiés est en augmentation constante, cependant ceux-ci ne représentent qu'une faible proportion des petits ARNnc clonés. Afin de pouvoir identifier les micro-ARNs parmi les centaines de séquences de petits ARN clonés, un certain nombre de caractéristiques de la famille des micro-ARNs ont été définies (Ambros *et al.* 2003a).

2.1.1. Qu'es-ce qu'un micro-ARN?

Les petits ARN *lin-4* et *let-7* furent d'abord appelés "small temporal RNAs" (stRNAs) (Pasquinelli *et al.* 2000), parce qu'ils étaient exprimés spécifiquement à un stade donné et tous deux impliqués dans les transitions de stades au cours du développement. Cependant, quelques années plus tard, plusieurs centaines de petits ARN partageant les mêmes caractéristiques intrinsèques que les

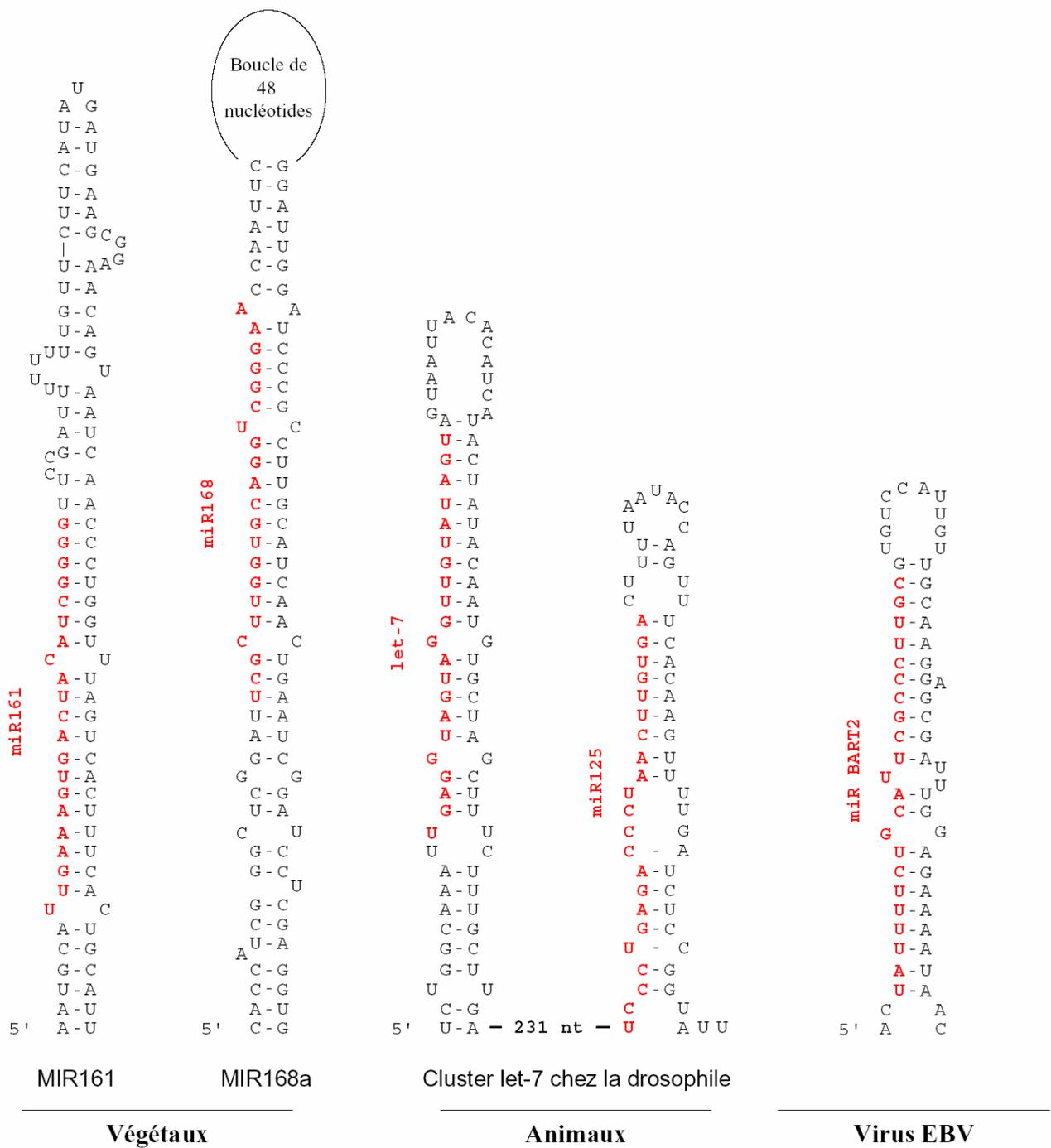


Figure 6 : Structures de différents précurseurs de micro-ARNs.

Structures en épingles à cheveux de précurseurs de micro-ARNs (en rouge) données par les logiciels d'appariements. La structure du précurseur de micro-ARNs animaux présente plusieurs micro-ARNs (dont 2 sont représentés ici) et sont donc dits organisés en cluster.

stRNAs purent être identifiés grâce au clonage massif des petits ARN chez plusieurs espèces (Lagos-Quintana *et al.* 2001; Lau *et al.* 2001; Lee et Ambros 2001; Lagos-Quintana *et al.* 2002; Llave *et al.* 2002a; Mourelatos *et al.* 2002; Park *et al.* 2002; Reinhart *et al.* 2002; Ambros *et al.* 2003b; Lagos-Quintana *et al.* 2003; Lim *et al.* 2003a; Lim *et al.* 2003c). Contrairement aux deux archétypes *lin-4* et *let-7*, la majorité de ces petits ARN ne semblaient pas être exprimés à un stade de développement spécifique mais plutôt dans des types cellulaires particuliers, le terme de micro-ARN fut donc adopté afin de désigner l'ensemble de ces petits ARN aux caractéristiques communes mais aux fonctions encore inconnues (Lagos-Quintana *et al.* 2001; Lau *et al.* 2001; Lee et Ambros 2001).

Les cinq caractéristiques retenues pour différencier les micro-ARNs des siRNAs et de tout autre petit ARN de fonction encore inconnue sont une combinaison de caractéristiques de leur biosynthèse et de validation de leur expression. Elles ont été définies par Ambros et col. (2003a).

Au niveau de la biosynthèse des micro-ARNs, 1) les précurseurs doivent se replier sur eux-mêmes en une structure en épingle à cheveux contenant sur un des bras au moins 16 des 22 nucléotides du micro-ARN (Figure 6). Cette structure ne doit pas présenter de grandes boucles non appariées, et seule la structure possédant la plus basse énergie libre donnée par les logiciels d'appariements peut être retenue, 2) les micro-ARNs et la structure secondaire de leurs précurseurs doivent être conservées d'un point de vue phylogénétique. Ce second point suggère une fonction essentielle de ces ARNnc.

Au niveau de l'expression des micro-ARNs, il est essentiel que 3) l'on soit en mesure de détecter le petit ARN par des techniques d'hybridation et que 4) la séquence précise du petit ARN soit obtenue par des techniques de clonage. Enfin, 5) le dernier critère stipule que les précurseurs des micro-ARNs doivent s'accumuler chez des mutants incapables de réaliser les étapes de maturation du micro-ARN. (Ambros *et al.* 2003a).

Il était également essentiel d'uniformiser le système d'annotation des micro-ARNs et de leurs gènes. En effet, les séquences des micro-ARNs sont souvent conservées entre différentes espèces et, au

sein d'une même espèce, plusieurs précurseurs d'un même micro-ARN peuvent exister. Il était donc nécessaire de pouvoir nommer de façon différente des micro-ARNs distincts clonés chez deux espèces différentes, mais également de donner le même nom à des micro-ARNs orthologues clonés chez des espèces différentes ou à des micro-ARNs présentant peu de différences au niveau de leurs séquences au sein d'une même espèce.

Ainsi, les micro-ARNs sont nommés en utilisant le préfixe "miR" suivi d'un chiffre permettant de les identifier. Les gènes de ces micro-ARNs sont notés de la même façon mais en utilisant les conventions de notations des gènes de l'espèce chez qui il est décrit. Les chiffres permettant d'identifier les micro-ARNs leurs sont donnés de façon séquentielle. Cependant des micro-ARNs possédant la même séquence porteront le même nom quelque soit l'organisme de provenance. Les micro-ARNs possédant des séquences variant de quelques bases peuvent posséder le même nom auquel sera ajouté un suffixe (une lettre ou un chiffre).

Les gènes de micro-ARNs pourront être notés *mir-125* et le micro-ARN sera noté miR-125. Ce gène de micro-ARN, conservé chez l'homme et la souris, est cependant en triple exemplaire. Ces gènes sont donc notés *mir-125a*, *mir-125b-1* et *mir-125b-2* et les micro-ARNs correspondants sont notés miR-125a, miR-125b-1 et miR-125b-2. Les séquences des micro-ARNs miR-125b-1 et miR-125b-2 sont identiques, alors que celle de miR-125a varie de quelques bases par rapport aux précédents. Chez les plantes, les micro-ARNs et leurs gènes sont notés respectivement miR161 et *MIR161*. De la même façon, les gènes permettant la production du micro-ARN miR162 sont dupliqués, à la fois chez *Arabidopsis* et le riz. Les gènes sont donc notés *MIR162a* et *MIR162b*. La séquence du micro-ARN miR162b rencontré chez le riz varie d'une base par rapport aux micro-ARNs miR162a et miR162b d'*A. thaliana*. Les séquences de tous les micro-ARNs, ainsi que celles de leurs précurseurs sont disponibles sur le site Rfam database (<http://www.sanger.ac.uk/Software/Rfam/mirna>).

2.1.2. Biosynthèse

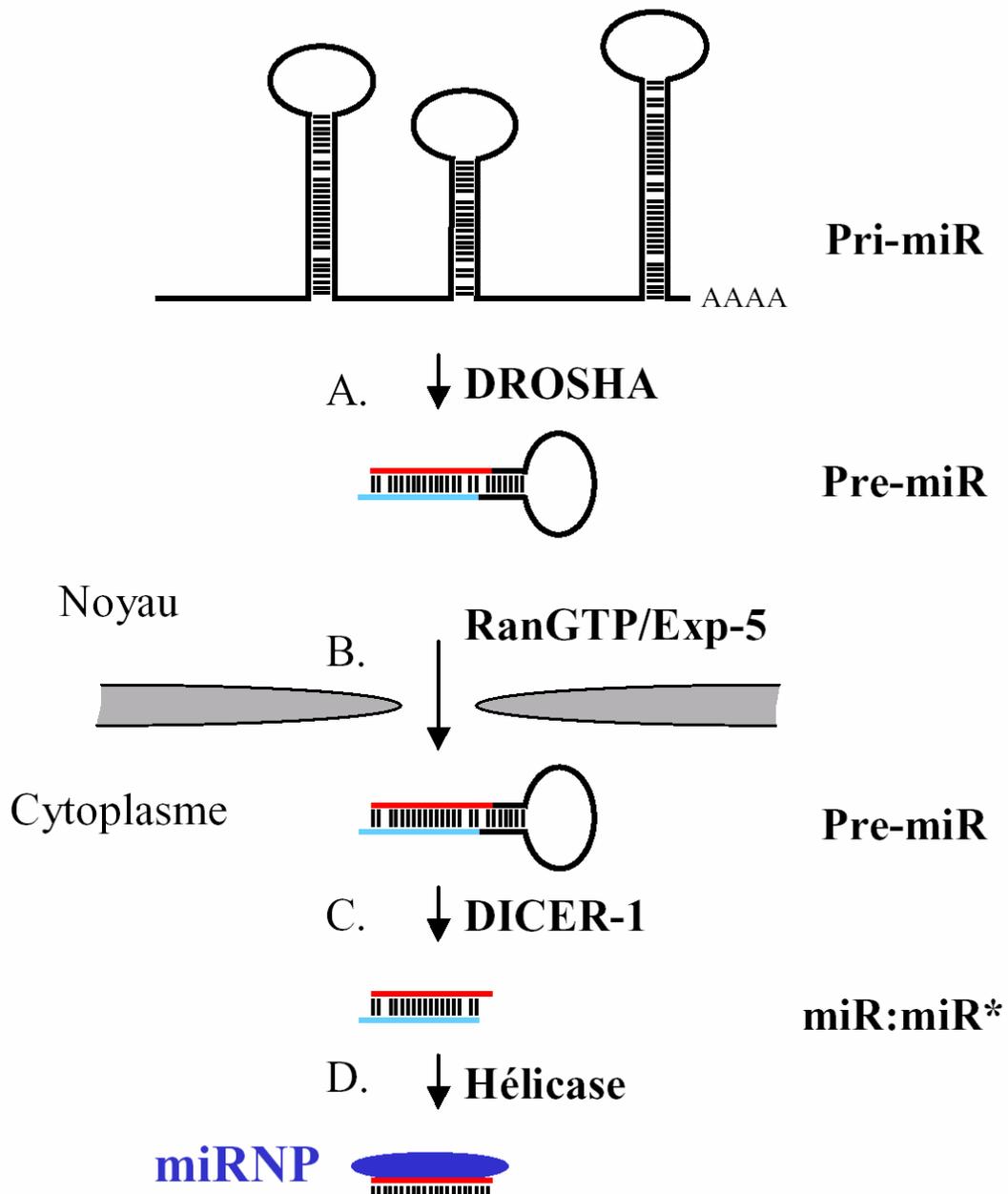


Figure 7 : Mécanisme de biosynthèse des micro-ARNs chez les animaux.

Les transcrits précurseurs de micro-ARNs (pri-miRs), possédant des structures en épingle à cheveux, subissent une première étape de maturation par la RNase III DROSHA (A) qui libère ainsi un pre-miR plus court. Ce pre-miR possède les caractéristiques d'un clivage par une RNase III c'est-à-dire une extrémité 5' phosphate et une extrémité 3' hydroxyle.

Après export du pre-miR vers le cytoplasme (B), le pre-miR est clivé par DICER-1 à 22 nucléotides de l'extrémité ne possédant pas la boucle (C) et libère ainsi un duplex miR:miR* présentant sur un brin la séquence du micro-ARN (miR en rouge) et sur l'autre brin la séquence complémentaire qui était appariée au micro-ARN (miR* en bleu) au sein du précurseur. Le brin possédant l'appariement le moins stable à son extrémité 5' (miR) va être pris en charge par les protéines du complexe miRNP, l'autre brin (miR*) sera rapidement détruit (D). Dans certains cas, la stabilité de l'appariement à l'extrémité 5' est équivalente pour les deux brins et on peut donc avoir stabilisation de chaque brin dans le complexe miRNP. D'après Bartel (2004)

Chez les animaux mais également chez les plantes, la plupart des gènes de micro-ARNs sont situés à une distance des gènes annotés qui suggère qu'ils possèdent leurs propres promoteurs et régulateurs de transcription (Lagos-Quintana *et al.* 2001; Lau *et al.* 2001; Lee et Ambros 2001; Reinhart *et al.* 2002). Cependant, chez les animaux, certains micro-ARNs semblent produits à partir des introns de gènes codants (près d'1/4 des micro-ARNs connus chez l'homme et chez *C. elegans*) (Bartel 2004; Ohler *et al.* 2004). Dans 80% des cas, les introns sont dans la même orientation que les micro-ARNs (Ohler *et al.* 2004). Cependant, il n'est pas clairement établi que les micro-ARNs présents dans les introns soient transcrits indépendamment ou qu'ils soient produits après épissage des transcrits primaires. Certains précurseurs chez les animaux sont également capables de produire plusieurs micro-ARNs. Dans ce cas, les micro-ARNs sont disposés les uns à la suite des autres sur un même précurseur (en cluster), formant chacun une boucle intramoléculaire nécessaire à leur production (Figure 6 : cluster *let-7*) (Lagos-Quintana *et al.* 2001; Lau *et al.* 2001). La co-expression des micro-ARNs présents au sein d'un même précurseur a été vérifiée (Lee *et al.* 2002; Aravin *et al.* 2003; Bashirullah *et al.* 2003). Ce type d'arrangement concerne près de la moitié des micro-ARNs de *Drosophila melanogaster*, alors qu'il est très peu rencontré chez *C. elegans* ou l'homme (Aravin *et al.* 2003; Lagos-Quintana *et al.* 2003; Lim *et al.* 2003a; Lim *et al.* 2003b). Chez les plantes, la taille des précurseurs est plus grande et plus variable (Reinhart *et al.* 2002). Certains micro-ARNs présentent parfois les mêmes profils d'expression, mais ce type d'arrangement en poly-cistron (cluster) n'a encore jamais été mis en évidence chez les plantes.

La transcription des gènes de micro-ARNs par une ARN polymérase ADN dépendante (à identifier) donne naissance à un précurseur de micro-ARN nommé pri-miR (primary precursor for miR) (Lee *et al.* 2002) (Figure 7). A partir de ce pri-miR, un précurseur plus court appelé pre-miR (precursor for miR) est obtenu à partir duquel, après son passage dans le cytoplasme, est généré un micro-ARN mature et fonctionnel (Figure 7).

Chez l'homme, la première étape de maturation du pri-miR est réalisée dans le noyau par une enzyme de type RNase III nommée DROSHA (A) (Lee *et al.* 2002; Lee *et al.* 2003). Celle-ci permet la libération du pre-miR (replié en épingle à cheveux), passant ainsi d'un précurseur pouvant

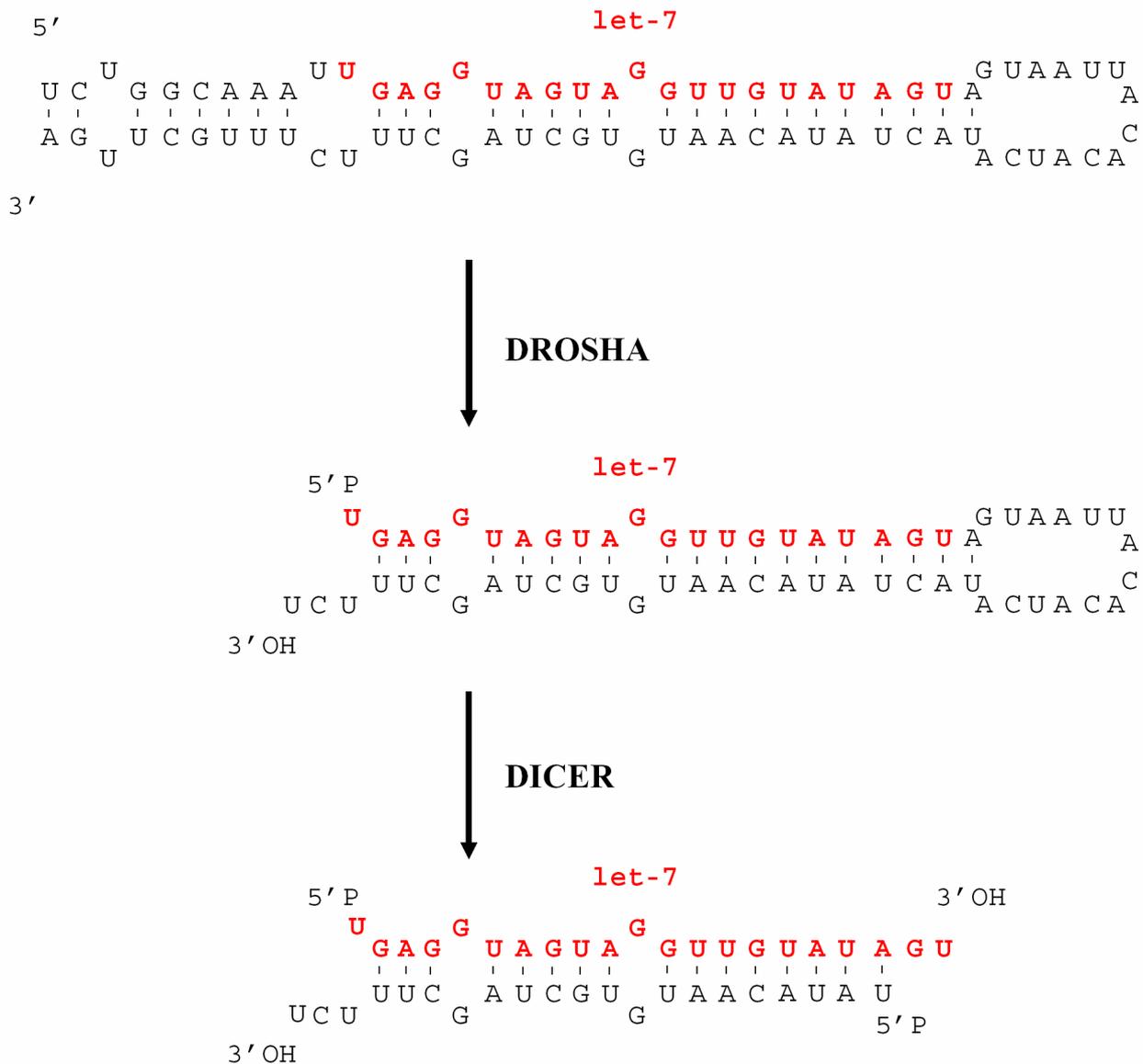


Figure 8 : Mécanisme de biosynthèse des micro-ARNs chez les animaux.

Exemple du micro-ARN *let-7* chez la drosophile. Le pre-miR est libéré du pri-miR par la protéine DROSHA. Le pre-miR est ensuite clivé par DICER-1 à 21-22 nucléotides de l'extrémité ne possédant pas la boucle et libère ainsi un duplex miR:miR* possédant sur un brin la séquence du micro-ARN (en rouge) et sur l'autre brin la séquence complémentaire qui était appariée au micro-ARN au sein du précurseur. Le site de coupure de DICER est donc totalement dépendant du site de coupure par DROSHA.

atteindre le kilobase à un précurseur de 60-80 nucléotides (Lee *et al.* 2002; Zeng et Cullen 2003) (Figure 7A). La libération du pre-miR se fait par coupure endonucléolytique à chacune des extrémités de la boucle, ce dernier étant replié et possédant les caractéristiques d'un clivage par une RNase de type III, c'est-à-dire une extrémité 5' phosphorylée et de 1 à 4 nucléotides débordants à l'extrémité 3' hydroxylée (Basyuk *et al.* 2003). Les pre-miRs sont ensuite activement exportés vers le cytoplasme par l'EXPORTINE-5 qui est un transporteur au travers du pore nucléaire et qui ne fonctionne qu'en présence de la GTPase RAN (Figure 7B) (Yi *et al.* 2003; Bohnsack *et al.* 2004; Lund *et al.* 2004). Cette EXPORTINE-5 n'est pas spécialisée dans l'export des pre-miRs, elle est également capable de réaliser le transport, du noyau vers le cytoplasme, d'ARN viraux hautement repliés et des ARNt (Bohnsack *et al.* 2002; Gwizdek *et al.* 2003). Une fois dans le cytoplasme, les pre-miRs subissent une seconde étape de maturation réalisée par une autre endonucléase de type III possédant un domaine ARN hélicase, nommée DICER (Figure 7C) (Grishok *et al.* 2001; Hutvagner *et al.* 2001; Ketting *et al.* 2001; Knight et Bass 2001; Wienholds *et al.* 2003). Cette enzyme, fut d'abord identifiée comme essentielle à la production des siRNAs impliqués dans le processus de RNAi chez les animaux (Bernstein *et al.* 2001), puis de nombreuses équipes notèrent que le mutant *dcr-1* de *C. elegans* présentait des anomalies développementales sévères (Grishok *et al.* 2001; Knight et Bass 2001), ce qui conduisit à montrer l'implication de DICER dans la maturation des micro-ARNs.

L'activité de DICER a été étudiée de façon plus approfondie dans la production des siRNAs que dans celle des micro-ARNs. Il est cependant probable que la libération du micro-ARN mature soit réalisée de la même façon que la libération des siRNAs: Après dimérisation intramoléculaire des deux domaines RNase III, DICER reconnaît la portion ARNdb du précurseur pre-miR, probablement par une affinité accrue pour ce type d'ARN (Zhang *et al.* 2002; Zhang *et al.* 2004). DICER ne possède pas d'affinité pour les boucles formées par les repliements intramoléculaires des précurseurs, celle-ci réalise donc une coupure typique des RNases III à environ deux tours d'hélices du début de la tige (Figure 8). Cette coupure libère ainsi un ARNdb imparfait possédant, tout comme les siRNAs, une extrémité 5' phosphorylée et deux nucléotides débordants en 3'

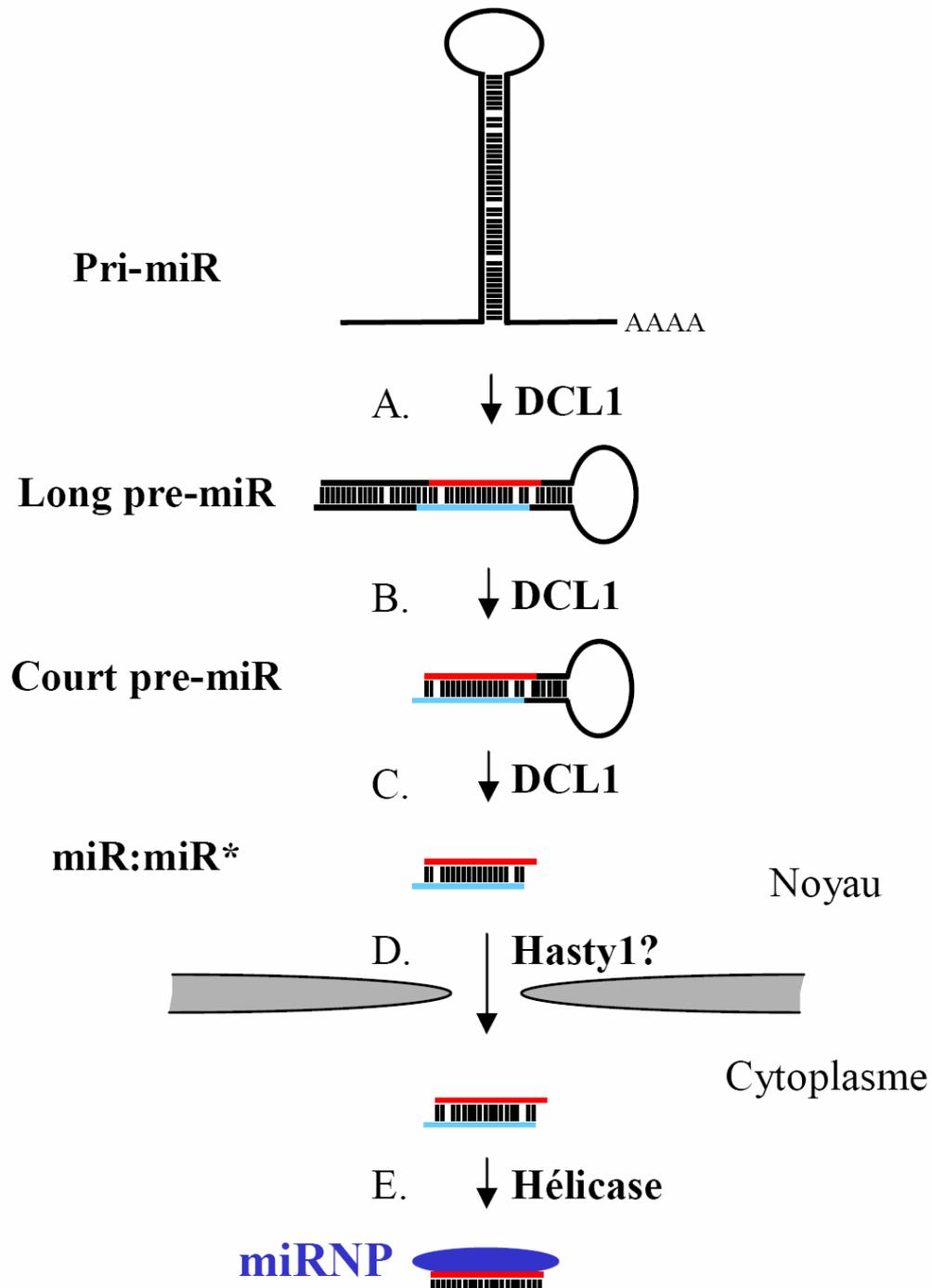


Figure 9 : Mécanisme de biosynthèse des micro-ARNs chez *A. thaliana*.

La protéine DCL1, l'orthologue de DICER-1, est impliquée dans trois étapes de maturation du micro-ARN dans le noyau, allant du pri-miR vers un long pre-miR (A), de ce long pre-miR vers un pre-miR plus court (B), puis de ce pre-miR vers le micro-ARN mature (C). Le duplex miR:miR* serait exporté vers le cytoplasme par une protéine homologue de l'EXPORTINE-5 animale, HASTY. Dans le cytoplasme, grâce à une hélicase l'un des brins du duplex miR:miR* est ensuite intégré dans le complexe effecteur miRNP (suivant la règle énoncée dans la légende de la Figure 7) qui sera ainsi guidé vers les ARNm complémentaires.

(Zhang *et al.* 2002) (Figure 8). Cet ARNdb imparfait possède donc sur un brin la séquence du micro-ARN mature et sur l'autre brin la séquence complémentaire qui était appariée au micro-ARN au sein du précurseur, appelé micro-ARN complémentaire et noté miR* (Lau *et al.* 2001). Ce duplex miR:miR* est une molécule de faible durée de vie. Il est en effet très difficile de détecter le miR* alors que le micro-ARN est beaucoup plus abondant. Ceci indique que le brin correspondant au micro-ARN est stabilisé par l'association de celui-ci avec les protéines du complexe ribonucléoprotéique, nommé miRNP, qui vont permettre la reconnaissance de l'ARNm cible (Lau *et al.* 2001). Comme nous l'avons vu, il semble que le clivage par DICER soit conditionné par le début de l'extrémité libre de la structure en épingle à cheveux, et que le site de coupure de DICER soit donc totalement dépendant du site de coupure par DROSHA. C'est donc la coupure réalisée par DROSHA qui impose la séquence finale des micro-ARNs en déterminant la première extrémité du micro-ARN et en ne laissant aucune autre possibilité à DICER que de couper 22 nucléotides plus loin (Lee *et al.* 2003). La spécificité du clivage par DROSHA serait déterminée par la structure secondaire du pri-miR (Lee *et al.* 2003). Récemment, deux séquences consensus au sein des gènes de micro-ARNs ont pu être identifiées chez *C. elegans* par des analyses bioinformatiques (Ohler *et al.* 2004). La détection de la plus conservée des deux, située 200bp en amont du micro-ARN, a permis d'identifier 9 nouveaux gènes de micro-ARNs chez *C. elegans* (Ohler *et al.* 2004). L'étude de ces séquences consensus permettra de déterminer si elles sont requises pour recruter des facteurs de transcription ou positionner une enzyme impliquée dans la maturation des micro-ARNs (Ohler *et al.* 2004).

Chez les plantes, l'orthologue de la protéine animale DICER-1, nommée DICER-LIKE1 (DCL1), est nécessaire à la production des micro-ARNs actuellement connus (Park *et al.* 2002; Reinhart *et al.* 2002) (Figure 9). La fonction remplie par DCL1 chez les plantes est plus complexe. En effet, contrairement à ce que l'on observe chez les animaux mutés dans *DICER-1*, les pre-miRs ne s'accumulent pas chez les mutants *dcl1* (Grishok *et al.* 2001; Hutvagner *et al.* 2001; Lee et Ambros 2001; Park *et al.* 2002; Reinhart *et al.* 2002; Lim *et al.* 2003a). De plus, DCL1 est localisée dans le noyau. Ces deux éléments indiquant que DCL1 remplirait la fonction de la protéine DROSHA

(Papp *et al.* 2003). Il a ainsi été montré récemment que DCL1 est nécessaire à toutes les étapes de maturation conduisant le pri-miR à la production du micro-ARN mature (Kurihara et Watanabe 2004) (Figure 9). Ainsi chez les plantes, DCL1 est impliquée dans trois étapes de maturation, allant du pri-miR vers un long pre-miR (Figure 9A), de ce long pre-miR vers un pre-miR plus court (Figure 9B), puis de ce pre-miR vers le micro-ARN mature (Figure 9C) (Kurihara et Watanabe 2004).

2.1.3. Mécanisme d'action des micro-ARNs

2.1.3.1. Déterminants du mécanisme d'action

Chez les animaux et chez les plantes, les micro-ARNs s'associent à un complexe protéique appelé miRNP (miR containing RiboNucléoProteic complex) et vont permettre de guider ce complexe jusqu'à l'ARN cible auxquels ils s'apparient.

Chez les animaux, la majorité des micro-ARNs identifiés va induire, *via* le complexe miRNP, l'inhibition de la traduction (Olsen et Ambros 1999). Cependant, un cas de clivage d'ARN cible par un micro-ARN a été rapporté chez les mammifères (Yekta *et al.* 2004). En effet, contrairement à la majorité des micro-ARNs identifiés chez les mammifères, le site de complémentarité de miR-196 avec les ARNm de la famille HOX est parfait. La régulation de ces ARNm par miR-196 est réalisée non pas par inhibition de la traduction, mais par clivage de la cible au milieu de la zone de complémentarité (Yekta *et al.* 2004).

Chez les plantes, la majorité des micro-ARNs induisent le clivage de l'ARN cible (Llave *et al.* 2002b; Kasschau *et al.* 2003; Xie *et al.* 2003; Jones-Rhoades et Bartel 2004) et seul un cas d'inhibition de la traduction a été rapporté. Il concerne la régulation par le miR172 du gène *APETALA2 (AP2)* (Aukerman et Sakai 2003; Chen 2004). Ainsi, aucune variation de la quantité de transcrit *AP2* n'a été observée chez des plantes surexprimant miR172 alors que la quantité de protéine AP2 diminue très fortement (Aukerman et Sakai 2003; Chen 2004).

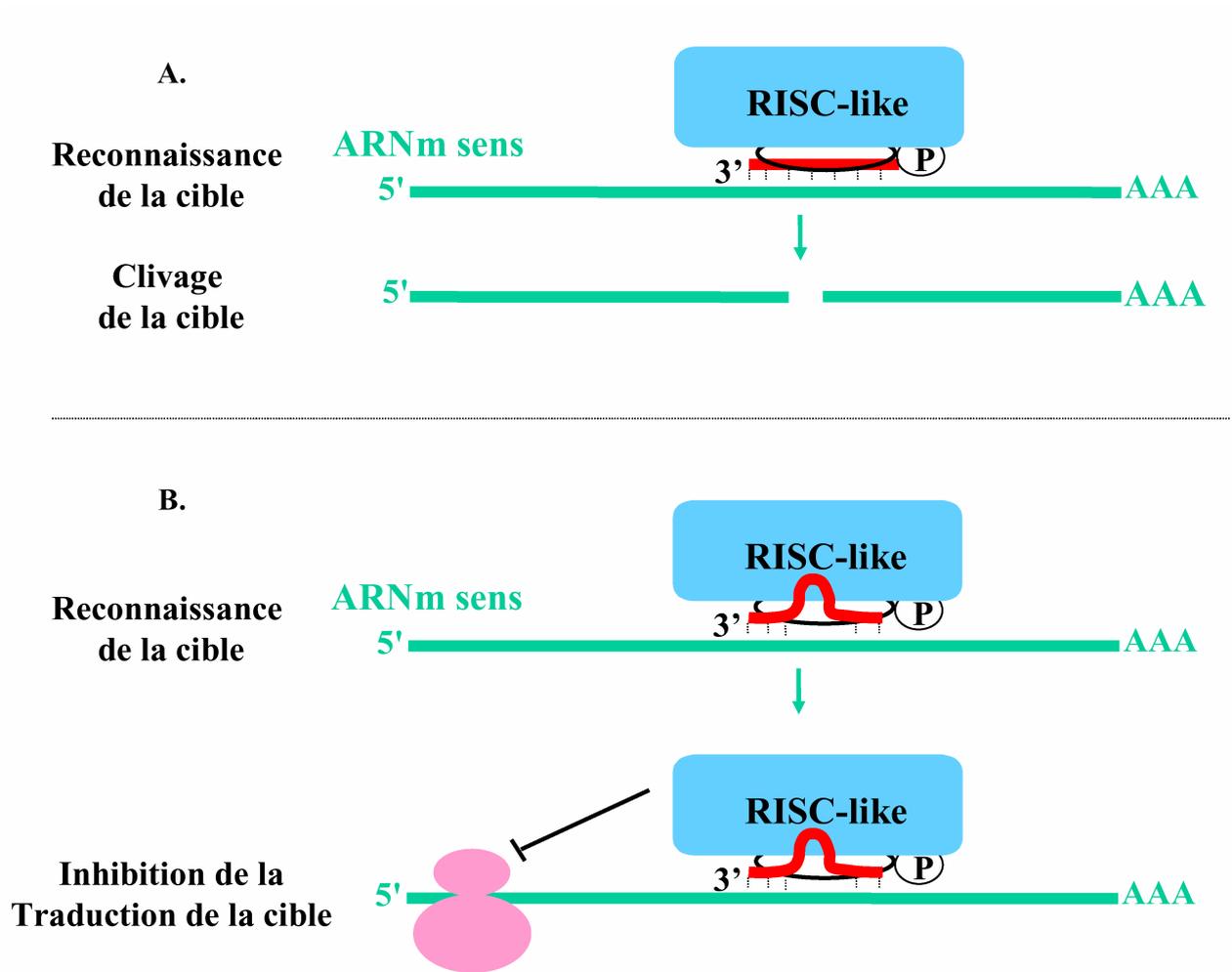


Figure 10 : Mécanismes d'action des micro-ARNs chez les animaux.

A. Un appariement quasi-parfait entre le micro-ARN et l'ARNm cible conduit à un clivage de l'ARNm par l'activité SLICER du complexe RISC-like.

B. Un appariement entre le micro-ARN et l'ARNm cible possédant une boucle centrale non appariée conduit à une inhibition de la traduction de l'ARNm ciblé.

Chez les plantes et les animaux, l'appariement de la première moitié située en 5' du micro-ARN est essentiel pour la reconnaissance de la cible (Doench et Sharp 2004; Mallory *et al.* 2004b). Chez les plantes, l'appariement quasiment parfait entre le micro-ARN et sa cible a, en majorité, lieu au niveau de la séquence codante des ARNm cibles. Chez les animaux, cet appariement est réalisé majoritairement au niveau de la région 3' non traduite de l'ARNm cible, et on observe toujours au niveau de cet appariement une zone centrale non appariée formant une boucle qui semble capitale pour les étapes de régulation par inhibition de la traduction (Lewis *et al.* 2003). On pourrait donc conclure schématiquement qu'un appariement micro-ARN/ARNm quasiment parfait conduit à un clivage de l'ARNm cible (Figure 10A) et qu'un appariement micro-ARN/ARNm avec une zone centrale de plusieurs bases non appariées conduit à une inhibition de la traduction (Figure 10B). Chez les animaux, les travaux réalisés jusqu'à présent vont dans ce sens. Par contre, chez les plantes, le mécanisme d'action des micro-ARNs semble plus complexe.

En effet, miR172, en plus d'une inhibition de la traduction de l'ARNm *AP2*, est capable d'induire le clivage des transcrits codants les protéines *TOE1*, *TOE2* et *TOE3*, dont les deux premières sont des inhibiteurs de la floraison (Aukerman et Sakai 2003; Kasschau *et al.* 2003; Chen 2004). En effet, des sites de clivage au milieu du duplex formé par l'appariement miR172/ARNm *TOE* ont pu être validés pour chacun des transcrits *TOE1*, *TOE2* et *TOE3* (Aukerman et Sakai 2003; Kasschau *et al.* 2003). De plus, contrairement à l'ARNm cible *AP2*, dont l'abondance ne varie pas chez le mutant *dcl1-9* (affecté dans la production des micro-ARNs), on observe des variations dans l'accumulation des transcrits *TOE1*, *TOE2* et *TOE3* chez ce mutant (Kasschau *et al.* 2003). miR172 utiliserait donc les deux mécanismes d'action pour la régulation de ses différentes cibles. Il semble, de plus, exister une compétition pour la régulation du gène *AP2* entre clivage de l'ARNm et inhibition de sa traduction, puisque des produits de clivage ont également été identifiés par 5' RACE pour le transcrit *AP2* (Aukerman et Sakai 2003; Kasschau *et al.* 2003).

Chez les animaux, l'utilisation préférentielle d'un des deux mécanismes d'action serait régit par le niveau de complémentarité entre le micro-ARN et les ARNm cibles. Chez les plantes, vu les

résultats obtenus avec miR172, le choix du mode de régulation apparaît beaucoup plus complexe et ne semble pas uniquement basé sur le niveau de complémentarité.

2.1.3.2. Les complexes miRNPs

Actuellement, la nature des composants des complexes miRNPs est encore mal connue (Schwarz et Zamore 2002; Bartel 2004; Mallory et Vaucheret 2004).

Cependant, en se basant sur la nature très proche des siRNAs et des micro-ARNs et sur les similitudes de leurs mécanismes d'actions, nous pouvons penser que les complexes miRNP et siRNP (encore appelé RISC) sont très semblables voir identiques.

2.1.3.2.A. Caractéristiques communes des siRNAs et des micro-ARNs

D'un point de vue biochimique, les micro-ARNs et les siRNAs présentent les mêmes caractéristiques. Ils possèdent une taille comparable de 20 à 25 nucléotides et sont tous deux phosphorylés en 5' et hydroxylés en 3'. Il est donc impossible de distinguer ces deux types de petits ARNnc par leurs caractéristiques intrinsèques. De plus, le micro-ARN mature (miR) semble provenir d'un duplex miR:miR* très semblable au duplex de siRNA (Figure 8). Le miR* est défini comme le brin complémentaire du micro-ARN avec 2 bases débordantes non appariées à l'extrémité 3'. L'existence des miRs* a été montrée par le clonage de ceux-ci même s'ils sont beaucoup moins abondants que les micro-ARNs (Lagos-Quintana *et al.* 2002; Aravin *et al.* 2003; Lim *et al.* 2003c). Chez *C. elegans*, 3400 clones correspondaient de façon redondante à 80 micro-ARNs alors que seuls 38 clones correspondaient à 13 miRs* (Lim *et al.* 2003c). Cette faible abondance des miRs* serait due à une très faible durée de vie du duplex miR:miR* par rapport au duplex siRNA et à la stabilisation du micro-ARN par association avec les protéines du complexe miRNP tandis que le miR* ne serait pas stabilisé et donc rapidement détruit.

L'entrée des micro-ARNs et des siRNAs dans les complexes RNPs est identique. Tout comme les siRNAs, l'un des deux brins du duplex miR:miR* va entrer dans le complexe RNP. Dans les deux cas, les règles d'entrée de l'un des deux brins dans le complexe RNP ont été élucidées. Une étude statistique appuyée par une étude biochimique ont montré que la stabilité relative de l'appariement de chacune des extrémités des micro-ARNs et des siRNAs est le facteur principal qui détermine le brin entrant dans le complexe RNP (Khvorova *et al.* 2003; Schwarz *et al.* 2003). Le brin du duplex possédant son extrémité 5' du côté où l'appariement est le moins stable (sur les 5 premières bases) est toujours celui qui entre dans le complexe RNP (Khvorova *et al.* 2003; Schwarz *et al.* 2003). Cette observation a permis de poser l'hypothèse de l'existence d'une protéine à fonction hélicase initiant la séparation des deux brins à l'extrémité 5' la moins stable (Schwarz *et al.* 2003). La règle énoncée s'applique à la fois aux siRNAs et aux micro-ARNs de drosophile mais également aux micro-ARNs chez les plantes (Schwarz *et al.* 2003). Cette règle permet également d'expliquer que certains miRs*, tels que miR-2a-2*, miR-4*, miR-8*, miR-10* ou miR-13a* aient pu être clonés, et que leur expression, ainsi que celle de leur micro-ARN respectif, ait pu être vérifiée *in vivo* (Aravin *et al.* 2003; Schwarz *et al.* 2003). En effet, dans ces cas on constate qu'au sein du duplex miR:miR* les extrémités 5' de chaque duplex possèdent une stabilité comparable (Aravin *et al.* 2003; Schwarz *et al.* 2003). Il semble donc que dans quelques rares cas, chacun des deux brins puisse être fonctionnel et réguler des cibles différentes.

D'autres éléments indiquent que les micro-ARNs et les siRNAs intègreraient des complexes RNP très similaires voire identiques. Chez des cellules humaines, si l'on crée un ARN cible contenant un site de complémentarité parfait avec le micro-ARN *let-7*, inhibiteur traductionnel, et que l'ARN cible modifié est mis en présence de la particule immuno-purifiée miRNP-*let7*, on obtient un clivage de cette cible plutôt qu'une inhibition de sa traduction (Hutvagner et Zamore 2002).

De façon réciproque, il a pu être montré chez les mammifères que des siRNAs pouvaient, non pas agir par clivage de l'ARNm cible, mais par inhibition de sa traduction lorsque le site de complémentarité entre le siRNA et l'ARNm cible était imparfait (Doench *et al.* 2003). Ces résultats montrent que les siRNAs et les micro-ARNs peuvent indifféremment intégrer des complexes RNPs

dirigeant le clivage de l'ARN cible ou l'inhibition de sa traduction. De plus, les tailles des complexes miRNPs et siRNPs sont toutes deux estimées à 500-550 kilodaltons (Hammond *et al.* 2001; Mourelatos *et al.* 2002; Schwarz et Zamore 2002). L'hypothèse la plus probable est donc que ces complexes miRNPs et siRNPs sont très proches voir identiques et que leur fonctionnement dépend principalement du niveau de complémentarité entre le petit ARNnc régulateur et sa cible.

2.1.3.2.B. Composants identifiés du complexe siRNP

Les arguments présentés précédemment mettent en évidence la forte probabilité d'identité entre les complexes siRNP et miRNP. Nous allons donc décrire les travaux issus des expériences de RNAi ayant permis d'identifier des éléments du complexe siRNP encore appelé complexe RISC dans la littérature.

Le complexe RISC, impliqué dans le RNAi chez les animaux, a été purifié à partir de cellules de *D. melanogaster* et de cellules humaines. Dans les deux cas, ce complexe siRNP contient un membre de la famille ARGONAUTE, AGO2 (Hammond *et al.* 2001; Carmell *et al.* 2002; Martinez *et al.* 2002; Okamura *et al.* 2004). A noter que chez l'homme, le complexe miRNP possède également cette protéine AGO2 donnant un argument supplémentaire en faveur d'une identité entre le complexe RISC et le complexe miRNP (Mourelatos *et al.* 2002).

Les protéines de la famille ARGONAUTE (AGO) sont connues sous le nom de protéines à domaine PAZ et PIWI (PPD protéines), du fait que tous les membres AGO possèdent ces deux domaines (Cerutti *et al.* 2000). Nous reviendrons dans le chapitre suivant sur la fonction de ces deux domaines. Le complexe RISC, chez *D. melanogaster*, est également composé de la protéine VIG (Vasa Intronic Gene) probablement capable de lier l'ARN, de dFXR (drosophila Fragile X mental Retardation protein), protéine qui a également été détectée comme composant du complexe miRNP, et de la nuclease TUDOR-SN (TUDOR Staphylococcal Nuclease) (Figure 11: siRNP; Figure 12 : miRNP) (Caudy *et al.* 2002; Ishizuka *et al.* 2002; Caudy *et al.* 2003). Les fonctions de ces dernières protéines au sein du complexe RISC ne sont pas connues. Ce complexe RISC est également

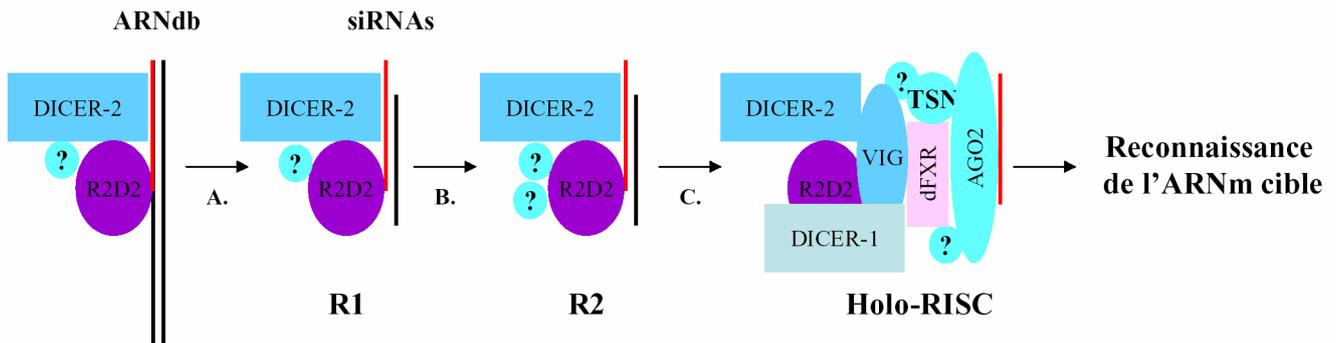


Figure 11 : Entrée des siRNAs dans le complexe siRNP appelé RISC chez la drosophile.

Le complexe R2D2/DICER-2 permet la production des siRNAs à partir de longs ARNdb mais peut également s'associer aux siRNAs préformés (A). Ce complexe est converti en un complexe de plus haut poids moléculaire en recrutant des protéines non identifiées (B). Le complexe holo-RISC, possédant les propriétés de liaison à l'ARNm cible et ne possédant plus qu'un des deux brins du siRNA est formé par recrutement des protéines DICER-1, VIG, TSN, dFXR et AGO2 (C). D'après Pham *et al.* (2004).

composé de la protéine R2D2, possédant des homologies avec HYL-1 et RDE-4, ainsi que des protéines DICER-1 et DICER-2 (Liu *et al.* 2003; Pham *et al.* 2004).

2.1.3.2.C. Formation du complexe RISC

Chez la drosophile, l'entrée des siRNAs dans le complexe RISC est réalisée par un couplage protéique entre la production des siRNAs et leur mise en place dans le complexe (Figure 11) (Pham *et al.* 2004; Tomari *et al.* 2004). Il existe deux complexes intermédiaires (R1 et R2) avant la formation du complexe effecteur de clivage nommé holo-RISC (correspondant au RISC). L'entrée des siRNAs dans le complexe RISC est réalisée par la formation d'un premier complexe de faible poids moléculaire (R1). Celui-ci est composé des protéines R2D2/DICER-2, prenant en charge les siRNAs (Pham *et al.* 2004) (Figure 11). DICER-2 est impliquée dans le clivage de longs ARNdb en de nombreux siRNAs (Figure 11A). Ce complexe est également capable de se fixer à des siRNAs préformés probablement grâce au domaine de liaison à l'ARNdb de R2D2. Enfin, ce complexe est nécessaire au transfert des siRNAs vers les étapes effectrices du RNAi (Liu *et al.* 2003). Le complexe R2D2/DICER-2/siRNAs est ensuite converti en un complexe de poids moléculaire plus élevé (R2), puis enfin en complexe holo-RISC qui possède les propriétés de fixation à l'ARNm cible et de clivage (Figure 11B et C) (Pham *et al.* 2004). Ce complexe holo-RISC est formé des protéines AGO2, VIG, dFXR et TUDOR-SN. Le domaine PIWI d'AGO2 est le domaine qui permet l'interaction entre AGO2 et DICER (Tahbaz *et al.* 2004). De plus, le domaine PAZ d'AGO2 possède une structure en feuillet β qui, de concert avec un appendice latéral, est capable de fixer les ARN simples brins d'une taille d'au moins 5 nucléotides, mais également de fixer les ARN double brin (Lingel *et al.* 2003; Song *et al.* 2003b; Yan *et al.* 2003). Ceci suggère que les protéines de la famille PPD pourraient être directement associées aux siRNAs, avant et après la reconnaissance de l'ARNm cible.

Au sein de ce complexe RISC mature, où les siRNAs sont désormais à l'état de simple brin, les protéines R2D2, DICER-1 et DICER-2 sont également présentes. Ces données ont donc permis de

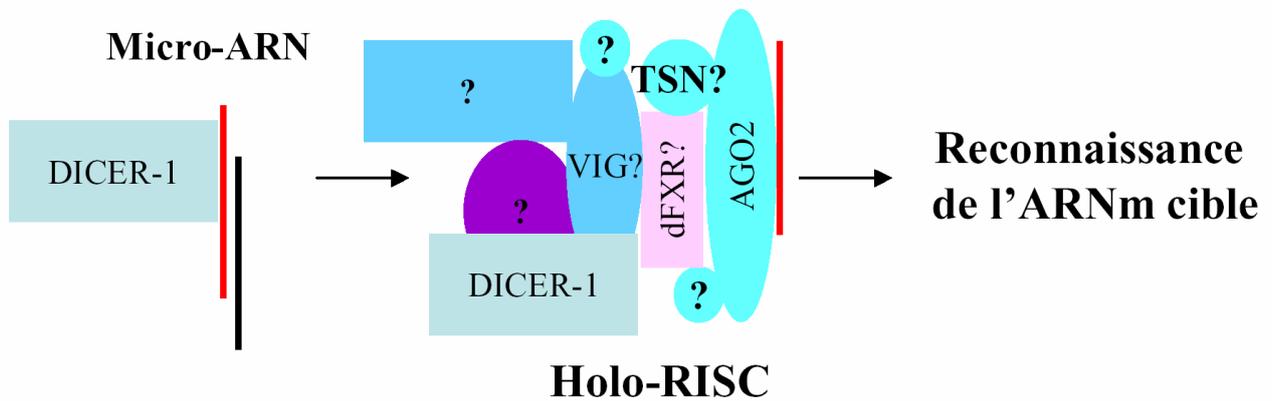


Figure 12 : Entrée des micro-ARNs dans le complexe miRNP appelé RISC chez la drosophile.

A. L'entrée des micro-ARNs, produits par DICER-1, dans le complexe miRNP appelé RISC se ferait à une étape intermédiaire par recrutement des mêmes protéines que celles du complexe intégrant les siRNAs siRISC. Les protéines R2D2 et DICER-2 sont absentes de ce complexe miRNP.

postuler que le complexe initial ne fait pas que transférer les siRNAs vers les étapes effectrices du RNAi, mais qu'il recrute les autres protéines afin de former un complexe de haut poids moléculaire capable de produire le clivage d'un ARNm cible (Figure 11) (Pham *et al.* 2004). De plus, une autre étude chez la drosophile a permis de montrer que la protéine ARMITAGE était essentielle à la formation du complexe final holo-RISC formé autour d'un autre membre de la famille des protéines PPD, la protéine AUBERGINE (Tomari *et al.* 2004). Cette protéine ARMITAGE, une hélicase ATP dépendante putative, semble impliquée dans la transition des complexes de faibles poids moléculaires vers le complexe holo-RISC, et interviendrait en facilitant le désappariement des siRNAs afin que ne soit conservé que l'un des deux brins au sein du complexe RISC (Tomari *et al.* 2004). S'agit-il du même complexe RISC décrit dans l'étude précédente ou d'un complexe RISC spécifique d'un autre type cellulaire? Seules des études croisées permettront de le déterminer.

L'identité entre RISC et miRNP est également renforcée dans cette étude par le fait qu'en absence de siRNAs synthétiques dans le milieu réactionnel, un complexe composé de toutes les protéines décrites précédemment, sauf R2D2 et DICER-2, est observé et attribué à la présence des micro-ARNs endogènes (Pham *et al.* 2004). Il est donc possible que les différences de composition entre miRNP et RISC correspondent aux étapes spécifiques de mise en place des siRNAs et des micro-ARNs dans le complexe RNP (Figure 11 et 12). Ceci est corrélé avec l'observation que la protéine DICER-2, requise pour la formation du complexe RISC, n'est pas requise pour la formation du complexe miRNP. Par contre, la protéine DICER-1 est nécessaire à la formation de chacun des deux complexes (Lee *et al.* 2004). DICER-1 possède une affinité pour chacun des deux types de petits ARN et co-précipite avec les micro-ARNs (Pham *et al.* 2004). La protéine DICER-1 étant nécessaire au passage du complexe RISC intermédiaire vers le complexe holo-RISC, il est envisageable que la formation du complexe miRNP débute à cette étape (Lee *et al.* 2004).

2.1.3.2.D. Activité "slicer" du complexe RISC

Au sein de ce complexe RISC, la protéine portant l'activité de clivage des ARNm cibles, nommée

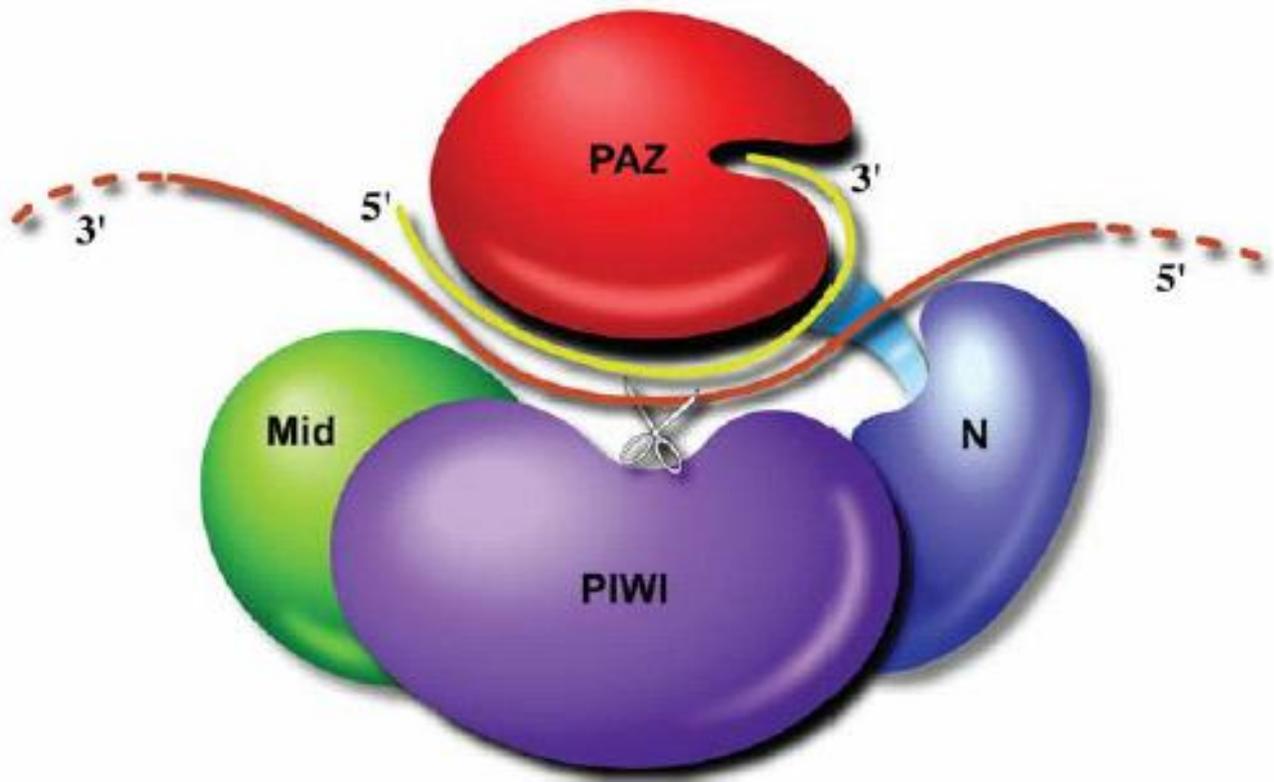


Figure 13 : Modélisation du mécanisme d'action de l'activité "slicer" d'AGO2 chez l'homme.

L'extrémité 3' des siRNAs serait liée au domaine PAZ d'AGO2 plaçant le reste du siRNA (jaune) à l'intérieur du tunnel formé par le repliement du domaine N terminal, PAZ, PIWI et intermédiaire (Mid), et dans lequel est également placé l'ARNm cible en orientation complémentaire (rouge). La modélisation prédit que le domaine catalytique SLICER, présent au niveau du domaine PIWI, couperait ainsi à 9 nucléotides de l'extrémité 5' du siRNA. Illustration : Song *et al.* 2004.

"activité slicer" a été identifiée très récemment chez les mammifères (Liu *et al.* 2004; Song *et al.* 2004). Cette activité de clivage des ARNm cibles est portée, non pas par un composant recruté autour du complexe RISC néo-formé, mais par l'un des composants essentiel de ce complexe et appartenant à la famille des protéines ARGONAUTE (PPD). Chez les mammifères, quatre membres de cette famille existent. Les travaux réalisés par Hannon et col. (2004) montrent clairement que l'activité slicer responsable du clivage des ARNm au cours du RNAi, est portée par la protéine AGO2, alors que les deux autres protéines sont incapables de réaliser cette étape de clivage. Cette activité slicer d'AGO2 est portée par un domaine RNase H cryptique situé à l'intérieur du domaine PIWI (Figure 13) (Liu *et al.* 2004; Song *et al.* 2004). Cette activité RNase H requiert la présence d'un cation divalent et libère des produits possédant des extrémités 5' phosphate et 3' OH, tout comme l'activité slicer de RISC (Martinez et Tuschl 2004; Schwarz *et al.* 2004). Ce résultat est conforté par une étude parallèle décrivant la structure cristalline de la protéine AGO de *Pyrococcus furiosus*. Cette structure montre de façon claire que le repliement de la protéine sur elle-même permet la formation d'un tunnel où peut être réalisé le clivage de l'ARNm cible (Song *et al.* 2004) (Figure 13).

Le domaine PAZ dont l'affinité pour l'extrémité 3' des ARN simple brin est reconnue, est placé de telle façon que le reste du petit ARN pénètre à l'intérieur du tunnel (Figure 13) (Song *et al.* 2003b; Lingel *et al.* 2004; Song *et al.* 2004). Ce tunnel peut également accueillir l'ARNm cible placé en orientation complémentaire. Des modélisations prédisent que le domaine catalytique présent au niveau du domaine PIWI serait ainsi placé à 9 nucléotides de l'extrémité 5' du petit ARN. Ceci coïncide de façon précise avec les clivages généralement observés pour des siRNAs de 21 nucléotides (Liu *et al.* 2004). Le mutant *ago2* est le seul des quatre mutants *ago* chez les mammifères à présenter des anomalies phénotypiques. Les mutants *ago2* sont létaux au stade embryonnaire et présentent des défauts au niveau de la fermeture du tube neural, des malformations cardiaques ainsi qu'une très grande différence de taille des embryons au stade 10,5 jours comparés aux embryons sauvages (Liu *et al.* 2004). AGO2, présente dans le complexe RISC lors du RNAi, a également été mise en évidence dans le complexe miRNP, inhibiteur de la traduction, chez les

Micro-ARN	Organismes	Gènes cibles	Fonction
<i>lin-4</i>	Nématodes	<i>lin-14, lin-28</i>	Transition de stades larvaires précoces ^{1, 2, 3}
<i>let-7</i>	Nématodes	<i>lin-41, hbl-1</i> (FT putatif)	Transition de stades larvaires tardifs ^{4, 5, 6}
<i>lsy-6</i>	Nématodes	<i>cog-1</i> (FT), <i>unc-37</i> (FT)	Développement asymétrique neuronal ⁷
miR-273	Nématodes	<i>die-1</i>	Développement asymétrique neuronal ⁸
<i>Bantam</i>	Insectes	<i>hid</i> (Pro-apoptotique)	Apoptose ⁹
miR-14	Insectes	nd	Apoptose et métabolisme des acides gras ¹⁰
miR-181	Mammifères	nd	Différenciation hématopoïétique ¹¹

¹Lee *et al.* 1993; ²Wightman *et al.* 1993; ³Moss *et al.* 1997; ⁴Reinhart *et al.* 2000; ⁵Abrahante *et al.* 2003; ⁶Lin *et al.* 2003; ⁷Johnston and Hobert 2003; ⁸Chang *et al.* 2004; ⁹Brennecke *et al.* 2003; ¹⁰Xu *et al.* 2003; ¹¹Chen *et al.* 2004.

Tableau 4 : micro-ARNs impliqués dans différents processus cellulaires et physiologiques chez les animaux.

mammifères. Il semble donc qu'un appariement insuffisant entre le petit ARN et sa cible modifie l'arrangement spatial du duplex dans le complexe miRNP. L'activité slicer serait alors bloquée et l'activité d'inhibition de la traduction serait alors privilégiée. Ce blocage de la traduction serait réalisé par interaction directe du complexe RISC avec les ribosomes (Pham *et al.* 2004).

Chez les plantes, la protéine PPD portant l'activité de clivage "slicer" n'est actuellement pas identifiée. Suite au clivage des cibles de micro-ARNs chez les plantes, les deux parties de l'ARNm cibles sont rapidement éliminées. Chez *A. thaliana*, la partie 3' de certaines cibles de micro-ARNs est dégradée dans le cytoplasme par une protéine nommée AtXRN4 (Souret *et al.* 2004). En effet, les mutants *xrn4* accumulent les fragments 3' générés par le clivage de certains micro-ARNs. Le devenir du fragment 5' généré lors du clivage est encore à élucider. Chez les plantes, la régulation des ARNm cibles de micro-ARNs se faisant majoritairement par clivage, l'identification de ce mutant *xrn4* fournit un outil intéressant pour la validation des sites de clivage de cibles peu abondantes ou celles dont les extrémités 3' sont rapidement dégradées. En effet, certains micro-ARNs, tels que miR169 ou miR399, n'ont toujours pas de cible validée.

2.1.4. Implication des micro-ARNs dans différents processus cellulaires et physiologiques

Chez les plantes, la majorité des cibles de micro-ARNs sont des ARNm codant des facteurs de transcriptions, alors que chez les animaux les cibles sont plus variées. Malgré cette divergence, les études précises de la fonction des micro-ARNs ont dans tous les cas permis de montrer que ces derniers remplissent des fonctions essentielles pour un développement correct (Tableau 4).

Chez *C. elegans*, le micro-ARN *let-7* est essentiel à la transition du dernier stade larvaire vers le stade adulte. Il régule l'expression du gène hétérochronique *lin-41* dont le produit est une protéine clé qui a pour fonction de réprimer l'apparition des caractères adultes (Reinhart *et al.* 2000; Slack *et al.* 2000). Les mutants *let-7* présentent ainsi des caractères cellulaires larvaires au stade adulte.

D'autres micro-ARNs impliqués dans les transitions de stades ont été mis en évidence chez *C. elegans*. Le micro-ARN *lin-4* est impliqué dans la transition du premier stade larvaire vers le second (Moss *et al.* 1997). Lorsque le site de complémentarité du micro-ARN *lin-4* est muté au niveau de la région 3' non traduite du gène cible *lin-28*, la progression de la différenciation des lignées cellulaires somatiques ne se fait pas correctement.

Récemment, il a été montré que les micro-ARNs chez *C. elegans* ne sont pas seulement impliqués dans les transitions de stades. Ainsi, le micro-ARN *lisy-6* (laterally symmetrical), dont l'expression est restreinte à moins de dix neurones, permet de bloquer l'expression du gène *cog-1* au niveau du neurone sensoriel gauche, nommé ASEL. L'expression du micro-ARN *lisy-6* n'est pas détectée au niveau du neurone sensoriel droit (ASER). Ceci permet l'expression de *cog-1* au niveau de ce neurone et conduit à une polarisation des neurones droit et gauche. La perte de restriction spatiale de l'expression du gène *cog-1* par le micro-ARN *lisy-6* conduit à une perte de polarisation gauche/droite des neurones sensoriels conduisant à une perte de discrimination de certains signaux chimiques (Johnston et Hobert 2003).

Chez *D. melanogaster*, plusieurs exemples de régulations de processus cellulaires par des micro-ARNs ont également été décrits. Les mutants affectés dans la production du micro-ARN miR-14 sont viables mais sont hypersensibles aux stress, présentent une forte augmentation en certains acides gras et ont une espérance de vie réduite (Xu *et al.* 2003). miR14 a ainsi été identifié comme un suppresseur de mort cellulaire programmée et est également impliqué dans la régulation du métabolisme des acides gras. Parmi les cibles potentielles de ce micro-ARN, l'ARNm codant la caspase pro-apoptotique DRICE est un candidat particulièrement intéressant puisque cette protéine est impliquée dans certains processus de mort cellulaire. Un autre micro-ARN, nommé *Bantam* a été identifié pour son action anti-apoptotique. Tout comme miR-14, le micro-ARN *Bantam* est impliqué dans la régulation spatio-temporelle d'un gène pro-apoptotique, nommé *hid* chez la drosophile (Brennecke *et al.* 2003). Il semble que de nombreux autres micro-ARNs impliqués dans la suppression des processus de mort cellulaire programmée aient été identifiés, cependant les cibles

de ces micro-ARNs ne sont pas validées (Enright *et al.* 2003; Xu *et al.* 2003). La compréhension de l'implication des micro-ARNs dans la régulation des processus apoptotiques est d'un intérêt particulier pour la compréhension des mécanismes de survie cellulaire au cours des processus de tumorigenèse. En effet, les cribles réalisés pour identifier les oncogènes impliqués dans certains cancers par comparaison des transcriptomes de cellules saines et cancéreuses ont probablement écarté d'éventuels précurseurs de micro-ARNs ou autres transcrits non codants (Xu *et al.* 2003). Chez l'homme, des dérégulations de l'expression des micro-ARNs sont dès à présent associées à certains types de cancers (Calin *et al.* 2002; Michael *et al.* 2003; Metzler *et al.* 2004). Les travaux de Calin et collègues (2002) ont ainsi pu associer une leucémie lymphocytaire chronique (B-CLL) à une absence ou diminution de l'abondance des micro-ARNs miR-15 et miR-16 provoquée par une délétion d'une région du chromosome 13 contenant les gènes *mir-15* et *mir-16*. Une autre étude, après clonage de petits ARN à partir de muqueuse colorectale, a permis de montrer que les micro-ARNs miR-143 et miR-145 ne sont plus détectés dans les cellules de différents types de carcinomes (Michael *et al.* 2003). Chez les mammifères, une autre fonction clairement attribuée aux micro-ARNs concerne la différenciation hématopoïétique. Grâce à des expériences d'expression ectopique de micro-ARNs, il a pu être montré que miR-181 est impliqué dans la différenciation des lymphocytes B, alors que miR-142 et miR-223 semblent impliqués dans la différenciation des lymphocytes T. Le profil d'expression de chacun de ces micro-ARNs coïncide avec la fonction qui leur est attribuée par les expériences d'expression ectopique (Chen *et al.* 2004a).

Récemment, des micro-ARNs ont également été mis en évidence au sein du génome du virus d'Epstein-Barr (Pfeffer *et al.* 2004). Tout comme les micro-ARNs animaux, les micro-ARNs viraux décrits se présentent sous la forme de clusters. Les transcrits viraux sont capables de former une structure en épingle à cheveux et utilisent la machinerie de biosynthèse des micro-ARNs de la cellule hôte. Ces micro-ARNs sont capables de cibler des gènes des cellules hôtes impliqués dans la prolifération cellulaire, des gènes codant des régulateurs de transcription ou des composants de certaines voies de transduction des signaux. L'un de ces micro-ARNs est capable de produire le

Micro-ARN	Gènes cibles	Fonction
miR159	Facteurs de transcription MYB (MYB33, MYB65)	Développement des feuilles contrôle de la transition florale ¹
miR-JAW	Facteurs de transcription TCP (TCP2, TCP3, TCP4, TCP10, TCP24)	Architecture des feuilles ²
miR164	Facteurs de transcription à domaine NAC (CUC1, CUC2...)	Restriction de la zone frontière du méristème ⁴ Séparation des organes embryonnaires, végétatifs et floraux ^{3,4}
miR165/166	Facteurs de transcription HD-ZIP II (REV, PHV, PHB)	Polarisation abaxiale/adaxiale des feuilles ^{5,6}
miR168	AGO1	Contrôle le ciblage des micro-ARNs, Axialisation des feuilles ^{7,8}
miR172	Facteurs de transcription AP2-like (AP2, TOE1, TOE2, TOE3)	Développement des fleurs, Détermination des organes floraux, contrôle de la floraison ^{9,10}

¹Achard *et al.* 2004; ²Palatnik *et al.* 2003; ³Mallory *et al.* 2004a; ⁴Laufs *et al.* 2004; ⁵Mallory *et al.* 2004b; ⁶Juarez *et al.* 2004 ; ⁷Vaucheret *et al.* 2004; ⁸Kidner et Martienssen 2004; ⁹Aukerman et Sakai 2003; ¹⁰Chen 2004.

Tableau 5 : micro-ARNs impliqués dans différents processus cellulaires et physiologiques chez les plantes.

clivage de l'ARNm codant l'ADN polymérase virale au cours du stade de latence et du stade lytique (Pfeffer *et al.* 2004). Les micro-ARNs identifiés chez les plantes semblent principalement dédiés à la régulation de l'expression de gènes impliqués dans des fonctions développementales (Tableau 5). D'ailleurs, les mutants de gènes impliqués dans la biosynthèse ou l'action des micro-ARNs, tels que *hen1* ou *dc11* présentent des anomalies développementales sévères (Park *et al.* 2002; Reinhart *et al.* 2002; article annexe n°1 : Boutet *et al.* 2003). Le premier exemple de micro-ARN chez les plantes a surtout permis de mettre en évidence le mécanisme d'action prédominant des micro-ARNs par clivage des ARNm cibles chez les plantes (Llave *et al.* 2002b). Cette étude posait cependant les prémisses des régulations développementales décrites ci-après puisque ces travaux montrèrent également que miR171 était exprimé préférentiellement dans les feuilles et les inflorescences chez *A. thaliana* et que ce profil d'expression est conservé chez le riz. Ce micro-ARN permet la régulation de trois membres de la famille des facteurs de transcription SCARECROW-LIKE ou GRAS, dont certains membres sont impliqués dans des voies de développement, telles que le développement radial des racines ou la signalisation hormonale (Di Laurenzio *et al.* 1996; Helariutta *et al.* 2000).

Un autre micro-ARN, miR164 permet la régulation de la famille des facteurs de transcription à domaine NAC (Mallory *et al.* 2004a). Cette famille de facteurs de transcription spécifique des plantes compte plus d'une centaine de membres et tient son nom de trois de ses membres les plus illustres: NAM (No Apical Meristem) du pétunia, ATAF1 et ATAF2 d'*Arabidopsis* et CUP-SHAPED COTYLEDON2 d'*Arabidopsis* (CUC2) (Aida *et al.* 1999; Ooka *et al.* 2003). Les membres CUC1 et CUC2 sont impliqués dans la formation du méristème apical ainsi que dans la séparation des cotylédons, des sépales et des étamines (Takada *et al.* 2001; Aida *et al.* 1999). L'expression d'un transcrit *CUC1* résistant à la reconnaissance et donc à la régulation par son micro-ARN a permis de mettre en évidence que la régulation de *CUC1* par miR164 est essentielle pour un développement correct. Les plantes exprimant cette copie modifiée du transcrit *CUC1* présentent en effet de nombreuses anomalies: plusieurs pétales supplémentaires, de un à deux sépales manquants, une diminution importante de la fertilité, un défaut d'orientation des cotylédons et une réduction de

la taille des pétioles des feuilles (Mallory *et al.* 2004a). Le point majeur de cette étude a été de montrer que l'expression ectopique de miR164 récapitule le phénotype observé chez le double mutant *cuc1 cuc2* (fusion des pétioles des cotylédons, des étamines et des sépales), montrant ainsi sans ambiguïté que la principale régulation de l'expression des gènes à domaine NAC est réalisée par miR164.

De la même façon, l'expression d'une copie mutée du gène *CUC2* au niveau du site d'appariement de miR164, lorsque son expression est induite dans le méristème, conduit à une sévère inhibition du développement des feuilles (Laufs *et al.* 2004). Des analyses microscopiques de plantes transformées avec le gène rapporteur de la GFP sous le contrôle d'un promoteur inductible par l'éthanol et portant les séquences régulatrices du gène *STM*, permettant de visualiser la zone frontière des méristèmes végétatifs et apicaux, ont permis de montrer que la régulation de *CUC2* par miR164 permet de restreindre la taille de cette zone frontière. Chez des plantes exprimant le transcrit résistant au clivage par miR164, ainsi que chez les mutants *hen1-5*, *hyl1-1* et *dcl1-9*, on observe un élargissement progressive de ce zone frontière méristématique, expliquant les fusions d'organes décrites précédemment (Laufs *et al.* 2004).

D'autres études ont permis de montrer que le micro-ARN miR-JAW grâce à la régulation de 5 membres de la famille des facteurs de transcription TCP est essentiel à la morphologie des feuilles (Palatnik *et al.* 2003). Les micro-ARNs miR165 et miR166, dont les séquences varient d'un seul nucléotide, sont eux essentiels pour la polarité radiaire des plantules et la polarisation adaxiale-abaxiale des feuilles et régulent les gènes de la famille des facteurs de transcription HD-ZIP de classe III (Revoluta, Phabulosa et Phavoluta) (Emery *et al.* 2003; Tang *et al.* 2003; Juarez *et al.* 2004; Mallory *et al.* 2004b). Les études concernant la régulation des gènes de la famille HD-ZIP III par miR165 et miR166 indiquent que cette régulation permet de restreindre l'expression de ces gènes au niveau du méristème apical et à la surface adaxiale des primordiums foliaires (Emery *et al.* 2003; Juarez *et al.* 2004; Kidner et Martienssen 2004). La perte de régulation par ces micro-ARNs conduit à l'expression ectopique des gènes de la famille HD-ZIP III dans le domaine abaxial des primordiums foliaires, changeant ainsi la position de la jonction adaxial/abaxial. Cette régulation

Micro-ARN	Famille des gènes ciblés	Cibles validées
miR156 miR157	Facteurs de transcriptions Squamosa Binding Protein	SBP-like/At5g43270 ¹ , SPL10/At1g27370 ² , SPL3/At2g33810 ³
miR159	Facteurs de transcription MYB	MYB33/At5g06100 ⁴ , MYB65/At3g11440 ⁴
miR-JAW	Facteurs de transcription TCP	TCP2/At4g18490 ⁴ , TCP3/At1g53230 ⁴ , TCP4/At3g15030 ⁴ , TCP10/At2g31070 ⁴ , TCP24/At1g30210 ⁴
miR160	Auxin Response Factor	ARF10/At2g28350 ¹ , ARF17/At1g77850 ¹
miR161	Pentatricopeptide Repeat Protein	PPR/At1g06580 ²
miR162	DCL1	DCL1/At1g01040 ⁵
miR164	Facteurs de transcription à domaine NAC	CUC2/At5g53950 ¹ , CUC1/At3g15170 ¹ , At5g61430 ⁶ , At5g07680 ⁶ , NAC1/At1g56010 ⁶
miR165 miR166	Facteurs de transcription HD-ZIP III	PHB/At2g34710 ⁷ , PHV/At1g30490 ⁷ , REV/At5g60690 ⁸
miR167	Auxin Response Factor	ARF8/At5g37020 ¹
miR168	Argonaute	AGO1/At1g48410 ²
miR169	CCAAT Binding Factor	At3g05690 ⁹
miR170 miR171	Facteurs de transcription Scarecrow-like	SCL6-III/At3g60630 ¹⁰ , SCL6-IV/At4g00150 ¹⁰
miR172	Facteurs de transcription APETALA2-Like	AP2/At4g36920 ^{1,11,12} , TOE1/At2g28550 ^{1,12} , TOE2/At5g60120 ^{1,12} , TOE3/At5g67180 ¹
miR393	Protéines F Box	At1g12820 ⁹ , At3g26810 ⁹ , TIR1/At3g62980 ⁹ , At4g03190 ⁹
miR394	Facteurs de transcription bHLH	At3g23690 ⁹
miR395	Protéines F Box	At1g27340 ⁹
miR395	ATP sulfurylases	APS4/At5g43780 ⁹
miR396	Growth regulating Factor	GRL1/At2g22840 ⁹ , GRL2/At4g37740 ⁹ , GRL3/At2g36400 ⁹ , GRL7/At5g53660 ⁹ , GRL8/At4g24150 ⁹ , GRL9/At2g45480 ⁹
miR397	laccases	At2g29130 ⁹ , At2g38080 ⁹ , At5g60020 ⁹
miR398	Copper Superoxyde Dismutase Cytochrome C oxydase sous-unités V	CSD1/At1g08830 ⁹ , CSD2/At2g28190 ⁹ , At3g15640 ⁹

¹Kasschau *et al.* 2003; ²Vazquez *et al.* 2004; ³Chen *et al.* 2004; ⁴Palatnik *et al.* 2003; ⁵Xie *et al.* 2003; ⁶Mallory *et al.* 2004a; ⁷Tang *et al.* 2003; ⁸Emery *et al.* 2003; ⁹Jones-Rhoades and Bartel 2004; ¹⁰Llave *et al.* 2002b; ¹¹Chen 2004; ¹²Aukerman and Sakai 2003.

Tableau 6 : Cibles de micro-ARNs dont le clivage a été validé par des expériences de 5'RACE.

par miR165 et miR166 est également impliquée dans la formation correcte du système vasculaire et la structuration des méristèmes (Emery *et al.* 2003; Mallory *et al.* 2004b; McHale et Koning 2004). Le micro-ARN miR172 est essentiel pour l'inhibition de la cible APETALA2 ainsi que des deux répresseurs floraux AP2-like nommés TOE1 et TOE2 (Target of EAT-D). La régulation de ces cibles par miR172 est essentielle à l'induction florale ainsi qu'à la spécification des organes floraux (Aukerman et Sakai 2003; Chen 2004).

Le micro-ARN miR162 est, quant à lui, un cas particulier. En effet, il est impliqué dans la régulation de l'ARNm codant la protéine DCL1. La protéine DCL1 étant impliqué dans la production des micro-ARNs, il a été proposé une autorégulation de la machinerie de régulation contrôlée par les micro-ARNs (Xie *et al.* 2003).

D'une manière générale, chez les plantes, les cibles de micro-ARNs sont principalement des facteurs de transcription (Tableau 6). Il semble que les régulations par les micro-ARNs aient pour principale fonction d'arrêter rapidement et complètement la production d'une protéine en un site ou groupe de cellules donné, afin que la différenciation de ces cellules puisse se faire de façon restreinte.

Chez les animaux, les cibles de micro-ARNs semblent impliquées à des niveaux de régulation beaucoup plus divers. Peu de cibles ont été validées et l'appariement micro-ARN/ARNm cible n'est pas parfait. Les cibles prédites chez les mammifères appartiennent aux facteurs de transcription, mais également à des gènes impliqués dans la transduction des signaux ou encore au métabolisme du collagène (Lewis *et al.* 2003). Seule la validation de ces cibles permettra d'assigner une fonction aux régulations par les micro-ARNs chez les animaux.

L'implication des micro-ARNs dans les différents processus cellulaires et physiologiques décrits précédemment soulève la question de la régulation de l'expression de ces micro-ARNs.

2.1.5. Régulation de l'expression des gènes de micro-ARNs?

Chez les animaux, les profils d'expression des micro-ARNs ont été étudiés de façon plus approfondie que chez les plantes. Ainsi, les co-expressions des micro-ARNs miR-310, miR-311 et miR-312 ou des micro-ARNs miR-100, miR-125 et *let-7* chez la drosophile ont été vérifiées (Aravin *et al.* 2003). Ces deux arrangements de micro-ARNs en cluster présentent chacun un profil d'expression particulier. Les micro-ARNs miR-310, miR-311 et miR-312 sont très abondants au cours des 4 premières heures du développement de l'embryon, et leur niveau d'accumulation diminue progressivement jusqu'à ne plus être détectable dès le premier stade larvaire, alors que les micro-ARNs miR-100, miR-125 et *let-7* ne sont détectés ni au cours du développement de l'embryon ni aux cours des quatre stades larvaires, et s'expriment exclusivement au stade puppe et adulte (Aravin *et al.* 2003). L'expression de ce dernier polycistron est régulée par des hormones. Il a, en effet, été clairement montré qu'une hormone stéroïde (ecdysone) produite en fin de troisième stade larvaire permettait l'activation de la production des micro-ARNs de ce polycistron, alors qu'elle inhibe la production d'un autre micro-ARN: miR-34. De façon réciproque, l'expression de miR-34 est activée et celle des micro-ARNs miR-100, miR-125 et *let-7* est inhibée par une hormone antagoniste, l'hormone juvénile (Sempere *et al.* 2003).

D'autres micro-ARNs animaux présentent des profils d'expression particuliers. Ainsi, chez la souris, miR-223 et miR-181, impliqués dans la différenciation des lignées hématopoïétiques, s'accumulent principalement dans la moelle osseuse et le thymus, respectivement (Chen *et al.* 2004a). Chez *C. elegans*, miR-200b est exclusivement détecté au niveau du poumon et le polycistron des micro-ARNs miR-34 à miR-41 est principalement exprimé au stade embryonnaire (Lau *et al.* 2001; Grad *et al.* 2003). L'observation que les micro-ARNs présentent des sites préférentiels d'expression a ainsi conduit à de nouvelles campagnes de clonage de micro-ARNs à partir de tissus spécifiques ou à des analyses de l'expression des micro-ARNs dans différents tissus. Ainsi, des micro-ARNs spécifiquement exprimés dans le cerveau ont pu être identifiés chez la souris (Krichevsky *et al.* 2003; Kim *et al.* 2004). Le micro-ARN miR-19b est détecté dans le cerveau avant la naissance, et n'est plus détecté ensuite (Krichevsky *et al.* 2003), les miR-9, miR-131, miR-129-2* et miR-124a sont exclusivement détectés au niveau du cerveau (Lagos-Quintana *et al.* 2002; Krichevsky *et al.*

2003; Kim *et al.* 2004). Le clonage de micro-ARNs à partir de 8 tissus différents chez la souris a également permis de montrer que miR-1a est exclusivement exprimé au niveau du coeur (Lee et Ambros 2001; Lagos-Quintana *et al.* 2002) et que miR-122a est exclusivement exprimé dans le foie (Lagos-Quintana *et al.* 2002).

Chez les plantes, l'expression des gènes de micro-ARNs est également régulée de façon spatio-temporelle. Ainsi le micro-ARN miR1 s'accumule préférentiellement dans des plantules de riz au stade première feuille (Wang *et al.* 2004), tandis que miR167 d'*Arabidopsis* dont l'accumulation est quasiment identique dans tous les tissus adultes testés ne semble pas exprimé dans des plantules de moins de 14 jours (Park *et al.* 2002; Reinhart *et al.* 2002). L'expression des micro-ARNs chez les plantes est également dépendante de facteurs hormonaux et environnementaux. Ainsi, l'accumulation de miR395 diminue en présence de fortes concentrations en sulfate. Cette diminution résulte ainsi en une accumulation de ses cibles, parmi lesquelles *APSI* qui est impliquée dans les premières étapes de l'assimilation du sulfate inorganique (Jones-Rhoades et Bartel 2004). De même, miR159 s'accumule en présence d'acide gibbérellique (AG). Les cibles de miR159 sont des facteurs de transcription de la famille MYB impliqués dans l'induction de la floraison en jours courts, l'apport d'AG conduit ainsi à une floraison tardive (Achard *et al.* 2004).

L'expression spatio-temporelle des micro-ARNs suggère fortement que leur production est dépendante, comme n'importe quel autre gène codant, d'un promoteur contrôlé par des éléments de régulation en cis. De plus, la description de précurseurs de micro-ARNs polyadénylés chez les animaux (Lagos-Quintana *et al.* 2002) et les plantes (Kurihara et Watanabe 2004), laisse entendre que l'ARN polymérase ADN dépendante II est responsable de leur transcription. L'expression des micro-ARNs serait donc régulée comme n'importe quel autre gène au niveau transcriptionnel et peut-être même au niveau post-transcriptionnel par une régulation des étapes de maturation des micro-ARNs.

2.2. La méthylation induite par des siRNAs endogènes

Comme nous l'avons vu dans le paragraphe 1.1.4 consacré aux composants du mécanisme de RdDM, ce processus est impliqué dans le maintien de l'état inactif des rétroéléments chez les plantes. L'état transcriptionnellement inactif du rétrotransposon AtSN1 est clairement associé à la présence de siRNAs (Xie *et al.* 2004).

Lors de campagnes de clonage de petits ARN endogènes chez *A. thaliana*, il a été rapporté que 27% des siRNAs endogènes clonés correspondaient à des transposons (Xie *et al.* 2004). Il a d'ailleurs été montré récemment que l'état inactif des transposons présents au niveau d'un site interstitiel hétérochromatique (des régions hétérochromatiques différentes des régions centromériques et péri-centromériques), caractérisé par une méthylation de l'ADN et des méthylations de la lysine 9 des histones H3, est perdu chez le mutant *ddm1* (Lippman *et al.* 2004).

Les transposons présents au niveau de ce site, pour lesquels des siRNAs ont été décrits, perdent les marques hétérochromatiques (méthylation de l'ADN et méthylation des lysines 9 des histones H3) et sont transcrits chez le mutant *ddm1* (Lippman *et al.* 2004). Il semble que l'état hétérochromatique de ce site interstitiel soit ainsi largement sous le contrôle du processus de RdDM. Comme certains des siRNAs clonés par Xie et col. (2004) correspondent à des transposons présents au niveau de régions hétérochromatiques centromériques et péri-centromériques, il est probable que la formation de ces régions dépende également de ce processus de RdDM.

Chez *A. thaliana*, d'autres loci sont inactivés par le processus de RdDM. Ainsi le gène *FWA*, décrit précédemment, le gène *PAI2* (codant une phosphoribosyl anthranilate isomerase) impliqué dans la synthèse de tryptophane (Jeddeloh *et al.* 1998), le gène *MEDEA* qui est transmis dans un état inactif par le chromosome mâle (Vielle-Calzada *et al.* 2000) et *TSI* (transcriptionally silent information) une copie dégénérée du transposon Athilla dont la fonction du transcrit qu'elle produit est inconnue (Steimer *et al.* 2000). Tous ces loci sont réactivés chez le mutant *met1*, et sont des cibles potentielles de RdDM.

Chez d'autres espèces, de nombreuses évidences montrent que le RdDM existe et est dirigé contre les éléments transposables, même si les mécanismes sous-jacents sont moins élucidés. Ainsi, les mutants *mut-2*, *mut-7*, *rde-2* et *rde-3* impliqués dans le RNAi chez *C. elegans* présentent une forte réactivation des éléments transposables Tc1 et Tc4 (Ketting *et al.* 1999; Tabara *et al.* 1999) (Grishok *et al.* 2000). Le mutant *mut6*, affecté dans le gène codant une DEAH box RNA hélicase également impliquée dans la PTGS chez *C. reinhardtii*, montre une réactivation du rétroélément TOC1 et du transposon Gulliver (Wu-Scharf *et al.* 2000).

De plus, chez la drosophile, il a été montré que le processus de Position effect Variegation (PEV), processus capable d'induire la formation d'un état inactivé d'un gène lorsqu'il est placé dans un domaine hétérochromatique, nécessite la fonction des protéines PIWI, AUBERGINE et HOMELESS, également requises pour le RNAi (Pal-Bhadra *et al.* 2004). Ceci suggère que la formation d'un état hétérochromatique chez la drosophile pourrait être du à des siRNAs.

La démonstration du ciblage de promoteurs par des siRNAs exogènes a été faite récemment chez les mammifères suggérant que le processus pourrait également avoir lieu avec des siRNAs endogènes (Kawasaki et Taira 2004).

2.3. Conclusion

L'implication des ARN non codants dans la régulation de l'expression des gènes impliqués dans de nombreux processus et chez de très nombreux organismes est actuellement en plein essor. Ceux-ci peuvent intervenir à tous les niveaux de l'expression génique en influençant la transcription, la stabilité des ARNm ou la traduction. Les cas de régulations géniques par des ARNnc se multiplient et des programmes de recherche visant à identifier les ARNnc à grande échelle voient le jour. Certains sont des projets bioinformatiques (Lehner *et al.* 2002; Shendure et Church 2002; Yelin *et al.* 2003; Bonnet *et al.* 2004), d'autres visent à cloner l'ensemble des ADNc pleine longueur chez plusieurs espèces et permettent ainsi d'identifier les ARNnc et paires d'ARN sens-antisens (Kiyosawa *et al.* 2003; Osato *et al.* 2003).

Ces approches permettront d'identifier des candidats d'intérêt dont l'étude approfondie permettra probablement de généraliser les régulations de l'expression des gènes par les ARN non codants.

Chez les plantes, plusieurs mutants affectés dans le processus d'inactivation épigénétique post-transcriptionnelle ont été identifiés. Les nombreux points communs entre les siRNAs issus de transgènes et les micro-ARNs nous ont donc conduit, en collaboration avec l'équipe d'Hervé Vaucheret, à tester l'implication de ces différentes protéines dans la voie de régulation génique par les micro-ARNs. Parallèlement à ces études, nous avons initié un crible moléculaire dans le but d'identifier des ARNm cibles de siRNAs endogènes putatifs. Nous l'avons vu, les siRNAs endogènes décrits à ce jour chez les plantes sont produits par une voie impliquant une ARN polymérase ARN dépendante nommée RDR2 et une RNase III nommée DCL3 (Xie *et al.* 2004). Ces siRNAs, contrairement aux siRNAs issus de transgènes, agissent en cis en induisant une inactivation transcriptionnelle des séquences à leur origine.

Afin d'identifier des cibles de siRNAs endogènes putatifs agissant de façon comparable à ceux impliqués dans la voie de PTGS sens, nous avons comparé les transcriptomes des mutants *sgs2* et *sgs3* à celui d'une plante sauvage. Nous avons pu mettre en évidence l'existence d'une nouvelle classe de siRNAs endogènes, dont la production dépend notamment de la protéine RDR6, et de la protéine SGS3. L'utilisation des mutants identifiés des voies des micro-ARNs et de PTGS sens ont permis d'identifier plusieurs composants nécessaires à l'accumulation de ces nouveaux siRNAs. Nous avons également montré que ces siRNAs régulent l'accumulation d'ARNm cibles produits à partir de loci indépendants du locus produisant ces siRNAs. Les travaux réalisés sur les voies de biosynthèse et d'action des micro-ARNs et des siRNAs endogènes ont contribué à la compréhension des régulations géniques post-transcriptionnelles chez les plantes.

-

II. RESULTATS & DISCUSSION

-

PREMIÈRE PARTIE: Contribution à l'identification de protéines impliquées dans la voie de régulation des gènes par les micro-ARNs chez *Arabidopsis thaliana*.

1.1. Introduction

La découverte des micro-ARNs chez les Eucaryotes est une révolution puisqu'elle correspond à la mise en évidence de nouveaux mécanismes de régulation de l'expression des gènes. Chez les plantes, l'identification de plusieurs acteurs du mécanisme de biosynthèse et du mécanisme d'action des micro-ARNs a pu être réalisée grâce à l'étude du processus d'inactivation épigénétique post-transcriptionnelle impliquant des transgènes. L'identification d'acteurs dont certains sont impliqués dans la voie de biosynthèse des siRNAs de transgènes et d'autres dans la voie de biosynthèse des micro-ARNs permet de séparer ces deux voies. En effet, si certains sont communs aux deux voies, d'autres sont spécifiques à l'une ou à l'autre. Ceci s'explique par leur mode de production: les siRNAs sont produits à partir d'un long ARN double brin découpé en de multiples siRNAs alors que les micro-ARNs sont générés à partir d'un repliement intramoléculaire créant localement une structure double brin qui couvre la séquence du micro-ARN mature. Il existe donc une différence au niveau des acteurs nécessaires pour la stabilisation des zones ARNdb formées ainsi que pour la production des petits ARN matures. Il a par exemple été montré chez les plantes, que la protéine DCL1 est nécessaire à la maturation des micro-ARNs à partir de la structure en épingle à cheveux, alors qu'elle n'est pas requise pour la production des siRNAs de transgènes (Reinhart *et al.* 2002; Finnegan *et al.* 2003; Kurihara et Watanabe 2004).

Comme nous l'avons vu en introduction (revue bibliographique paragraphe 2.1.3.2.A), les micro-ARNs et les siRNAs chez les plantes possèdent des caractéristiques communes. En effet, la majorité des micro-ARNs identifiés à ce jour chez les plantes agissent tels des siRNAs impliqués dans le RNAi, par clivage de leurs cibles au milieu de la zone de complémentarité.

Les protéines nécessaires au processus de PTGS induit par les siRNAs chez les plantes ont donc été testées afin de savoir si certaines étaient également impliquées dans les voies de biosynthèse et d'action des micro-ARNs.

1.2. Contribution à la mise en évidence de l'implication des protéines HEN1 et HYL1 dans la voie de biosynthèse des micro-ARNs et de la protéine AGO1 dans les étapes effectrices de régulation des gènes par les micro-ARNs.

Les travaux réalisés par l'équipe d'Hervé Vaucheret ont permis d'identifier plusieurs protéines impliquées dans la PTGS induite par des transgènes sens (S-PTGS) (Elmayan *et al.* 1998; Mourrain *et al.* 2000; Fagard *et al.* 2000; Morel *et al.* 2002). L'introduction d'un transgène possédant le gène rapporteur *uidA* codant la protéine GUS (β -glucuronidase) sous le contrôle du promoteur fort et constitutif de l'ARN 35S du virus de la mosaïque du chou fleur (35S CaMV) a permis à cette équipe d'obtenir plusieurs lignées transgéniques inactivant le gène rapporteur au niveau post-transcriptionnel. L'une de ces lignées, la lignée L1 possédant le locus 23b1, présente l'inactivation post-transcriptionnelle du transgène la plus efficace parmi les quatre lignées obtenues et a donc été retenue pour le crible d'identification de mutants de S-PTGS (Elmayan *et al.* 1998). La mutagenèse chimique (EMS) de graines de cette lignée L1 a permis d'identifier 30 plantes mutantes, présentant une réactivation de l'expression du transgène *uidA*, classées en quatre groupes de complémentation *sgs1*, *sgs2/rdr6*, *sgs3* et *ago1* (Elmayan *et al.* 1998; Fagard *et al.* 2000; Morel et Vaucheret 2000; Mourrain *et al.* 2000; Morel *et al.* 2002).

La poursuite de ce crible a permis d'identifier un nouvel acteur de la voie de S-PTGS, HEN1, n'appartenant pas à l'un de ces quatre groupes. Boutet et col. (2003) (article annexe n°1) ont ainsi pu montrer que le mutant 23-2 est allélique des mutants *hen1-1*, *hen1-2* et *hen1-3* caractérisés précédemment (Chen *et al.* 2002; X. Chen non publié). Ce mutant 23-2, appelé *hen1-4*, comme les mutants de S-PTGS (Mourrain *et al.* 2000; Beclin *et al.* 2002) présente une hypersensibilité au virus CMV, se traduisant par une accumulation des ARNs viraux (Figure 1 article annexe n°1: Boutet *et*

al. 2003), et une accumulation des ARNm *uidA* comparable à celle observée chez les autres mutants (Figure 2A article annexe n°1: Boutet *et al.* 2003). Les travaux de Vaucheret et col. ont également permis de montrer que les mutants *hen1-4* et *hen1-5* sont affectés dans l'accumulation de 8 des 15 micro-ARNs identifiés au moment de ces travaux (Figure 2 article annexe n°2: Vazquez *et al.* 2004a).

La production des micro-ARNs chez les plantes requiert donc l'intervention de HEN1 mais également de DCL1 (Park *et al.* 2002; Reinhart *et al.* 2002; article annexe n°1 : Boutet *et al.* 2003). Ces deux protéines semblent toutes deux localisées au sein du noyau, comme l'indique la présence de séquences de localisation nucléaire, et elles possèdent toutes deux des domaines de liaison aux ARNdb. Afin de tester l'implication d'autres protéines d'*Arabidopsis* dans la voie de régulation génique par les micro-ARNs, l'équipe d'Hervé Vaucheret a étudié des lignées transgéniques possédant une insertion dans des gènes codant des protéines ayant un domaine de fixation aux ARNdb et un domaine d'adressage au noyau. L'une de ces lignées, homozygote pour l'insertion SALK_064863 dans le gène At1g09700, présente des anomalies de développement rappelant partiellement celles observées chez les mutants *hen1* et les mutants partiels *dcl1*. Ces travaux ont permis de montrer que la protéine HYL1 (Hyponastic leaves) n'est pas requise pour le processus de S-PTGS (Figure 14B; Han *et al.* 2004), ni pour la PTGS induite par des ARNdb, mais que le mutant *hyl1-2* est affecté de façon partielle dans l'accumulation des 8 micro-ARNs testés précédemment (Figure 2 article annexe n°2: Vazquez *et al.* 2004a).

Au cours de la collaboration avec l'équipe d'Hervé Vaucheret, mon travail a donc consisté en une analyse détaillée de l'accumulation des cibles des micro-ARNs chez les différentes plantes mutantes produites et caractérisées par cette équipe. J'ai également validé de nouvelles cibles de microARNs.

1.2.1. Quantification des ARNm *uidA* par la technique de PCR quantitative en temps réel

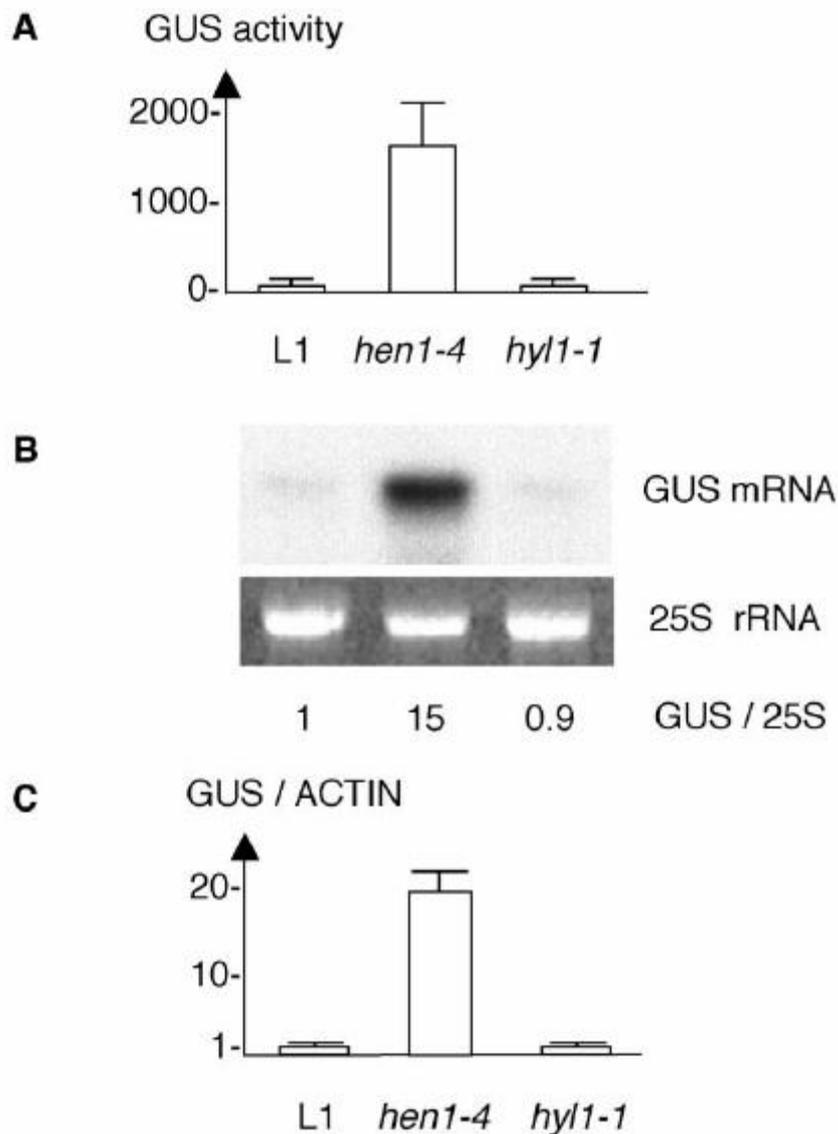


Figure 4. GUS Activity and *GUS* mRNA Accumulation in Wild-Type and Mutant Plants

(A) GUS activity in leaves of L1 plants triggering S-PTGS and of *hen1-4* and *hyl1-1* mutants homozygous for the L1 locus. The average activity of eight measurements in eight independent plants is given. Deviation bars correspond to standard errors.

(B) *GUS* mRNA relative accumulation estimated by Northern blot. mRNA extracted from leaves was hybridized with a *GUS* DNA probe. Ethidium bromide-stained gel is shown for standardization. The ratio between *GUS* and 25S signals is indicated.

(C) *GUS* mRNA relative accumulation estimated by RT-qPCR. Quantifications are normalized with *ACTIN2*. The L1 value is arbitrarily fixed to one.

Figure 14: Figure 4 de l'article annexe n°2 (Vazquez et al. 2004a).

J'ai tout d'abord confirmé l'effet des différentes mutations présentées ci-dessus sur l'accumulation des ARNm *uidA* par quantification des ARNm par PCR quantitative en temps réel. Pour cela j'ai désigné un couple d'amorces de PCR répondant aux exigences de la PCR quantitative (voir matériel et méthode de l'article annexe n°1: Boutet *et al.* 2003) et amplifié les ADNc *uidA*, afin de quantifier l'accumulation relative du transcrit *uidA* chez les mutants *hen1-4* et *hyl1-1* (Figure 14C). Les ADNc ont été synthétisés à partir d'ARNs extraits du même matériel végétal que celui ayant servi à réaliser le northern blot présenté dans la figure 14B (Vazquez *et al.* 2004a), c'est-à-dire de feuilles de plantes ayant poussé en serre. Ces données montrent une accumulation 20 fois supérieure des ARNm *uidA* chez la plante *hen1-4* comparée à celle observée chez la plante sauvage isogénique de L1. Chez le mutant *hyl1-1*, possédant le locus 23b1, le transcrit *uidA* présente le même niveau d'accumulation que celui observé chez la plante sauvage isogénique L1 (Figure 14C). Ces données confirment ainsi l'implication de la protéine HEN1 dans la voie de S-PTGS alors que la protéine HYL1 n'est pas impliquée dans cette voie.

De plus, la quantité des transcrits *uidA* mesurée par PCR quantitative en temps réel chez les mutants *hen1-4* et *hyl1-1* et la plante L1 est du même ordre de grandeur que celle observée par northern blot (Figure 14B et 14C). Ces données confirment également que les conditions d'utilisation de la technique de PCR quantitative en temps réel au laboratoire permettent une quantification relative de l'abondance de transcrits de façon sensible et fiable.

1.2.2. Caractérisation, par la technique modifiée de 5' RLM (RNA ligase-mediated)-RACE, de sites de clivage sur des ARNm cibles potentielles de micro-ARNs.

Au cours de ce travail d'analyse de l'accumulation des cibles de micro-ARNs et afin de réaliser une étude permettant de généraliser l'implication des protéines HEN1, HYL1 et AGO1 dans la voie de régulation de l'expression des gènes par les micro-ARNs, il était indispensable que des cibles potentielles de micro-ARNs soient validées avant leur quantification chez les différents mutants. Au

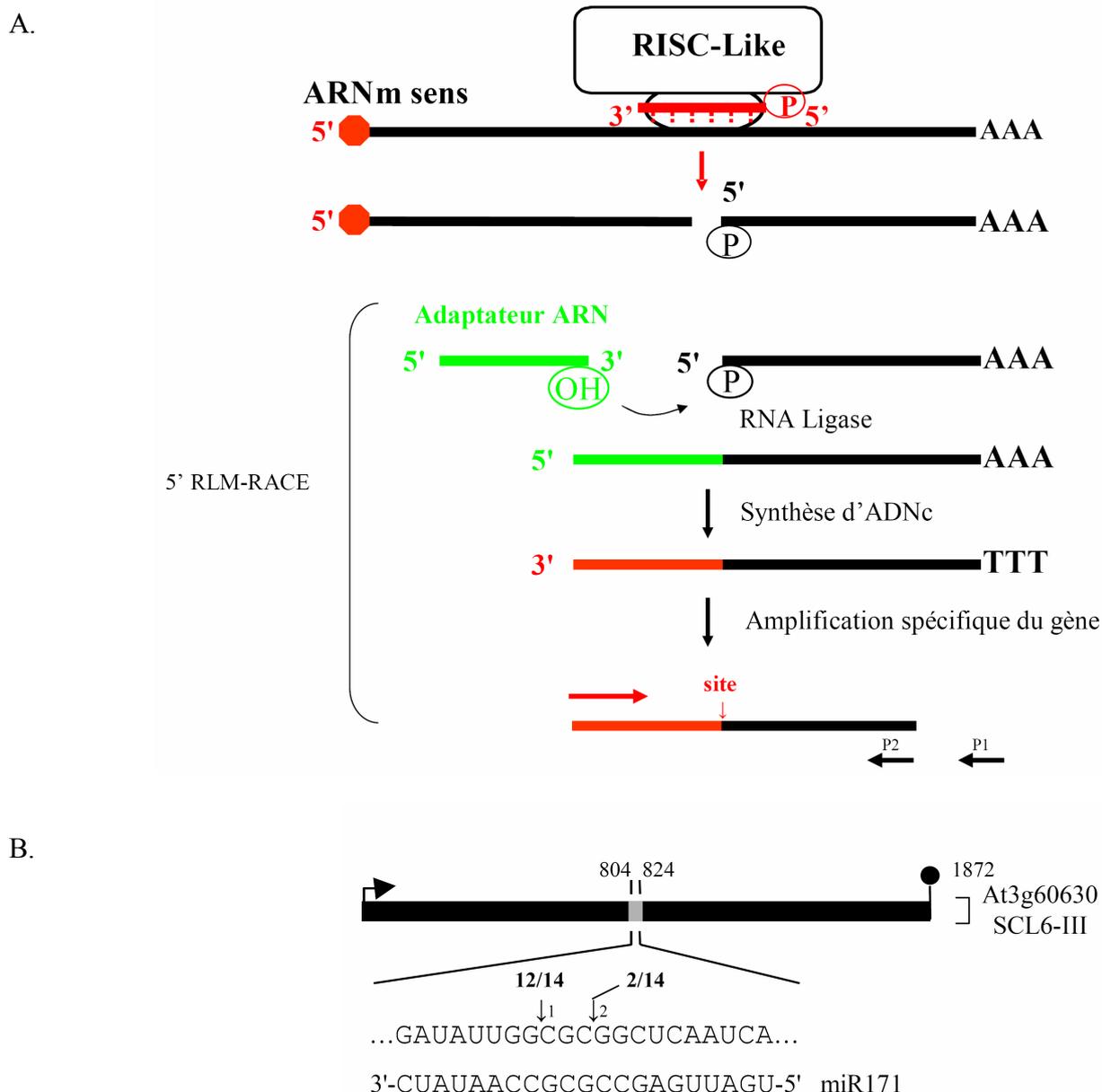


Figure 15: Principe de la technique de 5' RLM-RACE et position des sites de clivage sur le transcrit *SCL6-III* guidé par le micro-ARN miR171.

A. Représentation schématique de la technique modifiée de 5' RNA Ligase-Mediated RACE (5' RLM-RACE). Contrairement à la technique classique, les ARN totaux ne sont pas décapés, ce qui permet la ligature de l'adaptateur uniquement au niveau du site de clivage de l'ARNm cible par le complexe RISC. P1 et P2 indiquent les deux amorces utilisées lors de deux PCR successives.

B. Représentation schématique des sites de clivage sur la séquence codante du transcrit *SCL6-III*. La flèche et le rond indiquent le début et la fin du cadre de lecture du transcrit. La séquence du transcrit *SCL6-III* correspondant à la zone grisée est la séquence complémentaire au micro-ARN. Les sites de clivage 1 et 2 identifiés sont indiqués par des flèches et le nombre de clones correspondant à chaque site sur le nombre total de clones séquencés est indiqué.

moment de ce travail, sur les 15 familles de micro-ARNs identifiées chez *Arabidopsis*, 9 possédaient des cibles dont le clivage avait été validé par des expériences de 5' RLM-RACE modifiée (Llave *et al.* 2002; Kasschau *et al.* 2003). Les micro-ARNs miR161 et miR168 possédaient des cibles prédites (Rhoades *et al.* 2002) mais non validées.

Afin de valider le clivage de cibles de micro-ARNs, j'ai donc préparé des ADNc à partir d'ARNs extraits d'un mélange de feuilles et de fleurs à différents stades de développement (voir matériel et méthode article annexe n°2 : Vazquez *et al.* 2004a). Le choix de ce mélange de matériel végétal a été fait puisque les sites d'action tissulaires des micro-ARNs n'étaient pas connus et offraient ainsi une plus grande chance d'être en présence d'ARNm présentant la marque du clivage guidé par un micro-ARN. Cette marque est la présence d'une extrémité 5' phosphate sur le fragment correspondant à la partie 3' de l'ARNm (Figure 15A). Afin d'identifier le site de clivage sur l'ARNm, les ARN totaux ont été ligaturés à un adaptateur ARN présentant une extrémité 3'OH. Seuls les ARNs possédant une extrémité 5' phosphate, libérée suite au clivage par le complexe miRNP, peuvent être ligaturés à l'adaptateur ARN alors que les transcrits pleine longueur possédant une extrémité 5' capée (7mGuanine) ne peuvent pas être ligaturés à l'adaptateur ARN (Figure 15A). Suite à la transcription inverse avec une amorce polydT de ces ARNs ligaturés, j'ai vérifié que la qualité des matrices produites permettait de retrouver les résultats obtenus par Llave et col. (2002) de validation de la première cible de micro-ARN identifiée chez les plantes, à savoir l'ARNm *SCL6-III* ciblé par le micro-ARN miR171. Après une PCR avec un couple d'amorce, dont une est spécifique du transcrit à caractériser et l'autre spécifique de l'adaptateur ARN, et une électrophorèse en gel d'agarose, une bande unique a été observée, éluée du gel, clonée et séquencée. Les sites de clivage ainsi obtenus pour l'ARNm *SCL6-III* sont localisés au milieu de la zone de complémentarité avec miR171 (Figure 15B) et sont donc bien les mêmes que ceux obtenus par Llave et col. (2002). A noter que des proportions différentes du nombre de clones ont été obtenues pour chacun des sites, 12 clones et 2 clones sur les 14 séquencés, respectivement pour le site 1 et le site 2 (Figure 15B) contre 4 clones sur les 8 séquencés pour chaque site dans l'étude de Llave et col. (2002). Ceci peut

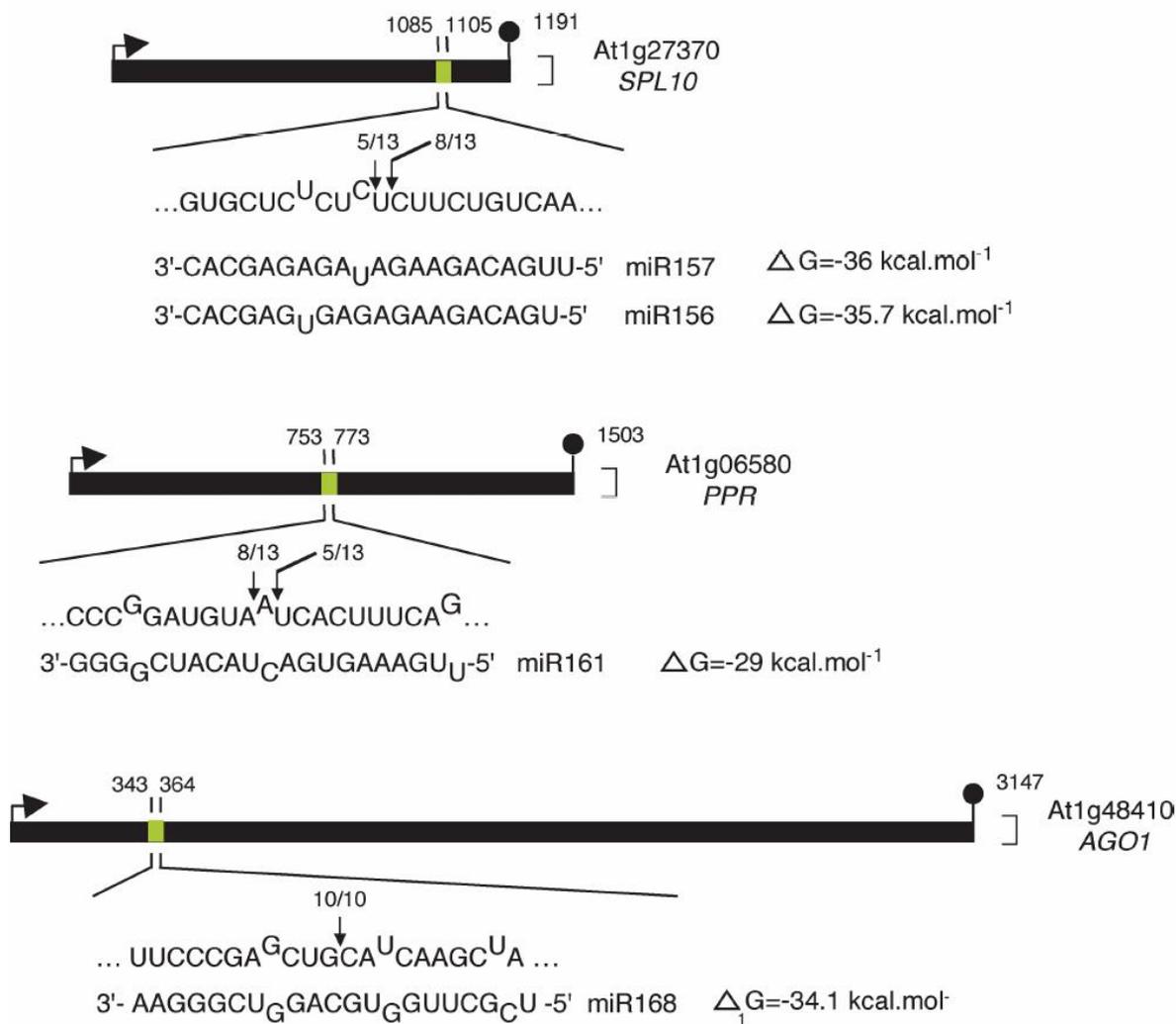


Figure S1. Diagrammatic Representation of At1g27370/*SPL10*, At1g06580/*PPR*, and At1g48410/*AGO1* mRNA Cleavage Sites Determined by Modified 5' RLM-RACE

Thick, black lines represent open reading frames. Arrows represent transcription starts. Filled circles represent polyadenylation signals. Gray boxes represent the putative miRNA/mRNA interaction site with the nucleotide position within the ORF indicated. mRNA partial sequence from 5' to 3' and miRNA sequence from 3' to 5' are shown in the expanded regions. The positions inferred as 5' ends of miR-guided cleavage products and the number of sequenced 5' RACE clones corresponding to each site are indicated by vertical arrowheads. Free energies of duplex structures as calculated with mfold "hybridization server" [S5] at 25°C are indicated between each duplex.

Figure 16: Figure S1 de "Supplemental data" de l'article annexe n°2 (Vazquez et al. 2004a).

être du au fait que le nombre de clones séquencés dans les deux études n'est pas suffisant pour faire une comparaison statistique.

La qualité de la matrice pour la 5' RLM-RACE modifiée étant satisfaisante, j'ai donc entrepris de caractériser les sites de clivage de certaines cibles de micro-ARNs n'ayant pas encore été validés. J'ai également confirmé la validation d'une cible des micro-ARNs miR156/157. En effet, l'ARNm de la cible des micro-ARNs miR156/157, At5g43270/*SPL2*, présente la même accumulation chez le mutant *dcl1-7* et la plante sauvage alors que l'accumulation du micro-ARNs miR156 est fortement diminuée chez le mutant *dcl1-7* (Kasschau *et al.* 2003). Ces résultats pouvaient laissés penser que la validation du site de clivage guidé par le miR156 était un artéfact. Après une première PCR avec un couple d'amorces, dont une est située sur le transcrit à caractériser et l'autre spécifique de l'adaptateur ARN (voir matériel et méthode de l'article annexe n°2: Vazquez *et al.* 2004a), aucune bande n'a pu être observée après électrophorèse des produits de PCR sur gel pour la cible At1g48410/*AGO1* du micro-ARN miR168 et la cible At1g06580/*PPR* du micro-ARN miR161, alors qu'une bande unique a été observée pour la cible At1g27370/*SPL10* des micro-ARNs miR156/157. Afin d'obtenir une quantité suffisante de fragment à cloner (visualisable sur gel), une seconde PCR a donc été réalisée sur le produit de la première PCR avec une amorce spécifique du transcrit située en 5' de la première amorce (nested PCR) et l'amorce spécifique de l'adaptateur ARN. A l'issue de l'électrophorèse en gel d'agarose de cette seconde PCR, une bande unique pour chacune des cibles At1g48410/*AGO1* et At1g06580/*PPR* a été observée (données non montrées). Les bandes obtenues ont été éluées des gels d'agarose, clonées et séquencées. Les résultats obtenus sont présentés dans la figure 16.

Dans un souci d'éviter de prendre en compte des amplifications non spécifiques, l'amorce spécifique de l'adaptateur ARN couvre seulement les 23 premiers nucléotides en partant du 5' de l'adaptateur qui compte 44 nucléotides au total. Cela permet donc, après séquençage, de vérifier la présence des 44 nucléotides appartenant à l'adaptateur ARN en 5' du fragment d'ARN ligaturé. La

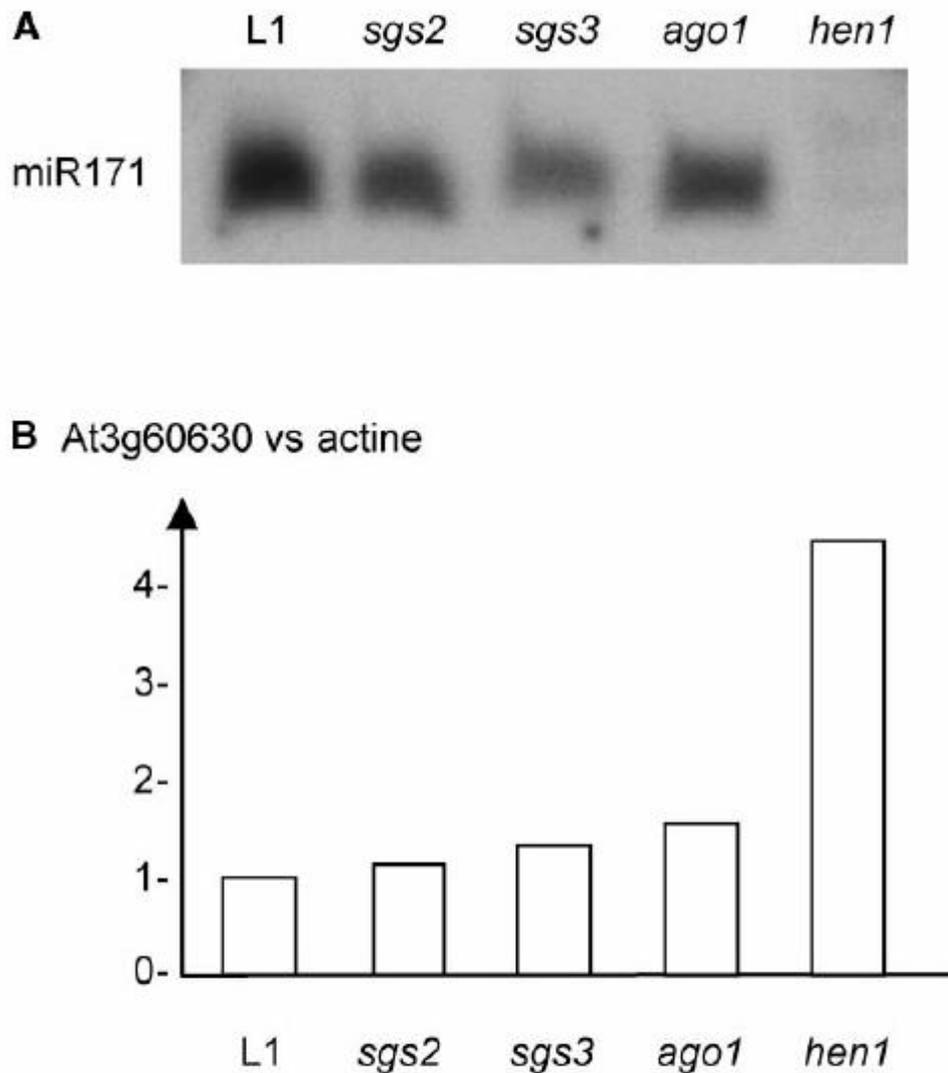


Figure 3. *miR171* miRNA and SCL mRNA Accumulation in Wild-Type and Mutant Plants

(A) Small RNA extracted from wild-type and mutant flowers was hybridized with a probe complementary to *miR171*.

(B) Total RNA extracted from wild-type and mutant flowers was quantified for SCL6-III mRNA relative accumulation by real-time PCR by using primers surrounding the cleavage site. Quantifications are normalized with actine2 transcript. The wild-type value is 1. AU, arbitrary unit.

Figure 17: Figure 3 de l'article annexe n°1 (Boutet et al. 2003).

présence de ces 44 nucléotides témoigne ainsi de la ligature de l'adaptateur ARN au site de clivage de l'ARNm étudié et élimine donc toute PCR non spécifique due à l'appariement de l'amorce adaptateur ARN de 21 nucléotides sur une autre région du transcrit à étudier.

Ainsi, j'ai pu caractériser le site de clivage des ARNm cibles des micro-ARNs, miR168 et miR161, respectivement At1g48410/*AGO1* et At1g06580/*PPR* et confirmer le clivage guidé par miR156/157 grâce à la caractérisation du site de clivage de l'ARNm At1g27370/*SPL10* (Figure 16).

1.2.3. Quantification des ARNm cibles de micro-ARNs par PCR quantitative en temps réel

Afin d'analyser l'effet des mutations *hen1-4* et *hyl1-2* sur la quantité des ARNm cibles de micro-ARNs, j'ai désigné un couple d'amorces de PCR de part et d'autre du site de clivage des ARNm, cibles de chacun des 8 micro-ARNs dont l'expression a été analysée chez les mutants. Ainsi, seuls les ARNm non clivés seront amplifiés. La quantité des ARNm cibles a également été mesurée chez le mutant *dcl1-9*, affecté dans la production des micro-ARNs, et chez la plante sauvage. Les résultats obtenus sont présentés dans les figures 17 et 18. Les résultats obtenus montrent que la mutation *hen1-4* conduit à une accumulation des ARNm cibles des micro-ARNs. Cela reflète la diminution du clivage des ARNm cibles en corrélation avec une absence d'accumulation des micro-ARNs chez le mutant *hen1-4* par rapport à la plante sauvage. L'augmentation de la quantité des ARNm cibles chez le mutant *hen1-4* est comparable à celle observée chez le mutant *dcl1-9*. Les données obtenues sur les cibles de micro-ARNs et celles obtenues sur la quantification des micro-ARNs ont permis de conclure que la protéine HEN1 est impliquée dans la production des micro-ARNs (article annexe n°1: Boutet et *al.* 2003).

L'effet de la mutation *hyl1-2* est plus discret. En effet, chez ce mutant certains micro-ARNs sont toujours détectables par hybridation moléculaire (figure 2 de l'article annexe n°2: Vazquez et *al.* 2004a).

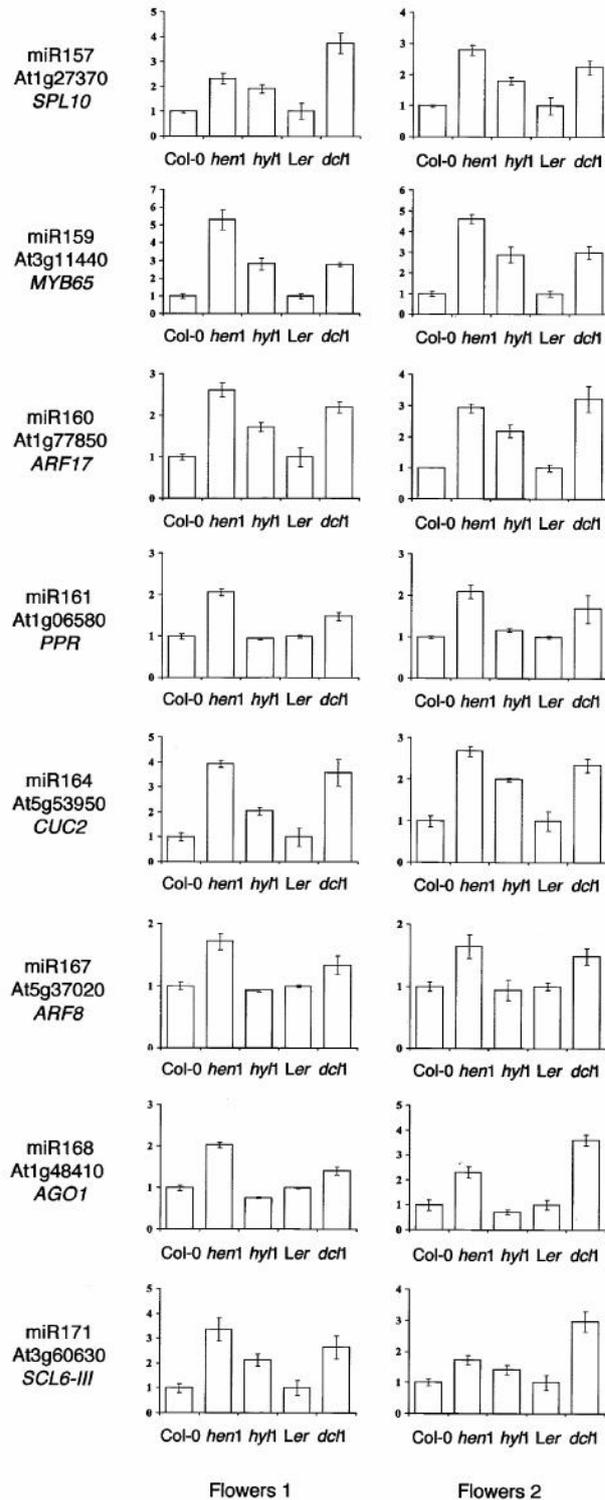


Figure 3. miRNA Target mRNA Accumulation in *hen1*, *hyf1*, and *dcl1* Mutants

HMW RNA extracted from flowers of Col, *hen1-4*, *hyf1-2*, *Ler*, and *dcl1-9* was quantified for *SPL10*, *MYB65*, *ARF17*, *PPR*, *CUC2*, *ARF8*, *AGO1*, and *SCL6-III* mRNA relative accumulation by real-time quantitative PCR by using primers surrounding the cleavage site on two cDNA synthesis (Flowers 1 and Flowers 2). Quantifications are normalized with *ACTIN2*. The Col and Ler values are arbitrarily fixed to one. miRNA numbers and gene names corresponding to each pair are given.

Figure 18: Figure 3 de l'article annexe n°2 (Vazquez et al. 2004a)

L'effet de la mutation *hyl1-2* sur les cibles des micro-ARNs est moins important que celui de la mutation *hen1-4*. En effet, la plupart des cibles de micro-ARNs que j'ai testées chez *hyl1-2* présentent une accumulation moins conséquente que chez le mutant *hen1-4*, ce qui est cohérent avec l'observation que certains micro-ARNs sont toujours détectables et que ce mutant *hyl1-2* présente des anomalies développementales moins sévères que les mutants *hen1-4* ou *dcl1-9*. Pour certaines cibles de micro-ARNs, telles que At1g06580/*PPR*, At5g37020/*ARF8* ou At1g48410/*AGO1*, aucune différence d'accumulation des ARNm cibles n'est observée entre le mutant *hyl1-2* et la plante sauvage. Ces données, mises en corrélation avec celles des hybridations des micro-ARNs, suggèrent que la faible quantité de micro-ARNs détectée chez *hyl1-2* est suffisante pour produire un clivage des ARNm cibles comparable à celui d'une plante sauvage et que ces micro-ARNs sont donc produits en excès chez la plante sauvage. L'accumulation observée chez le mutant *hyl1-2* de l'ARNm At1g77850/*ARF17*, impliqué dans la réponse aux auxines, fournit des éléments pouvant expliquer que ce mutant ait été identifié précédemment comme un mutant déficient dans la réponse aux hormones (Lu et Fedoroff 2000). J'ai donc participé avec l'équipe d'Hervé Vaucheret à la caractérisation de l'implication des protéines HEN1 et HYL1 dans la voie de production des micro-ARNs. Les résultats obtenus, notamment le fait que les mutants nuls *hen1-5* et *hyl1-2* ne sont pas létaux, suggèrent que des protéines possédant des fonctions redondantes avec celles des protéines HEN1 et HYL1 doivent exister ou que ces protéines interviennent seulement pour faciliter la production des micro-ARNs.

L'équipe d'Hervé Vaucheret avait précédemment montré que la protéine AGO1 était nécessaire à la mise en place de la S-PTGS (Fagard *et al.* 2000; Béclin *et al.* 2002). Vaucheret et col. (2004) ont montré que l'ARNm *AGO1* subit une régulation par le micro-ARN miR168 et que cette régulation est essentielle pour un développement correct des plantes (article annexe n°3: Vaucheret *et al.* 2004). Chez une plante sauvage, l'introduction d'une copie supplémentaire du gène *AGO1* sous le contrôle de son propre promoteur, et dont le site d'appariement avec le micro-ARN miR168 a été muté, confère à la plante des anomalies de développement. Les anomalies observées sont dues à une

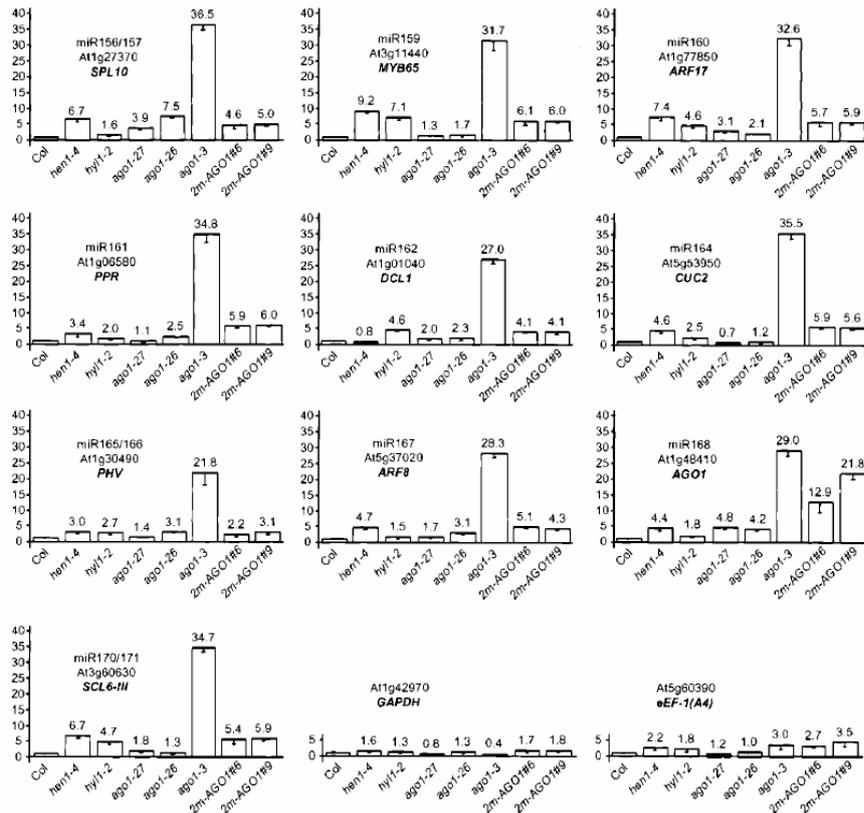


Figure 2. *ago1*, *hen1*, and *hyl1* mutants have increased steady-state levels of miRNA targets. RNA extracted from rosettes of isogenic wild-type or mutant siblings deriving from heterozygote parents and of untransformed plants or *2m-AGO1* transformants was quantified for the indicated mRNA by real-time quantitative PCR using primers surrounding the cleavage site. *GAPDH* and *eEF-1(A4)* were used as nontarget controls. Quantifications were normalized to that of *ACTIN2*, then to the value of the wild-type plants or wild-type siblings, which was arbitrarily fixed to 1.

Figure 19: Figure 2 de l'article annexe n°3 (Vaucheret et al. 2004)

production ectopique de la protéine AGO1. L'introduction d'un gène *MIR168a* modifié de telle façon que le micro-ARN miR168a produit soit complémentaire du site muté introduit dans le gène *AGO1*, permet de restaurer un phénotype sauvage. Ceci montre que la régulation de l'ARNm *AGO1* par le micro-ARN miR168 est essentielle pour un développement normal chez *Arabidopsis thaliana* (article annexe n°3: Vaucheret *et al.* 2004).

Contrairement à ce qui est observé chez les mutants *hen1* et *hyl1*, les mutants faibles *ago1-26* et *ago1-27* ne présentent aucune différence d'accumulation des micro-ARNs par rapport à la plante sauvage alors que la mutation nulle *ago1-3* conduit à une perte d'accumulation de 8 des 10 micro-ARNs testés (Figure 3 de l'article annexe n°3: Vaucheret *et al.* 2004). Certains éléments indiquent cependant que la protéine AGO1, impliquée dans la S-PTGS, pourrait intervenir dans le processus de régulation de l'expression des gènes par les micro-ARNs. En effet, les mutants *ago1* présentent des anomalies de développement sévères (Bohmert *et al.* 1998). Ainsi, le mutant nul *ago1-3* possède des feuilles de très petite taille, parfois succulentes et les quelques plantes arrivant à floraison ne possèdent qu'une inflorescence terminale (Bohmert *et al.* 1998). La protéine AGO1 est donc un bon candidat pour intervenir dans la voie de régulation de l'expression des gènes par les micro-ARNs, et ainsi remplir la même fonction que la protéine AGO2 humaine au sein du complexe RISC. De plus les mutants *ago-1* présentent des anomalies développementales rappelant celles observées chez les mutants *hen1* et *hyl1* (Figure 1 article annexe n°3: Vaucheret *et al.* 2004). D'autre part, l'existence d'un micro-ARN, miR168, capable de réaliser le clivage de l'ARNm AGO1 (Figure 16) suggère qu'AGO1 est impliquée dans la voie des micro-ARNs (Reinhart *et al.* 2002; Rhoades *et al.* 2002; article annexe n°2 : Vazquez *et al.* 2004a; article annexe n°3 : Vaucheret *et al.* 2004). En effet, cela suggère une autorégulation d'un composant de la voie des micro-ARNs, semblable à celle décrite pour la régulation de DCL1 par miR162 (Xie *et al.* 2003). De façon comparable aux expériences que j'ai réalisées chez les mutants *hen1* et *hyl1*, j'ai quantifié les ARNm d'une cible de chacun des 10 micro-ARNs analysés (Figure 19).

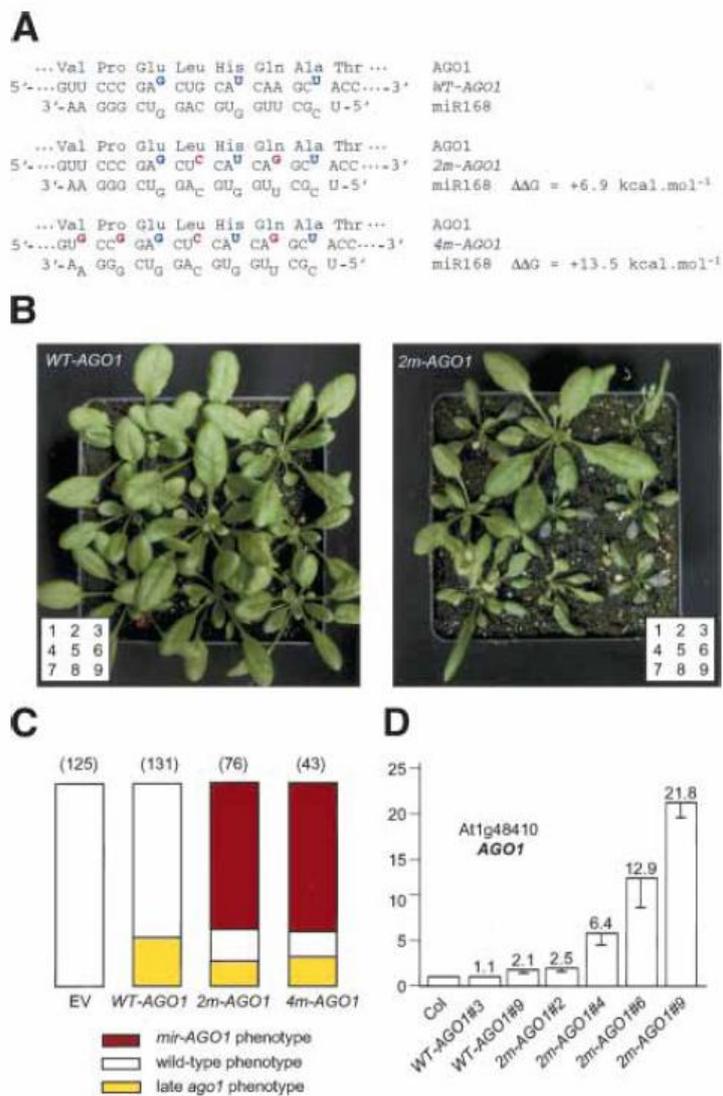


Figure 4. Silent mutations in the *miR168* complementary site of the *AGO1* mRNA induce developmental defects. (A) The *WT-AGO1* mRNA naturally contains three mismatches with *miR168* (in blue), including a G:U wobble pair. Silent mutations in *2m-AGO1* and *4m-AGO1* constructs introduce two and four additional mismatches (in red), reducing complementarity with *miR168*. $\Delta\Delta G$ was calculated using mfold. (B) Representative sets of transformants carrying the *WT-AGO1* or *2m-AGO1* construct. (C) Proportion of transformants showing a wild-type phenotype (open bar), an *ago1* phenotype caused by late cosuppression (yellow bar), or a *mir-AGO1* phenotype caused by *AGO1* overexpression (red bar). Plants were transformed with either an empty vector (EV) or the *WT-AGO1*, *2m-AGO1*, or *4m-AGO1* constructs. The number of transformants analyzed is indicated in parentheses. (D) *AGO1* mRNA accumulation determined by real-time quantitative PCR in untransformed plants (Col) or plants transformed with the *WT-AGO1* or *2m-AGO1* constructs. Quantifications were normalized to that of *ACTIN2*. The value in Col was arbitrarily fixed to 1. Numbers (#) correspond to the plants shown in B.

Figure 20: Figure 4 de l'article annexe n°3 (Vaucheret et al. 2004)

Au cours de ce travail, les quantifications des ARNm cibles des micro-ARNs ont été réalisées avec des ADNc synthétisés à partir d'ARNs extraits de feuilles puisque les mutants nuls *ago1-3* ne fournissent pas suffisamment de matériel floral. Afin de pouvoir comparer les résultats obtenus chez les mutants *hen1* et *hyl1* à ceux obtenus chez les mutants *ago1*, les quantifications des ARNm cibles ont également été réalisées chez les mutants *hen1-4* et *hyl1-2*. De façon intéressante, chez les mutants faibles *ago1-26* et *ago1-27*, ne présentant pas de diminution d'accumulation des micro-ARNs testés, la plupart des ARNm cibles s'accumulent (Figure 19). Ces résultats montrent que la protéine AGO1 est impliquée dans les étapes effectrices de clivage des ARNm cibles de micro-ARNs.

Cependant, une diminution de l'accumulation des micro-ARNs est observée chez les mutants nuls *ago1-3* (Figure 3 de l'article annexe n°3: Vaucheret *et al.* 2004). L'association des micro-ARNs avec la protéine AGO1 chez les mutants faibles *ago1-26* et *ago1-27* serait donc suffisante pour la stabilisation des micro-ARNs alors que l'absence de la protéine chez le mutant nul *ago1-3* ne permettrait pas leur stabilisation et conduirait à leur dégradation plus ou moins rapide.

L'analyse de l'accumulation des ARNm cibles chez les plantes possédant un transcrite *AGO1* résistant au clivage par miR168 (*2m-AGO1*) (Figure 20D) montre que chacun des ARNm cibles des micro-ARNs, s'accumule de façon plus importante et à un niveau comparable à l'accumulation détectée chez les mutants *hen1* et *hyl1* (Figure 19). Ces données peuvent expliquer la ressemblance phénotypique des plantes *2m-AGO1* avec les mutants *hen1* et *hyl1*. Ces résultats permettent de poser des hypothèses concernant la participation d'AGO1 aux étapes effectrices. En effet, l'excès de protéine AGO1 générée chez les plantes *2m-AGO1* interfère avec le clivage des ARNm cibles de micro-ARNs. Ceci pourrait avoir lieu soit par titration des micro-ARNs par un excès de protéine AGO1 libre ou par titration d'autres composants du complexe RISC conduisant à la formation d'un complexe RISC incomplet. Le fait que les mutants nuls *ago1-3* ne soient pas létaux, contrairement à ce qui est observé pour le mutant nul *dcl1* suggère qu'une ou plusieurs protéines de la famille PPD peuvent partiellement remplir la fonction d'AGO1 dans les étapes effectrices de clivage des ARNm cibles de micro-ARN.

Mon travail de thèse, initié au sein du laboratoire de Physiologie de la Différenciation Végétale avant la mise en évidence des micro-ARNs chez les plantes et de leur rôle au cours du développement, a consisté à analyser les mutants de S-PTGS, *sgs2* et *sgs3*, présentant des anomalies au niveau de la morphologie foliaire. Notre hypothèse de travail était que ce défaut phénotypique correspondait à une dérégulation de gènes régulés par un mécanisme d'inactivation épigénétique post-transcriptionnelle chez la plante sauvage, similaire au mécanisme de S-PTGS. Afin d'identifier ces gènes, j'ai mis au point la technique de cDNA-AFLP au laboratoire puis comparé les transcriptomes des deux mutants à celui d'une plante sauvage. Les résultats de ces travaux sont présentés dans la partie suivante.

DEUXIÈME PARTIE: ARTICLE: Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs.

1.1. Introduction

L'inactivation épigénétique post-transcriptionnelle présente une fonction clairement établie dans la lutte contre les virus. Chez les plantes, les mutants *sgs2/sde1/rdr6* et *sgs3/sde2* présentent une sensibilité supérieure au virus CMV qu'une plante sauvage (Dalmay *et al.* 2001; Mourrain *et al.* 2000). Les protéines RDR6 et SGS3 sont essentielles à la production de siRNAs au cours de la PTGS induite par des transgènes sens, par contre elles ne sont pas requises pour la PTGS induite par des transgènes en répétitions inversées produisant directement des ARNdb (Mourrain *et al.* 2000; Béclin *et al.* 2002; article annexe n°1 : Boutet *et al.* 2003). Les protéines RDR6 et SGS3 permettraient donc la formation d'un ARN double brin conduisant à la production des siRNAs lors de la PTGS sens (Dalmay *et al.* 2000b; Mourrain *et al.* 2000). En plus de ces rôles dans la résistance aux virus et la PTGS sens, de nombreux éléments semblent indiquer un rôle de ce processus dans la régulation de gènes endogènes. En effet, les mutants *sgs2/sde1/rdr6* et *sgs3/sde2* présentent un

même phénotype, différent de celui d'une plante sauvage. A savoir une incurvation des feuilles vers la face inférieure en conditions optimales de croissance (données non montrées et Vazquez *et al.* 2004b, Figure 6). Les protéines RDR6 et SGS3, impliquées dans la production de siRNAs de transgènes sens, pourraient donc être impliquées dans une même voie de production de siRNAs endogènes qui régulerait l'expression de gènes endogènes, notamment ceux contrôlant la formation des feuilles, par clivage des ARNm correspondants. Au début de notre étude, l'existence de siRNAs endogènes régulant l'expression des gènes avait été décrite chez la drosophile. En effet, Aravin *et col.* (2001) avaient mis en évidence la régulation post-transcriptionnelle des ARNm du gène *Stellate* par des siRNAs endogènes de 25-27 nucléotides, essentielle pour assurer la fertilité des drosophiles mâles. Ces siRNAs endogènes sont produits à partir d'un ARNdb formé par transcription convergente du locus nommé *Suppressor of Stellate, Su(Ste)*.

Les ARNm des transgènes inactivés chez la plante sauvage sont accumulés chez les mutants *sgs2/rdr6* et *sgs3*, notre hypothèse de travail était donc que les ARNm cibles de siRNAs endogènes devaient s'accumuler chez les mutants *sgs2/rdr6* et *sgs3* par rapport à la plante sauvage. Afin de caractériser les ARNm cibles régulées par des siRNAs endogènes chez *A. thaliana*, nous avons comparé les transcriptomes des mutants *sgs2/rdr6* et *sgs3* à celui d'une plante sauvage. Pour cela, nous avons utilisé la technique de cDNA AFLP (cDNA Amplified Fragment Length Polymorphism) qui est une technique basée sur l'amplification sélective et quantitative de fragments de restriction de fragments d'ADNc double brin. En bref, les ADNc double brin sont soumis à une double digestion enzymatique, dont une enzyme possède un site de restriction à 6 bases (AseI) et l'autre un site à quatre bases (TaqI). Afin de pouvoir analyser les produits d'amplification une sous population de ces fragments est amplifiée avec des amorces possédant deux nucléotides supplémentaires, dits sélectifs, en 3'. Il existe ainsi 16 amorces sélectives spécifiques de chacun des adaptateurs et donc 256 combinaisons d'amorces permettent l'amplification de l'ensemble des fragments TaqI-AseI. Cette amplification sélective est réalisée après radio-marquage de l'une des deux amorces au phosphore 33.

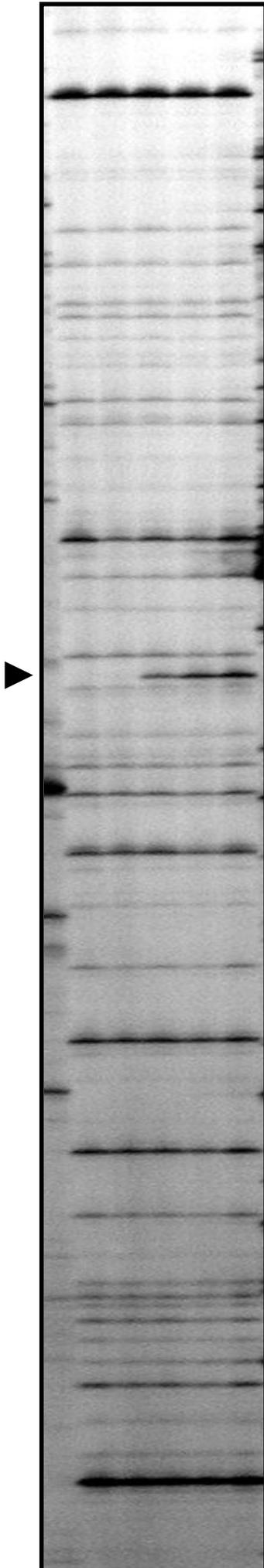
1 2 3 4 5

Figure 21 : Profil de cDNA- AFLP réalisé avec les couples d'amorces AseICA-TaqAG pour les plantes sauvages (1et 2), *rdr6* (3 et 4) et *sgs3* (5). Nous observons une forte conservation générale de l'intensité des bandes pour tous les échantillons La flèche indique la position d'une bande d'intensité différentielle. Le profil présenté représente la longueur totale de migration. Les bandes de faible intensité ne sont pas visibles pour le temps d'exposition présenté (4 jours).

Couple d'amorce	Taille de la séquence obtenue	Gène correspondant	Fonction prédite
AseI AT- TaqI CC	242bp	At4g38620	Facteur de transcription MYB4
AseI AT- TaqI GG	379bp	At1g62930	Protéine de la famille PPR
AseI TA- TaqI GT	123bp	At2g27400	ND

Tableau 7 : Gènes correspondants aux séquences obtenues pour les couples d'amorces de cDNA AFLP présentant une bande plus intense chez les mutants *rdr6* et *sgs3* comparé à la plante sauvage.

Afin d'estimer la partie du transcriptome d'*A. thaliana* que nous étions capable de visualiser dans nos conditions de restriction, 250 transcrits ont été analysés aléatoirement - sans discrimination de taille, ni de fonction - dans la banque de donnée Genbank. Nous avons déterminé que 32% de ces transcrits possédaient les deux sites de restriction à des distances permettant de produire des fragments observables sur gel de séquence à 6% (50-500 pb). Nous avons réalisé deux répétitions pour 80 couples d'amorces, permettant de visualiser environ 4.000 transcrits.

Les plantes analysées au stade 4 feuilles étaient issues de culture *in vitro* afin de limiter au maximum les variations dues aux conditions de culture. Cela explique que la grande majorité des bandes observées ne présentent pas de variations d'intensité entre les mutants *sgs2/rdr6*, *sgs3* et la plante sauvage (Figure 21). Pour les 80 couples d'amorces testés, seulement trois bandes présentaient une accumulation reproductible chez les deux mutants *sgs2/rdr6* et *sgs3* par rapport à la plante sauvage. Ces trois bandes ont été éluées du gel, clonées et séquencées. Les dix clones séquencés pour chacune des bandes correspondaient à une seule séquence (Table 7). L'une de ces séquences correspondait à un transcrit issu d'une région intergénique comprise entre les gènes At2g27395 et At2g27410 que nous avons notée At2g27400. Par analyses informatiques, nous avons trouvé que plusieurs petits ARN non codants de 21 nucléotides étaient potentiellement produits par le gène At2g27400. Nous avons donc caractérisé en partie la voie de production de ces petits ARN endogènes et recherchés s'ils avaient un rôle dans la régulation de l'expression de gènes endogènes.

1.2. ARTICLE

Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs.

Publié dans Molecular Cell

1.3. Conclusion

Comme nous l'avions montré précédemment (article annexe n°1 : Boutet *et al.* 2003), nous avons confirmé que les protéines RDR6 et SGS3 ne sont pas requises pour la production des micro-ARNs. Ces données permettent donc de séparer d'un point de vue génétique les différentes voies de production des petits ARN chez les plantes.

Le crible de cDNA AFLP entre la plante sauvage et les mutants *sgs2/rdr6* et *sgs3* a permis de mettre en évidence un ARN produisant une nouvelle classe de siRNAs endogènes chez les plantes. Cet ARN est produit par le gène At2g27400 situé dans une région initialement définie comme une région intergénique. Cet ARN est à l'origine de la production de plusieurs siRNAs contigus (cluster de siRNAs), identifiés par clonage, comme le montre l'absence de ces siRNAs chez une plante homozygote pour une insertion T-DNA dans le gène At2g27400. L'accumulation de ces siRNAs requiert l'intervention des protéines RDR6, SGS3, HEN1, HYL1 et DCL1. L'ARN At2g27400, dans sa forme non épissée, serait converti en un long ARNdb par la RdRP RDR6, et clivés en siRNAs par une RNase de type III, DCL1. Ceci définit donc une nouvelle voie de production de petits ARN endogènes qui partage certains éléments, dont DCL1, avec la voie de production des micro-ARNs. Les nouveaux siRNA endogènes identifiés dans notre étude sont différents des siRNAs hétérochromatiques. En effet, nous avons montré que les protéines RDR6 et SGS3 ne sont pas requises pour l'accumulation des siRNAs hétérochromatiques impliqués dans le maintien inactif des séquences répétées (Xie *et al.* 2004). Tout comme la majorité des micro-ARNs chez les plantes, les siRNAs produits par l'ARN At2g27400 sont capables de guider le clivage d'ARNm endogènes au milieu de la zone d'appariement entre le siRNA et l'ARNm cible. En effet, l'absence de ces siRNAs chez les mutants *dcl1*, *hen1*, *hyl1*, *sgs3*, et *sgs2/rdr6* est corrélée à l'augmentation de l'accumulation des ARNm cibles possédant une zone d'appariement avec ces siRNAs. D'autre part, nous avons validé les sites de clivage par des expériences de 5'RACE. Chez le mutant *ago1*, la quantité des siRNAs

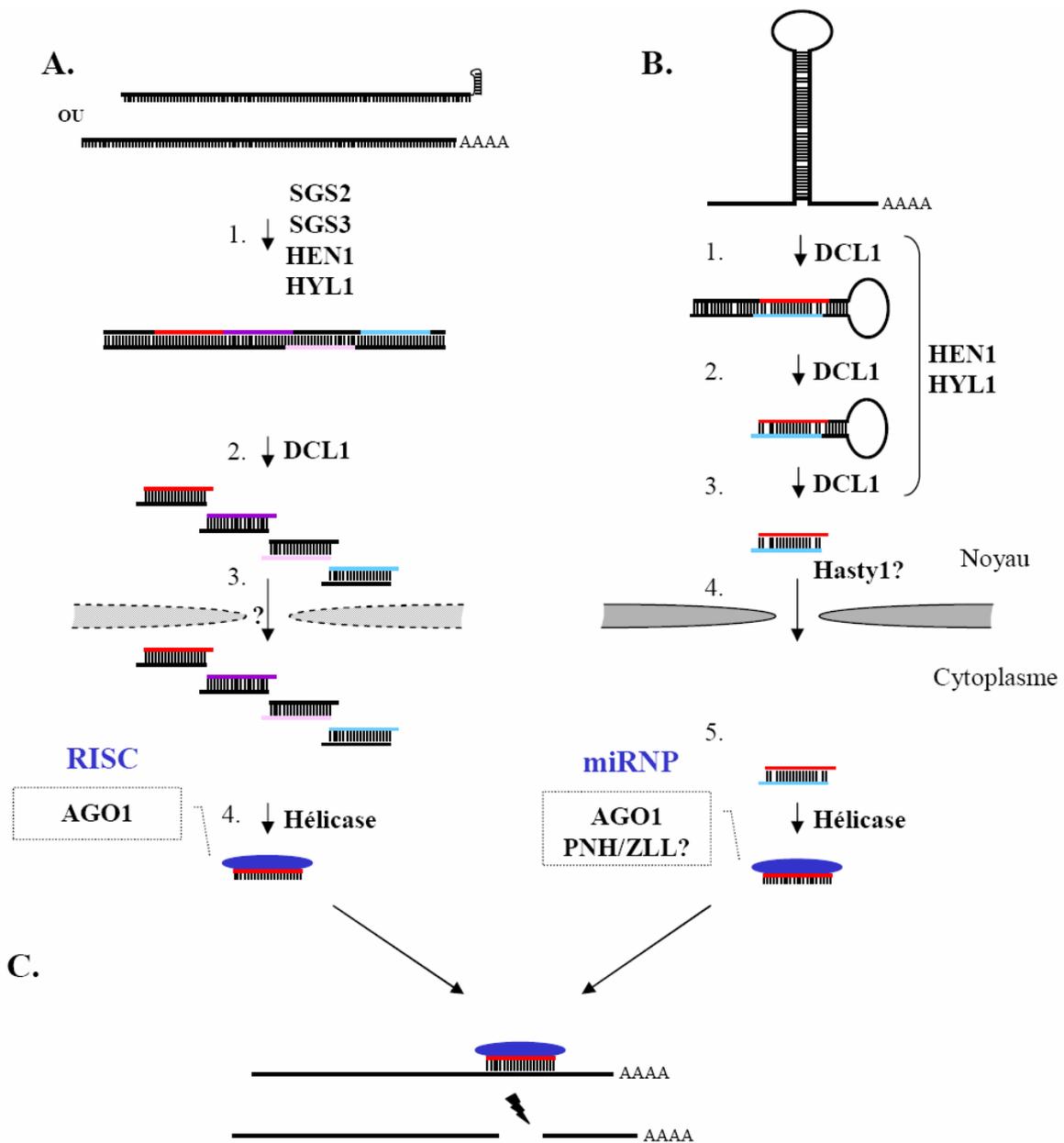


Figure 22 : Mécanismes de biosynthèse et d'action des petits ARNc endogènes régulateurs de l'expression des gènes au niveau post-transcriptionnel chez *A. thaliana*.

A. Mécanisme de biosynthèse des siRNAs dépendants de RDR6/SGS2.

Les loci produisant les siRNAs dépendants de RDR6 produiraient un ARN présentant des caractéristiques particulières telles qu'une structure permettant l'initiation de la synthèse d'un brin complémentaire par RDR6/SGS2. L'étape de production d'un ARNdb requiert l'intervention des protéines SGS3, HEN1 et HYL1 (1). Cet ARNdb serait clivés en de multiples siRNAs par DCL1 (2), qui, dans le cas d'une production au niveau du noyau, seraient exportés vers le cytoplasme (3). Le brin du duplex possédant l'extrémité 5' la moins stable va alors être intégré dans le complexe RISC, contenant probablement la protéine AGO1 (4).

B. Mécanisme de biosynthèse des micro-ARNs (voir figure 14).

C. Les siRNAs et les micro-ARNs, intégrés dans le complexe RISC guident le clivage de l'ARNm complémentaire.

issus de At2g27400 ne varie pas, cependant on observe une augmentation de l'accumulation des ARNm cibles. Il apparaît donc que la protéine AGO1, présente dans le complexe miRNP, est également présente au sein du complexe siRNP, effecteur du clivage des ARNm ciblés par les siRNAs produits par At2g27400.

Nos résultats mettent donc en évidence une nouvelle voie de régulation post-transcriptionnelle de l'expression des gènes possédant des caractéristiques communes avec la majorité des micro-ARNs chez les plantes. Les incurvations des feuilles observées chez les mutants *sgs2/rdr6* et *sgs3* ne sont cependant pas causées par la perte de cette voie de régulation génique puisque chez le mutant *ago7*, présentant le même phénotype, la production de ces siRNAs et le clivage de leurs cibles ne sont pas affectés.

Nous avons donc montré que les voies de production des siRNAs endogènes se diversifient d'un point de vue génétique mais également d'un point de vue fonctionnel. Les anomalies développementales des mutants *sgs2/rdr6*, *sgs3* et *ago7* semblent indiquer l'existence d'une voie supplémentaire de production et/ou d'action d'autres siRNAs endogènes qu'il reste à identifier.

1.4. Discussion générale et Perspectives

Il existe actuellement trois voies distinctes identifiées de régulation génique par des petits ARN endogènes chez *A. thaliana*: la voie de régulation génique par les micro-ARNs (Bartel 2004), la voie de régulation génique par les siRNAs dépendants de RDR6 (Vazquez *et al.* 2004b) et la voie de maintien des éléments transposables et des éléments répétés par les siRNAs dépendants de RDR2 (Chan *et al.* 2004) (Figure 22 et revue bibliographique : Figure 5). Ces voies peuvent être séparées les unes des autres d'un point de vue génétique mais possèdent des éléments communs.

Le mécanisme précis de production des siRNAs dépendants de RDR6 reste à élucider. En effet, nous avons rapporté l'identification de deux lignées transgéniques possédant une insertion dans le gène At2g27400 à l'origine de ces siRNAs. L'insertion située dans l'intron affecte la production des

siRNAs alors que celle présente dans la région 3' du transcrit n'affecte pas leur production. Ceci pourrait indiquer la présence entre ces deux insertions d'une séquence nécessaire à l'initiation de la synthèse du brin complémentaire par RDR6. L'un des repliements intramoléculaire présent entre ces deux insertions pourrait servir de site d'initiation par la RdRP RDR6 (Vazquez *et al.* 2004b Figure S1). En effet, l'initiation de la synthèse d'un brin complémentaire par la RdRP grâce à la présence d'une boucle en 3' des transcrits pourrait être une propriété commune à toutes les RdRP (Song et Simon 1995). L'obtention de plantes transgéniques produisant des transcrits At2g27400 plus ou moins tronqués dans la partie 3' pourraient permettre d'identifier la séquence minimale nécessaire à la production de ces siRNAs et de déterminer le site d'initiation de la synthèse du brin complémentaire par la RdRP RDR6.

Comme nous l'avons mentionné dans la revue bibliographique, chez les plantes après clivage des ARNm cibles de micro-ARNs, l'extrémité 3' de l'ARNm est rapidement dégradée par la protéine AtXRN4 (Souret *et al.* 2004). Si l'implication de cette protéine dans la dégradation des extrémités 3' des ARNm cibles des siRNAs dépendants de RDR6 se vérifie, le mutant *xrn4* sera un bon outil pour la confirmation des sites de clivage des cibles de siRNAs peu abondantes.

Nous avons rapporté l'accumulation d'un ARN pré-messager du gène At2g46740 chez les différents mutants de la voie de biosynthèse et d'action des siRNAs dépendants de RDR6, indiquant que le clivage des cibles pourrait avoir lieu dans le noyau ou permettre de dégrader des transcrits alternatifs non épissés exportés vers le cytoplasme. Ce type d'observation a également été rapporté dans le cas de la régulation de l'ARNm DCL1 par miR162 (Xie *et al.* 2003). Il serait donc intéressant de tester la présence des éléments du complexe RISC dans le noyau, notamment AGO1. L'implication de la production des siRNAs dépendants de RDR6 dans un processus cellulaire reste à démontrer.

Des études bioinformatiques nous ont permis de détecter la présence d'une TATA box et d'une CCAAT box situées, respectivement, à -34 et à -255 nucléotides du début du transcrit At2g27400 (résultats non présentés). La présence de ces séquences consensus et la polyadénylation du transcrit

At2g27400 indiquent que ce transcrit est produit par l'ARN polymérase ADN dépendante II et que la transcription du gène est probablement régulée de façon spatio-temporelle. En effet, il a été montré récemment que l'expression d'un micro-ARN pouvait être régulée par une hormone végétale: l'acide gibbérellique. L'introduction de gènes rapporteurs *uidA* ou GFP sous le contrôle du promoteur du gène At2g27400 pourrait également permettre de poser des hypothèses concernant le rôle de ces siRNAs dans un processus cellulaire. Ces transformants seraient un outil intéressant pour tester l'activation ou la répression de la transcription de ce gène dans différentes conditions de culture (photopériode, stress biotiques et abiotiques).

L'identification d'autres régions produisant des siRNAs dépendants de RDR6 pourrait permettre d'assigner un rôle à ces siRNAs. Certaines banques de siRNAs clonés chez les plantes sont disponibles en ligne (<http://cgrb.orst.edu/smallRNA/db/>). L'identification de clusters de siRNAs produits tous les 21 nucléotides à partir d'une même région sont de bons candidats à tester lors d'hybridations moléculaires de siRNAs sur de l'ARN de plantes sauvages et de mutants, tels que *rdr6*.

Lors de notre analyse de cDNA AFLP, nous avons également identifiés deux autres candidats potentiellement cibles de siRNAs dépendants de RDR6. Chez les mutants de la voie de régulation génique par ces siRNAs, la vérification de l'accumulation des deux transcrits, correspondants à un facteur MYB putatif (At4g38620) et à une PPR (At1g62930), les hybridations de petits ARN avec des sondes ARN correspondants à ces gènes et la détermination de sites de clivage dans ces transcrits pourraient également permettre d'identifier de nouveaux siRNAs dépendants de RDR6 et de leur attribuer un rôle biologique.

La fonction biologique des cibles des siRNAs dépendants de RDR6 est un point très important à élucider. En effet, l'organisation en cluster de ces siRNAs indique que les siRNAs d'un même cluster sont produits en même temps et sont donc capables de co-réguler l'expression de plusieurs types de gènes. Il sera donc également très important de déterminer et de valider de nouvelles cibles des différents siRNAs produits par le gène At2g27400 afin de pouvoir identifier des cibles dont les protéines possèdent des fonctions prédites. Nous avons montré que les siRNAs produits par le gène

At2g27400 ne sont pas conservés chez d'autres espèces (Vazquez *et al.* 2004b). Ceci pourrait indiquer que chaque espèce végétale possède des régulations spécifiques par des siRNAs. Il sera donc intéressant de tester l'existence des siRNAs du gène At2g27400 dans d'autres écotypes d'*Arabidopsis thaliana* et de vérifier si on observe des différences de régulations des diverses cibles de ces siRNAs. Si des différences sont observées on pourrait faire l'hypothèse que les régulations géniques par des siRNAs RDR6-dépendants participent aux variabilités phénotypiques et/ou adaptatives observées entre les écotypes. En effet, chacun des écotypes d'*Arabidopsis* présente une forme particulière de la feuille, un nombre de feuille différent ou un temps de floraison particulier (<http://seeds.nottingham.ac.uk/Nasc/action.lasso>).

Plusieurs approches pourraient également permettre de poser des hypothèses concernant la fonction biologique des cibles régulées par les siRNAs provenant du gène At2g27400. L'analyse de différentes lignées de mutants d'insertion de T-DNA, placées en condition de culture optimum ou en condition de stress pourrait apporter des éléments intéressants, telles que les lignées SALK_008144 et SALK_039307 dont l'insertion est prédite, respectivement, dans les gènes cibles At5g18040 et At4g29760 (cibles de ces siRNAs). De même, il pourrait être envisagé l'étude de lignées transgéniques exprimant les gènes cibles résistants à la régulation par modification du site d'appariement avec ces siRNAs. Enfin, il serait également intéressant d'étudier la lignée EYK112 affectée dans la production des siRNAs provenant du gène At2g27400 et des plantes exprimant ce gène sous le contrôle du promoteur fort et constitutif de l'ARN 35S du CaMV, surexpresseurs potentiels des siRNA produits par At2g27400.

-

III. ANNEXES

-

ARTICLE annexe N°1:

***Arabidopsis HEN1*: a genetic link between endogenous micro-RNAs controlling development and siRNA controlling transgene silencing and virus resistance**

Publié dans Current Biology

ARTICLE annexe N°2

The nuclear dsRNA binding protein HYL1 is required for micro-RNAs accumulation and plant development, but not posttranscriptional transgene silencing

Publié dans Current Biology

ARTICLE annexe N°3

The action of *ARGONAUTE1* in the miR pathway and its regulation by the miR pathway are crucial for plant development

Publié dans Genes and development

-

IV. RÉFÉRENCES BIBLIOGRAPHIQUES

-



Aboobaker, A. A. and M. L. Blaxter (2003). "Use of RNA interference to investigate gene function in the human filarial nematode parasite *Brugia malayi*." Mol Biochem Parasitol **129**(1): 41-51.

Abrahante, J. E., A. L. Daul *et al.* (2003). "The *Caenorhabditis elegans* hunchback-like gene *lin-57/hbl-1* controls developmental time and is regulated by microRNAs." Dev Cell **4**(5): 625-37.

Achard, P., A. Herr *et al.* (2004). "Modulation of floral development by a gibberellin-regulated microRNA." Development **131**(14): 3357-3365.

Adams, M. D., S. E. Celniker *et al.* (2000). "The genome sequence of *Drosophila melanogaster*." Science **287**(5461): 2185-95.

Aida, M., T. Ishida and M. Tasaka (1999). "Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes." Development **126**(8): 1563-70.

Al-Anouti, F. and S. Ananvoranich (2002). "Comparative analysis of antisense RNA, double-stranded RNA, and delta ribozyme-mediated gene regulation in *Toxoplasma gondii*." Antisense Nucleic Acid Drug Dev **12**(4): 275-81.

Altuvia, S., A. Zhang *et al.* (1998). "The *Escherichia coli* OxyS regulatory RNA represses *fh1A* translation by blocking ribosome binding." Embo J **17**(20): 6069-75.

Ambros, V., B. Bartel *et al.* (2003a). "A uniform system for microRNA annotation." Rna **9**(3): 277-9.

Ambros, V., R. C. Lee *et al.* (2003b). "MicroRNAs and Other Tiny Endogenous RNAs in *C. elegans*." Curr Biol **13**(10): 807-18.

Anandalakshmi, R., G. J. Pruss *et al.* (1998). "A viral suppressor of gene silencing in plants." Proc Natl Acad Sci U S A **95**(22): 13079-84.

Andersen, A. A. and B. Panning (2003). "Epigenetic gene regulation by noncoding RNAs." Curr Opin Cell Biol **15**(3): 281-9.

Andersen, J. and N. Delihias (1990). "micF RNA binds to the 5' end of ompF mRNA and to a protein from *Escherichia coli*." Biochemistry **29**(39): 9249-56.

Angell, S. M. and D. C. Baulcombe (1997). "Consistent gene silencing in transgenic plants expressing a replicating potato virus X RNA." Embo J **16**(12): 3675-84.

Aoki, Y., D. P. Cioca *et al.* (2003). "RNA interference may be more potent than antisense RNA in human cancer cell lines." Clin Exp Pharmacol Physiol **30**(1-2): 96-102.

Aravin, A. A., M. Lagos-Quintana *et al.* (2003). "The small RNA profile during *Drosophila melanogaster* development." Dev Cell **5**(2): 337-50.

Aravin, A. A., N. M. Naumova *et al.* (2001). "Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline." Curr Biol **11**(13): 1017-27.

Argaman, L., R. Hershberg *et al.* (2001). "Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*." Curr Biol **11**(12): 941-50.

Aufsatz, W., M. F. Mette *et al.* (2002). "RNA-directed DNA methylation in *Arabidopsis*." Proc Natl Acad Sci U S A **99**(Suppl 4): 16499-506.

Aukerman, M. J. and H. Sakai (2003). "Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes." Plant Cell **15**(11): 2730-41. Epub 2003 Oct 10.

B

Bartel, D. P. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." Cell **116**(2): 281-97.

Bashirullah, A., A. E. Pasquinelli *et al.* (2003). "Coordinate regulation of small temporal RNAs at the onset of *Drosophila* metamorphosis." Dev Biol **259**(1): 1-8.

Basyuk, E., F. Suavet *et al.* (2003). "Human let-7 stem-loop precursors harbor features of RNase III cleavage products." Nucleic Acids Res **31**(22): 6593-7.

- Baulcombe, D. C. (1999). "Fast forward genetics based on virus-induced gene silencing." Curr Opin Plant Biol **2**(2): 109-13.
- Béclin, C., R. Berthome *et al.* (1998). "Infection of tobacco or *Arabidopsis* plants by CMV counteracts systemic post-transcriptional silencing of nonviral (trans)genes." Virology **252**(2): 313-7.
- Béclin, C., S. Boutet *et al.* (2002). "A branched pathway for transgene-induced RNA silencing in plants." Curr Biol **12**(8): 684-8.
- Bernstein, E., A. A. Caudy *et al.* (2001). "Role for a bidentate ribonuclease in the initiation step of RNA interference." Nature **409**(6818): 363-6.
- Bohmert, K., I. Camus *et al.* (1998). "AGO1 defines a novel locus of *Arabidopsis* controlling leaf development." Embo J **17**(1): 170-80.
- Bohnsack, M. T., K. Czaplinski and D. Gorlich (2004). "Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs." Rna **10**(2): 185-91.
- Bohnsack, M. T., K. Regener *et al.* (2002). "Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm." Embo J **21**(22): 6205-15.
- Bollman, K. M., M. J. Aukerman *et al.* (2003). "HASTY, the *Arabidopsis* ortholog of exportin 5/MSN5, regulates phase change and morphogenesis." Development **130**(8): 1493-504.
- Bonnet, E., J. Wuyts *et al.* (2004). "Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes." Proc Natl Acad Sci U S A **101**(31): 11511-6. Epub 2004 Jul 22.
- Bosher, J. M. and M. Labouesse (2000). "RNA interference: genetic wand and genetic watchdog." Nat Cell Biol **2**(2): E31-6.
- Boutet, S., F. Vazquez *et al.* (2003). "*Arabidopsis* HEN1. A Genetic Link between Endogenous miRNA Controlling Development and siRNA Controlling Transgene Silencing and Virus Resistance." Curr Biol **13**(10): 843-8.

Boyle, J. P., X. J. Wu *et al.* (2003). "Using RNA interference to manipulate endogenous gene expression in *Schistosoma mansoni* sporocysts." Mol Biochem Parasitol **128**(2): 205-15.

Brennecke, J., D. R. Hipfner *et al.* (2003). "*bantam* encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*." Cell **113**(1): 25-36.

C

Calin, G. A., C. D. Dumitru *et al.* (2002). "Frequent deletions and down-regulation of micro- RNA genes miR-15 and miR-16 at 13q14 in chronic lymphocytic leukemia." Proc Natl Acad Sci U S A **99**(24): 15524-9.

Campalans, A., A. Kondorosi and M. Crespi (2004). "Enod40, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*." Plant Cell **16**(4): 1047-59. Epub 2004 Mar 22.

Cao, X., W. Aufsatz *et al.* (2003). "Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation." Curr Biol **13**(24): 2212-7.

Cao, X. and S. E. Jacobsen (2002). "Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing." Curr Biol **12**(13): 1138-44.

Carmell, M. A., Z. Xuan *et al.* (2002). "The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis." Genes Dev **16**(21): 2733-42.

Catalanotto, C., G. Azzalin *et al.* (2002). "Involvement of small RNAs and role of the *qde* genes in the gene silencing pathway in *Neurospora*." Genes Dev **16**(7): 790-5.

Catalanotto, C., M. Pallotta *et al.* (2004). "Redundancy of the two *DICER* genes in transgene-induced posttranscriptional gene silencing in *Neurospora crassa*." Mol Cell Biol **24**(6): 2536-45.

Caudy, A. A., R. F. Ketting *et al.* (2003). "A micrococcal nuclease homologue in RNAi effector complexes." Nature **425**(6956): 411-4.

- Caudy, A. A., M. Myers *et al.* (2002). "Fragile X-related protein and VIG associate with the RNA interference machinery." Genes Dev **16**(19): 2491-6.
- Cerutti, L., N. Mian and A. Bateman (2000). "Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain." Trends Biochem Sci **25**(10): 481-2.
- Chalker, D. L. and M. C. Yao (2001). "Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in *Tetrahymena thermophila*." Genes Dev **15**(10): 1287-98.
- Chan, S. W., D. Zilberman *et al.* (2004). "RNA silencing genes control de novo DNA methylation." Science **303**(5662): 1336.
- Chang, S., R. J. Johnston, Jr. *et al.* (2004). "MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode." Nature **430**(7001): 785-9.
- Chen, C. Z., L. Li *et al.* (2004a). "MicroRNAs modulate hematopoietic lineage differentiation." Science **303**(5654): 83-6. Epub 2003 Dec 4.
- Chen, J., W. X. Li *et al.* (2004b). "Viral virulence protein suppresses RNA silencing-mediated defense but upregulates the role of microRNA in host gene expression." Plant Cell **16**(5): 1302-13. Epub 2004 Apr 20.
- Chen, X. (2004). "A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development." Science **303**(5666): 2022-5. Epub 2003 Jul 31.
- Chen, X., Liu, J. *et al.* (2002). "HEN1 functions pleiotropically in *Arabidopsis* development and acts in C function in the flower." Development **129**(5):1085-94.
- Chi, J. T., H. Y. Chang *et al.* (2003). "Genomewide view of gene silencing by small interfering RNAs." Proc Natl Acad Sci U S A **100**(11): 6343-6. Epub 2003 May 2.
- Chuang, C. F. and E. M. Meyerowitz (2000). "Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*." Proc Natl Acad Sci U S A **97**(9): 4985-90.

Cogoni, C., J. T. Irelan *et al.* (1996). "Transgene silencing of the *al-1* gene in vegetative cells of *Neurospora* is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation." Embo J **15**(12): 3153-63.

Cogoni, C. and G. Macino (1999a). "Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase." Nature **399**(6732): 166-9.

Cogoni, C. and G. Macino (1999b). "Posttranscriptional gene silencing in *Neurospora* by a RecQ DNA helicase." Science **286**(5448): 2342-4.

Cogoni, C. and G. Macino (2000). "Post-transcriptional gene silencing across kingdoms." Curr Opin Genet Dev **10**(6): 638-43.

Colaiacovo, M. P., G. M. Stanfield *et al.* (2002). "A targeted RNAi screen for genes involved in chromosome morphogenesis and nuclear organization in the *Caenorhabditis elegans* germline." Genetics **162**(1): 113-28.

Crété, P., S. Leuenberger *et al.* (2001). "Graft transmission of induced and spontaneous post-transcriptional silencing of chitinase genes." Plant J **28**(5): 493-501.

D

Dalmay, T., A. Hamilton *et al.* (2000a). "Potato virus X amplicons in *Arabidopsis* mediate genetic and epigenetic gene silencing." Plant Cell **12**(3): 369-79.

Dalmay, T., A. Hamilton *et al.* (2000b). "An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus." Cell **101**(5): 543-53.

Dalmay, T., R. Horsefield *et al.* (2001). "SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in *Arabidopsis*." Embo J **20**(8): 2069-78.

Di Laurenzio, L., J. Wysocka-Diller *et al.* (1996). "The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root." Cell **86**(3): 423-33.

Doench, J. G., C. P. Petersen and P. A. Sharp (2003). "siRNAs can function as miRNAs." Genes Dev **17**(4): 438-42.

Doench, J. G. and P. A. Sharp (2004). "Specificity of microRNA target selection in translational repression." Genes Dev **18**(5): 504-11. Epub 2004 Mar 10.

Doi, N., S. Zenno *et al.* (2003). "Short-interfering-RNA-mediated gene silencing in mammalian cells requires DICER and eIF2C translation initiation factors." Curr Biol **13**(1): 41-6.

Domeier, M. E., D. P. Morse *et al.* (2000). "A link between RNA interference and nonsense-mediated decay in *Caenorhabditis elegans*." Science **289**(5486): 1928-31.

Durand-Dubief, M., L. Kohl and P. Bastin (2003). "Efficiency and specificity of RNA interference generated by intra- and intermolecular double stranded RNA in *Trypanosoma brucei*." Mol Biochem Parasitol **129**(1): 11-21.

E

Elbashir, S. M., J. Harborth *et al.* (2001a). "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells." Nature **411**(6836): 494-8.

Elbashir, S. M., W. Lendeckel and T. Tuschl (2001b). "RNA interference is mediated by 21- and 22-nucleotide RNAs." Genes Dev **15**(2): 188-200.

Elmayan, T., S. Balzergue *et al.* (1998). "*Arabidopsis* mutants impaired in cosuppression." Plant Cell **10**(10): 1747-58.

Emery, J. F., S. K. Floyd *et al.* (2003). "Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes." Curr Biol **13**(20): 1768-74.

Enright, A. J., B. John *et al.* (2003). "MicroRNA targets in *Drosophila*." Genome Biol **5**(1): R1. Epub 2003 Dec 12.

F

Fagard, M., S. Boutet *et al.* (2000). "AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals." Proc Natl Acad Sci U S A **97**(21): 11650-4.

Fagard, M. and H. Vaucheret (2000). "(TRANS)GENE SILENCING IN PLANTS: How Many Mechanisms?" Annu Rev Plant Physiol Plant Mol Biol **51**: 167-194.

Finnegan, E. J., R. Margis and P. M. Waterhouse (2003). "Posttranscriptional gene silencing is not compromised in the *Arabidopsis* CARPEL FACTORY (DICER-LIKE1) mutant, a homolog of DICER-1 from *Drosophila*." Curr Biol **13**(3): 236-40.

Fire, A., D. Albertson *et al.* (1991). "Production of antisense RNA leads to effective and specific inhibition of gene expression in *C. elegans* muscle." Development **113**(2): 503-14.

Fire, A., S. Xu *et al.* (1998). "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*." Nature **391**(6669): 806-11.

Fraser, A. G., R. S. Kamath *et al.* (2000). "Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference." Nature **408**(6810): 325-30.

G

Garnier, O., V. Serrano *et al.* (2004). "RNA-mediated programming of developmental genome rearrangements in *Paramecium tetraurelia*." Mol Cell Biol **24**(17): 7370-9.

Gitlin, L., S. Karelsky and R. Andino (2002). "Short interfering RNA confers intracellular antiviral immunity in human cells." Nature **418**(6896): 430-4.

Glazov, E., K. Phillips *et al.* (2003). "A gene encoding an RNase D exonuclease-like protein is required for post-transcriptional silencing in *Arabidopsis*." Plant J **35**(3): 342-9.

Goff, S. A., D. Ricke *et al.* (2002). "A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*)." Science **296**(5565): 92-100.

Goffeau, A., B. G. Barrell *et al.* (1996). "Life with 6000 genes." Science **274**(5287): 546, 563-7.

- Gonczy, P., C. Echeverri *et al.* (2000). "Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III." Nature **408**(6810): 331-6.
- Grad, Y., J. Aach *et al.* (2003). "Computational and experimental identification of *C. elegans* microRNAs." Mol Cell **11**(5): 1253-63.
- Griffiths-Jones, S. (2004). "The microRNA Registry." Nucleic Acids Res **32**(Database issue): D109-11.
- Grishok, A., A. E. Pasquinelli *et al.* (2001). "Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing." Cell **106**(1): 23-34.
- Grishok, A., H. Tabara and C. C. Mello (2000). "Genetic requirements for inheritance of RNAi in *C. elegans*." Science **287**(5462): 2494-7.
- Gwizdek, C., B. Ossareh-Nazari *et al.* (2003). "Exportin-5 mediates nuclear export of minihelix-containing RNAs." J Biol Chem **278**(8): 5505-8. Epub 2002 Dec 30.
- H**
- Hamilton, A., O. Voinnet *et al.* (2002). "Two classes of short interfering RNA in RNA silencing." Embo J **21**(17): 4671-9.
- Hamilton, A. J. and D. C. Baulcombe (1999). "A species of small antisense RNA in posttranscriptional gene silencing in plants." Science **286**(5441): 950-2.
- Hammond, S. M., E. Bernstein *et al.* (2000). "An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells." Nature **404**(6775): 293-6.
- Hammond, S. M., S. Boettcher *et al.* (2001). "Argonaute2, a link between genetic and biochemical analyses of RNAi." Science **293**(5532): 1146-50.
- Han, M. H., S. Goud *et al.* (2004). "The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation." Proc Natl Acad Sci U S A **101**(4): 1093-8. Epub 2004 Jan 13.

Hayashi, T., K. Makino *et al.* (2001). "Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12." DNA Res **8**(1): 11-22.

Helariutta, Y., H. Fukaki *et al.* (2000). "The SHORT-ROOT gene controls radial patterning of the Arabidopsis root through radial signaling." Cell **101**(5): 555-67.

Hogan, N. C., F. Slot *et al.* (1995). "Stability of tandem repeats in the *Drosophila melanogaster* Hsr-omega nuclear RNA." Genetics **139**(4): 1611-21.

Hutvagner, G., J. McLachlan *et al.* (2001). "A cellular function for the RNA-interference enzyme DICER in the maturation of the *let-7* small temporal RNA." Science **293**(5531): 834-8.

Hutvagner, G. and P. D. Zamore (2002). "A microRNA in a multiple-turnover RNAi enzyme complex." Science **297**(5589): 2056-60.

I

Ishizuka, A., M. C. Siomi and H. Siomi (2002). "A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins." Genes Dev **16**(19): 2497-508.

J

Jackson, A. L., S. R. Bartz *et al.* (2003). "Expression profiling reveals off-target gene regulation by RNAi." Nat Biotechnol **21**(6): 635-7.

Jackson, J. P., A. M. Lindroth *et al.* (2002). "Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase." Nature **416**(6880): 556-60.

Jacque, J. M., K. Triques and M. Stevenson (2002). "Modulation of HIV-1 replication by RNA interference." Nature **418**(6896): 435-8.

Jeddeloh, J. A., J. Bender and E. J. Richards (1998). "The DNA methylation locus DDM1 is required for maintenance of gene silencing in *Arabidopsis*." Genes Dev **12**(11): 1714-25.

Jensen, S., M. P. Gassama *et al.* (2002). "Regulation of I-Transposon Activity in *Drosophila*. Evidence for cosuppression of nonhomologous transgenes and possible role of ancestral i-related pericentromeric elements." Genetics **162**(3): 1197-209.

Johnston, R. J. and O. Hobert (2003). "A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*." Nature **426**(6968): 845-9. Epub 2003 Dec 14.

Jones, L., F. Ratcliff and D. C. Baulcombe (2001). "RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance." Curr Biol **11**(10): 747-57.

Jones-Rhoades, M. W. and D. P. Bartel (2004). "Computational Identification of Plant MicroRNAs and Their Targets, Including a Stress-Induced miRNA." Mol Cell **14**(6): 787-99.

Juarez, M. T., J. S. Kui *et al.* (2004). "microRNA-mediated repression of rolled leaf1 specifies maize leaf polarity." Nature **428**(6978): 84-8.

K

Kamath, R. S., A. G. Fraser *et al.* (2003). "Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi." Nature **421**(6920): 231-7.

Kasschau, K. D., Z. Xie *et al.* (2003). "P1/HC-Pro, a viral suppressor of RNA silencing, interferes with *Arabidopsis* development and miRNA unction." Dev Cell **4**(2): 205-17.

Kawasaki, H. and K. Taira (2004). "Induction of DNA methylation and gene silencing by short interfering RNAs in human cells." Nature **15**: 15.

Ketting, R. F., S. E. Fischer *et al.* (2001). "DICER functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*." Genes Dev **15**(20): 2654-9.

Ketting, R. F., T. H. Haverkamp *et al.* (1999). "Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD." Cell **99**(2): 133-41.

Ketting, R. F. and R. H. Plasterk (2000). "A genetic link between co-suppression and RNA interference in *C. elegans*." Nature **404**(6775): 296-8.

Khvorova, A., A. Reynolds and S. D. Jayasena (2003). "Functional siRNAs and miRNAs exhibit strand bias." Cell **115**(2): 209-16.

Kidner, C. A. and R. A. Martienssen (2004). "Spatially restricted microRNA directs leaf polarity through ARGONAUTE1." Nature **428**(6978): 81-4.

Kim, J., A. Krichevsky *et al.* (2004). "Identification of many microRNAs that copurify with polyribosomes in mammalian neurons." Proc Natl Acad Sci U S A **101**(1): 360-5. Epub 2003 Dec 22.

Kiyosawa, H., I. Yamanaka *et al.* (2003). "Antisense transcripts with FANTOM2 clone set and their implications for gene regulation." Genome Res **13**(6B): 1324-34.

Klahre, U., P. Crété *et al.* (2002). "High molecular weight RNAs and small interfering RNAs induce systemic posttranscriptional gene silencing in plants." Proc Natl Acad Sci U S A **99**(18): 11981-6.

Kloc, M., G. Spohr and L. D. Etkin (1993). "Translocation of repetitive RNA sequences with the germ plasm in *Xenopus* oocytes." Science **262**(5140): 1712-4.

Knight, S. W. and B. L. Bass (2001). "A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*." Science **293**(5538): 2269-71.

Krichevsky, A. M., K. S. King *et al.* (2003). "A microRNA array reveals extensive regulation of microRNAs during brain development." Rna **9**(10): 1274-81.

Kurihara, Y. and Y. Watanabe (2004). "From The Cover: *Arabidopsis* micro-RNA biogenesis through DICER-Like 1 protein functions." Proc Natl Acad Sci U S A **101**(34): 12753-8. Epub 2004 Aug 16.

Kuwabara, T., J. Hsieh *et al.* (2004). "A small modulatory dsRNA specifies the fate of adult neural stem cells." Cell **116**(6): 779-93.

L

- Lagos-Quintana, M., R. Rauhut *et al.* (2001). "Identification of novel genes coding for small expressed RNAs." Science **294**(5543): 853-8.
- Lagos-Quintana, M., R. Rauhut *et al.* (2003). "New microRNAs from mouse and human." Rna **9**(2): 175-9.
- Lagos-Quintana, M., R. Rauhut *et al.* (2002). "Identification of tissue-specific microRNAs from mouse." Curr Biol **12**(9): 735-9.
- Lai, E. C., P. Tomancak *et al.* (2003a). "Computational identification of *Drosophila* microRNA genes." Genome Biol **4**(7): R42. Epub 2003 Jun 30.
- Lander, E. S., L. M. Linton *et al.* (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.
- Lanz, R. B., N. J. McKenna *et al.* (1999). "A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex." Cell **97**(1): 17-27.
- Lau, N. C., L. P. Lim *et al.* (2001). "An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*." Science **294**(5543): 858-62.
- Laufs, P., A. Peaucelle *et al.* (2004). "MicroRNA regulation of the CUC genes is required for boundary size control in *Arabidopsis* meristems." Development **131**(17): 4311-22. Epub 2004 Aug 4.
- Lee, R. C. and V. Ambros (2001). "An extensive class of small RNAs in *Caenorhabditis elegans*." Science **294**(5543): 862-4.
- Lee, R. C., R. L. Feinbaum and V. Ambros (1993). "The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*." Cell **75**(5): 843-54.
- Lee, Y., C. Ahn *et al.* (2003). "The nuclear RNase III DROSHA initiates microRNA processing." Nature **425**(6956): 415-9.
- Lee, Y., K. Jeon *et al.* (2002). "MicroRNA maturation: stepwise processing and subcellular localization." Embo J **21**(17): 4663-70.

Lee, Y. S., K. Nakahara *et al.* (2004). "Distinct roles for *Drosophila* DICER-1 and DICER-2 in the siRNA/miRNA silencing pathways." Cell **117**(1): 69-81.

Lehner, B., G. Williams *et al.* (2002). "Antisense transcripts in the human genome." Trends Genet **18**(2): 63-5.

Lewis, B. P., I. H. Shih *et al.* (2003). "Prediction of mammalian microRNA targets." Cell **115**(7): 787-98.

Lewis, D. L., J. E. Hagstrom *et al.* (2002). "Efficient delivery of siRNA for inhibition of gene expression in postnatal mice." Nat Genet **32**(1): 107-108.

Liang, F., I. Holt *et al.* (2000). "Gene index analysis of the human genome estimates approximately 120,000 genes." Nat Genet **25**(2): 239-40.

Lim, L. P., M. E. Glasner *et al.* (2003a). "Vertebrate microRNA genes." Science **299**(5612): 1540.

Lim, L. P., N. C. Lau *et al.* (2003b). "The microRNAs of *Caenorhabditis elegans*." Genes Dev **17**(8): 991-1008. Epub 2003 Apr 2.

Lin, S. Y., S. M. Johnson *et al.* (2003). "The *C. elegans* hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target." Dev Cell **4**(5): 639-50.

Lindroth, A. M., X. Cao *et al.* (2001). "Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation." Science **292**(5524): 2077-80. Epub 2001 May 10.

Lingel, A., B. Simon *et al.* (2003). "Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain." Nature **426**(6965): 465-9. Epub 2003 Nov 16.

Lippman, Z., A. V. Gendrel *et al.* (2004). "Role of transposable elements in heterochromatin and epigenetic control." Nature **430**(6998): 471-6.

Liu, J., M. A. Carmell *et al.* (2004). "Argonaute2 Is the Catalytic Engine of Mammalian RNAi." Science **29**: 29.

Liu, Q., T. A. Rand *et al.* (2003). "R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway." Science **301**(5641): 1921-5.

Llave, C., K. D. Kasschau *et al.* (2002a). "Endogenous and silencing-associated small RNAs in plants." Plant Cell **14**(7): 1605-19.

Llave, C., Z. Xie *et al.* (2002b). "Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA." Science **297**(5589): 2053-6.

Lu, C. and N. Fedoroff (2000). "A mutation in the *Arabidopsis* HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin." Plant Cell **12**(12): 2351-2366.

Lund, E., S. Guttinger *et al.* (2004). "Nuclear export of microRNA precursors." Science **303**(5654): 95-8. Epub 2003 Nov 20.

Lynn, K., A. Fernandez *et al.* (1999). "The PINHEAD/ZWILLE gene acts pleiotropically in *Arabidopsis* development and has overlapping functions with the ARGONAUTE1 gene." Development **126**(3): 469-81.

M

Majdalani, N., C. Cunning *et al.* (1998). "DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription." Proc Natl Acad Sci U S A **95**(21): 12462-7.

Malhotra, P., P. V. Dasaradhi *et al.* (2002). "Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of *Plasmodium falciparum*." Mol Microbiol **45**(5): 1245-54.

Mallory, A. C., D. V. Dugas *et al.* (2004a). "MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs." Curr Biol **14**(12): 1035-46.

Mallory, A. C., L. Ely *et al.* (2001). "HC-Pro suppression of transgene silencing eliminates the small RNAs but not transgene methylation or the mobile signal." Plant Cell **13**(3): 571-83.

Mallory, A. C., B. J. Reinhart *et al.* (2004b). "MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region." Embo J **23**(16): 3356-64. Epub 2004 Jul 29.

Mallory, A. C. and H. Vaucheret (2004). "MicroRNAs: something important between the genes." Curr Opin Plant Biol **7**(2): 120-5.

Martens, H., J. Novotny *et al.* (2002). "RNAi in *Dictyostelium*: the role of RNA-directed RNA polymerases and double-stranded RNase." Mol Biol Cell **13**(2): 445-53.

Martinez, J., A. Patkaniowska *et al.* (2002). "Single-stranded antisense siRNAs guide target RNA cleavage in RNAi." Cell **110**(5): 563-74.

Martinez, J. and T. Tuschl (2004). "RISC is a 5' phosphomonoester-producing RNA endonuclease." Genes Dev **18**(9): 975-80. Epub 2004 Apr 22.

Masse, E., F. E. Escorcía and S. Gottesman (2003). "Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*." Genes Dev **17**(19): 2374-83. Epub 2003 Sep 15.

Mattick, J. S. (2001). "Non-coding RNAs: the architects of eukaryotic complexity." EMBO Rep **2**(11): 986-91.

McCaffrey, A. P., L. Meuse *et al.* (2002). "RNA interference in adult mice." Nature **418**(6893): 38-9.

McHale, N. A. and R. E. Koning (2004). "MicroRNA-directed cleavage of *Nicotiana sylvestris* PHAVOLUTA mRNA regulates the vascular cambium and structure of apical meristems." Plant Cell **16**(7): 1730-40. Epub 2004 Jun 11.

McRobert, L. and G. A. McConkey (2002). "RNA interference (RNAi) inhibits growth of *Plasmodium falciparum*." Mol Biochem Parasitol **119**(2): 273-8.

Meins, F., Jr. (2000). "RNA degradation and models for post-transcriptional gene-silencing." Plant Mol Biol **43**(2-3): 261-73.

- Meins, F., Jr. and C. Kunz (1995). "Gene silencing in transgenic plants: a heuristic autoregulation model." Curr Top Microbiol Immunol **197**: 105-20.
- Mette, M. F., W. Aufsatz *et al.* (2000). "Transcriptional silencing and promoter methylation triggered by double-stranded RNA." Embo J **19**(19): 5194-201.
- Metzler, M., M. Wilda *et al.* (2004). "High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma." Genes Chromosomes Cancer **39**(2): 167-9.
- Michael, M. Z., O. C. SM *et al.* (2003). "Reduced accumulation of specific microRNAs in colorectal neoplasia." Mol Cancer Res **1**(12): 882-91.
- Mochizuki, K., N. A. Fine *et al.* (2002). "Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*." Cell **110**(6): 689-99.
- Mochizuki, K. and M. A. Gorovsky (2004). "Conjugation-specific small RNAs in *Tetrahymena* have predicted properties of scan (scn) RNAs involved in genome rearrangement." Genes Dev **18**(17): 2068-73. Epub 2004 Aug 16.
- Morel, J. B., C. Godon *et al.* (2002). "Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance." Plant Cell **14**(3): 629-39.
- Morel, J. B. and H. Vaucheret (2000). "Post-transcriptional gene silencing mutants." Plant Mol Biol **43**(2-3): 275-84.
- Moss, E. G., R. C. Lee and V. Ambros (1997). "The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the lin-4 RNA." Cell **88**(5): 637-46.
- Mourelatos, Z., J. Dostie *et al.* (2002). "miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs." Genes Dev **16**(6): 720-8.
- Mourrain, P., C. Béclin *et al.* (2000). "Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance." Cell **101**(5): 533-42.
- Moussian, B., H. Schoof *et al.* (1998). "Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during *Arabidopsis* embryogenesis." Embo J **17**(6): 1799-809.

Murphy, S., F. Altruda *et al.* (1984). "DNA sequences complementary to human 7 SK RNA show structural similarities to the short mobile elements of the mammalian genome." J Mol Biol **177**(4): 575-90.

N

Napoli, C., C. Lemieux and R. Jorgensen (1990). "Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans." Plant Cell **2**(4): 279-289.

Ngo, H., C. Tschudi *et al.* (1998). "Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*." Proc Natl Acad Sci U S A **95**(25): 14687-92.

Nirenberg, M. W. and J. H. Matthaei (1961). "The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides." Proc Natl Acad Sci U S A **47**: 1588-602.

O

Ogawa, Y. and J. T. Lee (2002). "Antisense regulation in X inactivation and autosomal imprinting." Cytogenet Genome Res **99**(1-4): 59-65.

Ohler, U., S. Yekta *et al.* (2004). "Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification." Rna **10**(9): 1309-1322.

Okamura, K., A. Ishizuka *et al.* (2004). "Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways." Genes Dev **18**(14): 1655-66. Epub 2004 Jul 1.

Olsen, P. H. and V. Ambros (1999). "The *lin-4* regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation." Dev Biol **216**(2): 671-80.

Ooka, H., K. Satoh *et al.* (2003). "Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*." DNA Res **10**(6): 239-47.

Osato, N., H. Yamada *et al.* (2003). "Antisense transcripts with rice full-length cDNAs." Genome Biol **5**(1): R5. Epub 2003 Dec 11.

P

Palatnik, J. F., E. Allen *et al.* (2003). "Control of leaf morphogenesis by microRNAs." Nature **425**(6955): 257-63. Epub 2003 Aug 20.

Palauqui, J. C., T. Elmayan *et al.* (1997). "Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions." Embo J **16**(15): 4738-45.

Pal-Bhadra, M., U. Bhadra and J. A. Birchler (1997). "Cosuppression in *Drosophila*: gene silencing of Alcohol dehydrogenase by white-Adh transgenes is Polycomb dependent." Cell **90**(3): 479-90.

Pal-Bhadra, M., B. A. Leibovitch *et al.* (2004). "Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery." Science **303**(5658): 669-72.

Papp, I., M. F. Mette *et al.* (2003). "Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors." Plant Physiol **132**(3): 1382-90.

Park, W., J. Li *et al.* (2002). "CARPEL FACTORY, a DICER homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*." Curr Biol **12**(17): 1484-95.

Pasquinelli, A. E., B. J. Reinhart *et al.* (2000). "Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA." Nature **408**(6808): 86-9.

Pfeffer, S., M. Zavolan *et al.* (2004). "Identification of virus-encoded microRNAs." Science **304**(5671): 734-6.

Pham, J. W., J. L. Pellino *et al.* (2004). "A DICER-2-dependent 80s complex cleaves targeted mRNAs during RNAi in *Drosophila*." Cell **117**(1): 83-94.

Plath, K., S. Mlynarczyk-Evans *et al.* (2002). "Xist RNA and the mechanism of X chromosome inactivation." Annu Rev Genet **36**: 233-78. Epub 2002 Jun 11.

R

Reinhart, B. J., F. J. Slack *et al.* (2000). "The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*." Nature **403**(6772): 901-6.

Reinhart, B. J., E. G. Weinstein *et al.* (2002). "MicroRNAs in plants." Genes Dev **16**(13): 1616-26.

Rhoades, M. W., B. J. Reinhart *et al.* (2002). "Prediction of plant microRNA targets." Cell **110**(4): 513-20.

Riggs, A. D., R. A. Martienssen and V. E. A. Russo (1996). "Epigenetic mechanisms of gene regulation." NY: Cold spring Harbor Laboratory Press **1-4**.

Rivas, E., R. J. Klein *et al.* (2001). "Computational identification of noncoding RNAs in *E. coli* by comparative genomics." Curr Biol **11**(17): 1369-73.

Robinson, K. A. and S. M. Beverley (2003). "Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite *Leishmania*." Mol Biochem Parasitol **128**(2): 217-28.

Rouze, P., N. Pavy and S. Rombauts (1999). "Genome annotation: which tools do we have for it?" Curr Opin Plant Biol **2**(2): 90-5.

S

Schwarz, D. S., G. Hutvagner *et al.* (2003). "Asymmetry in the assembly of the RNAi enzyme complex." Cell **115**(2): 199-208.

Schwarz, D. S., Y. Tomari and P. D. Zamore (2004). "The RNA-induced silencing complex is a Mg²⁺-dependent endonuclease." Curr Biol **14**(9): 787-91.

Schwarz, D. S. and P. D. Zamore (2002). "Why do miRNAs live in the miRNP?" Genes Dev **16**(9): 1025-31.

Semizarov, D., L. Frost *et al.* (2003). "Specificity of short interfering RNA determined through gene expression signatures." Proc Natl Acad Sci U S A **100**(11): 6347-52.

Sempere, L. F., N. S. Sokol *et al.* (2003). "Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and broad-Complex gene activity." Dev Biol **259**(1): 9-18.

Shendure, J. and G. M. Church (2002). "Computational discovery of sense-antisense transcription in the human and mouse genomes." Genome Biol **3**(9): RESEARCH0044. Epub 2002 Aug 22.

Sijen, T., J. Fleenor *et al.* (2001a). "On the role of RNA amplification in dsRNA-triggered gene silencing." Cell **107**(4): 465-76.

Sijen, T., I. Vijn *et al.* (2001b). "Transcriptional and posttranscriptional gene silencing are mechanistically related." Curr Biol **11**(6): 436-40.

Silhavy, D. and J. Burgyan (2004). "Effects and side-effects of viral RNA silencing suppressors on short RNAs." Trends Plant Sci **9**(2): 76-83.

Skelly, P. J., A. Da'dara and D. A. Harn (2003). "Suppression of cathepsin B expression in *Schistosoma mansoni* by RNA interference." Int J Parasitol **33**(4): 363-9.

Slack, F. J., M. Basson *et al.* (2000). "The *lin-41* RBCC gene acts in the *C. elegans* heterochronic pathway between the *let-7* regulatory RNA and the LIN-29 transcription factor." Mol Cell **5**(4): 659-69.

Sleutels, F., R. Zwart and D. P. Barlow (2002). "The non-coding *Air* RNA is required for silencing autosomal imprinted genes." Nature **415**(6873): 810-3.

Smardon, A., J. M. Spoerke *et al.* (2000). "EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in *C. elegans*." Curr Biol **10**(4): 169-78.

Song, C. and A. E. Simon (1995). "Requirement of a 3'-terminal stem-loop in in vitro transcription by an RNA-dependent RNA polymerase." J Mol Biol **254**(1): 6-14.

Song, E., S. K. Lee *et al.* (2003a). "RNA interference targeting Fas protects mice from fulminant hepatitis." Nat Med **9**(3): 347-51.

Song, J. J., J. Liu *et al.* (2003b). "The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes." Nat Struct Biol **10**(12): 1026-32.

Song, J. J., S. K. Smith *et al.* (2004). "Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity." Science **29**: 29.

Souret, F. F., J. P. Kastenmayer and P. J. Green (2004). "AtXRN4 Degrades mRNA in Arabidopsis and Its Substrates Include Selected miRNA Targets." Mol Cell Biol **15**(2): 173-83.

Sousa, C., C. Johansson *et al.* (2001). "Translational and structural requirements of the early nodulin gene *enod40*, a short-open reading frame-containing RNA, for elicitation of a cell-specific growth response in the *alfalfa* root cortex." Mol Cell Biol **21**(1): 354-66.

Steimer, A., P. Amedeo *et al.* (2000). "Endogenous targets of transcriptional gene silencing in *Arabidopsis*." Plant Cell **12**(7): 1165-78.

Storz, G., J. A. Opdyke and A. Zhang (2004). "Controlling mRNA stability and translation with small, noncoding RNAs." Curr Opin Microbiol **7**(2): 140-4.

Sunkar, R. and J. K. Zhu (2004). "Novel and Stress-Regulated MicroRNAs and Other Small RNAs from *Arabidopsis*." Plant Cell **16**(8): 2001-19. Epub 2004 Jul 16.

Szymanski, M., M. Z. Barciszewska *et al.* (2003). "Noncoding RNA transcripts." J Appl Genet **44**(1): 1-19.

T

Tabara, H., M. Sarkissian *et al.* (1999). "The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*." Cell **99**(2): 123-32.

Tahbaz, N., F. A. Kolb *et al.* (2004). "Characterization of the interactions between mammalian PAZ PIWI domain proteins and DICER." EMBO Rep **5**(2): 189-94. Epub 2004 Jan 16.

Takada, S., K. Hibara *et al.* (2001). "The CUP-SHAPED COTYLEDON1 gene of *Arabidopsis* regulates shoot apical meristem formation." Development **128**(7): 1127-35.

Tang, G., B. J. Reinhart *et al.* (2003). "A biochemical framework for RNA silencing in plants." Genes Dev **17**(1): 49-63.

The *C. elegans* Sequencing Consortium. (1998) "Genome sequence of the nematode *C. elegans*: a platform for investigating biology". " Science **282**(5396): 2012-8.

Tijsterman, M., R. F. Ketting *et al.* (2002a). "RNA helicase MUT-14-dependent gene silencing triggered in *C. elegans* by short antisense RNAs." Science **295**(5555): 694-7.

Tijsterman, M., K. L. Okihara *et al.* (2002b). "PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of *C. elegans*." Curr Biol **12**(17): 1535-40.

Tomari, Y., T. Du *et al.* (2004). "RISC assembly defects in the *Drosophila* RNAi mutant armitage." Cell **116**(6): 831-41.

Tufarelli, C., J. A. Stanley *et al.* (2003). "Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease." Nat Genet **34**(2): 157-65.



Van der Krol, A. R., L. A. Mur *et al.* (1990). "Flavonoid genes in *Petunia*: addition of a limited number of gene copies may lead to a suppression of gene expression." Plant Cell **2**(4): 291-9.

Vaucheret, H., F. Vazquez *et al.* (2004). "The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development." Genes Dev **18**(10): 1187-97. Epub 2004 May 6.

Vazquez, F., V. Gascioli *et al.* (2004a). "The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing." Curr Biol **14**(4): 346-51.

Vazquez, F., H. Vaucheret *et al.* (2004b). "Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs." Mol Cell sous presse.

Venter, J. C., M. D. Adams *et al.* (2001) "The sequence of the human genome."

- Vielle-Calzada, J. P., R. Baskar and U. Grossniklaus (2000). "Delayed activation of the paternal genome during seed development." Nature **404**(6773): 91-4.
- Voinnet, O. (2001). "RNA silencing as a plant immune system against viruses." Trends Genet **17**(8): 449-59.
- Voinnet, O., C. Lederer and D. C. Baulcombe (2000). "A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*." Cell **103**(1): 157-67.
- Voinnet, O., Y. M. Pinto and D. C. Baulcombe (1999). "Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants." Proc Natl Acad Sci U S A **96**(24): 14147-52.
- Voinnet, O., P. Vain *et al.* (1998). "Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA." Cell **95**(2): 177-87.
- W**
- Walbot, V. and D. A. Petrov (2001). "Gene galaxies in the maize genome." Proc Natl Acad Sci U S A **98**(15): 8163-4.
- Wang, H., A. Iacoangeli *et al.* (2002). "Dendritic BC1 RNA: functional role in regulation of translation initiation." J Neurosci **22**(23): 10232-41.
- Wang, J. F., H. Zhou *et al.* (2004). "Identification of 20 microRNAs from *Oryza sativa*." Nucleic Acids Res **32**(5): 1688-95. Print 2004.
- Wassarman, K. M., F. Repoila *et al.* (2001). "Identification of novel small RNAs using comparative genomics and microarrays." Genes Dev **15**(13): 1637-51.
- Watson, J. D. and F. H. Crick (1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid." Nature **171**(4356): 737-8.
- Wesley, S. V., C. A. Helliwell *et al.* (2001). "Construct design for efficient, effective and high-throughput gene silencing in plants." Plant J **27**(6): 581-90.

Wianny, F. and M. Zernicka-Goetz (2000). "Specific interference with gene function by double-stranded RNA in early mouse development." Nat Cell Biol **2**(2): 70-5.

Wienholds, E., M. J. Koudijs *et al.* (2003). "The microRNA-producing enzyme DICER1 is essential for zebrafish development." Nat Genet **35**(3): 217-8.

Wightman, B., T. R. Burglin *et al.* (1991). "Negative regulatory sequences in the *lin-14* 3'-untranslated region are necessary to generate a temporal switch during *Caenorhabditis elegans* development." Genes Dev **5**(10): 1813-24.

Wightman, B., I. Ha and G. Ruvkun (1993). "Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*." Cell **75**(5): 855-62.

Wilda, M., U. Fuchs *et al.* (2002). "Killing of leukemic cells with a BCR/ABL fusion gene by RNA interference (RNAi)." Oncogene **21**(37): 5716-24.

Winston, W. M., C. Molodowitch and C. P. Hunter (2002). "Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1." Science **295**(5564): 2456-9.

Wong, G. K., D. A. Passey and J. Yu (2001). "Most of the human genome is transcribed." Genome Res **11**(12): 1975-7.

Wu-Scharf, D., B. Jeong *et al.* (2000). "Transgene and transposon silencing in *Chlamydomonas reinhardtii* by a DEAH-box RNA helicase." Science **290**(5494): 1159-62.

X

Xie, Z., L. K. Johansen *et al.* (2004). "Genetic and functional diversification of small RNA pathways in plants." PLoS Biol **2**(5): E104. Epub 2004 Feb 24.

Xie, Z., K. D. Kasschau and J. C. Carrington (2003). "Negative feedback regulation of DICER-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation." Curr Biol **13**(9): 784-9.

Xu, P., S. Y. Vernooy *et al.* (2003). "The *Drosophila* microRNA Mir-14 suppresses cell death and is required for normal fat metabolism." Curr Biol **13**(9): 790-5.

Y

Yamasaki, K., K. Joh *et al.* (2003). "Neurons but not glial cells show reciprocal imprinting of sense and antisense transcripts of *Ube3a*." Hum Mol Genet **12**(8): 837-47.

Yan, K. S., S. Yan *et al.* (2003). "Structure and conserved RNA binding of the PAZ domain." Nature **426**(6965): 468-74. Epub 2003 Nov 16.

Yao, M. C., P. Fuller and X. Xi (2003). "Programmed DNA deletion as an RNA-guided system of genome defense." Science **300**(5625): 1581-4.

Yekta, S., I. H. Shih and D. P. Bartel (2004). "MicroRNA-directed cleavage of HOXB8 mRNA." Science **304**(5670): 594-6.

Yelin, R., D. Dahary *et al.* (2003). "Widespread occurrence of antisense transcription in the human genome." Nat Biotechnol **21**(4): 379-86.

Yi, R., Y. Qin *et al.* (2003). "EXPORTIN-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs." Genes Dev **17**(24): 3011-6. Epub 2003 Dec 17.

Yokota, T., N. Sakamoto *et al.* (2003). "Inhibition of intracellular hepatitis C virus replication by synthetic and vector-derived small interfering RNAs." EMBO Rep **4**(6): 602-8.

Z

Zamore, P. D., T. Tuschl *et al.* (2000). "RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals." Cell **101**(1): 25-33.

Zeng, Y. and B. R. Cullen (2003). "Sequence requirements for micro RNA processing and function in human cells." Rna **9**(1): 112-23.

Zhang, H., F. A. Kolb *et al.* (2002). "Human DICER preferentially cleaves dsRNAs at their termini without a requirement for ATP." Embo J **21**(21): 5875-85.

Zhang, H., F. A. Kolb *et al.* (2004). "Single processing center models for human DICER and bacterial RNase III." Cell **118**(1): 57-68.

Zilberman, D., X. Cao and S. E. Jacobsen (2003). "ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation." Science **299**(5607): 716-9.

Analyse de mutants affectés dans les contrôles épigénétiques post-transcriptionnels chez *Arabidopsis thaliana*.

Zilberman, D., X. Cao *et al.* (2004). "Role of *Arabidopsis* ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats." Curr Biol **14**(13): 1214-20.