Les AtNSRs, protéines régulatrices de l'épissage alternatif et du silencing post transcriptionnel

Florian Bardou

▶ **To cite this version:**

Florian Bardou. Les AtNSRs, protéines régulatrices de l'épissage alternatif et du silencing post transcriptionnel. Sciences agricoles. Université Paris Sud - Paris XI, 2013. Français. NNT : 2013PA112054 . tel-01139502

HAL Id: tel-01139502
https://theses.hal.science/tel-01139502

Submitted on 6 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE PARIS-SUD

ÉCOLE DOCTORALE : Science du végétal

DISCIPLINE BIOLOGIE

THÈSE DE DOCTORAT

par

Florian BARDOU

Les AtNSRs, protéines régulatrices de l'épissage alternatif et du silencing post transcriptionnel

Soutenance le 04 avril 2013

Directeur de thèse : Martin Crespi DR2 (ISV CNRS, Gif sur Yvette)

Composition du jury :

Rapporteurs : DERAGON Jean-marc PR1 (Université de Perpignan UPVD)
GAGLIARDI Dominique DR2 (IBMP CNRS, Strasbourg)

Examinateurs : DELARUE Marianne MCF (IBP, U-Psud, Orsay)
NUSSAUME Laurent DR1 (CEA, Cadarache)

Remerciements

Tout d'abord, je souhaite remercier Jean-Marc Deragon et Dominique Gagliardi d'avoir accepté d'êtres rapporteurs de cette thèse. De même, je remercie les autres membres du jury, Marianne Delarue et Laurent Nussaume d'avoir accepté de juger mes travaux.

Bien entendu, je souhaite remercier Martin Crespi tout particulièrement pour m'avoir donné l'opportunité d'intégrer son équipe et pour le temps, la patience et l'attention qu'il m'à porté durant ces 4 ans. J'insiste sur le fait que sa passion contagieuse pour la science et sa capacité à encourager dans les moments difficile m'ont permis de travailler dans des conditions exceptionnelles.

Je remercie Allison Mallory, Cécile Anthoneli et Jeff Leung pour leur excellents conseils durant les comités de thèse.

Je remercie mes collègues de labo du 210 et du 236 (Fede, Ana, Virginie, Aurélie, Nathalie, le prince khan, Jéremie, Céline, Christine, Philippe, Alexis et tout les autres) avec qui j'ai pu apprendre mais aussi partager, ils ont prouvé que travail et divertissement peuvent être réunis pour une bonne ambiance de boulot.

Je souhaite aussi remercier mes collègues de la team Frugier avec qui on passe de bon moment au repas mais aussi dans l'interminable RER. Je souhaite remercier particulièrement Anouck qui m'a expliqué l'art d'enseigner à la Diet. De même, je souhaite remercier tout ceux qui ont permis que mon monitorat à l'université P7 ce soit bien déroulé.

L'ISV fourmille de personnes compétentes et intéressantes, je remercie tous les membres qui permettent de faire vivre cet institut.

Je souhaite remercier Caroline Hartmann pour la correction de l'introduction de mon manuscrit.

Merci à mes deux stagiaires Quentin et Marine, les encadrer a été une expérience formidable.

Je voudrais aussi remercier sincèrement mes parents, ma sœur et ma famille ainsi que tout mes proches. Une grosse pensée pour mon chat∞.

Enfin merci à mes amis d'ici ou de là pour tout le reste. Merci aussi au « non lieu » pour les bon moment passés durant les derniers mois de ma thèse.

Table des matières

I- Introduction ……1

1- L'Architecture de la racine ……1

1.1- Formation de la racine latérale et le rôle de la signalisation par l'auxine...2

1.2- Les ARN non-codant régulateurs de l'architecture racinaire ……...7

2- Les Protéines de liaison à l'ARN ……...13

2.1 Les RBPs dans les corps cytoplasmiques ……14

2.2.1 Les RBPs dans les p-bodies ……15

2.2.2 Les RBPs dans les « silencing bodies » ……16

2.2.3 Les RBPs dans les stress granules ……17

2.2- Les RBPs dans les corps nucléaires ……17

2.1.1 Le Nucléole ……19

2.1.2 Les RBPs dans les Cajal bodies ……20

2.1.3 Les RBPs dans les Dicing bodies ……...21

2.1.4 Les RBPs dans les Speckles ……22

2.1.5 npcARN et RNA binding proteins dans le noyau ……..24

3- L'épissage alternatif ……33

3.1. L'épissage du pré-ARNm chez les plantes ……...33

3.2- Régulation de l'AS ……37

4- Objectifs ……...41

II- Résultats ……43

II.1- Modulation of nuclear alternative splicing regulators by long ncRNA ……44

1.1 Introduction ……44

1.2 Publication ……45

Materials and methods ……....52

References ……...57

Figures ……..……..61

1.3 Résultats complémentaires ……..77

II.2- SGS14, SGS15, NSRa and NSRb encode alternative splicing-related factors that facilitate post-transcriptional gene silencing mediated by intron-containing transgenes ……79

2.1 Introduction sur le silencing post transcriptionnel ……79

2.2 Les mutants déficients en PTGS ……..80

2.2.2 La lignée *p35S::uidA* « L1 »

2.2.3La lignée *p35S::NIA2* « *2a3* »

2.2.4 La lignée *pSUC2*::*antisensPDS* « JAP3 »

2.3 la voie du S-PTGS des transgènes et propagation du silencing ……82

2.4 la voie du IR-PTGS des transgènes ……84

II.3- Publication ……..87

Experimental procedures…..……101
References ……..…...103
Figures ……..……106

II.4- Résultats complémentaires ……..117

III- Discussion ……119

1- Localisation des NSR et interaction avec les lncARN ……..121

2- Les AtNSRs, des régulateurs nucléaires de l'épissage alternatif ……123

3- Les AtNSRs et le silencing ……..125

IV- Perspectives ……129

IV- Matériels et méthodes ……131

V- Bibliographie ……137

VI- annexes ……165

Abréviations

ARN: Acide ribonucléique (English: RNA)

ADNc: AND codant

ncARN: ARN non codant (non coding RNA)

npcARN/ncARN: ARN non codant pour des protéines (non protein coding RNA)

siARN: small interfering RNA

miARN: microARN

RL: Racine latérale

P-bodies: Processing-bodies

RISC: RNA-Induced silencing complexe

dsARN: ARN double brin

ssARN: ARN simple brin

RDR: RNA-dependante RNA polymerase

AS: Epissage alternative (Alternative splicing)

FT: Facteur de transcription

GFP: Green fluorescente protein

NAA: Naphthalene Acetic Acid

At: *Arabidopsis thaliana*

Mt: *Medicago truncatula*

RBP: Protéine de liaison à l'ARN (RNA binding protein)

NSR: Nuclear speckles RNA binding protein

RNP: Ribonucléoprotéine

PTGS: Silencing post-transcriptionnel des gènes

TGS: Silencing transcriptionnel des gènes

sORF: Short open reading frame

ENOD: early nodulin factor

NMD: Non sens mediated decay

RIP: RNA immunoprecipitation

nat-siRNA: naturally occurring antisens siRNA

ARF: Auxin response factor

AGO: ARGONAUTE

I- Introduction

I- Introduction

1. L'architecture de la racine

Les plantes sont des organismes sessiles, elles subissent donc leur environnement et se doivent de s'adapter au mieux de leur habitat. La ramification au niveau des racines permet l'absorption en eau et en nutriments, ainsi que l'ancrage de la plante dans le sol. La racine diffère de la tige par plusieurs caractères : sa structure interne, son géotropisme positif, la présence d'une coiffe terminale et de poils absorbants et la présence de racines latérales (RL). Les racines sont aussi le siège de symbioses avec les bactéries et les champignons du sol, en particulier pour le métabolisme de l'azote chez les légumineuses. Les racines peuvent présenter des adaptations afin de faciliter le développement de la plante dans un environnement particulier (exemple des racines du palétuvier). Dans certains cas les racines servent aussi à stocker des nutriments (comme chez le radis par exemple).

Le système racinaire des plantes est une structure dynamique dont l'architecture est déterminée par la modulation de la croissance de la racine principale et la capacité de ramification en réponse à l'environnement du sol. Cette plasticité repose donc sur l'intégration continue de facteurs environnementaux et de programmes endogènes contrôlant le développement des racines latérales.

Les plantes comme les animaux possèdent plusieurs molécules signales endogènes, nommées phytohormones, qui contrôlent leur développement. Plusieurs de ces hormones ont été rapportées comme ayant un rôle dans la régulation de l'architecture des racines, notamment l'auxine, les cytokinines, l'acide abscissique, l'éthylène et les gibbérellines. L'auxine (ou acide indole 3-acétique) est une phytohormone et une régulatrice majeure de la croissance de la plante et notamment du développement racinaire. En outre, Il existe de nombreux facteurs environnementaux qui contrôlent l'architecture de la racine, la disponibilité en nutriments comme le nitrate, le phosphate ou le sulfate (figure 1), la disponibilité en eau, la température, ou encore le potentiel hydrogène (ou pH). De même, la croissance en conditions de stress abiotique comme par exemple la présence de NaCl ou de métaux entraîne une adaptation morphologique du système racinaire qui permet à la plante de survivre dans ces conditions environnementales (Osmond et al., 2007).

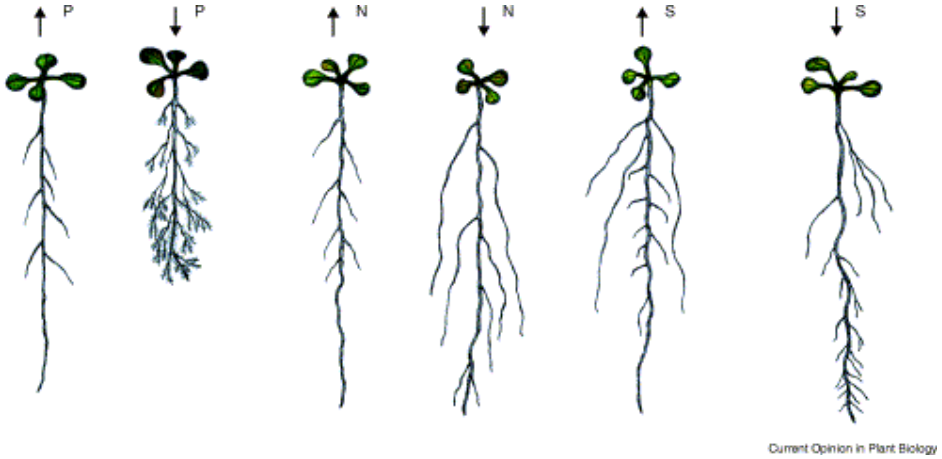


Figure 1 La plasticité de l'architecture de la racine chez *Arabidopsis thaliana*

L'architecture racinaire des plantes comme celle *d'Arabidopsis thaliana* présente de faibles variations d'un individu à l'autre. En revanche, l'architecture racinaire est fortement modifiée en réponse à la disponibilité en nutriments (P=Phosphate; N=nitrate; S=Sulfate) représentant la capacité d'adaptation de la plante pour son environnement. (Lopez-Bucio et al., 2003)

1.1- Formation de la racine latérale et le rôle de la signalisation par l'auxine

Nous limiterons ce chapitre aux données concernant la plante modèle *Arabidopsis thaliana.* La formation des racines latérales chez *Arabidopsis*, comme chez la plupart des espèces, est un processus développement qui combine plusieurs étapes. Il est bien établi que les cellules du péricycle, situées au pôle de xylème sont à l'origine des cellules fondatrices des racines latérales (Dolan et al., 1993 ; Figure 2). L'étape de pré-initiation de ces cellules particulières du péricycle est dépendante du flux d'auxine provenant du méristème de la racine lors des étapes précoces du développement de la plante (De Smet et al., 2007). Cette première étape se déroule rythmiquement à la base du méristème apical de la racine. Ensuite, au fur et à mesure de la croissance de la racine primaire, l'étape d'initiation des racines latérales est déclenchée par la réponse des cellules fondatrices à un second signal auxinique provenant cette fois-ci de la tige (Dubrovsky et al., 2008). Cette initiation débute par des divisions successives des cellules fondatrices (Malamy and Benfey, 1997 ; Dubrovsky et al., 2001 ; Figure 3a), résultant dans la formation d'un primordium constitué d'une simple couche de cellules. De nouvelles divisions sont ensuite à l'origine d'un primordium en forme de dôme qui va émerger des tissus de la racine principale. Bien que ces étapes se déroulent de façon séquentielles tout au long de la racine primaire, elles peuvent être interrompues (Dubrovsky et al., 2006 ; Lucas et al., 2008). Les différentes études menées principalement chez *Arabidopsis*, ont permis d'identifier et de caractériser en détail les mécanismes génétiques contrôlant ce processus développement. Toutes ces études convergent vers un rôle essentiel de l'auxine dans la ramification de la racine et

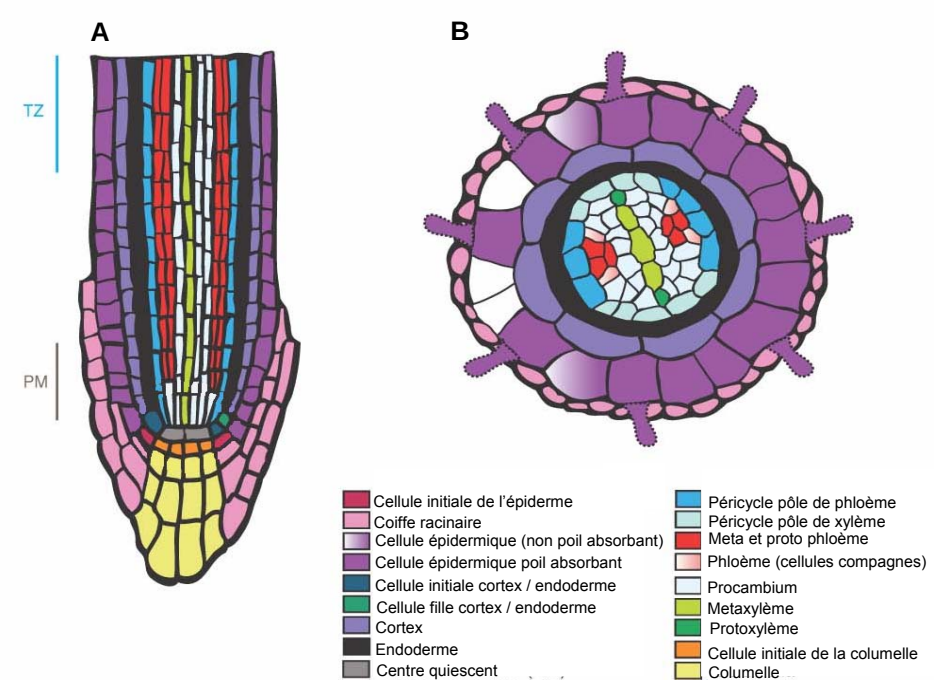


Figure 3 Anatomie d'une racine d'*Arabidopsis thaliana*

(A) Schéma d'une coupe longitudinale de racine, les différents types cellulaires sont représentés par des couleurs distinctes. La niche de cellules souches (ou STN pour stem cell niche) se situe au dessus de la coiffe et comprend le centre quiescent et les cellules initiales. Le méristème proximal (PM pour proximal meristem) est alimenté par des divisions des cellules initiales ; après quelques divisions supplémentaires ces cellules commencent à s'allonger dans la zone de transition (TZ pour transition zone). (B) Schéma d'une section transversale de racine montrant une symétrie radiale autour de l'axe pour les couches externes des cellules. Il y a environ 15 types de cellules différentes dans la racine. L'épiderme comporte des files de cellules atrichoblastiques et trichoblastiques respectivement incapables et capable de former des poils absorbants. Les tissus vasculaires, xylème et phloème, sont dans la stèle situés à des pôles opposés. L'axe radial est composé du cylindre vasculaire entouré par le péricycle (ensemble, ils constituent la stèle), puis l'endoderme, le cortex, et enfin l'épiderme (Heidi et al., 2012).

établissent un parallèle avec le rôle de cette hormone dans les parties aériennes (dominance apicale). Les mouvements d'auxine au sein de la racine et les réponses spécifiques de certaines cellules à l'hormone sont des éléments essentiels qui régulent l'architecture de la racine. L'utilisation d'inhibiteur de transport d'auxine comme le NPA (N-1-Naphtylphtalamic-acid) a permis d'établir que le transport basipète (venant du méristème racinaire) de l'auxine est important pour la pre-initiation (Casimiro et al., 2001). A ce stade, l'auxine serait critique pour la transition G1/S des cellules du péricycle (Himanen et al., 2002).

Des résultats récents commencent à mettre en lumière les effecteurs de l'auxine intervenant dans la formation de racines latérales et plus précisément lors de la pré-initiation. D'une manière générale, la signalisation par l'auxine dépend de l'action de couples de facteurs de transcription (FT) : IAA et ARF qui sont fixés sur des motifs AUXRE (Auxin response element) des promoteurs de gènes répondant à l'auxine (Figure 3C). En absence d'hormone, au niveau du promoteur, chaque protéine IAA forme un dimère avec une protéine ARF (Auxin response factors) spécifique, ce complexe empêchant la transcription

des gènes répondant à la phytohormone. En présence d'auxine, la dégradation des protéines IAA est induite, via le protéasome, ce qui permet la dimùérisation de l'ARF et ainsi la transcription du gène de réponse à l'auxine. Plusieurs couples ARF et IAA existent chez les plantes ; certains d'entre eux participent à la signalisation racinaire.

Le FT GATA23, est le marqueur moléculaire le plus précoce du développement des racines latérales; il est en effet détecté dans les cellules fondatrices (De Rybel et al., 2010 ; Figure 2). Les plantes qui surexpriment le gène GATA23 présentent un nombre plus important de cellules fondatrices mais également plus de primordia de racines latérales par rapport aux plantes sauvages. Ce phénotype semble donc indiquer que GATA23 est suffisant pour déterminer l'identité des cellules fondatrices. L'expression du gène GATA23 est réprimée par IAA28. Ce FT de type IAA interagit avec 5 ARFs différentes : ARF5, ARF6, ARF7, ARF8 et ARF19 (De Rybel et al., 2010). IAA28/ARFs/GATA23 représente donc le premier module contrôlant l'identité des cellules fondatrices. La première preuve visible de l'initiation de racine latérale est la migration des noyaux de deux cellules du péricycle vers leur paroi cellulaire commune (Figure 3A, 3B, 3D) (De Smet et al., 2007). Cette migration nucléaire est dépendante du module SOLITARY ROOT/IAA14 (SLR/IAA14), IAA28, ARF7, ARF19 (Fukaki et al., 2002 ; Fukaki et al., 2005 ; De Rybel et al., 2010 ; Figure 3D). Dans les conditions normales, IAA14 se lie est réprime l'action des FT ARF7 et ARF 19 (Okushima et al., 2005). La migration des noyaux est suivie rapidement par une première division anticline. Le récepteur de type kinase ARABIDOPSIS CRINKLY4 (ACR4) contrôle cette division. Les phénotypes du mutant nul *acr4* et des plantes surexprimant ce gène sont opposés. En effet, il a été observé une plus faible densité de primordia de racines latérales chez le mutant et une densité plus forte chez le surexpresseur.

De plus, les primordia formés chez le mutant *acr4* présentent des caractéristiques anormales. Ces résultats indiquent un rôle d' ACR4 dans la répression de la division cellulaire prolifération. ACR4 agit donc à la fois de façon autonome dans les cellules fondatrices, afin d'assurer la spécification de celles-ci et de façon non autonome dans les cellules voisines pour empêcher des divisions à l'origine de racines latérales trop proches les unes des autres (De Smet et al., 2008 ; Figure 3B). Ensuite, le primordium « monocouche» (stage1) subit une succession de divisions périclines et anticlines. La première division péricline génère un primordium à deux feuillets (stage 2), qui se divise pour aboutir à une forme caractéristique en dôme (stage 3 à 8) qui pourra ou non émerger de la racine principale (Malamy and Benfey, 1997). Récemment, un second module de signalisation par l'auxine, constitué de BODENLOSS (BDL/IAA12) et MONOPTEROS (MP/ARF5), a été

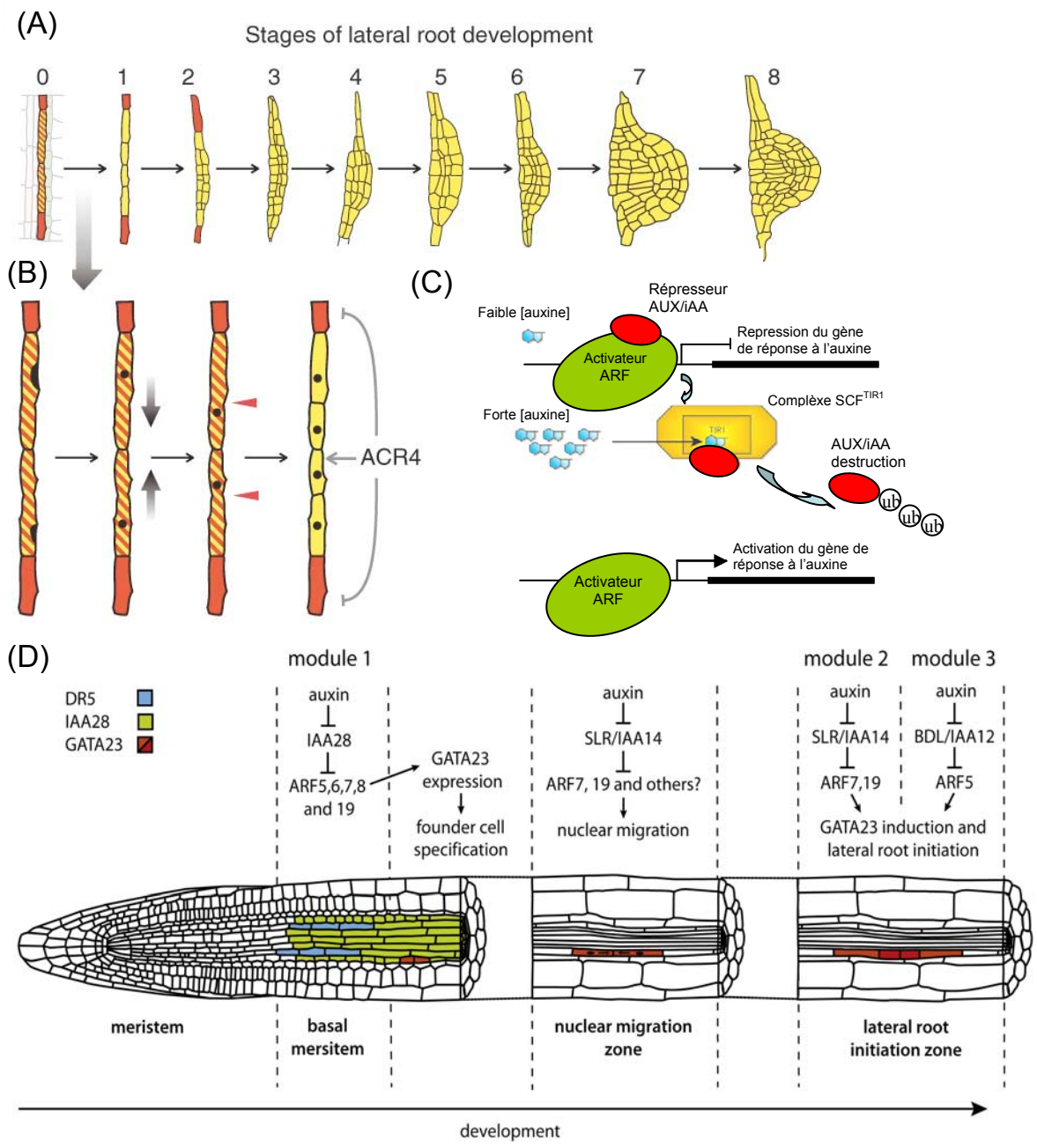


Figure 3 Les événements cellulaire du développement de la racine latérale chez *Arabidospsis* (Malamy and Benfey, 1997 (a) ; De Rybel et al., 2010 (c)).

(A) Chaque étape du développement des racines latérales est représenté, de la pré-initiation (stade 0) à l'émergence du primordium (étape 8). Les étapes sont numérotées selon (Malamy and Benfey, 1997). (B) Détails des événements cellulaires précoces au cours de l'initiation des racines latérales. Les noyaux des deux cellules du péricycle répondant à l'auxine s'arrondissent et migrent vers leur paroi cellulaire commune (flèches grises). Cette étape est suivie par deux divisions asymétriques (flèches rouges) créant deux cellules plus petites dans le centre et deux grandes à la périphérie. Le récepteur ACR4 contrôle ces divisions tout en réprimant celles des cellules voisines. (C) Schéma représentant le principe de la signalisation par l'auxine via les FT ARFs (Auxin response elements) et les répresseurs Aux/IAA. Le Complexe SCF^{TIR1} est composé de (Skp1/Cullin/F-box protein Transport Inhibitor Response 1). ub represente l'ubiquitination des Aux/IAA qui entraine leur dégradation par le protéasome. (D) Représentation schématique des modules de signalisation d'auxine agissant successivement pendant l'acquisition de la spécification des cellules fondatrices dans le méristème basal (module 1) et au cours de la migration nucléaire ainsi que lors des processus d'initiation des racines latérales (modules 2 et 3). L'action du module 1 permet l'expression de GATA23 qui est le premier marqueur des RL. GATA 23 pourrait intervenir dans la détermination de l'identité cellulaire des cellules fondatrices.

identifié ; il agit en effet après le module SLR/ARF7/ARF19 lors de l'initiation des racines latérales (De Smet et al., 2010, Figure 3D).

La phase d'émergence correspond au moment où cette racine latérale traverse le cortex et l'épiderme de la racine primaire (Malamy and Benfey, 1997, Figure 3a, 3b). La progression du primordium induit la reprogrammation des cellules voisines qui synthétisent les enzymes nécessaires aux remodelages de la paroi cellulaire (Gonzalez-Carranza et al., 2007). Ces modifications pariétales sont indispensables à l'émergence des racines latérales. En effet, chez le mutant *lax3* (un transporteur d'influx d'auxin) le nombre de racines latérales émergées est affecté alors que le nombre d'initiations de primordia est identique à celui observé chez les plantes sauvages. LAX3 est spécifiquement exprimée dans les cellules corticales et épidermiques, juste en face du point d'émergence de la racine latérale en formation. L'influx de l'auxine médié par LAX3 dans ces cellules conduit à la production d'enzymes de remodelage de la paroi cellulaire afin de faciliter la séparation des cellules et donc l'émergence de la racine nouvellement formée (Swarup et al., 2008). Une boucle de rétroaction positive est alors formée : le module SLR-ARF7/19 et LAX3, afin de contrôler l'expression de LAX3 dans les cellules proches des primordia, ce qui restreint les modifications de paroi cellulaire sur les cellules en contact avec le primordium émergent.

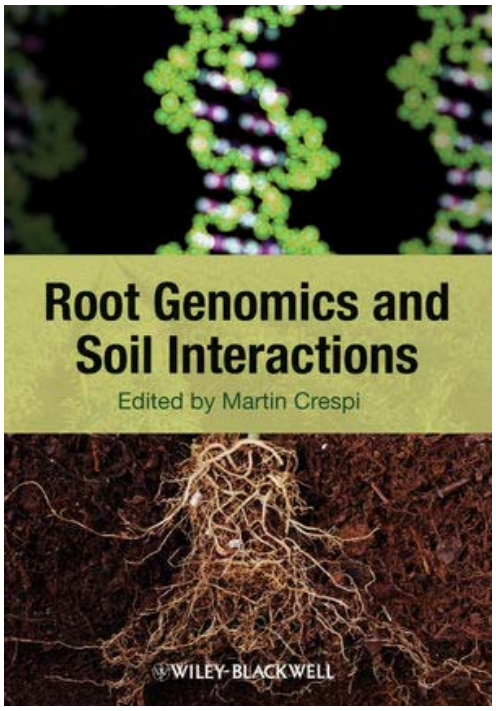
Après l'émergence, le méristème de la nouvelle racine est activé ; dans certains cas, la croissance de la racine latérale peut dépasser celle de la racine principale. Le moment exact où le nouveau méristème s'active n'est pas clairement défini. Des travaux, obtenus après culture de primordia isolés, indiquent qu'un méristème est fonctionnel et autonome quand le primordium auquel il appartient possède une épaisseur de 3-5 couches cellulaires (Laskowski et al., 1995). De façon intéressante, ce stade de développement correspond au moment ou le méristème de la racine latérale nouvellement formée commence à synthétiser de l'auxine (Ljung et al., 2005). La capacité à synthétiser de l'auxine pourrait donc coïncider avec l'activation méristématique.

Outre les mécanismes moléculaires décrits ici, directement liés à l'auxine et régulant le développement racinaire, plusieurs autres mécanismes de régulation existent. Ces dernières années, une nouvelle famille de régulateurs de l'expression des gènes impliquée dans la modulation de l'architecture racinaire à été mis en lumière: la famille des ARN non codant pour des protéines (npcARN). Plusieurs de ces ARN non-codants régulent de manière post-transcriptionnelle l'expression des ARFs mais également d'autres régulateurs de l'action des auxines.

1.2 Les ARN non-codants régulateurs de l'architecture racinaire

Lors de ma thèse j'ai eu l'occasion de participer à la rédaction du chapitre n°2 "The complex eukaryotic transcriptome: non-protein-coding RNAs and root dévelopement" (Ariel et al., 2012) inclut dans le livre "Root genomics and soil interactions". Dans ce chapitre, nous avons rapporté les données récentes illustrant l'action des ARN non-codants sur le développement de la racine. J'ai contribué à la rédaction de ce chapitre, notamment en décrivant le rôle des long npcARN dans la régulation de l'architecture racinaire et les « RNA binding protéines » reliées à l'action et au métabolisme des npcARN, tel que celui du npcARN *ENOD40*. De plus, j'ai réalisé la figure de ce chapitre ainsi qu'une table qui permet de référencer les différents micro-ARNs ou miRNAs en relation avec le développement de la racine.

J'ai placé dans cette introduction l'alinéa 3 du chapitre 2 du livre « Root Genomics and Soil Interactions » (Ariel et al., 2012). Cet alinéa décrit les principaux ARN non codant impliqués dans la régulation de l'architecture racinaire Le Chapitre du livre complet est présenté en annexe 1.



Aniléa 3 du chapitre 2 du livre « Root Genomics and Soil Interactions » (Ariel et al., 2012 ; annexe 1).

➔ Non protein coding RNAs in root development

In this section, we will discuss more specific npcRNAs linked to root development. The root architecture of the plant constitutes an important model to study how developmental plasticity is translated into growth responses under stress conditions. Indeed, primary root development and the formation of de novo meristems to generate lateral roots (LRs) are conditioned by the soil environment (Osmont et al. 2007). The remarkable developmental plasticity called the attention of Charles Darwin and his son Francis. In their monograph on the Power of Movements they referred to the behavior (i.e. gravitropism) of the growing root, postulating that the root tip acts like a plant brain: “We believe that there is no structure in plants more wonderful, as far as its functions are concerned, than the tip of the radicle. It is hardly an exaggeration to say that the tip of the radicle thus endowed, and having the power of directing the movements of the adjoining parts, acts like the brain of the lower animals” (cited by Kutschera and Niklas 2009).

Several of the best described riboregulators in plant biology, the miRNAs, have been linked to root development, such as the control of primary root growth or the formation of organs from de novo meristems, such as lateral and adventitious roots, or the legume-specific nitrogen-fixing nodules, through symbiotic interactions with soil bacteria (Khan et al. 2011).

External cues influence plant growth by modulating hormone levels and signaling. Auxin is one of the main phytohormones regulating root growth and architecture. Auxin activity is mediated by the Auxin Response Factor (ARF) genes, a plant-specific family of transcriptional regulators (Okushima et al. 2005). TAS3 is a trans-acting siRNA, whose biogenesis requires the initial miR390-mediated cleavage of the TAS3 precursor (Marin et al. 2010). The cleavage product is then converted to double-stranded RNA through the RdR6 activity and sequential DCL4-mediated cleavage events (Peragine et al. 2004; Vazquez et al. 2004; Allen et al. 2005; Gasciolli et al. 2005; Xie et al. 2005; Yoshikawa et al. 2005; Adenot et al. 2006). Of the four tasiRNAs precursors identified (TAS1-4) in Arabidopsis, cleavage of TAS3 is unique since it requires the specific action of the miR390/AGO7 complex for ta-siRNA production (Montgomery et al. 2008). These tasiRNAs inhibit ARF2, ARF3, and ARF4, thus releasing repression of lateral root growth (Marin et al. 2010). In addition, ARF2, ARF3 and ARF4 affect auxin-induced miR390 accumulation. Positive and negative feedback regulation of miR390 by ARF2, ARF3, and ARF4 thus ensures the proper definition of the

miR390 expression pattern resulting in the adaptation of the root system architecture (RSA) by auxin. In Arabidopsis, it was shown that auxin-associated miRNAs tightly control adventitious root formation through a complex regulation that involved various ARFs (Gutierrez et al. 2009). Whereas miR160 positively regulates adventitious root formation by controlling ARF17, the auxin-related miR167, which targets ARF6 and 8, is a negative regulator of adventitious root development. In Arabidopsis, ARF8 and ARF17 play antagonist roles in auxin homeostasis (Tian et al. 2004).

Other molecular mechanisms involved in root plasticity in response to the environment and endogenous signals are the natural antisense RNA (NATs), which are transcripts complementary to cis or trans-mRNAs and exert a repressive activity on them. Considering that NATs may encode proteins, some of them can be classified as dual RNAs because of their double function: both NAT and mRNA (Bardou et al. 2011). NATs have been described in several organisms like yeast, human, mice and plants. We can distinguish two different classes of NATs: cis-NATs, which are generated by antisense transcription at the same genomic locus, and trans-NATs, which are generated from different loci. Interestingly, a large number of transcripts were predicted to have both cis- and trans-NATs, suggesting that antisense transcripts can form a complex regulatory network (Henz et al. 2007). In Arabidopsis, a cis-NAT pair encoding SRO5 and P5CDH, was shown to repress translation, and improving Arabidopsis salt tolerance (Borsani et al. 2005). High salt concentration is toxic for plants when up-taken from the soil by the root (Ariel et al. 2010), strongly affecting root metabolism and development. In response, plants have developed biological mechanisms that prevent NaCl accumulation or absorption (Munns et al. 2010). P5CDH is constitutively expressed and encodes the D1-pyrroline-5-carboxylate dehydrogenase, which prevents proline accumulation, whereas SRO5 is induced by salt stress and encodes an unknown protein. Under high levels of NaCl, both genes form a natural double-stranded pair of transcripts proved to be cleaved by DCL2 and DCL1 generating 24- and 21-nucleotide nat-siRNAs that will repress the constitutively expressed P5CDH mRNA and lead to increased salt tolerance (Borsani et al. 2005).

In many cases, the function or coding capacity of NAT-RNAs remain largely unknown. PHO genes participate in phosphate (Pi) transport in the cellular response to Pi starvation in plants. Phosphate is normally up-taken from the soil and is essential for plant development. In rice, three PHO1 genes can form a cis-NAT pair with npcRNAs. Surprisingly, the cis-NAT associated with the OsPHO1-3 gene does not appear to be regulated either developmentally or in response to Pi deficiency, in contrast to the cis-NATs of OsPHO1-1 and OsPHO1-2, which are strongly up-regulated by Pi starvation, while the expression of the complementary sense transcript remains relatively stable (Secco et al. 2010). However, considering that the analysis was performed using whole roots, a

hypothesis is that the expression of the sense and antisense OsPHO1 transcripts may not occur in the same cells or tissues in all cases. In Arabidopsis, a search for npcRNAs (Ben Amor et al. 2009) allowed to identify 13 antisense npcRNAs complementary to protein-coding transcripts. One of these (npc536) forms a cis-NAT with AT1G67930, and its over-expression allowed plants to grow under salt stress without modifying AT1G67930 mRNA accumulation. Furthermore, npc536 mutants do not show any misregulation of the antisense transcript. As this NAT contains a short open reading frame (ORF) conserved in rice, npc536 may act through this encoded peptide. Alternatively, npc536 may regulate translation of the AT1G67930 mRNA or act as a trans-NAT, with an unidentified complementary target that plays a role in the salt stress response.

Apart from npcRNAs that lead to small RNAs such as the TAS or the NAT genes, only few npcRNAs have been implicated in root developmental processes (Charon et al. 2010). The ENOD40 genes code for highly structured plant mRNAs that contain a series of short ORFs without any long ORF (Charon et al. 1999; Gultyaev et al. 2007) and are involved in legume-specific root nodule organogenesis. Root nodules are nitrogen-fixing symbiotic plant organs that result from the interaction of soil bacteria of the genus Rhizobium with the root cells of host legumes (Oldroyd and Downie 2008; Crespi et al. 2008). This process initiates with cell-specific division in the roots, where ENOD40 is strongly expressed (Yang et al. 1993; Crespi et al. 1994). The ENOD40 gene is characterized by specific conserved nucleotide sequences that can be also found in some non-leguminous plants (Gultyaev et al. 2007). Furthermore, transgenic lines with increased or decreased levels of ENOD40 exhibit accelerated or reduced nodulation, respectively (Charon et al. 1999). It was first proposed that ENOD40 was an npcRNA due to its highly stable RNA secondary structure, a characteristic of known npcRNAs (Crespi et al. 1994; Hofacker et al. 2002) however other authors proposed that this transcript may encode a small primary oligopeptide of around 13 aminoacids (Charon et al. 2010). Translational analysis identified two short ORFs (sORF I and II; 13 and 27 amino acids long, respectively) that could be translated from this transcript in Medicago truncatula (Sousa et al. 2001) and sORF I contains a conserved nucleotide region across legumes, but not other plants, in contrast to the highly conserved stem-loops of the ENOD40 RNA throughout all known plant species (Girard et al. 2003). A cell-specific assay for the action of ENOD40 in Medicago sativa was developed using a biolistic process and suggested that translation of these sORF may be biologically relevant. Interestingly, mutations in the predicted structured RNA region also strongly inhibited this biological activity (Sousa et al. 2001). These results confirm the importance of both the sORF peptides and the RNA secondary structure of ENOD40 in its activity and suggest that ENOD40 encodes a bi-functional or dual RNA. To gain further insight into the action of ENOD40, molecules that interact with the peptides or RNA were identified. A novel RNA-

binding protein MtRBP1 (for M. truncatula RNA-Binding Protein 1), which interacts with the ENOD40 RNA, was identified using a yeast three-hybrid screen. Immunolocalization studies and the use of an MtRBP1-DsRed2 fusion construct showed that MtRBP1 localized to nuclear "speckles," which are nuclear ribonucleoprotein complexes known to house the splicing machinery in plant cells (Cioce and Lamond, 2005; Handwerger and Gall, 2006; Li et al. 2006). These nuclear speckles (or inter- chromatin granule clusters) are spotted shapeless structures containing elevated concentrations of splicing snRNPs and other splicing-related proteins that participate in the co-transcriptional splicing of mRNAs at the chromosomes (Shaw and Brown, 2004). Interestingly, MtRBP1 was located in the cytoplasm of ENOD40 expressing cells in M. truncatula nodules. The direct involvement of the ENOD40 RNA in MtRBP1 relocalization into cytoplasmic granules was confirmed using a transient expression assay and an MS2 bacteriophage system to tag the ENOD40 RNA (Campalans et al. 2004). This in vivo approach to monitor RNA-protein interactions demonstrated that the cytoplasmic relocalization of MtRBP1 was mediated by ENOD40 and suggested that the relocalization of nuclear RNA-binding proteins during specific developmental processes could be a new function mediated by npcRNAs (Campalans et al. 2004). On the other hand, the ENOD40 peptides expressed in soybeans were shown to bind to sucrose synthase (SUC1) suggesting a potential role of these peptides in the regulation of sucrose utilization in the nodules (Rohrig et al. 2002). These results further highlight that npcRNA genes may act as bi-functional RNAs in plants as many genes contain potentially active sORF-encoded peptides. In Arabidopsis, more than 3000 sORFs are transcribed, suggesting that huge numbers of sORF-encoded peptides are still hidden in genomic regions that have not been annotated yet (Hanada et al. 2007). For example, the POLARIS (PLS) gene in Arabidopsis was identified experimentally using a promoter-trap approach and shown to have a root-specific expression pattern (Topping et al. 1997). pls mutant plants have short roots with radially expanded cells and reduced leaf vascularization (Casson et al. 2002). The PLS gene is transcribed as a relatively short, 500-nucleotide mRNA, which contains three short ORFs that encode putative peptides of 8, 9 and 36 amino acids. Over-expression of the ORF encoding the 36-amino acid peptide partially rescues the short-root phenotype. Although the function of PLS has not been fully elucidated, a role in hormonal homeostasis, including ethylene signaling and auxin transport, and in the regulation of microtubule cytoskeletal dynamics was proposed (Chilley et al. 2006). These dual RNA further highlights the large variety of mechanisms that npcRNAs may trigger in root tissues.

2- Les Protéines de liaison à l'ARN:

Chez les eucaryotes, les protéines de liaison à l'ARN (RBPs pour RNA-binding proteins) ont un rôle crucial dans la régulation post-transcriptionnelle des gènes. Elles gouvernent divers processus développementaux en modulant l'expression des transcrits spécifiques. Le génome d'*Arabidopsis* code plus de 200 RBPs différentes, dont la plupart sont spécifiques des plantes (Lorkovic, 2009). Les études concernant des RBPs de plantes ont clairement montré qu'en plus de leur rôle dans divers processus de développement, elles sont aussi impliquées dans l'adaptation des plantes aux diverses conditions environnementales (Ambrosone et al., 2012). Les RBPs agissent en régulant l'épissage du pré-ARNm, la mise en place de la coiffe, la polyadénylation, la stabilité et l'exportation des ARN. Plus récemment, il a été démontré qu'elles peuvent aussi intervenir sur les modifications de la chromatine (He et al., 2011) et qu'elles sont des protéines clefs des mécanismes liés au silencing *via* les petits ARN (Xie and Qi, 2008).

La vie d'un ARNm commence par la transcription du gène par l'ARN polymérase II (RNAPol II). De manière concomitante, le recrutement de plusieurs machineries sur les transcrits naissants (ou pre-ARNm) permet de produire l'ARNm fonctionnel qui est prêt pour l'exportation du noyau puis la traduction. Les pré-ARNm ont une coiffe (composée d'une guanosine méthylée en position N7) à l'extrémité 5 ', des introns qui sont éliminés par le processus d'épissage ; Les pré-ARNm sont clivés et polyadénylés à l'extrémité 3'. La mise en œuvre correcte de ces modifications affecte les événements en aval, comme l'exportation de l'ARNm, la stabilité de celui-ci et sa traduction (Guthrie and Steitz, 2005 ; Lejeune and Maquat, 2005). Les transcrits de l'ARN Pol II sont recouverts par des RBPs dès le début de la transcription jusqu'à leur dégradation dans le cytoplasme (Aguilera, 2005 ; Dreyfuss et al., 2002 ; Moore, 2005, Lorković et al., 2009).

Les RBPs sont caractérisées par la présence d'un ou plusieurs domaines de liaison à l'ARN ; le motif le plus représenté, chez Arabidopsis, est le motif de reconnaissance de l'ARN RRM (RNA Recognition Motif) et le domaine KH (K-homologie domain). Les protéines qui possèdent un motif RRM représentent 70% des RBPs de plantes. Le RRM est un domaine de 80-90 acides aminés formé de quatre brins β en orientation antiparallèle et de deux hélices α dans l'ordre β1α1β2β3α2β4. Les brins centraux, β1 et β3 abritent un ou deux motifs conservés : un octamère nommé RNP1 et un hexamère nommé RNP2, certains RRM ne possèdent qu'un des deux motifs RNP.

Les domaines de liaison à l'ARN sont souvent combinés dans une même protéine ou bien présents avec d'autres domaines protéiques impliqués dans l'interaction protéine-protéine (Dreyfuss et al., 2002 ; Lorkovic and Barta, 2002). Fait intéressant, de nombreux

gènes codant pour des RBPs sont exprimés de façon différentielle en fonction du type cellulaire ou du stade de développement (Licatalosi and Darnell, 2006 ; Terzi and Simpson, 2008 ; Ule and Darnell, 2006). En raison de l'absence de systèmes *in vitro* d'origine végétale pour étudier les événements post-transcriptionnels, comme l'épissage du pré-ARNm et la polyadénylation, les RBPs ont été moins étudiées chez les plantes, que chez d'autres organismes. Toutefois, des travaux concernant des RBPs de plantes sont en train d'émerger ; elles sont de plus en plus étudiées notamment lors de la régulation de la transition florale (Lim et al., 2004 ; Macknight et al., 1997 ; Mockler et al., 2004 ; Schomburg et al., 2001 ; Streitner et al., 2008), la signalisation par l'acide abscissique (ABA) (Cao et al., 2006 ; Riera et al., 2006), la réponses au stress (Kim and Kang, 2006 ; Kim et al., 2007 ; Kim et al., 2010 ; Kwak et al., 2005), le cycle circadien (Staiger, 2002 ; Zhao et al., 2004) et la modification de la chromatine (Baurle et al., 2007 ; Liu et al., 2007 ; Baurle and Dean, 2008 ; Veley and Michaels, 2008).

L'assemblage des complexes RNPs a donc une importance cruciale pour comprendre la régulation de l'expression des gènes. Les interactions entre les RBPs et l'ARN hautement structuré en complexe ribonucléoproteiques (RNPs) dans différents granules subcellulaire dans le noyau et le cytoplasme des cellules eucaryotes nous révèlent plusieurs aspects de la régulation de l'expression des gènes se déroulant dans le contexte des RNPs. Différents « granules à ARN » existent chez les cellules eucaryotes et je vais donc résumer leurs rôles putatifs dans l'organisation cellulaire. De plus, pour chaque compartiment subcellulaire les RBPs et les mécanismes qui leur sont associés seront décrits.

2.1 Les RBP dans les corps cytoplasmiques

La cellule eucaryote possède, hormis son principal organite le noyau, différent organites cytoplasmiques dont le réticulum endoplasmique, l'appareil de golgi, plusieurs vésicules d'endocytose ou de transports, les lysosomes et les peroxysomes.

En plus de ces organites, des granules associant des ARN et des protéines ont été caractérisés (Anderson and Kedersha, 2006). Ces compartiments sans membrane ou granules se forment dans le cytoplasme grâce à des interactions ARN/protéines. Ils semblent être le siège de régulations fines lors de la traduction et la dégradation des transcrits d'ARN. Généralement, L'ARNm organisé en granules RNP permet la régulation précise de sa traduction dans de larges complexes de ribosomes ou polysomes. Ces granules sont essentiellement composées d'ARN, de ribosomes, de facteur d'initiation de la traduction (eIF4E eIF2α) ainsi que d'autres RBPs. Après avoir été traduit certain ARNm sont

libérés depuis les polysomes et s'agrégent pour former des structures dynamiques appelées stress granules (SG) et processing bodies (P-bodies).

2.2.1 Les RBP dans les « Processing bodies »

Les « P-bodies » sont des agrégats de RNP contenant des ARNm dont la traduction est bloquée, un processus lié à la dégradation de l'ARNm. Les P-bodies contiennent des enzymes nécessaires à la voies de dégradation des ARNm comme des déadénylases (CCR4/POP2/NOT ; Sheth et Parker, 2003; Cougot et al, 2004 ; Andrei et al, 2005), le complexe enzymatique « decapping » comportant DCP1, DCP2 et, dans les cellules de mammifères, Helds, hEdc3, l'activateur du décaping p54/RCK (Fenger-Gron et al, 2005 ; Parker et Sheth, 2007) et l'exonucléase XRN1. Chez les plantes, la composante centrale du complexe de decapping est constituée d'au moins DCP1, DCP2 et VARICOSE (Xu et al, 2006 ; Goeres et al, 2007), dont les mutations sont létales au stade cotylédon. Les P-bodies contiennent également un heptamère LSM1-7 qui régule divers aspects de l'assemblage des RNPs (Ingelfinger et al, 2002 ; Sheth et Parker, 2003) et des éléments de la voie du NMD (Non sense Mediated Decay ; SMG5, SMG7 et UPF1 ; Unterholzner et Izaurralde, 2004 ; Fukuhara et al, 2005). La taille et le nombre de P-bodies sont généralement proportionnels au pool des ARNm non traduits. Lorsque les ARNm sont associés à des ribosomes bloqués, par un traitement à la cycloheximide, le nombre de P-bodies diminue considérablement. Il a été proposé qu'un flux de transcrits fixés sur des ribosomes étaient indispensables pour maintenir la structure en granules (Sheth et Parker, 2003 ; Cougot et al, 2004 ; Teixeira et al, 2005). A l'inverse, quand les ARNm sont dissociés des ribosomes en raison de l'inhibition de la traduction, la taille des P-bodies augmente (Kedersha et al, 2005 ; Teixeira et al, 2005 ; Koritzinsky et al, 2006). Les P-bodies et les granules de stress sont des granules ribonucléoprotéiques dynamiques qui interagissent entre eux. A l'intérieur, les ARNm peuvent y être dégradés, stockés ou re-rentrer dans un processus de traduction. En conséquence, la cellule peut rapidement réguler la traduction, en réponse à l'environnement, en utilisant plusieurs mécanismes de remobilisation des ARNm présents dans les compartiments nucléaires et cytoplasmiques (Kedersha et al, 2005 ; Wilczynska et al, 2005 ; Decker and Parker., 2012 ; Muench et al., 2012).

Durant ma thèse j'ai participé à des manipulations et à la rédaction d'un article intitulé « Cytoplasmic and nuclear quality control and turnover of single stranded RNA modulate post-transcriptional gene silencing in plants » dans lequel je suis troisième auteur et qui sera présenté en annexe 2 (Nucleic Acid research ; Moreno et al., 2013). Dans cet article nous avons prouvé que des sous unités du NMD, de la déadenylation ou de l'exosome suppriment le Sense-Post Transcriptional Gene Silencing (S-PTGS; cf introduction partie résultats 2 en relation avec le PTGS). De plus ces résultats indiquent que le RNA Quality Control (RQC)

ainsi que le PTGS sont entrelacés. En effet, des membres du RQC sont localisés dans le noyau et le cytoplasme et, dans la partie cytoplasmique ces protéines sont localisées dans deux types de granules : Les P-bodies (marqué par la protéine DCP1) ainsi que les siRNA-bodies (Corps cytoplasmiques associé à la localisation de protéines comme SGS3 et RDR6). Nous avons montré que ces deux types de granules cytoplasmiques ne colocalisaient pas. En réalité, ils sembleraient que les p-bodies et les si-RNA bodies peuvent entrer en interaction probablement afin d'échanger des composants entre ces deux corps cytoplasmiques impliqués dans la dégradation des ARN. Dans cet article j'ai principalement aidé aux clonages de ceratins gènes, j'ai aussi étudié la localisation de la protéine RRP4.

2.2.2 Les RBP dans les « silencing bodies »

Chez les plantes, certains petits ARN régulateurs (« small RNAs ») sont générés à partir de long dsARN synthétisés par la RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) assistée dans cette action par SGS3, une protéine de liaison à l'ARN (Peragine et al., 2004 ; Vazquez et al, 2004 ; Allen et al., 2005 ; Yoshikawa et al., 2005). RDR6 est une enzyme clef de la biosynthèse des dsARN impliqués dans les mécanismes régissant la S-PTGS ou l'interférence par des trans-acting small-interfering RNA (tasi-ARN). En revanche la fonction de la protéine SGS3 (qui est spécifique des plantes) reste obscure. La protéine SGS3 est composée de 3 domaines un domaine à doigt de zinc, un domaine XS ainsi qu'un domaine coiled-coil. Le domaine XS est impliqué dans la liaison aux ARN et le domaine coiled-coil dans l'homodimérisation de SGS3 (Elmayan et at., 2009 ; Fukunaga and Doudna, 2009). Les protéines SGS3 et RDR6 co-localisent dans les mêmes granules cytoplasmiques nommés « silencing bodies » et des résultats obtenus par BiFC suggèrent que ces deux protéines interagissent entre elles (Kumakura et al., 2009)., Les 3 domaines de SGS3 semblent importants pour la fonction de cette protéine. En revanche, seulement les domaines XS et coiled-coil sont importants pour permettre la localisation de SGS3 dans des corps cytoplasmiques ou siRNA bodies (Kumakura et al., 2009). Il est important de noter que ces granules cytoplasmiques sont distincts des P-bodies. Une étude récente de notre laboratoire (Jouannet et al. 2012) indique que la protéine AGO7 colocalise avec SGS3 et RDR6 dans les « silencing bodies ou siRNA bodies » compartiments sans membrane du cytoplasme. La protéine AGO7 intervient avec le miARN390 dans la formation de tasiRNA (ou trans acting siRNAs) à partir du long ARN non-codant TAS3 et grâce à l'intervention des protéines SGS3 et RDR6 (Vaucheret, 2005 ; Marin et al., 2010).

2.2.3 Les RBP dans les granules de stress

Les granules de stress (SGs) constituent des ensembles cytoplasmiques se formant dans les cellules eucaryotes exposées à un stress environnemental (UV, l'hypoxie, l'infection virale ...). Les premières descriptions de SGs ont été réalisées à partir de cultures cellulaires de tomate soumises à un choc thermique qui montraient des agrégats de petites protéines de choc thermique (HSP) dans leur cytoplasme (Nover et al., 1983). Plus tard, ces corps d'inclusions que l'on appelle granules de choc thermique (HSG) ont été observés chez la drosophile (Leicht et al, 1986 ; Arrigo et Welch, 1987), les vertébrés (Collier et Schlesinger, 1986; Arrigo et Welch, 1987) et chez les végétaux (Nover et al, 1983 ; Nover et Scharf, 1984). Les complexes de pré-initiation 48S constituent les composants centraux des SGs avec l'ARNm, les petites sous-unités ribosomales, les facteurs d'initiation de traduction (eIF3, eIF4E, eIF4G) et la PolyA-Binding Protein (PABP) (Kedersha et al., 2002). En outre, plusieurs protéines liant l'ARN et qui régulent la traduction ou la dégradation de l'ARNm font partie de ces SGs (Kim et al, 2005 ; Hofmann et al, 2006 ; Decker and Parker., 2012 ; Muench et al., 2012). Dans les cellules de mammifères, la formation de SG dans des conditions de stress nécessite la fixation de la RBP TIA-1 porteuse d'un domaine prion sur des ARNm polyadénylés et la présence de facteurs d'initiation de traduction (Gilks et al, 2004; Kedersha et al, 2000). Chez *Arabidopsis*, deux RBPs nucléaires (UBP1 et RBP47), dont la structure moléculaire est proche de la protéine TIA-1 des mammifères, sont impliqués dans l'épissage des pré-ARN (Lambermon et al, 2000; Lorkovic et al, 2000 ; Weber et al., 2008). Ces protéines possèdent un domaine « glutamine-rich prion-related domain (PRD)» (Gilks et al., 2004; Lorkovic and Barta, 2002), trois motifs de reconnaissance à l'ARN (RRM) et présentent une forte homologie de séquence avec la protéine TIA-1. En absence de stress, la protéine RBP47 étiquetée par une protéine fluorescente est localisée dans le noyau. En revanche, cette protéine est capable de re-localiser vers des granules cytoplasmiques lors d'uns stress thermique (Weber et al., 2008). Des immunoprécipitations avec des anticorps monoclonaux, ont montré que RBP47 semble fixée sur des ARNm polyadénylés. Enfin des résultats similaires ont été obtenus avec la protéine UBP1. Ces résultats montrent que les ARNm polyA, les facteurs d'initiation de la traduction et UBP1/RBP47 sont recrutés ensemble dans des SGs.

2.2 Les RBPs dans les corps nucléaires

Le noyau constitue le lieu de stockage du matériel génétique, il est entouré par une enveloppe composée d'une double membrane lipidique séparée par un espace périnucléaire en continuité avec le réticulum endoplasmique (Stewart et al., 2007). La communication

entre le noyau et le cytoplasme se déroule à travers des pores. Dans le noyau, chaque chromosome est localisé dans un territoire spécifique séparé par un compartiment interchromosomique où sont retrouvés les différents corps nucléaires. Au sein des territoires, les chromosomes sont hautement structurés avec des régions d'hétérochromatine condensée et des régions d'euchromatine ouverte qui connaissent une activité transcriptionnelle intense. Grâce au progrès dans le domaine de la microscopie, le noyau se dévoile comme une structure hétérogène composée de plusieurs sous-compartiments (Spector, 2011) (figure 4). Les corps nucléaires tel que, le nucléole, les Cajal bodies, les speckles, et les Dicer-bodies sont des domaines sub-nucléaires sans membrane. Les corps nucléaires compartimentent l'espace nucléaire et créent des sites distincts dans un volume confiné, concentrant ainsi réactifs et substrats afin de potentiellement faciliter les réactions biologiques. Des études récentes commencent à élucider les mécanismes moléculaires responsables de l'assemblage et du maintien de plusieurs de ces corps nucléaires (Kaiser et al., 2008 ; Mao et al., 2011 ; Shevtsov and Dundr, 2011). L'absence de membranes favorise les échanges entre ces corps nucléaires et l'environnement du nucléoplasme. En effet, la plupart des constituants protéiques des corps nucléaires sont également dispersés dans l'espace interchromatinien à des concentrations plus faibles. L'organisation des corps nucléaires est très finement régulée ; les interactions entre les protéines, les ARN et les RNPs sont responsables de la formation, du maintien, ainsi que de l'activité biologique de ces différents corps. De plus, il est maintenant admis que les ARN codants et non codants ont une capacité de régulation de l'architecture nucléaire ainsi que, dans certains cas, de la structure même de ces corps (Dundr and Misteli 2010 ; Caudron-Herger et al., 2011).

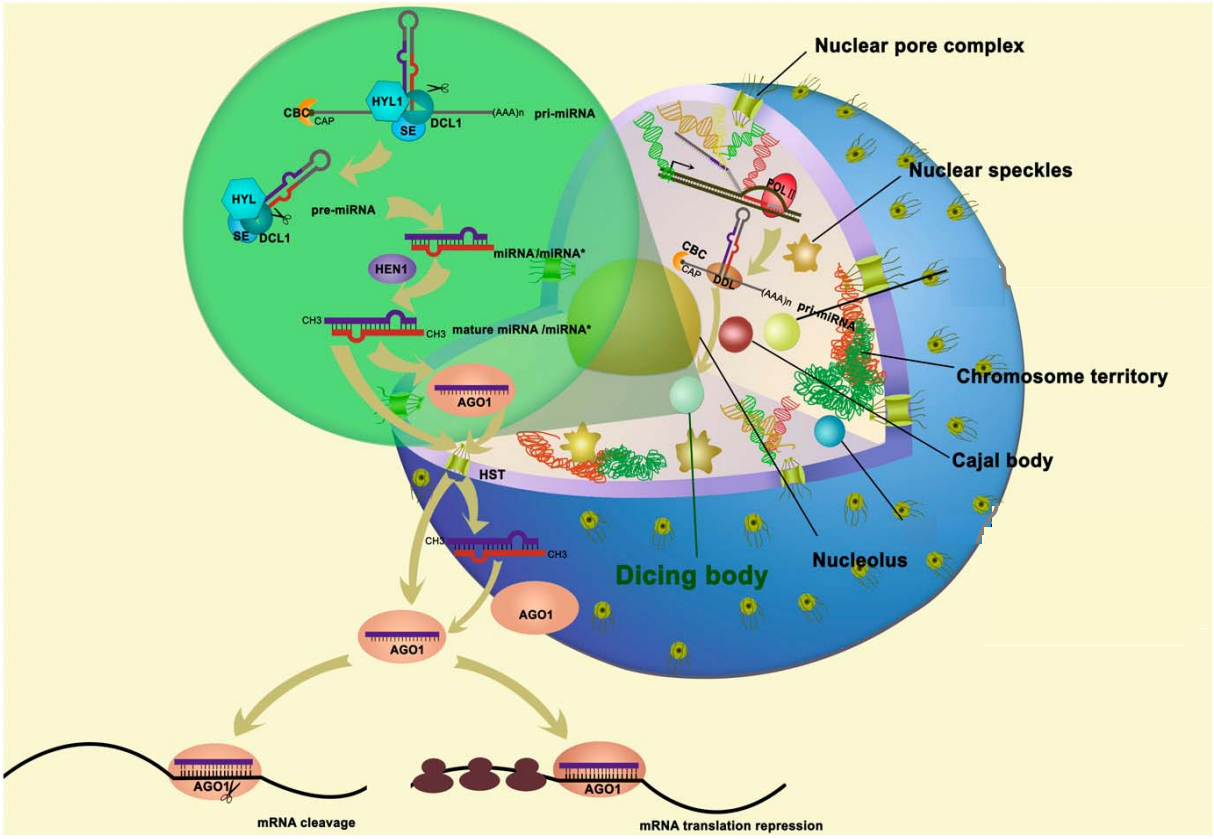


Figure 4 Vue d'ensemble des domaines nucléaires

Dans les cellules eucaryotes, le noyau est encapsulé dans deux couches de membranes dans lesquelles des complexes de pores nucléaires permettent le transport entre le noyau et le cytoplasme. Dans le noyau, les chromosomes sont organisés en territoires chromosomiques. La région interchromatinienne est très hétérogène, contenant divers domaines nucléaires. Dans une cellule végétale, ces corps nucléaires comprennent le nucléole, les CB, les speckles nucléaires et les D-bodies. Ces corps ont des tailles, des formes, des composants, une dynamique et des fonctions différentes. Les D-bodies jouent un rôle dans la biogenèse des miARN. Les gènes de miARN de plantes sont transcrits par l'ARN polymérase II afin de générer des pri-miARN. Les principales enzymes pour la biogenese et les mecanismes liés aux miARN sont représentées. (Liu et al., 2012).

2.1.1 Le nucléole

Le plus volumineux et le mieux caractérisé des compartiments nucléaires est le nucléole, qui est une structure dynamique formée autour de l'activité de transcription des ARNr par l'ARN Pol I (Figure 4, 5). Le nucléole est composé de 3 parties principales, un centre fibreux, qui contient des centaines de gènes codants des ARNr qu'on retrouve à plusieurs loci (aussi nommé région nucléolaire organisatrice) ; une composante dense et fibreuse, qui contient les gènes codant les ARNr en transcription et les ARNr naissant, et une composante granuleuse qui est le site de traitement post transcriptionnel des ARNr (Carmo Fonseca et al. 2000) (figure 4). Dans le nucléole des plantes, la composante dense et fibreuse semble moins dense et occupe la majeure partie du volume du nucléole (Shaw and Brown 2004). Le nucléole est donc le siège de transcription des ARNr mais également le site de modification des pre-ARNr par des « small nucleolar RNA » (snoRNA ; Brown et al., 2003) et de pre-assemblage des ribosomes (Olson et al, 2002, Pederson, 2010). Le nucléole

est aussi impliqué dans la biogenèse de RNP non ribosomale comme les « Signal Recognition Particles » ou SRPs (Handweger et al ; 2005 ; Sommerville et al ; 2005).

Les autres corps nucléaires sont moins bien caractérisés ; je vais résumer dans le chapitre suivant les connaissances actuelles concernant les « speckles » (Spector et al., 2010), les « Cajal bodies » (CB) (Mao et al., 2011) et les « Dicer bodies (D-bodies) » (Fang and Spector, 2007).

2.1.2 Les RBP dans les « Cajal bodies »

Les « Cajal bodies » (CB) sont généralement associés au nucléole avec lequel il partagent de nombreux constituants (Figure 4, 5, 6). Ces particules contiennent une variété d'ARN et de protéines impliqués dans l'assemblage et la modification des small nuclear RNPs (snRNPs) et des small nucleolar RNPs (snoRNPs). Il est clairement établi que des RNPs se déplacent (Figure 5) des CB jusqu'au nucléole (snoRNPs) ou jusqu'aux speckles (snRNPs) (Seleeman and Lamond, 1999 ; Gerbi et al., 2003 ; Nizami et al., 2010). Chez l' homme, les snARN du spliceosome sont transcrits dans le noyau puis exportés dans le cytoplasme où les snARN forment un complexe avec sept protéines Sm conservées chez la plupart des eucaryotes. Ces complexes snRNP sont ensuite hyper-méthylé aux extrémités 5' puis réimportés dans le noyau. Ils se concentrent d'abord dans CBs pour se rendre ensuite dans les speckles. Ces derniers rejoignent les régions actives du génome où ils jouent un rôle essentiel dans l'épissage de pré-ARNm (Matera et al., 2007). Les snRNPs du spliceosome U1, U2, U4/U6 et U5 sont aussi localisés dans les CBs avant leur exportation dans les speckles dans les cellules humaines. Les snoRNP U3 et U8, impliqués dans le « Processing » des pre-ARNr sont aussi localisés dans les CB (Gall et al., 2000). Li et al., 2006 ont d'autre part montrés que AGO4 et NRPD1b colocalisent avec DCL3, RDR2 et les siRNAs dans les CBs (Pontes et al., 2006). Ces résultats suggèrent que les CBs seraient le lieu d'assemblage d'un complexe AGO4/NRPD1b/siRNAs qui régulerait l'état chromatinien de certains loci (Pontes et Pikaard, 2008). Finalement les CBs jouent aussi un rôle primordial dans l'assemblage des télomèrases RNP ainsi que dans l'homéostasie de la longueur des télomères. Cependant, chez *Arabidopsis* et chez *Drosophila melanogaster* une absence de CB due à la déficience d'une protéine ubiquiste des CBs, nommée COILIN, n'affecte pas la formation d'individus parfaitement viables et se développant normalement (Collier et al., 2006 ; Liu et al., 2009) à la différence des souris mutantes *coilin* qui montrent une létalité d'environ 50% (Walker et al., 2009). Les CBs jouent donc un rôle dans la promotion de certaines étapes finales de la maturation de snRNP et / ou facilitent l'interaction individuel des snRNPs pour former des complexes d'ordre supérieur (Nizami et al., 2010).

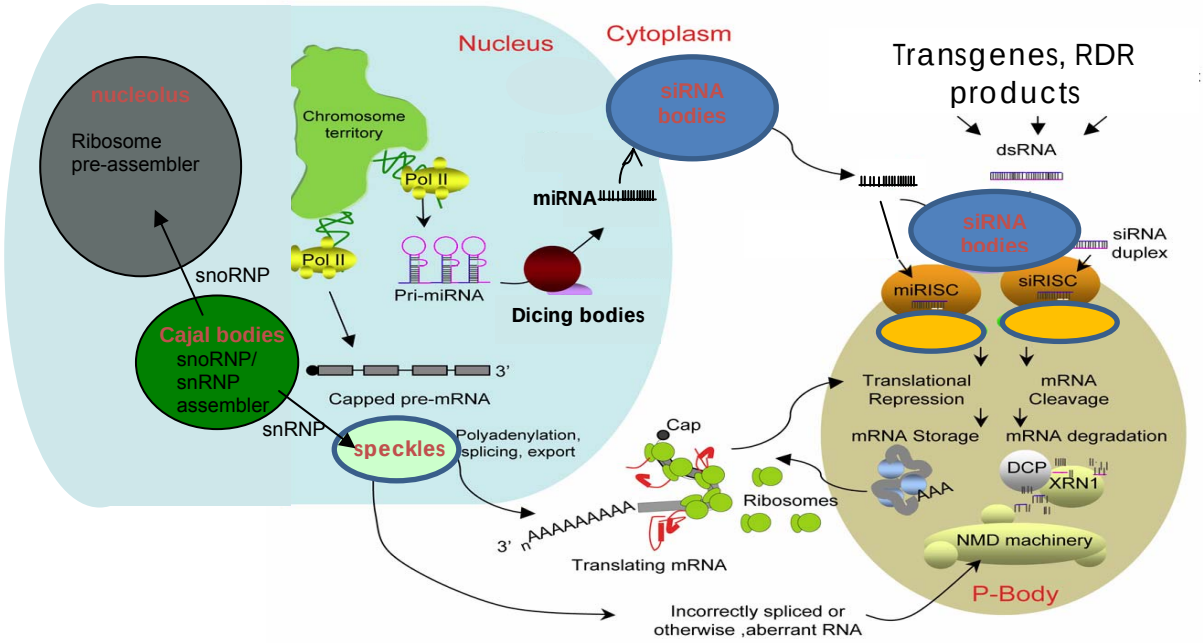


Figure 5 Flux d'ARN à travers les différents RNP bodies dans le noyau et le cytoplasme

Les ARNm sont transcrits dans le noyau puis sont maturés par mise en place d'une coiffe, d'une queue poly A et sont épissés dans le noyau. Les Cajal bodies permettent de redistribuer les protéines nécessaires à la maturation des ARN (vers les speckles) ainsi que les protéines impliquées dans l'assemblage des ribosomes (dans le nucléole). Les ARNm sont exportés vers le cytoplasme pour êtres traduits. Ils peuvent aussi êtres régulés par des npcARN comme les miARN (fabriqués dans les dicing bodies) ou les siARN dérivés de transgènes ou issus des produits des RDR. Les ARNm sont exportés dans le cytoplasme et clivés dans les siARN bodies via le complèxe RISC ou encore dégradés dans les P-bodies par les protéines DCP et XRN. Les P-bodies représentent aussi un lieu de stockage d'ARNm et abritent la machinerie NMD. (Adapté d'après Pontes and Pikard, 2008)

2.1.3 Les RBP dans les « Dicing bodies »

Chez Arabidopsis, le pre-miARN et le pri-miARN sont clivés dans le noyau par DCL1 (Papp et al., 2003). De nombreuses protéines de liaison aux ARN double brin (dsRBPs) participent à la biogenèse et la fonction de la voie métabolique des petits ARN (Vaucheret, 2006). La protéine HYL1 (Hyponastic Leaves 1) d'Arabidopsis est une dsRBP nucléaire impliquée dans la biogenèse des microARN (Lu and Fedoroff , 2000 ; Han et al., 2004). HYL1 comme SERRATE peuvent se lier à DCL1 (Hiraguri et al., 2005 ; Kurihara et al., 2006). Une mutation dans le gène HYL1 provoque une plus faible accumulation de miARN et une augmentation des ARNm cibles non clivées (Han et al., 2004). En utilisant des protéines de fusions fluorescentes, Song et al., 2007 ont observé une co-localisation entre les protéines DCL1 et HYL1 dans des corps nucléaires qui ressemblent à des CB. Ils ont par la suite été dénommés « Dicer bodies » ou D-Bodies car ils comportent l'ensemble des enzymes indispensables à la biogenèse des miARN et de certains siARN (Figure 4, 5). Fang and Spector 2007 ont également observé le recrutement de pri-miARN dans ces « Dicer

bodies. Les D-bodies seraient donc des corps de traitement des précurseurs des miARN ou un site de stockage et d'assemblage de la machinerie de la voie des miARN (Liu et al., 2011).

2.1.4 Les RBP dans les Speckles

La machinerie d'épissage du pre-ARNm qui inclut, les snRNPs, les sous-unités du spliceosome et les régulateurs de l'épissage, est internalisée dans des granules, les speckles nucléaires. Ces derniers sont localisés dans les canaux et les lacunes qui entourent les différents territoires chromosomiques dans le noyau. (Mintz and Spector, 2000 ; Spector 2001 ; Spector and Lamond 2010, Niedojadło et al., 2012) (Figure 4, 5, 6). Chez les plantes comme chez les animaux, ces petits corps nucléaires interchromatiniens sont le site de stockage des facteurs d'épissage et des protéines qui contiennent des motifs de liaison à l'ARN comme les protéines SR (serine/arginine rich) (Ali et al., 2003 ; Docquier et al., 2004; Fang et al., 2004 ; Lorkovic et al., 2004 ; Barta et al., 2008). Ces protéines SR sont nécessaires pour définir les jonctions exon/intron et pour l'assemblage du spliceosome (Barta et al., 2008). Chez les plantes, certaines protéines SR sont régulées par des stimuli environnementaux comme la variation de la températures par exemple (Ali et al., 2003). Chez *Arabidopsis,* des résultats récents montrent que certaines protéines SR en plus de leur localisation diffuse dans le nucléoplasme sont intégrées dans un réseau irrégulier qui semble correspondre aux « speckles ». Les protéines SR d'*Arabidopsis* sont classées en 7 sous-familles conservées. Les différentes protéines SR se distribuent dans des populations distinctes de « speckles » nucléaires ; cependant, il existe des cas de co-localisations partielles ou complètes de certains membres de la famille (Lorkovic et al., 2008). De plus, ces speckles nucléaires ont été montrées comme dynamiques en fonction de l'environnement de croissance d'une cellule particulière. Plusieurs kinases et phosphatases, qui peuvent respectivement phosphoryler et déphosphoryler des composants de la machinerie d'épissage, ont été trouvés dans les speckles, soutenant un rôle majeur de la modification post-traductionnelle des composants du spliceosome pour la régulation de l'épissage (Spector and Lamond 2011). Les speckles abritent donc le spliceosome qui est un grand complexe formé de nombreuses protéines et dont les interactions sont très dynamiques en fonction de l'activité cellulaire. Par conséquent, les speckles nucléaires ont été proposés comme servant de plate-forme pour les activités métaboliques impliquées dans la maturation et l'épissage des ARNm ainsi que dans leur exportation (Brown et al., 2006 ; Brown et al., 2008).

Les protéines du « spliceosome » ou les régulateurs de l'épissage comme les protéines hnRNP, snRNP, les protéines SR et les autres principales protéines liées à l'épissage seront présentées chapitre 3.2 de cette introduction.

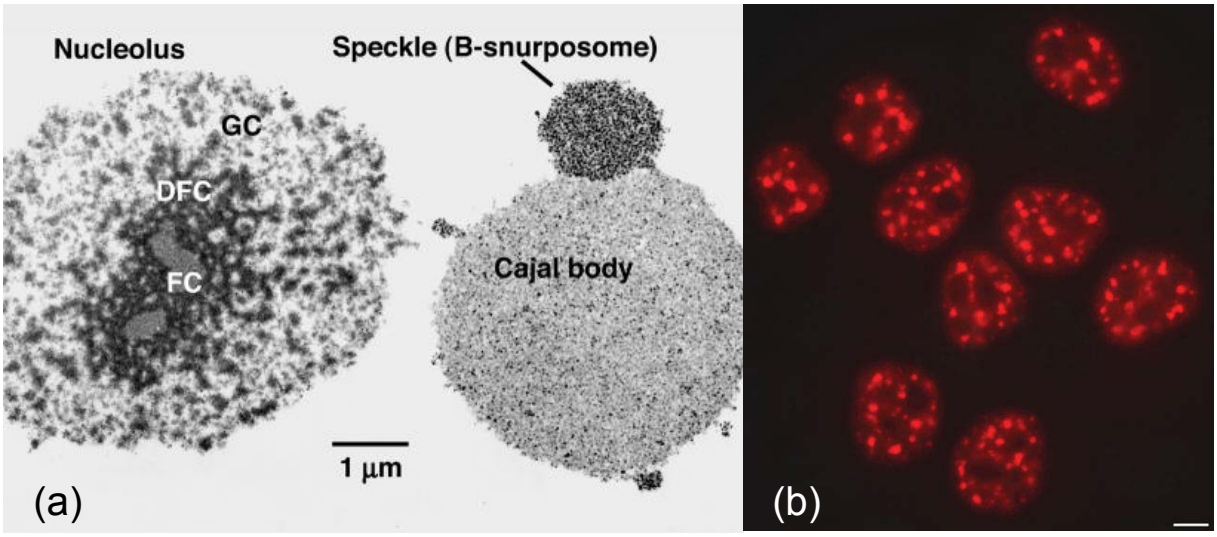


Figure 6 Détails structurels des speckles et d'autres organelles sub-nucléaires

(a) Détails structurels sub-nucléaire d'un noyau d'ovocyte de Xénope. Le Cajal bodie, le speckle et la partie granuleuse du nucléole contiennent des particules hétérogènes (CG granular compound ; DCF dense fibrilar component ; FC fibrilar center ; Bar 1µm) (Handwerger et al., 2005). (b) Speckles dans l'espace interchromatinien de noyau de cellule HeLa représentant des facteurs d'épissage. Bar 5µm (Spector and Lamond, 2011)

2.1.5 npcARN et RNA binding proteins dans le noyau

Lors de ma thèse j'ai eu l'opportunité de participer à l'écriture de deux autres revues qui traitent du rôle des RBPs en relation avec les npcARN et qui sont partiellement redondantes avec le chapitre de livre (Ariel et al., 2012, Annexe1) qui est plus focalisé sur la racine. Une première revue orientée ARN non codant dans le noyau sera présentée en annexe 3 « Non-Protein-Coding RNAs and their Interacting RNA-Binding Proteins in the Plant Cell Nucleus » (Charon et al., 2010). La seconde revue, « Dual RNAs in plants » (Bardou et al., 2011) sera proposée en annexe 4. Ici, je présente l'alinéa 2 du chapitre 2 du livre « Root Genomics and Soil Interactions » (Ariel et al., 2012 ; annexe 1) afin de réviser les connaissances qui lient les npcARN et les RBPs.

Alinéa 2 du chapitre 2 du livre « Root Genomics and Soil Interactions » (Ariel et al., 2012 ; annexe 1).

➔ The role of RNA binding proteins in npcRNA metabolism and activity

The post-transcriptional regulation of gene expression mediated by npcRNAs as well as all aspects of RNA metabolism is globally determined by a variety of RNA-binding proteins (RBPs; Lorkovic 2009). Most RBPs contain one or more conserved domains, as the RNA-recognition motif (RRM) and the K homology (KH) motif. Based on the characteristics of those conserved domains, the analysis of the A. thaliana genome revealed that plants express a complex set of RBPs, with 196 RRM- and 26 KH-containing proteins, the majority of them being plant-specific (Barta et al. 2008; Lorkovic and Barta 2002). Most of these proteins have not been characterized experimentally, and it is largely unclear how their action may control gene expression and development, primarily due to the difficulty in the identification of their RNA partners (Lorkovic 2009). Nevertheless, following forward and reverse genetic approaches, some studies begin to reveal the requirement of specific RBPs that have crucial roles in RNA metabolism during plant development.

Most RBPs are likely to have multiple RNA partners including mRNAs and npcRNAs (e.g. antisense RNAs, intergenic npcRNAs or small RNAs) that may integrate different RNPs (Ribonucleoproteins) to generate RNA networks in which npcRNAs can determine their localization, interfere or modulate their action (acting as competitors or activators against other substrates). Identification of the RBPs with which npcRNAs associate will contribute to understand their role in RNP networks in the cell. For example, several RBPs are involved in the biogenesis and action of small RNAs (e.g. DCLs, RDRs or AGOs, Valencia-Sanchez et al. 2006). The RNAi pathways have been largely diversified and several steps occur in the nucleus and/or the cytoplasm. The different small RNA silencing pathways differ mainly in the way of generation of the small RNA (Vaucheret et al. 2006). One RNA silencing mechanism is initiated by endogenous loci able to form double-stranded stem-loops that are processed by DICER (double-stranded RNAses) into small RNAs called miRNAs (miRNA pathway). In the siRNA pathway, a single-stranded RNA is targeted by RNA-dependent RNA polymerases to form long dsRNAs that then is cut by DICERs into small RNAs called siRNAs. In addition, long dsRNAs can be precursors of siRNAs when an npcRNA (a trans acting or TAS long non-coding RNAs) is targeted by a specific miRNA and cleaved, becoming a substrate of RDR polymerases to form a long dsRNA that will be processes in small siRNAs, called tasiRNAs (derived from TAS genes). Finally, two independent

transcripts may form a complementary dsRNA (natural antisense genes) and this region of the dsRNA molecule is processed by DICER into dsRNA small molecules or natsiRNAs. In all cases, one strand of the processed si/miRNAs duplex is incorporated into the RISC complex containing AGO proteins. The presence of this small RNA provides the RISC complex a sequence specificity to target a particular mRNA molecule. Gene Silencing can occur at both PostTranscriptional (PTGS, through mRNA cleavage and/or translation inhibition) and/or Transcriptional levels (TGS, through DNA methylation) (Vaucheret 2006).

In the cell there exist several sources of long npcRNAs that may lack of coding capacity because they are aberrant RNAs, e.g. mRNAs without either the 5' cap or the 3' polyA tail. Those aberrant RNAs can be converted in double stranded RNA by RDRs triggering silencing mechanisms through the action of DCLs or be eliminated by the normal degradation machinery, involving the exosome, XRN enzymes or the NMD pathway (Conti and Izaurralde 2005; Chekanova et al. 2007; Kurihara et al. 2009). These survey mechanisms conform what we previously defined as RQC mechinery and aim to ensure the correct expression of the genes (Houseley and Tollervey 2009).

One of those control survey mechanisms is the NMD, for non-sense-mediated mRNA decay, whose core RBPs are the UP-frameshift proteins (UPFs). This pathway recognizes premature nonsense or stop codons (PTC) within an mRNA by the action of the exon-junction complex (EJC) that marks correctly fused exons (Pontes and Pikaard 2008). Such PTC containing transcripts can arise as a result of genomic frameshifts, nonsense mutations or inefficiently spliced pre-mRNAs for example (Maquat 2004). In Arabidopsis there are three UPF genes (UPF1, UPF2 and UPF3), and a genome-wide analysis using a tiling array of the mutants upf1-1 and upf3-1 revealed that in addition to the expected NMD substrates, that is coding mRNAs, most npcRNAs, including large numbers of antisense RNAs, are degraded by this pathway (the tiling-array data can be viewed at http://omicspace.riken.jp/gps/group/psca3). In these mutants the percentage of up-regulated messenger-like-npcRNAs (mlnpcRNAs) to all expressed mlnpcRNAs was much higher than the percentage of up-regulated mRNAs to all expressed mRNAs. This suggests that in fact, besides the recognition of nonsense mRNAs, another very important role of NMD is the genome-wide suppression of the mlnpc-RNAs that are recognized as aberrant transcripts by this machinery (Kurihara et al. 2009). After recognition of an incorrectly positioned stop codon, the NMD system through the action of other RBPs, signals the elimination of this RNA by recruiting decapping enzymes followed by 5'→3' exonuclease activities (XRN enzymes), and/or deadenylation enzymes followed by 3'→5' exonucleolytic degradation (Amrani et al. 2006; Conti and Izaurralde 2005; Lejeune and Maquat 2005). This 3'→5' exonucleolytic

degradation is known to take place in the exosome complex. Hence, npcRNAs may affect the stability of other aberrant transcripts by interfering with this pathway.

The exosome is an evolutionary conserved macromolecular complex that mediates numerous reactions of 3'– 5' RNA processing and degradation, being essential for eukaryotic cell viability (Estevez et al. 2003; Mitchell et al. 1997). Some of its activities relay in the homeostatic mRNA turnover, decay of unstable mRNAs, nonsense-mediated mRNA decay and degradation of the mRNA fragments derived from endonucleolytic cleavage mediated by mi/siRNAs in RISC complexes (RNA-induced silencing complex) (Houseley et al. 2006). Combining genetic, proteomic, and whole-transcriptome analyses to investigate the function of the exosome complex Chekanova and co-workers (2007) found that individual subunits of the exosome are functionally specialized, ranging from being dispensable for growth and development (CSL4) to being essential for the development of female gametophytes (RRP41) or embryogenesis (RRP4). These findings demonstrate an unexpected degree of functional plasticity in the plant exosome core. Concerning the exosome targets they found multiple classes of stable structural RNAs, a select subset of mRNAs, primary miRNA (primiRNA) processing intermediates, tandem repeat-associated siRNA precursor species, as well as numerous long npcRNAs, as mentioned before, such as npcRNAs associated with heterochromatic regions in plants. In addition, the results revealed a novel layer of the transcriptome composed of intergenic npcRNAs that are tightly downregulated by constitutive exosome activity. This elegant work gave rise to a publicly available exosome-regulated transcriptome database (http://signal.salk.edu/cgi-bin/exosome) that will certainly help future work to elucidate regulatory mechanisms in complex eukaryotic transcriptomes.

Taking into account the roles of the NMD and the exosome, one could expect that the transcripts up-regulated in the upf mutants ought to overlap with the exosome substrates, as aberrant transcripts recognized by the NMD complex should be degraded from the 3' end by deadenylation and subsequent 3'→5' exonuclease activity in the exosome (Mitchell and Tollervey 2003; Lejeune et al. 2003). However it was not the case and only slight overlap was found between transcripts up-regulated in the upf mutants and exosome substrates. In addition, in the population identified as exosome substrates there were no natural antisense RNAs, or NAT-RNAs. It is possible that these differences are probably due to the experimental variations, such as age of plants, difference in growth conditions and statistical analysis (Kurihara et al. 2009) and certainly future work will need to address these discrepancies.

After recognition of an abnormal RNA by one of the cellular survey mechanisms, the RNA molecules enters the decay mechanisms usually starting by deadenylation of the 3' poly(A) tail and then by decapping of the 5' cap structure, followed by degradation in the 3'→5' and/or 5'→3' directions by the exosome or XRN exoribonucleases, respectively (Chiba

and Green 2009). In addition to aberrant transcripts derived from incorrect splicing or other cellular mechanisms, RNA substrates for this step also arrive from the silencing pathways after the cleavage of mRNA targets by si/miRNAs (Souret et al. 2004; Gy et al. 2007; Gregory et al. 2008; Rymarquis et al. 2011).

Despite that several components of the RQC machinery are required early on for plant growth and their mutations are lethal, inactivation of certain other components may lead to specific phenotypes. The Arabidopsis genome contains three XRN proteins (XRN2, XRN3 and XRN4) with different subcellular localizations and cellular functions. XRN2 and XRN3 are localized in the nucleus, are partially redundant and are required for primary cleavage and processing of pre-ribosomal RNAs (Zakrzewska-Placzek et al. 2010). On the other hand, XRN4, also called ETHYLENE INSENSITIVE 5 (EIN5), is cytoplasmic and necessary for a correct ethylene signalling in the plant and for the degradation of the 3' products resulting from the miRNA-mediated cleavage of target mRNAs (Roman et al. 1995; Olmedo et al. 2006; Gregory et al. 2008). Recent work pointed the association of XRN4 with transcripts encoding specific sequence motifs and select functional groups suggesting that these RNAs may be specifically targeted to the 5'→3' decay pathway for degradation in Arabidopsis. The RNA sequence thus would play a major role in this targeting (Rymarquis et al. 2011). In addition to these functions on survey mechanisms, the three XRN proteins are involved in PTGS acting as endogenous RNA silencing suppressors. The exoribonuclease XRN4 was shown to act as a PTGS suppressor, possibly through the degradation of RNA-dependent RNA polymerase (RdRp) templates. In this way, mutations in XRN4 lead to the accumulation of aberrant, uncapped RNAs derived from transgenes, which could enhance PTGS (Gazzani et al. 2004). Likewise, in another work XRN2 and XRN3, as well as FIERY1 (FRY1), were also shown to be suppressors of PTGS as the XRN activity is inhibited in a fry1 background (Gy et al. 2007). FRY1, also known as SAL1 and HOS2, is a dual function 3'(2'),5'-bisphosphate nucleotidase/inositol polyphosphate 1-phosphatase orthologous to Hal2 and CysQ in yeast and Escherichia coli, respectively (Neuwald et al. 1992; Glaser et al. 1993). It was isolated in a genetic screen based on ABA- and stress-inducible gene transcription, and the Arabidopsis mutation fry1 results in super-induction of ABA- and stress-responsive genes. Accordingly, fry1 mutants are more sensitive to ABA or stress inhibition, and present less tolerance to cold, drought, and salt stresses (Xiong et al. 2001). Accordingly to their activity as suppressors of PTGS, both fry1 and the xrn double and triple mutants accumulate RNA intermediates of miRNA-directed post-transcriptional regulation which are templates of XRN4 (Souret et al. 2004), and miRNA loops, which are templates of both XRN2 and XRN3 (Gy et al. 2007). To further analyze the role of XRN in the fry1 phenotype, they generated an xrn2 xrn3 xrn4 triple mutant that was fertile, unlike the sterile xrn2 xrn3 double mutant. Although the mechanism for the partial phenotypic rescue is unclear, it suggests that xrn4

mutations act to partially suppress the xrn2 xrn3 phenotypic effects. Neither XRN2 nor XRN4 is critical for proper plant development. However, the embryo-lethality of null xrn3 alleles and the developmental defects exhibited by hypomorphic xrn3-3 mutants suggest an essential role of XRN3. The xrn2 xrn3 xrn4 triple mutant displays the fry1 lateral root and drought tolerance phenotypes but does not affect primary root. Microscopical observations revealed that the altered root architecture in fry1 mutants was due to reduced meristem activity in the primary root and to a lateral root initiation defect. Altogether, these results suggest that the pleiotropic phenotype of the fry1 mutants results, at least in part, from a general perturbation in XRN activities.

Recently, in a genetic screen for Arabidopsis mutants deregulated in the expression of Phosphate High affinity Transporter 1;4 (PHT1;4), a novel fry1 allele was described. The authors identified a novel FRY1 function modulating the transcription of several Pi starvation markers in the root stele, however this fry1 mutant phenotype is independent of XRN activities. A transcriptomic analysis confirmed that the phenotype observed corresponded to a point mutation in the transcript corresponding to the fry gene.

The two periphery marks of the extreme borders of a eukaryotic mRNA are defined by the 5' m7G-cap structure and the 3' poly(A) tail, and in the nuclei of eukaryotic organisms the 5' cap is recognized by the cap-binding complex (CBC). In Arabidopsis there are two single genes for both the large and small subunits, ABH1 (ABA Hypersensitive 1) and CBP20 respectively, that form the CBC heterodimer and play numerous roles in RNA metabolism (Aguilera 2005; Hugouvieux et al. 2001). Mutant plants for the ABH1 RBP present ABA hypersensitivity in seed germination, stomata closure, reduced wilting during drought and ABA-induced guard cell calcium increases (Hugouvieux et al. 2001). ABH1 activity is necessary for the correct expression level of a subset of genes in the Arabidopsis genome (Hugouvieux et al. 2001), suggesting a specific connection between mRNA metabolism and ABA signalling. Consistent with their intimately linked activities, inactivation of CBP20 causes a similar serrated leaf phenotype and increased drought resistance as seen in abh1 mutants (Hugouvieux et al. 2001; Papp et al. 2004). Although the morphological and physiological effects of ABH1 and CBP20 have been quite well described (Hugouvieux et al. 2001; Bezerra et al. 2004; Papp et al. 2004), it is largely unclear how these proteins relate to the biochemical functions of the CBC. Nevertheless, through the analysis of the developmental defects on xrn4-abh1 double mutant plants, Gregory and co-workers (2008) demonstrated surprising roles in RNA silencing pathways for these two proteins, XRN4 and ABH1, involved in general RNA metabolism. They found that the loss of ABH1 decreases the levels of mature miRNAs, suggesting that this protein functions in the miRNA-mediated RNA silencing pathway. Indeed the Arabidopsis CBC would be important for proper pri-miRNA processing eventually providing a platform for recruitment of miRNA maturation factors (Laubinger et al.

2008). On the other hand, XRN4 affected the abundance of a distinct class of mainly 21 nt small RNAs, processed from both sense and antisense strands of some endogenous transcripts. Apparently those transcripts are converted to double-stranded RNA (dsRNA) and subsequently processed, and regularly accumulate in an uncapped form in xrn4 mutant plants. Taken together, these results suggest that an additional fate for endogenous uncapped transcripts is shuttling into an RNA silencing pathway where they become small RNA-biogenesis substrates.

Accordingly, a very recent work showed that there is a link between the Arabidopsis cap-binding protein ABH1 and the suppression of silencing (Christie et al. 2011). Their results indicate that genes containing introns are less susceptible to PTGS and that this intron suppression of gene silencing requires an efficient splicing that is dependent on ABH1. At the end of the 80s, it was been already published that endogenous genes generate much higher levels of gene expression than their cDNA counterparts (Callis et al. 1987). As well, in transgene-expressed viral RNA genomes, the addition of introns has also been shown to significantly enhance their accumulation (Marillonnet et al. 2005). Finally, genome-wide mRNA decay rates show that transcripts from intronless genes are significantly more unstable than those from intron-containing genes (Narsai et al. 2007). Based on a GFP-transgenic reporter system and varying the number of introns, this recent study provides a molecular basis to elucidate those evidences. They show that efficiently spliced introns may reduce RDR6 activity along spliced GFP transcripts via a mechanism requiring the cap-binding protein ABH1. Indeed the ABH1 protein has been previously correlated to pre-mRNA splicing in plants (Kuhn et al. 2007; Gregory et al. 2008; Laubinger et al. 2008) and various studies in yeast and animals have shown that the CBC is functionally and physically coupled to pre-mRNA splicing (Izaurralde et al. 1994). To investigate this hypothesis, they transformed abh1 mutants with the same GFP-transgenic reporter system and unexpectedly concluded that intron suppression of transgene silencing was lost in those mutants. They propose a model where an efficient intron splicing could decrease the quantity of aberrant RNA by-products produced by transcription (as uncapped or improperly terminated transcripts), or alternatively, facilitate recruitment of enzymes that degrade aberrant RNA as they are formed. As discussed before, various RQC pathways exist in the cell and compete against endogenous RdR polymerases activities to prevent extensive amplification of silencing. In other words, there is a constant fight between RQC and the silencing pathways competing for the same aberrant or endogenous npcRNA substrates produced in the cell. Thus it is possible that the large varieties of npcRNAs found in the cell may interfere with the RNA surveillance and degradation pathways and affect the processes of RNA maturation.

The total number of protein coding genes in diverse organisms varies much less than the number of different transcripts along evolutionary scales (Mattick and Makunin 2006;

Yasuda and Hayashizaki 2008). This supports the idea that npcRNAs are essential to understand the huge complexity of multi-cellular organisms. Indeed, the large diversity of npcRNAs identified up to now in eukaryotes, and their increasing number, may reflect the importance of riboregulation, mediated by npcRNA–RBPs interactions, in the determination of differentiation and adaptability in eukaryotes. Concretely, plants display a notable flexibility in their architecture and growing patterns in response to external stimuli, characterized by a great developmental plasticity. This quality allows higher plants to adapt to different environmental conditions, with individuals with the same genotype giving rise to different phenotypes. Thus, future challenges lie in understanding the implication of the various RNP networks to determine growth and developmental outcomes under different environmental conditions.

3- L'épissage alternatif

Dans les revues citées précédemment, j'ai mentionné les résultats obtenus dans le laboratoire sur l'identification d'une RBP qui était localisée dans les speckles nucléaires, et qui interagissait avec un lncARN. Ces granules contiennent la machinerie de l'épissage, dans cette partie je discuterai d'abord les mécanismes d'épissage et les rôles des RBPs impliquées dans l'épissage des ARN chez les plantes. Enfin, j'introduirai le principe de l'épissage alternatif et sa régulation.

.

3.1 L'épissage du pré-ARNm chez les plantes

La régulation de l'assemblage du spliceosome durant l'épissage d'un intron ainsi que la ligation des exons est dirigée par des séquences présentes sur le pre-ARNm. En plus des séquences directement impliquées dans l'élimination des introns il existe des séquences activatrices ou inhibitrices de l'épissage. De plus, chez les plantes, il est connu que la richesse en nucléotides UA d'un intron contribue à la reconnaissance du site de clivage. Le site d'épissage situé à la transition entre un intron UA riche et un exon CG riche serait choisi préférentiellement pour l'épissage (pour revue voir Lorkovic et al., 2000). La composition des introns riche en UA a souvent été considéré comme une spécificité propre des plantes; or, des études récentes indiquent que chez les animaux une concentration supérieure en CG existe également spécifiquement dans les exons (Schwartz et al., 2009 ; Tilgner et al., 2009).

La caractérisation biochimique de la machinerie de l'épissage chez les plantes a été entravée par le manque de systèmes d'épissage *in vitro*. Néanmoins, le clonage de l'ensemble de snARN du spliceosome de la classe principale (U1, U2, U4, U5 et U6 (Jakab et al., 1997 ; Shao et al., 2012) et une partie de la classe U12 (U6atac, U12 ; Shukla and Padgett, 1999 ; Lorkovic et al., 2000 ; Kwak et al., 2012) ont montré des similitudes dans leurs structures primaires et secondaires avec leurs homologues chez les métazoaires. Tous les éléments nécessaires à l'assemblage de snARN en RNP ou indispensables à l'appariement nucléotidique snARN-ARNm et qui ont été découvert chez les animaux sont conservés chez les plantes (Lorkovic et al., 2000). Des comparaisons de séquences, des reconstitutions *in vitro* et des analyses immunologiques ont également révélé que la plupart des composantes protéiques des snRNPs identifiées chez les animaux étaient conservées chez les plantes. (Simpson *et al.,* 1995 ; Golovkin and Reddy, 1996 ; Schwartz *et al.,* 2008). Fait intéressant, les génomes des plantes codent généralement plusieurs variants des snARN U1-U5. L'inspection de la base de données TAIR (http://www.arabidopsis.org), où sont rassemblés l'ensemble des séquences de la plante modèle Arabidopsis, indique que,

comme les U-snRNP, certains facteurs protéiques participant à l'assemblage du spliceosome et à la régulation de l'épissage sont hautement conservés. Cependant, très peu de ces protéines végétales ont été caractérisées expérimentalement. Les plus étudiées forment un groupe de facteurs nommés protéines Serine-Rich (SR). Chez les métazoaires, les protéines SR jouent un rôle important dans l'épissage alternatif ; elles favorisent les contacts entre les éléments du spliceosme et régulent la fonction des activateurs de l'épissage (figure 8 ; Barta et al., 2008). Les protéines SR sont composées d'un ou deux RRM (« RNA Recognition Motif ») placé en N-terminale interagissant avec des séquences spécifiques dans le pré-ARNm et un domaine riche en dipeptides Ser-Arg, (SR) impliqués dans des interactions protéine-protéine (Tacke and Manley, 1999 ; Reddy and Shad, 2011). Certaines des protéines SR végétales caractérisées sont des homologues de protéines de vertébrés et d'autres sont clairement spécifiques des plantes (Lazar *et al.,* 1995 ; Lopato *et al.,* 1996 ; Golovkin and Reddy, 1998 ; Golovkin and Reddy, 1999 ; Reddy and Shad., 2011 ; Duque 2011).

L'épissage alternatif (ou AS pour alternative splicing) est un mécanisme clef chez les eucaryotes, il permet l'augmentation de la diversité du transcriptome et du protéome et donc d'élargir la capacité codante du génome grâce à la formation d'ARNm différents à partir d'un même gène (Syed et al., 2012). La plupart des gènes eucaryotes sont alternativement épissés de manière tissu spécifique. Les défauts d'AS sont à l'origine de maladies chez les mammifères (Blencowe, 2006 ; Matlin et al, 2005; Wang et Cooper, 2007) et d'anomalies de la réponse des plantes aux conditions de l'environnement (James et al, 2012 ; Filichkin et al, 2010 ; Tanabe et al, 2007 ; Palusa et al, 2007 ; Yan et al, 2012). Des études récentes de séquençage haut débit indiquent qu'environ 60% des gènes qui contiennent des introns subissent un épissage alternatif chez les plantes (Marquez et al. 2012). Toutes les catégories fonctionnelles sont touchées, des gènes impliqués dans la réponse au stress à ceux impliqués dans le contrôle de la croissance, le développement ou la réponse au stress. L'AS peux aussi réguler le niveau de transcription d'un gène et il peut donner lieu à la formation d' ARNm non-codants mais aussi à des ARN non sens qui sont cibles du NMD (pour « Non-sense RNA Mediated Decay »). De plus, Les ARNm non-sens ou aberrants (qui présentent un codon stop suite à un décalage de phase de lecture par exemple) déclenchent une procédure de dégradation de l'ARNm appelé NMD. Le fait qu'une partie des ARNm produits par un transcrit puisse devenir des substrats de la machinerie NMD influence fortement la quantité de protéine produite (Kalyna et al., 2012 ; Lewis et al., 2003 ; Stamm et al., 2005 ; Maquat, 2004). Chez les mammifères et les plantes, les ARNm qui présentent une très grande distance entre le PTC (Premature stop codon) et la fin 3' de l'ARNm [long 3' UTRs (région non traduite)] et / ou les ARNm avec un complexe de jonction d'exon (EJC pour Exon junction complex) situé en aval d'un PTC sont les substrats de la dégradation

NMD (Maquat, 2004 ; Kertész et al., 2006 ; Figure 8). Le NMD est connu pour réduire les conséquences des mutations ou des erreurs survenues lors de la transcription afin de réduire le « bruit génomique » et d'empêcher la production de produits protéiques potentiellement tronqués. En outre, le NMD module les niveaux des isoformes d'AS qui contiennent un PTC. Environ 10% du transcriptome humain et de la levure non sujet à de l'AS est ciblé par le mécanisme de NMD. Pour l'homme, il faut rajouter les 30% des transcrits qui ont subi un AS et qui de ce faite contiennent un PTC (Lewis et al., 2003). Chez Arabidopsis, les gènes codant pour les protéines de l'horloge du cycle circadien, AtGRP7 et AtGRP8 (Arabidopsis thaliana riche en glycine protéines) sont des exemples de la régulation à la fois par l'AS et le NMD (Schoning et al., 2008). L'analyse des lignées mutantes pour les gènes clef du NMD : les up-frameshift1 (UPF1) et UPF3 (upf3-1 and upf1-5) ont permis de détecter des augmentations de l'abondance de certaines isoformes spécifiques d'AS dans un nombre important d'événement d'AS.

Il existe différent types d'épissage alternatif permettant de créer différentes isoformes d'ARNm (Figure 7). Tout d'abord un exon peut être inclus ou « sauté » (ou évité) (ES for Exon skipping/inclusion). Dans le premier cas, l'exon inclus nécessite deux événements d'épissage alors que le saut d'exon n'en demande qu'un seul. L'autre événement d'épissage alternatif est la rétention d'intron (IR pour intron rétention) ; un ou plusieurs introns peuvent êtres retenus dans l'ARNm mature. Enfin, la sélection alternative du site d'épissage en 3' ou en 5' de l'intron peut varier produisant des exons de longueur différents (Black, 2003 ; Stamm et al., 2005). Ces événements d'AS peuvent avoir lieux partout dans le pre-ARNm dans la partie codante comme dans la partie 3' ou 5' UTR. Enfin certains événements d'épissage peuvent aussi induire l'apparition d'un nouveau codon start induisant la formation d'une protéine commençant par une méthionine différente. L'AS produit de multiples ARNm à partir d'un même gène grâce à différents sites d'épissages. Il joue un rôle de régulation clef en modulant l'expression des gènes durant le développement notamment en réponse aux signaux environnementaux (Black, 2003 ; Graveley, 2001 ; Lareau et al., 2004 ; Stamm et al., 2005). La régulation de l'AS est complexe, elle dépend de l'interaction entre des séquences en -cis positives ou négatives dans le précurseur de l'ARNm (pre-ARNm) ainsi que de nombreux RBPs et autres facteurs qui agissent en -trans. Le niveau et l'activité de ces facteurs varient en fonction du tissu ou encore des conditions de cultures.

La quantité et l'activité des ces RBP est régulée au niveau cellulaire. L'épissage d'un transcrit sera donc dépendant de la régulation transcriptionnelle et post transcriptionnelle mais aussi traductionnelle et post traductionnelle de ces régulateurs dans la cellule ou le tissu considéré. Chez l'homme, 95% des gènes subissent au moins un épissage alternatif (Pan et al., 2008) ; le nombre largement supérieur de protéines par rapport au nombre de gènes codants ces protéines est principalement la conséquence de l'AS (Lewis et al., 2003).

Les nouvelles méthodes de séquençages sont en train de révolutionner la recherche dans le domaine de l'AS. Chez l'homme, la comparaison croisée de la séquence primaire des protéines et des ARNm correspondants a permis d'établir un code d'épissage qui est à l'origine de la prédiction des événements d'AS spécifiques dans un tissu précis (Pan et al., 2008). L'AS ne contribue pas exclusivement à augmenter la diversité protéique, il peut aussi générer des protéines tronquées qui ont potentiellement un rôle différent (activité, interaction protéine-protéine ou protéine-substrat, localisation, modification des protéines).

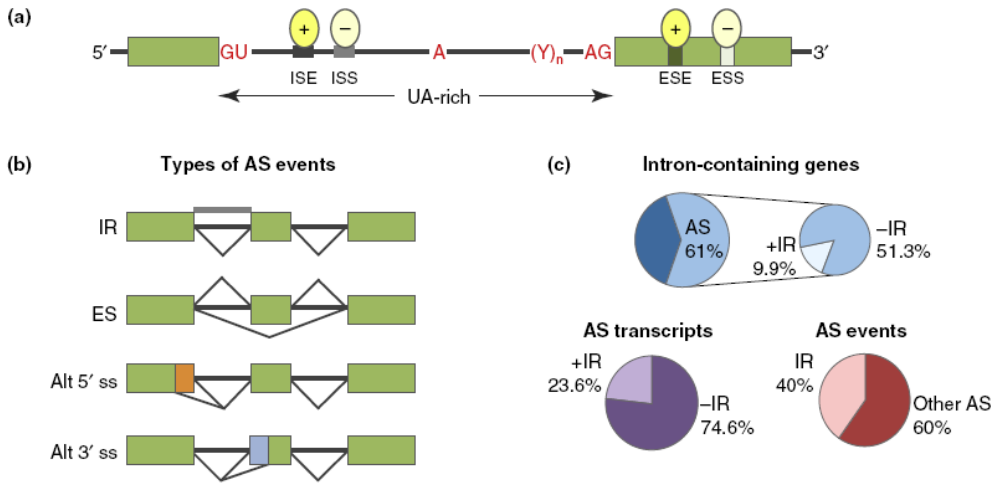


Figure 7 Les principaux types d'épissage alternatif (AS) et la fréquence des ces événements chez Arabidopsis (Syed et al., 2012).

(a) L'épissage du pré-ARNm est dirigé par des éléments en -cis qui comprennent des sites d'épissage, le point de branchement et les séquences polypyrimidine track. La sélection des sites d'épissage alternatif est affectée par des facteurs agissant en -trans qui se lient à des éléments en -cis exonique et introniques, appelés inducteur ou represseur de l'épissage. (b) Types d'événements d'AS. (c) Fréquence d'apparition de la rétention d'intron chez Arabidopsis. La rétention d'intron est la plus fréquente des cas d'AS chez Arabidopsis (40%), mais sa contribution à la diversité transcriptionnelle est beaucoup plus faible (Marquez et al., 2012). Parmi les 61% des gènes d'Arabidopsis contenant un intron qui subi l'AS, 51% des transcrits qui produisent de l'AS ne subisse pas de rétention d'intron (IR). Parmi les transcripts épissés alternativement, 23,6% contiennent un ou plusieurs introns non retenus (+ IR), tandis que le reste (74,6%) sont produits par d'autres événements d'AS. Abréviations: ESE, inducteur d'épissage exonique; ESS silenceur d'epissage exoniques; ISE, inducteur d'épissage introniques, ISS, silenceur d'épissage intronique; Alt 3' SS, sites d'épissage alternatif 3'; Alt 5' SS, des sites d'épissage alternatifs 5'; ES, exon sauté, IR, intron retenu.

L'AS peut toucher des gènes impliqués dans une grande variété de mécanismes moléculaires comme la transcription, la transduction de signaux, l'épissage, le développement, la réponse au stress biotique et abiotique. Ce mécanisme régule donc toutes les étapes de la vie des plantes du développement des graines à la germination en passant par la résistance aux maladies, le rythme circadien et le déterminisme de la floraison (Barbazuk et al., 2008 ; Reddy, 2007 ; Schoning et al.,2008 ; Simpson et al., 2010 ; Kalyna et al., 2012).

Récemment, de vastes analyses par ARN-seq du transcriptome chez *Arabidopsis* ont montré que la fréquence de l'AS est beaucoup plus importante que ce qui avait été décrit précédemment. En effet, ces études indiquent que plus de 61% de gènes contenant des introns subissent l'AS. Cette estimation de 61% de l'AS est basée sur l'analyse des plantes cultivées dans des conditions normales de croissance (Marquez et al., 2012 ; Syed et al., 2012 ; Figure 7). Il semble donc évident que ce niveau augmentera lorsque différents stades de développement ou bien des conditions environnementales différentes seront analysées (Marquez et al., 2012). Chez les plantes, la rétention d'intron (IR) est l'événement le plus fréquent (Filichkin et al., 2010 ; Barbazuk et al., 2008). Cependant, certains événements d'IR ont été récemment montré comme étant plus susceptibles de représenter des transcrits partiellement épissés en raison de la faible abondance de ces transcrits (Marquez et al., 2012).

3.2 Régulation de l'AS

La régulation de l'AS dépend de séquences présentes sur l'ARNm en –cis ainsi que de facteurs d'épissage (splicing factors) agissant en -trans. Les protéines Sérine / arginine-riches (SR) et les RNP hétérogènes nucléaires (hnRNP) sont des facteurs de l'épissage constitutif mais également de l'AS ; elles agissent de manière concentration dépendante dans le choix des sites d'épissage (Matlin et al., 2005) (figure 8). Les protéines SR sont très conservées chez les eucaryotes, elles présentent un ou deux motifs de reconnaissance d'ARN (RRMs) et un domaine C-terminale (CTD) riche en résidus sérine et arginine (SR domain ; Barta et al., 2008). Fait intéressant, les plantes possèdent deux fois plus de protéines SR que les organismes non photosynthétiques. Certaines SR sont spécifiques de la lignée verte (Richardson et al., 2011) et présentent des profils d'expression spatio-temporelle très différents (Kalyna and Barta, 2004). Les protéines hnRNP, en revanche, constituent un groupe structurellement divers de protéines liant l'ARN qui possèdent des rôles dans des processus moléculaires variés en addition de leur rôle dans L'AS (Martinez-Contreras et al., 2007). Les protéines SR et les protéines hnRNP se lient d'une part, aux signaux d'épissage et aux séquences promotrices/inhibitrices situées dans les introns ou dans les exons et, d'autre part, avec des facteurs d'épissage créant ainsi des interactions multiples qui dirigent le choix du site d'épissage où s'assemblera le spliceosome. En général, les protéines SR semblent promouvoir l'épissage et les protéines hnRNP inhibent la sélection du site d'épissage. Toutefois, les protéines SRs et hnRNP peuvent aussi avoir des fonctions inverses, par exemple dans les cellules HeLa, SRSF10 (également connu sous le nom SRP38) (Feng et al., 2008) est un régulateur négatif alors que la polypyrimidine tract-

binding protein (PTB; également connu sous le nom hnRNPI) est lui un régulateur positif (Xue et al.,2009). De plus, les variations tant de l'abondance que de l'activité de ces facteurs d'épissage déterminent le profile d'AS des gènes cibles. La régulation de l'expression de ces différents facteurs est donc fondamentale au cours des étapes de développement ou lors de la croissance chez les eucaryotes (Barta et al., 2008 ; Palusa et al., 2007 ; Tanabe et al., 2007). En outre, les conditions de croissance peuvent moduler l'AS d'un gène SR et/ou hnRNP (ces gènes sont d'ailleurs particulièrement sujet à l'AS), causant des changements dynamiques dans le profil d'expression de gènes codant des facteurs d'épissage et entraînant de ce fait des répercussions sur l'expression des gènes cibles. Par exemple, l'AS de nombreux gènes d'Arabidopsis codant pour des protéines SR est affecté par la température, la lumière, le sel et les hormones (Palusa et al., 2007 ; Tanabe et al., 2007; Duque, 2011). En outre, l'activité ou la localisation des protéines SR peuvent être affectées par la phosphorylation (Stamm et al., 2005 ; Stamm, 2008) ; en effet, des protéines kinases qui phosphorylent les protéines SR végétales ont été identifiés (de la Fuente van Bentem et al., 2006 ; Savaldi-Goldstein et al., 2003). Evidemment, la surexpression des protéines SR et les lignées knock-out montrent une variété de phénotypes de développement et de croissance, ce qui démontre l'importance de ces protéines dans la croissance et le développement des plantes (Barta et al., 2008).

Chez Arabidopsis, les hnRNP les mieux caractérisées à ce jour sont les protéines PTB, qui sont des orthologues de régulateurs d'épissage négatif chez les animaux (Rühl et al., 2012), ainsi que les « glycine-riches RBP » GRP7 et GRP8, des constituants d'un oscillateur couplé au rythme circadien (Schoning et al., 2007 ; Schuttpelz et al., 2008 ; Schoning et al., 2008). La famille des PTB d'Arabidopsis auto-régulent l'AS du gène qui les code ainsi que celui d'autres gènes cibles (Stauffer et al., 2010). GRP7 et GRP8 autorégulent également leur propre AS et régulent l'AS d'autre cibles, générant des ARNm improductifs qui sont ciblés par le NMD afin réduire post transcriptionellement le niveau d'expression des gènes (Schuttpelz et al., 2008 ; Schoning et al., 2008). Enfin, le Cap Binding Complex (CBC) se compose de deux sous-unités, AtCBP20 et AtCBP80. AtCBP20

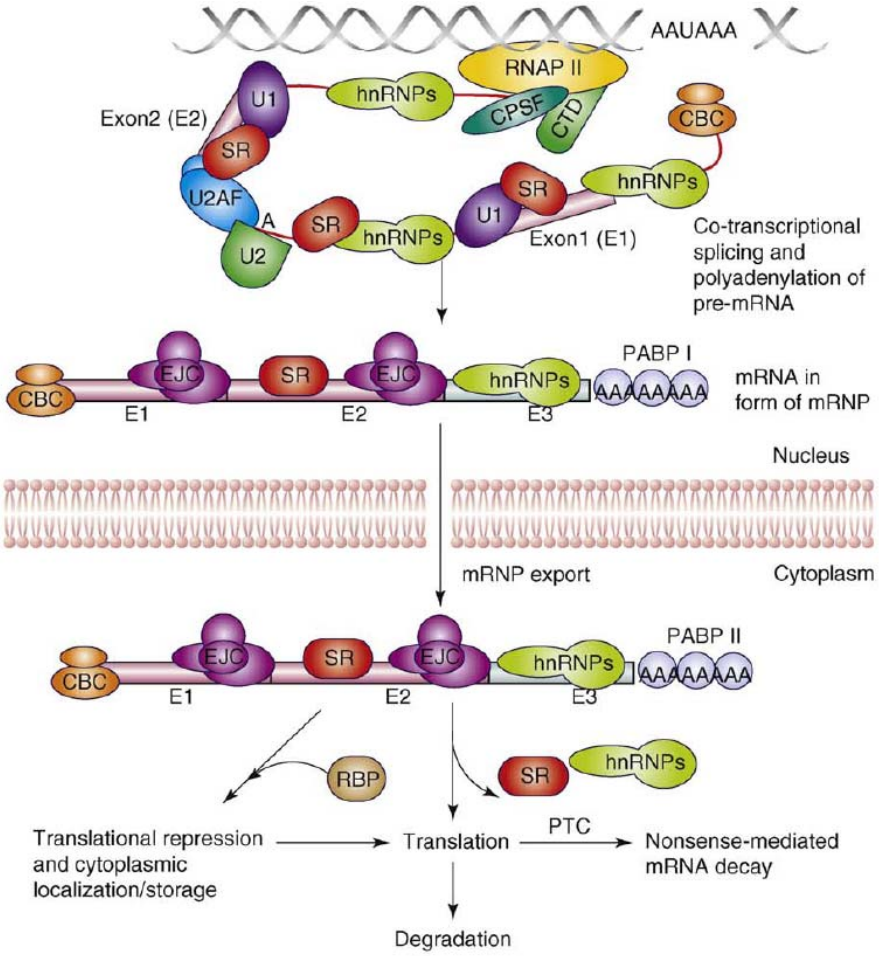


Figure 8 L'épissage alternatif et le destin des ARNm est déterminé par différents ensembles de protéines liant l'ARN (RBPs).

Les pré-ARNm naissant sont liés par les protéines hnRNP, ce qui facilite la liaison des autres RBPs Impliquées dans l'épissage (par exemple les protéines SR) ainsi que les protéines impliquées dans la polyadénylation ou qui participent activement aux mêmes processus. Les pré-ARNm sont co-transcriptionnellement épissés et polyadénylés par des machines complexes contenant des RBPs, comme les protéines SR like, la petite particule ribonucléoprotéique nucléaire U2 (snRNP), le facteur auxiliaire (U2AF) et les composants des snRNP U1 et U2 qui sont impliqués dans l'épissage. Au cours de l'épissage d'un intron, le spliceosome dépose les complexes de jonction d'exons (EJC) 20-24 nucléotides en amont de la jonction exon-exon. L'EJC comporte à nouveau un ensemble spécifique de RBPs (quatre sous-unités en complexe indiquées en violet foncé), et son dépôt sur l'ARNm affecte l'exportation de l'ARNm, la traduction et la surveillance du non sense-mediated decay (NMD). hnRNP et protéines SR accompagnent l'ARNm dans le cytoplasme, ce qui indique un rôle dans l'exportation des ARNm (Barta et al., 2008). Ces ARNm peuvent parfois être reconnus comme aberrants et sont donc ciblés par le NMD. Nomenclature: hnRNP, ribonucléoprotéine nucléaire hétérogène (un groupe de RBPs différentes qui se lient au pre-ARNm mais ne sont pas des composantes stables d'autres RNP); CBC, le nuclear cap binding complexe; SR, facteurs d'épissage formant des dipeptides, qui peuvent être phosphorylée. Cette modification régit les interactions avec d'autres facteurs d'épissage et pré-ARNm, ainsi que leur localisation sub-cellulaire; U1 et U2, U1 et U2 petites particules ribonucléoprotéines nucléaires. Elles sont impliquées dans l'épissage de pré-ARNm en reconnaissant et définissant les sites d'épissage 5' et les 3', respectivement. U2 snRNP peut s'apparier par complementarité de base avec la séquence de point de ramification (indiqué par un A dans l'intron) et contribue avec l'ensemble des autres facteurs importants à la définition du site 3' d'épissage; PTC, premature stop codon. Un codon stop est reconnu comme prématuré si il est situé à plus de 50 nucléotides en amont de la jonction exon-exon 3'; PABP I, nuclear poly(A) binding protein contient un domaine RRM unique; PABP II, cytoplasmic poly(A) binding protein.

contient un domaine de liaison à l'ARN canonique (RBD). La mutation de ces sous-unités a montré que le CBC est impliqué dans l'AS d'au moins certains gènes d'*Arabidopsis.* Ce complexe influencerait principalement l'AS du premier intron, en particulier le site 5' d'épissage (Raczynska et al., 2010). Le CBC joue plusieurs rôles dans la maturation des ARNm incluant l'épissage, l'export des ARNm, la protection des transcrits contre les dégradations par les nucléases et la stabilisation de la machinerie qui permet la formation de l'extrémité 3'.

Bien que l'abondance et l'activité des facteurs d'épissage détermine le profil d'expression des gènes en aval, très peu d'exemples ont été publiés sur le rôle biologique des différentes isoformes protéiques de régulateurs de l'épissage. La diversité fonctionnelle des isoformes d'AS à tout de même été démontré avec les deux isoformes de SR45 (qui ne diffèrent que de huit acides aminés et qui contiennent des sites putatifs de phosphorylation), ces isoformes complémentent différemment les phénotypes développementaux des pétales ou de la racine des mutant sr45 (Zhang and Mount, 2009). Ainsi, des isoformes avec des séquences très similaires peuvent avoir des résultats morphologiques différents, ces résultats reflètent l'importance que peut avoir l'AS sur l'expression génique et le développement chez les plantes.

4- Objectifs

Notre laboratoire a identifié une RBP qui interagissait avec l'ARN *ENOD40* ; cette interaction conduit la relocalisation de la protéine depuis les speckles nucléaires vers des particules cytoplasmiques (Campalans et al., 2004). Cette relocalisation a lieu lors de la formation des nodosités ainsi que lors de la formation de racines latérales, deux organes latéraux de la racine principale chez les légumineuses ; c'est à dire là où le gène *ENOD40* s'exprime fortement. L'objectif général de ma thèse a été de caractériser le rôle de ces protéines NSRs chez les plantes. De plus comme MtNSR1 interagit avec un npcARN, nous avons étudié l'action des ARN non codant sur la fonction de ces protéines.

A cette fin, nous avons d'abord identifiés des gènes homologues de *MTRBP1*/*MTNSR chez Arabidopsis*: *AtNSRa et AtNSRb.*, une espèce que ne contient pas une séquence nucléotidique clairement homologue à l'ARN *ENOD40* bien que cette plante modèle contienne plusieurs autres ARN non-codants (BenAmor et al, 2009). Dans le premier chapitre, vu le rôle de *ENOD40* dans la formation des organes latéraux chez les légumineuses, nous avons analysé le rôle des *AtNSRs* dans la régulation de l'architecture racinaire. Nous avons montré que ces gènes s'expriment dans les méristèmes racinaires (primaires et latéraux) et que les protéines correspondantes étaient localisées dans des « speckles » nucléaires chez Arabidopsis. Nous avons aussi étudié la physiologie de la racine dans les mutants, nsra et nsrb, ainsi que dans le double mutant nsra/nsrb, et dans des lignées qui sur-expriment AtNSRa ou AtNSRb. Un phénotype racinaire en réponse à l'auxine a ainsi été identifié. Nous avons aussi montré que les fusions des protéines fluorescentes aux NSRs localisait dans les « nuclear speckles » comme chez *Medicago* et co-localisait avec plusieurs protéines de référence qui marquent des territoires nucléaires, notamment en lien avec la machinerie de l'épissage. Ceci nous a mené à rechercher une éventuelle implication des AtNSRs dans l'épissage, notamment en réponse à l'auxine. Nous avons observé une perturbation important de l'épissage alternatif (mais pas de l'épissage en général) suggérant que les NSRs sont des nouveaux régulateurs de l'AS (collaboration avec J. Brown et C Simpson, SCRI Ecosse). Ensuite, nous avons recherché des ARNnc, comme *ENOD40* ou lncARN351 (un long ARN non codant qui se lie avec les NSRs *in planta* chez *Arabidopsis*) qui pourraient moduler l'épissage *via* ces AtNSRs.

Dans un deuxième chapitre, nous avons étudié le rôle des NSRs ainsi que de 2 autres protéines liées à l'épissage, dans les phénomènes de silencing. En effet, en parallèle à ces études, l'équipe d'Hervé Vaucheret avait identifié 2 protéines dont leurs homologues ont un lien avec l'épissage et qui présentent des effets inhibiteurs sur le

silencing (collaboration avec H. Vaucheret, INRA). Nous avons donc croisé des lignées qui sur-expriment ou des lignées mutantes pour les gènes AtNSRs avec les lignées affectées dans le silencing (tel que celles utilisé pour le crible « supressor of gene silencing » SGS) afin de voir l'effet de la mutation ou de la sur-expression sur les mécanismes du silencing. De plus, nous avons montré que plusieurs de ces protéines co-localisent avec les NSRs dans des particules nucléaires. De manière très intéressante, les mutants Atnsr sont perturbés dans la propagation du silencing et présentent des phénotypes semblables à ces autres mutants SGS. Nous avons donc étudiés l'épissage alternatif des différents mutants et essayer d'établir une liaison plus générale entre la régulation de l'épissage alternatif et le silencing (par les siRNAs) de transgènes contenant un intron.

II- Résultats

II- Résultats

II.1- Modulation of nuclear alternative splicing regulators by long ncRNA

1.1 Introduction

Dans cette première partie, qui sera principalement présenté sous la forme d'un article (en soumission) intitulé « Modulation of nuclear alternative splicing regulators by long ncRNA in Arabidopsis » ; je vais décrire les principaux résultats obtenus sur le rôle des protéines NSRs en relation avec l'épissage alternatif et les ARN non codant. Pour cela, nous avons collaboré avec le laboratoire de John Brown (SCRI, Scotland) et notamment avec l'aide de Craig Simpson. Leur laboratoire est spécialisé dans l'épissage alternatif chez les végétaux. Ils ont créé un panel composé de 288 gènes connus pour être épissés alternativement chez Arabidopsis. Ils disposent pour chaque gène d'un couple d'oligonucléotides qui permet d'amplifier toutes les isoformes épissées alternativement en une PCR. Ils peuvent ensuite mesurer la quantité de chaque isoformes présentes à l'aide d'un séquenceur. De plus Federico Ariel (ISV, CNRS) m'a donné une aide précieuse pour la mise au point et la réalisation des expériences d'immunoprécipitation d'ARN, il a aussi réalisé l'analyse statistique qui établie la relation entre le transcriptome du double mutant nsra/nsrb et la formation de la racine latérale. Philippe Laporte (ISV, CNRS) a quand à lui crée les lignées double mutantes nsra/nsrb pendant sa thèse avant mon arrivé au laboratoire. Enfin, Sandrine Balzergue (URGV, INRA) a réalisé les expériences de transcriptome CATMA ainsi que l'analyse statistique qui en découle.

1.2 Modulation of nuclear alternative splicing regulators by long ncRNA

Title: Modulation of nuclear alternative splicing regulators by long ncRNA in *Arabidopsis*

Florian Bardou[1], Federico Ariel[1], Craig G. Simpson[2], Philippe Laporte[1], Sandrine Balzergue[3], John W.S. Brown[2] and Martin Crespi[1]

Addresses:

[1] Institut des Sciences du Végétal, Centre National de la Recherche Scientifique, F-91198 Gif-sur-Yvette cedex, France

[2] Genetics Programme, SCRI, Invergowrie, Dundee DD2 5DA, Scotland, U.K

[3] Génomique Fonctionnelle d'Arabidopsis, Unité de Recherche en Génomique Végétale (URGV), UMR INRA 1165 – Université d'Evry Val d'Essonne – ERL CNRS 8196, Evry, France

Corresponding author: crespi@isv.cnrs-gif.fr

Keywords: Alternative splicing, long non-coding RNA, nuclear speckles, lateral root

Abstract:

Alternative splicing of pre-mRNA represents a major mechanism that underlies increased transcriptome complexity. Alternative splicing occurs in ribonucleoprotein particles that are generally localised within nuclear "speckles", but its regulation is poorly understood. Nuclear Speckle RNA-binding proteins (NSRs) interact with *ENOD40* RNA, a structured long non-coding RNA (lncRNA) that is induced in the *de novo* formed lateral root organs of legumes. Here, we show that AtNSRs act as alternative splicing regulators in *Arabidopsis thaliana*. AtNSRs-GFP translational fusions are expressed in root meristems and can be re-localised from nuclear speckles to the cytoplasm when co-expressed with *ENOD40* RNA. Double *Atnsr* mutants exhibit an altered ability to form lateral roots in response to the phytohormone auxin and modify auxin-induced alternative splicing patterns of one hundred genes (out of 288 analysed) within root tissues. RNA immunoprecipitation assays demonstrate that AtNSRs interact with their mRNA targets and with the lncRNA nc351. The expression of either the *ENOD40* RNA or nc351 in *Arabidopsis* affects the splicing pattern of several NSR-mRNA targets. Our results show that nuclear alternative splicing regulators can be hijacked by lncRNA, to modulate alternative splicing patterns during development.

Most eukaryotic genes are alternatively spliced in a cell type- and tissue-specific manner, and defects in alternative splicing (AS) can contribute to diseases in mammals (Blencowe, 2006 ; Matlin et al., 2005; Wang et al., 2008; Wang and Cooper, 2007) and to developmental plasticity in plants responding to environmental cues (James et al., 2012; Filichkin et al., 2010; Tanabe et al., 2007; Palusa et al., 2007). Whereas 90% of the Human genome is transcribed, only 1.2% of the genome encodes for proteins (Birney et al., 2007). Therefore, non-protein-coding RNAs (ncRNA) might represent a part of the transcriptome that can elicit new mechanisms of gene regulation, likely through their interaction with specific ribonucleoprotein complexes (Guttman and Rinn 2012). Long ncRNAs are predominantly involved in epigenetic patterning and chromatin remodelling or function as scaffolds that interfere or modulate the action of different RNA-related enzymatic complexes (Rinn and Chang, 2012). Interestingly, a large diversity of lncRNAs has exhibited tissue- or cell type-specific expression patterns (Djebali et al., 2012, Ben Amor et al., 2009, Liu et al., 2012), suggesting roles in specific cell types or developmental transitions. Here, we have identified new RNA-binding proteins that act as nuclear alternative splicing regulators in *Arabidopsis*

and that interact both with their mRNA targets and with lncRNAs. We propose that this latter interaction can modulate alternative splicing patterns in *Arabidopsis thaliana* roots.

In the model legume *Medicago truncatula, ENOD40* is a highly structured long RNA with poor protein coding potential and is involved in nodule organogenesis (Charon et al., 1999; Sousa et al., 2001). Although a peptide encoded within a very small ORF of this transcript can be translated *in vitro* (Rohrig et al., 2002), as has been recently shown for other lncRNAs using high throughput translational techniques (Ingolia et al., 2011), the evolutionary conservation of this gene family at the nucleotide sequence level expands beyond this small ORF. Furthermore, this peptide sequence is absent in several plants carrying *ENOD40* RNA-related sequences (Santi et al., 2003). By using a triple-hybrid approach with a conserved *ENOD40* RNA region, we previously identified an RNA-binding protein (RBP) that interacts with this lncRNA *in vivo* (Campalans et al., 2004). This protein contains an RNA recognition motif in the C-terminal region and a nuclear localisation signal. Because this RBP localises within nuclear speckles, we renamed it NSR1 (for Nuclear Speckles RNA-binding protein 1). NSR1 was shown to be re-localised from the nucleus to the cytoplasm specifically in root tissues where *ENOD40* is expressed, such as the nodule primordia and lateral roots initiation. This *ENOD40* RNA-induced re-localisation was confirmed using transient expression assays in a heterologous system (Campalans et al., 2004). Here, we have identified two different homologs of NSR1 in *Arabidopsis thaliana,* called NSRa and NSRb (Fig. S1).

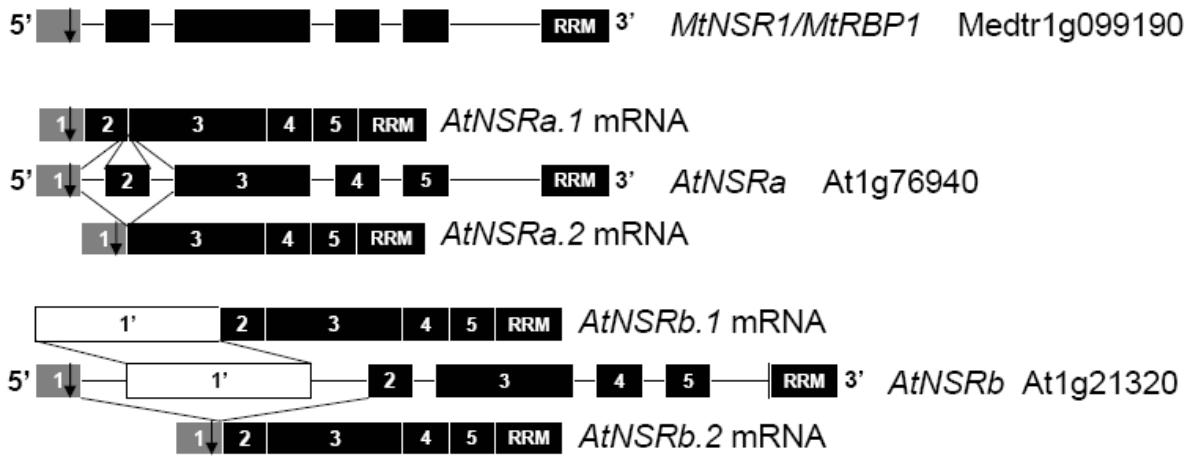


Fig. S1 *MtNSR1/MtRBP1* has two homologues in *Arabidopsis thaliana AtNSRa and AtNSRb.*

The NSRs from *Medicago truncatula* and *Arabidopsis thaliana* have a conserved gene structure with highly conserved domains (1 to 5), a RNA recognition motif in C-term part of the proteins (RRM), and a Nuclear Localisation Signal (NLS, represented by a black arrow in exon 1).The *AtNSRa* gene is alternatively spliced and produces two mRNAs differing by an intron retention between exons 2 and 3. On the other hand, the *AtNSRb* gene is alternatively spliced and produce two transcripts with different initial exons (1 or 1'; exon skipping) leading to two different start codons.

The *NSRb* locus yields two alternatively spliced transcripts with different start codons (initial exon skipping), whereas the *NSRa* locus produces two mRNAs due to an intron retention event (Fig. S1). These events are common for several splicing-related proteins in plants (Tanabe et al., 2007; Palusa et al., 2007). Interestingly, NSRa has been previously reported as a Serine Rich-related protein (SR proteins; Schindler et al., 2008), a class of proteins that is globally linked to splicing regulation in eukaryotes (Barta et al., 2008). Nevertheless, the NSRs belong to a plant-specific protein family, in contrast to SR proteins, which are conserved in all Eukaryota (Barta et al., 2008; Fig. S2).

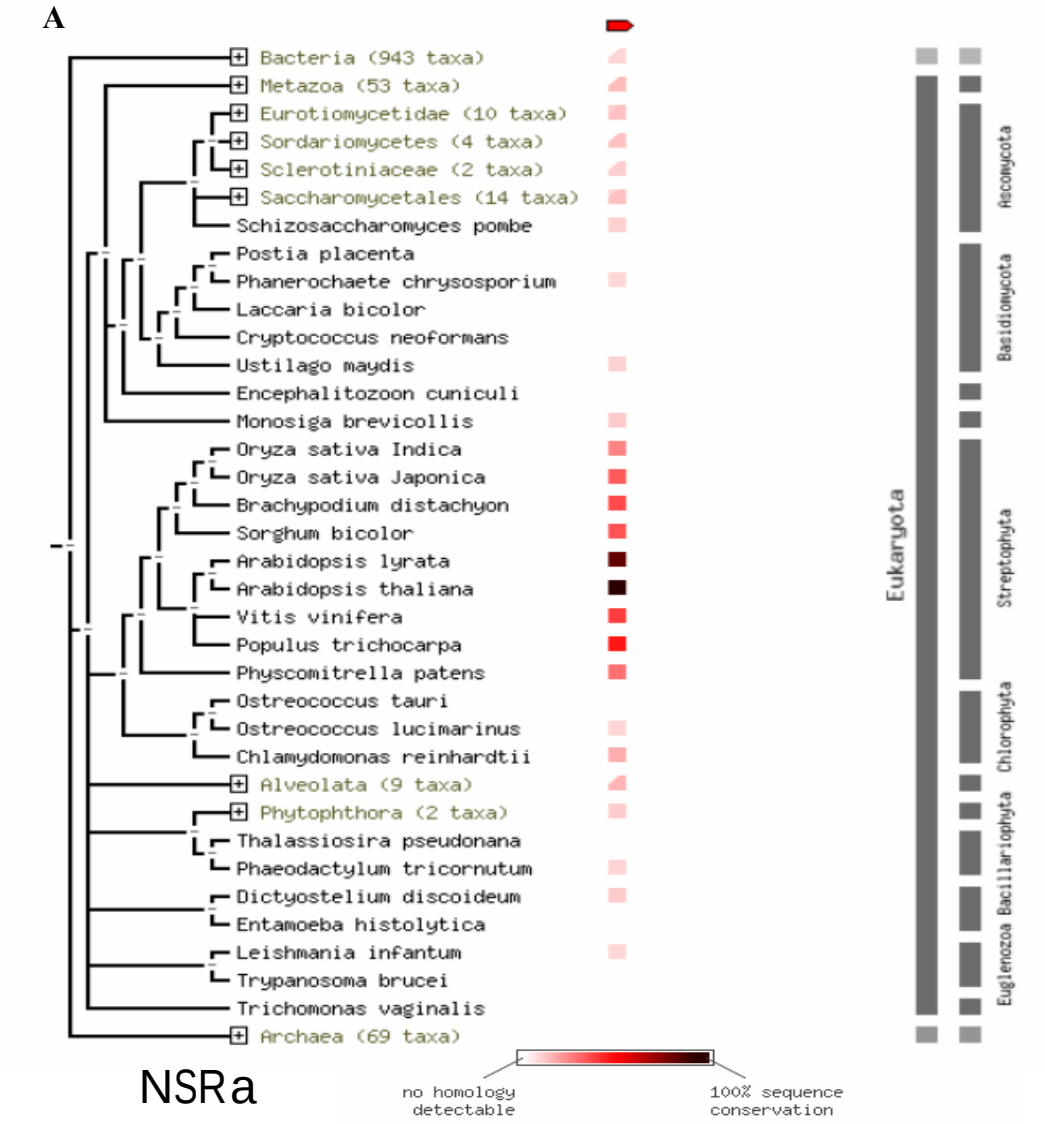


Fig. S2A AtNSRs are plant-specific RNA binding proteins.
Conservation of NSRa protein using STRING 9.0 (http://string-db.org). AtNSRa is conserved along the plant kingdom. The presence of an item in an organism is marked with a quantitative colour scale in Protein-mode (showing the amount of sequence conservation between your protein and its best hit in the other species, dark brown being 100% identity).

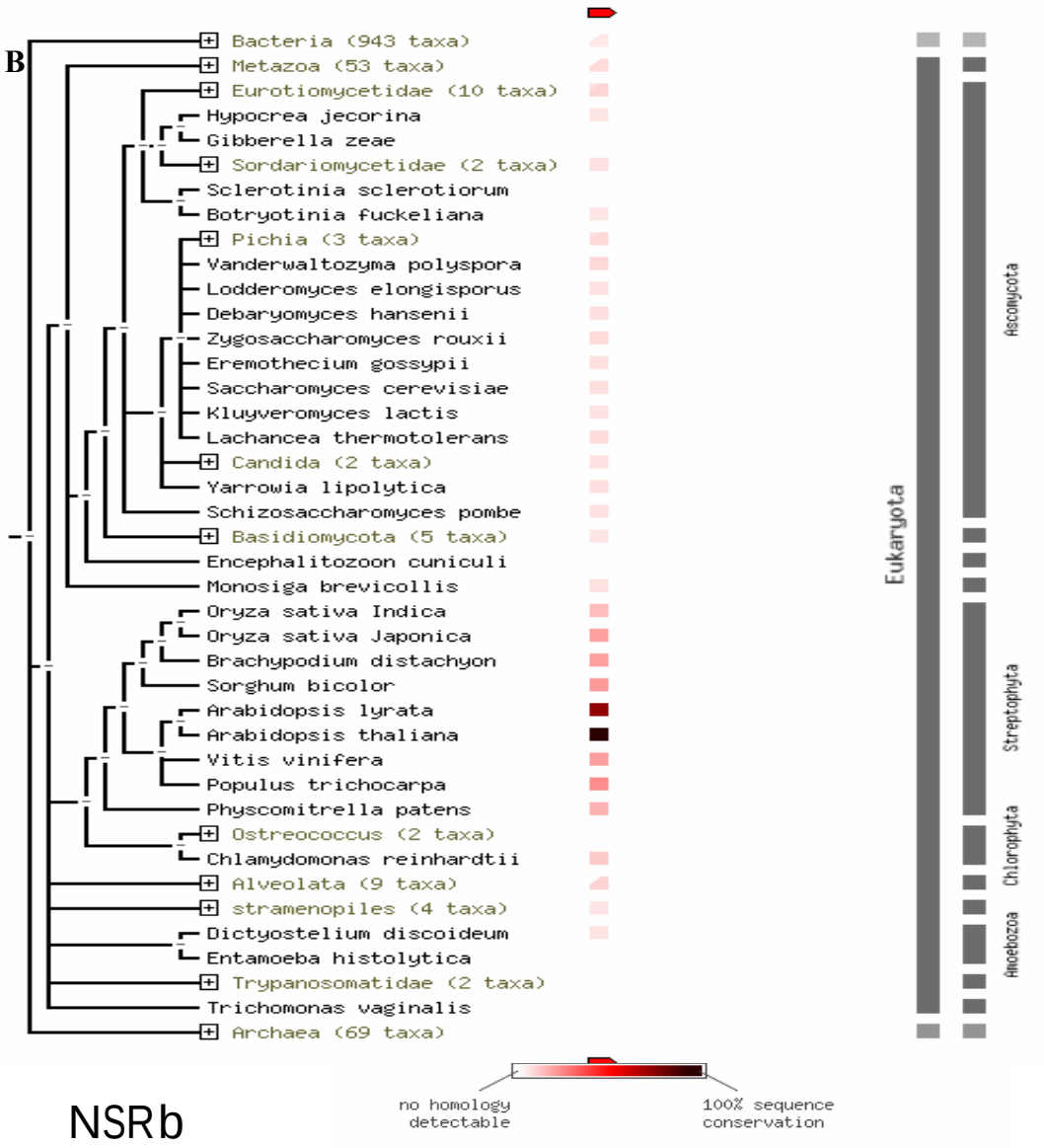


Fig. S2B AtNSRs are plant-specific RNA binding proteins.
Conservation of NSRb protein using STRING 9.0 (http://string-db.org). The new exon of the AtNSRb protein is only present in Arabidopsis. The presence of an item in an organism is marked with a quantitative colour scale in Protein-mode (showing the amount of sequence conservation between your protein and its best hit in the other species, dark brown being 100% identity).

The AtNSRs proteins also localise within nuclear dots or “speckles”, as demonstrated when translational GFP fusions are transiently expressed in tobacco leaves. Strikingly, when the AtNSRs are co-expressed with the *ENOD40* RNA, the particles are re-localised from such nuclear particles to cytoplasmic dots, indicating a conserved *ENOD40* re-localisation activity both for *Arabidopsis thaliana* and *Medicago truncatula* NSRs (Fig. 1A). We also confirmed the co-localisation of the NSR homologous proteins from *Arabidopsis thaliana* and *Medicago truncatula* within these nuclear particles (Fig. S3A). To further characterise these nuclear

particles containing AtNSRs, we co-localised these proteins with several nuclear markers. Both NSRa and NSRb exhibited co-localisation with the splicing-related proteins SRP34 and snRNP (Fig. 1B) and partial co-localisation with the silencing-related proteins UPF3 and DRB4, key factors for non-sense mediated decay (NMD) and silencing, respectively (Fig. S3B). Hence, AtNSRs mainly co-localise with RNPs housing the splicing machinery in dynamic speckles. Interestingly, recent data have highlighted a strong link between NMD and alternative splicing, as numerous alternatively spliced products are unproductive mRNAs targeted by NMD (Kalyna et al., 2012).

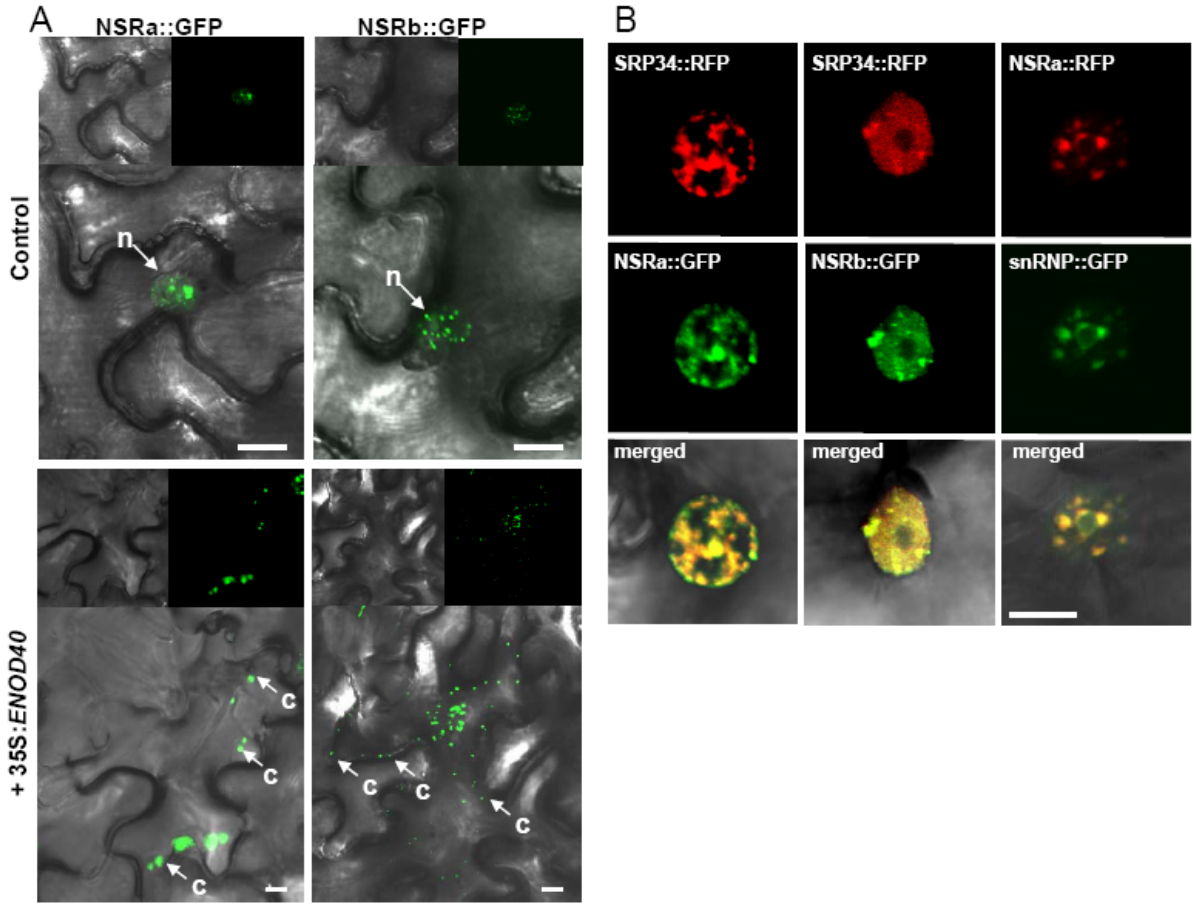


Fig. 1 AtNSRs localised in nuclear speckles and can relocalise into cytoplasmic dots when co-expressed with the lncRNA *ENOD40* .
(A) The AtNSRa:GFP and AtNSRb::GFP fusions localised in nuclear particles or "speckles" (control) and co-expression with a 35S:ENOD40 construct results in relocalisation from nuclear particles into cytoplasmic dots after transient expression in tobacco leaves. n= nucleus; c= cytoplasmic dots. (B) Both AtNSRs::GFP strongly co-localised with splicing-related proteins (SRP34::RFP; AT1G02840) and (snRNP::GFP; AT4G02840) in nuclear "speckles". Scale bars= 8µm

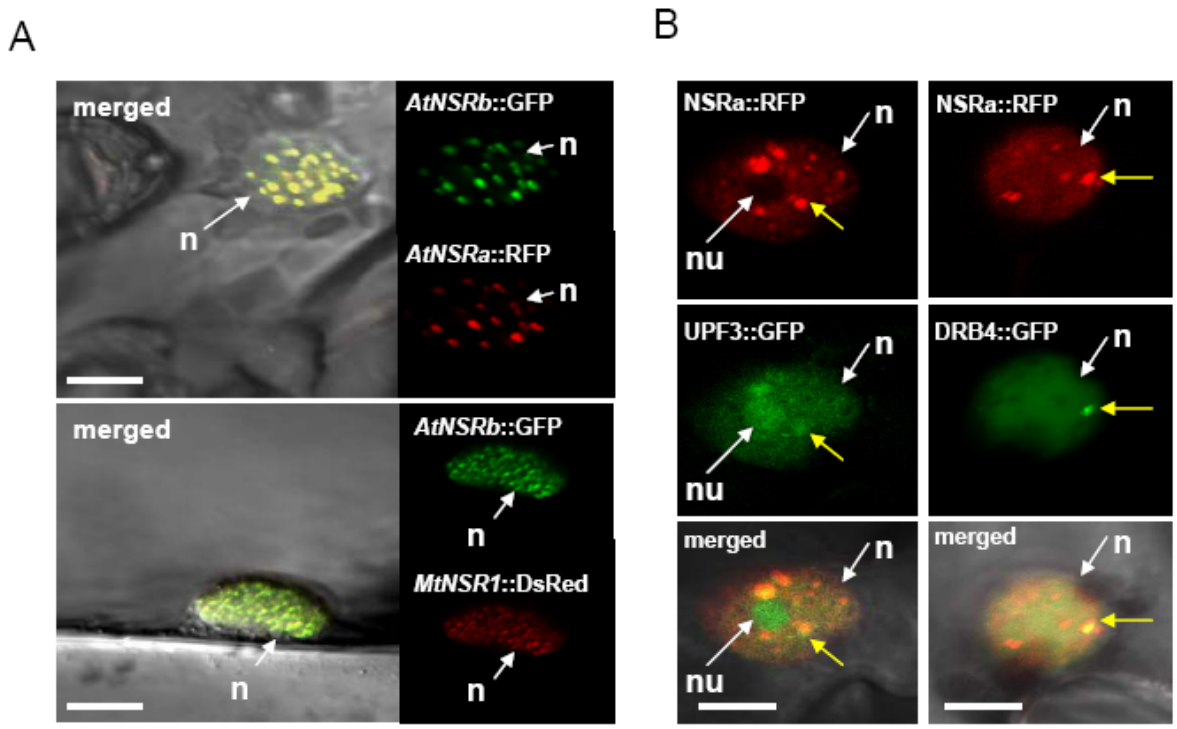


Fig. S3 The NSRs from *Arabidopsis* and *Medicago* co-localised in nuclear speckles and partially co-localised with UPF3 and DRB4 proteins.
(A) The 35S:AtNSRb::GFP and 35S:NSRa::RFP protein fusions co-localised in nuclear particles when transiently expressed in *N. benthamiana* leaves. Note the absence of signal in the cytoplasm. In addition, the 35S:MtNSR1::DsRed (Campalans et al., 2004) and 35S:AtNSRb::GFP fusions also co-localised in nuclear speckles when co-expressed in onion epidermal cells. BAR: 8µm. (B) AtNSRa::RFP partially co-localised with UPF3, a marker of NMD, in nuclear particles but not in the nucleolus whereas the nuclear particle observed with a DRB4-GFP fusion, a marker involved in siRNA biogenesis, also contains AtNSR1::RFP (yellow spot). n= nucleus ; nu= nucleolus ; yellow arrows indicate co-localisation ; scale bars= 5µm

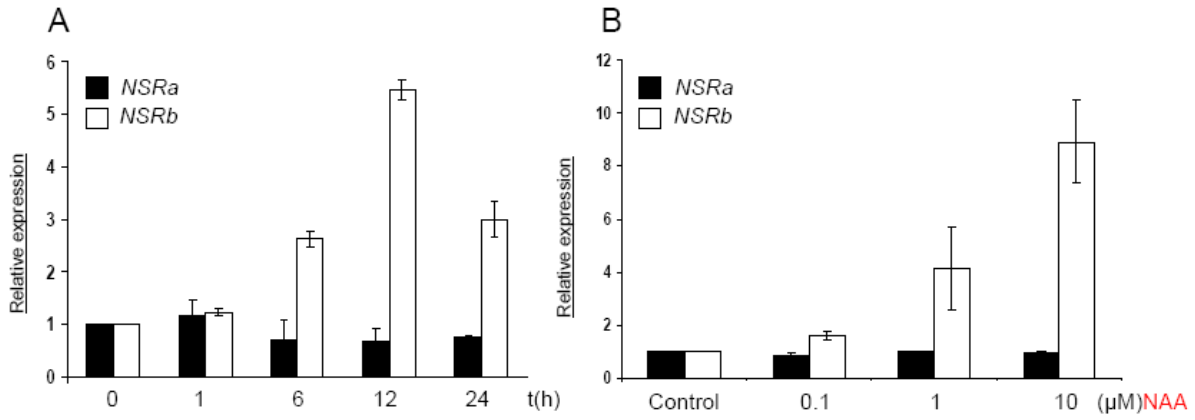


Fig. S4 *AtNSRb* is induced by auxin.
(A) Expression of AtNSR genes in response to auxin. (A) Kinetics of *AtNSRb* induction by auxin treatment (1µM NAA) is shown. In contrast, *AtNSRa* is constitutive. (B) The AtNSRb gene is induced by auxin in a dose-dependent manner. Errors bars indicate standard deviation.

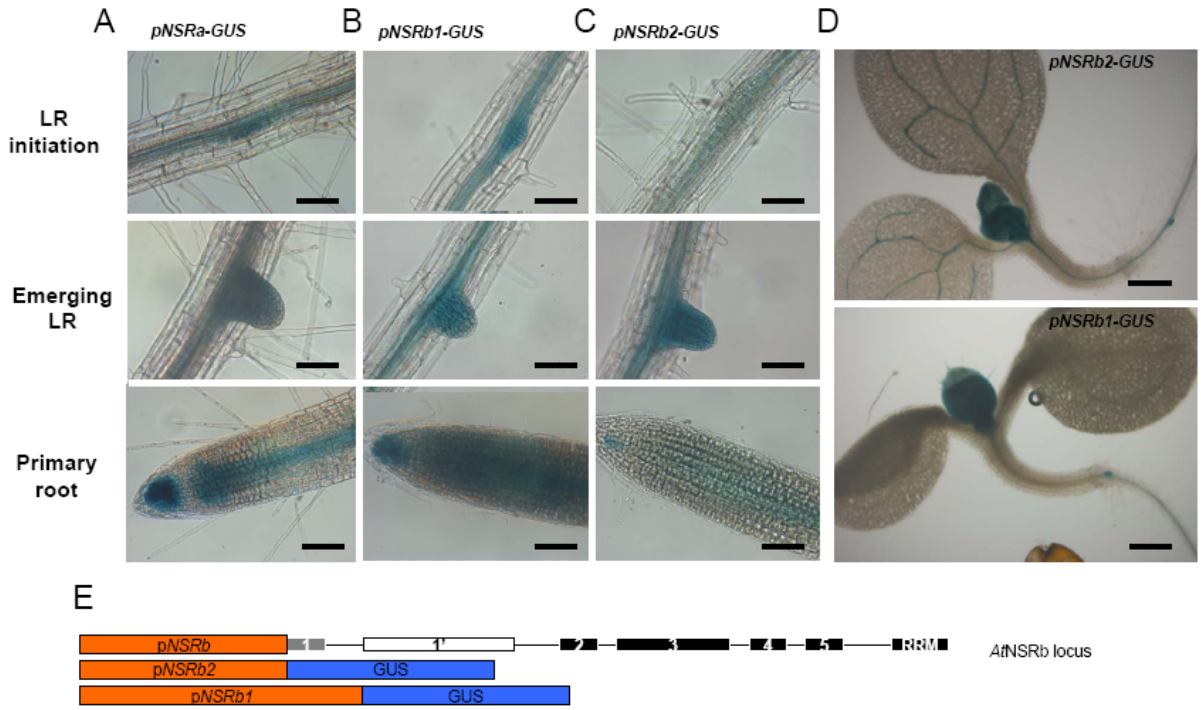


Fig. S5 Expression Analysis of *AtNSR* genes.
(A) The *AtNSRa* promoter fused to *GUS* (*pAtNSRa::GUS*) is active in vascular tissues and the collumella of the primary root and is induced during lateral root formation in dividing primordia. Scale bars = 100µm. (B) and (C) Expression analysis using the two 5' upstream promoters p*AtNSRb1*::GUS and p*AtNSRb2*::GUS (schematized in E) in the primary root and at sites of lateral root formation and emergence. Scale bars=100µm. (D) p*AtNSRb1*::GUS and p*AtNSRb2*::GUS show different expression patterns in vascular tissues of the cotyledons (panel D) whereas both are expressed in the shoot apical meristem. (E) Scheme of the two NSRb promoters fused to exon 1 or exon 1' used in (D). Scale bars=500µm

Arabidopsis databases (www.arabidopsis.org), as well as our previous results (BenAmor et al., 2009), have indicated that *AtNSRs* are induced by auxin. RT-qPCR analysis revealed that *NSRa* is constitutively expressed, whereas *NSRb* is induced by a 12 h treatment of auxin in roots in a dose-responsive manner (Fig. S4A, S4B). The *NSRa* and *NSRb* promoters were fused to the reporter gene *GUS* to determine their spatial expression pattern *in planta* (Fig. S5). Considering the two alternative transcription start sites identified for *NSRb* (Fig. S1), two *NSRb* promoter versions were used (Fig. S5E). The *NSRa* promoter drives *GUS* transcription in the root vascular tissues and in the shoot and root meristems and is induced during lateral root formation. The two *NSRb* promoters are active in vascular tissues, in root and shoot apical meristems and during lateral root formation (Fig. S5B, S5C). However, only the *NSRb2* promoter is active in the vascular tissues of the cotyledons (Fig. S5D). Similarly, GFP-tagged NSRa and NSRb controlled by the *NSRa* endogenous promoter (*NSRb* promoters were too weak to produce detectable GFP signal) were used to determine the localisation of both NSRa and NSRb proteins within the nuclear particles in root meristematic cells, vascular tissues and

all throughout the lateral root primordia (Fig. 2A). The nuclear speckle localisation was further confirmed using a constitutive 35S promoter, which produced more abundant and clearly detectable particles (Fig. 2B). Interestingly, in these plants, NSRs exhibited spontaneous cytoplasmic localisation within cells from the root vascular tissues (Fig. 2C), a tissue normally expressing the *ENOD40* RNA in *Medicago truncatula* (Charon et al., 1999).

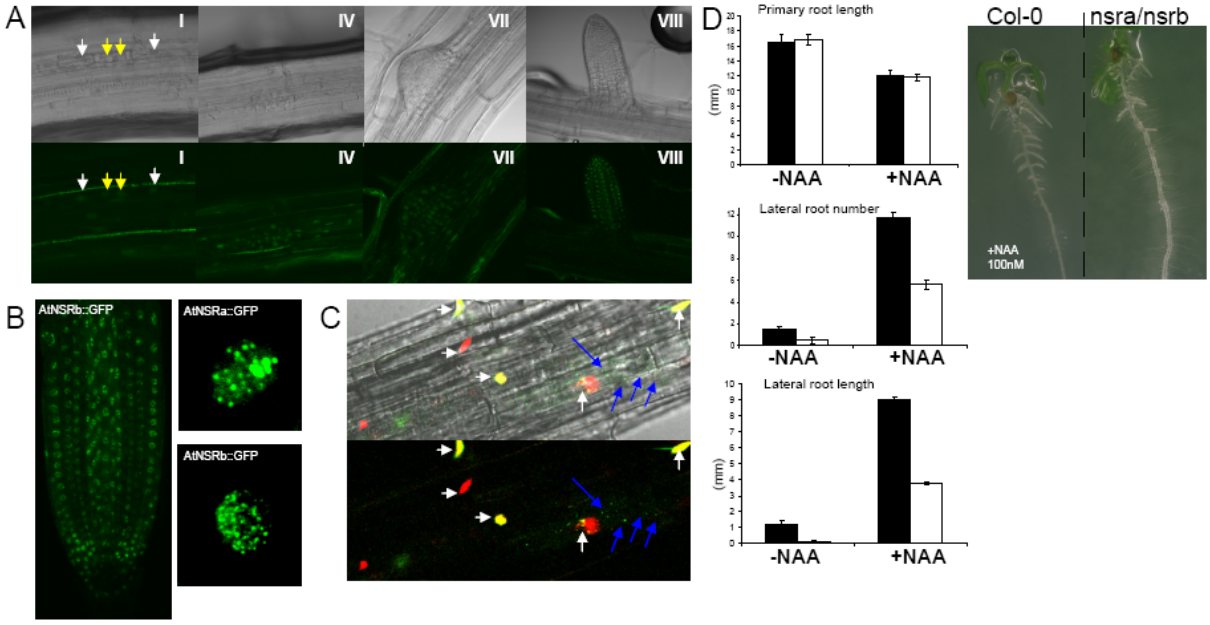


Fig. 2 NSRs are expressed during lateral root formation and the *nsra/nsrb* double mutant show a reduced number of auxin-induced lateral roots.
(A) Expression of *pNSRa:NSRa::GFP* labels nuclear particles during all steps of lateral root formation (I-II,IV,VII,VIII according to Malamy and Benfey 1997). Arrows indicate nuclei; yellow arrows show nuclei after migration during step I of lateral root formation. (B) NSRa::GFP and NSRb::GFP fusions localised in nuclei from primary and emerged lateral roots in several cell types when expressed under the control of its own promoter and showed a nuclear "speckle" pattern related to that observed using a strong promoter (Fig. 1). (C) Spontaneous relocalisation of NSRs from nuclear speckles to cytoplasmic particles is observed in vascular tissues. In these plants, the Histone H2B::RFP marker is also present to mark nuclei in red. White arrows indicate nuclei and blue arrows indicated cytoplasmic re-localisation of NSR particles. (D) The double *nsra*/*nsrb* mutants show reduced number and length of auxin-induced lateral roots (NAA, 100nM). At 7 days after germination under these conditions, the primary root length of the *nsra*/*nsrb* plants is not significantly different than WT in contrast to the lateral root number and total lateral root length. These phenotypes are minor but significant in the absence of NAA and become more accentuated in response to auxin, as shown in the images on the left (wt and double mutant 7 days after germination treated with auxin).

To address the physiological role of NSRs, we isolated simple *nsra* and *b Arabidopsis* mutants, which exhibited no remarkable phenotypes. However, 7-day-old double mutant *nsra/nsrb* plants exhibited lateral roots that were significantly reduced in number and length in response to auxin (fig. 2D), despite their similar main root length in relation to wild-type (WT) plants. This phenotype suggests that the double *nsra/nsrb* mutants are less sensitive to

auxin (Péret et al., 2009). In contrast, transgenic lines overexpressing *NSRa* or *NSRb* yielded dwarf plants, a phenotype that correlates with transgene expression levels (Fig. S6). This dwarf phenotype is likely attributed to the presence of smaller cells, which are clearly visible in the leaf epidermis (Fig. S6B).

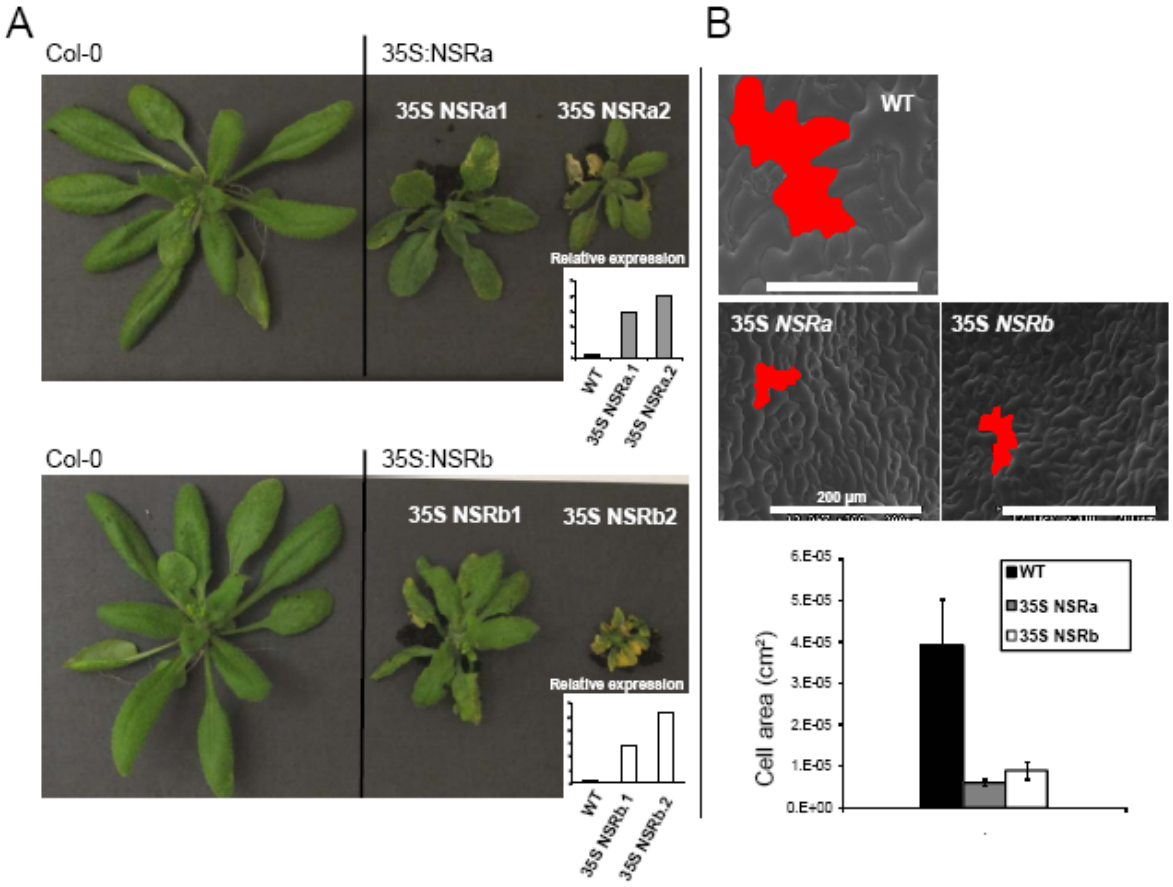


Fig. S6 Two independent transgenic lines over-expressing AtNSRa yield dwarf plants.
Over-expression of AtNSRa or AtNSRb using the strong 35S promoter resulted in dwarf plants, a phenotype that correlated with transgene expression level (Errors bars indicate standard deviation; expression levels in inset). (B) The 35S:NSRs plant have smaller cells in the leaf epidermis. Quantification of the area of at least 83 cells from epidermis leaves from WT and 35S:NSRa or 35S:NSRb. Errors bars indicate confidence interval 5%.

To unravel the mechanism of action of NSRs, we analysed the molecular response to auxin in the *nsra/nsrb* double mutants. A transcriptomic approach revealed that over 2200 genes were differentially regulated in the double mutants in comparison to WT plants after auxin treatment (12 h 10 µM NAA). In contrast, only 535 gene expressions change were detected in control conditions (Fig. S7A; accession no. GSE 39659). Interestingly, among the differentially expressed genes, we also identified 11 lncRNAs that were deregulated in double mutant plants, representing approximately 15% of the lncRNAs that we had identified previously in *Arabidopsis* (Ben Amor et al., 2009) (Fig. S7B). For at least two cases, the

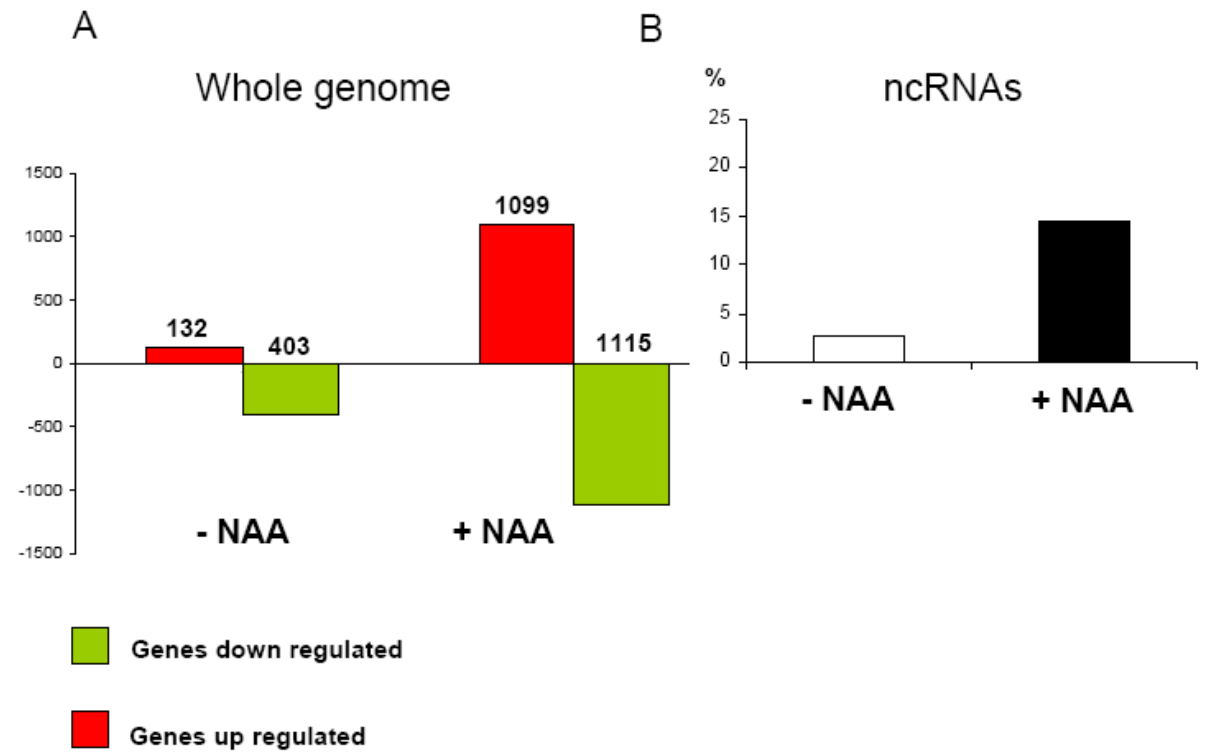


Fig. S7 Global transcriptome changes in response to auxin are perturbed in double *Atnsra/nsrb* mutants.

Transcriptome analysis of the double *Atnsra/nsrb* mutant showed (A) only 535 genes differentially regulated compared to WT whereas an auxin treatment revealed 2214 genes with expression changes under the same criteria when compared to WT (CATMA arrays V6). (B) Considering 76 ncRNAs from (Ben amor et al., 2009) around 2% are differently regulated in control condition between WT and nsra/nsrb double mutant. In contrast, after auxin treatment, 15% of these ncRNA are differently regulated between them ($p<0,0001$).

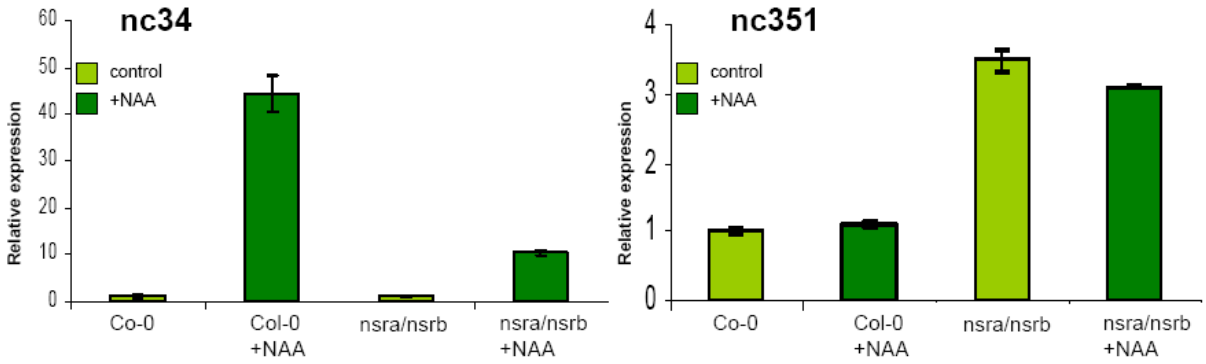


Fig. S8 Two lncRNA miss-regulated in the *nsra/nsrb* double mutant background.

The lncRNA nc34 is strongly induced by auxin, an accumulation significantly lost in the double mutant. The ncRNA 351 is up-regulated in the *nsra/nsrb* double mutant in both control and auxin conditions. Errors bars indicate standard deviation.

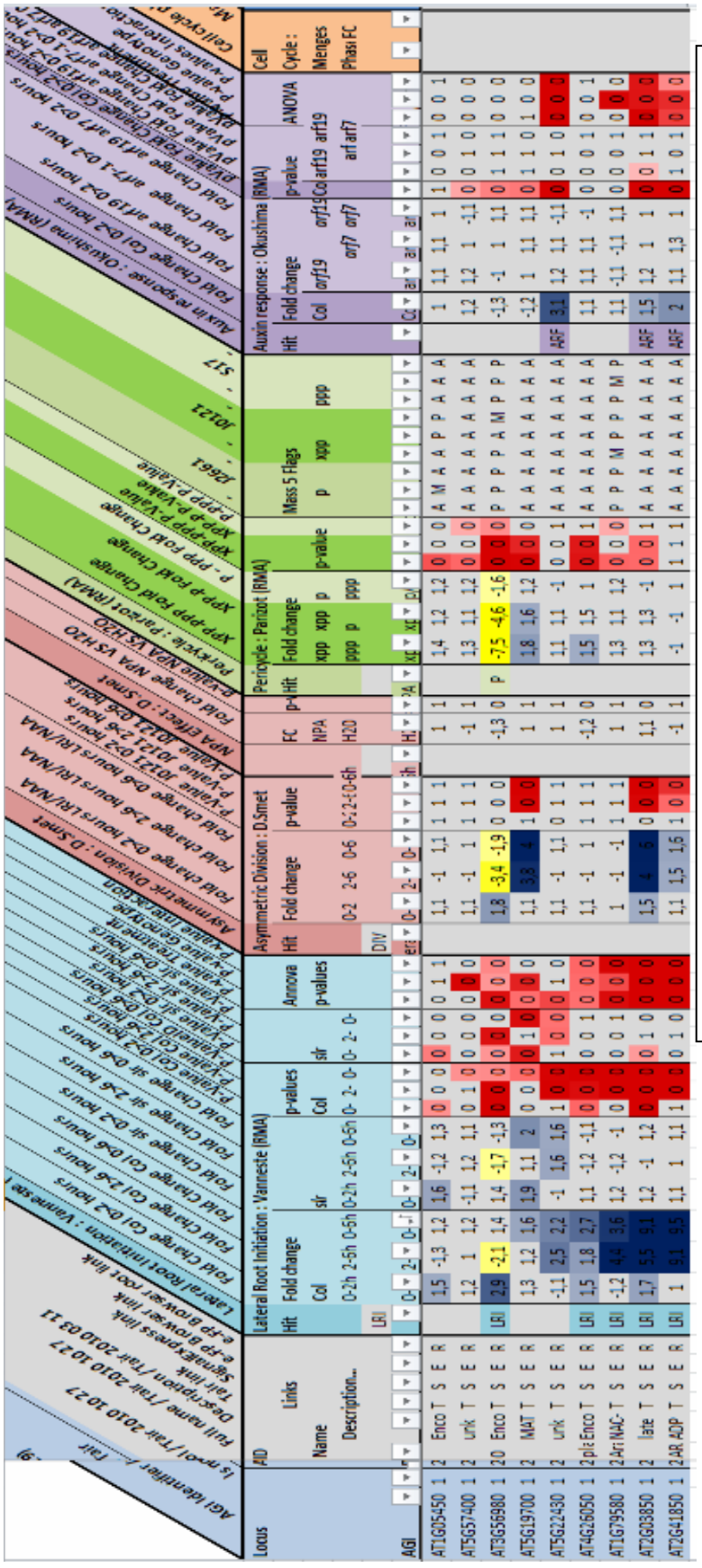
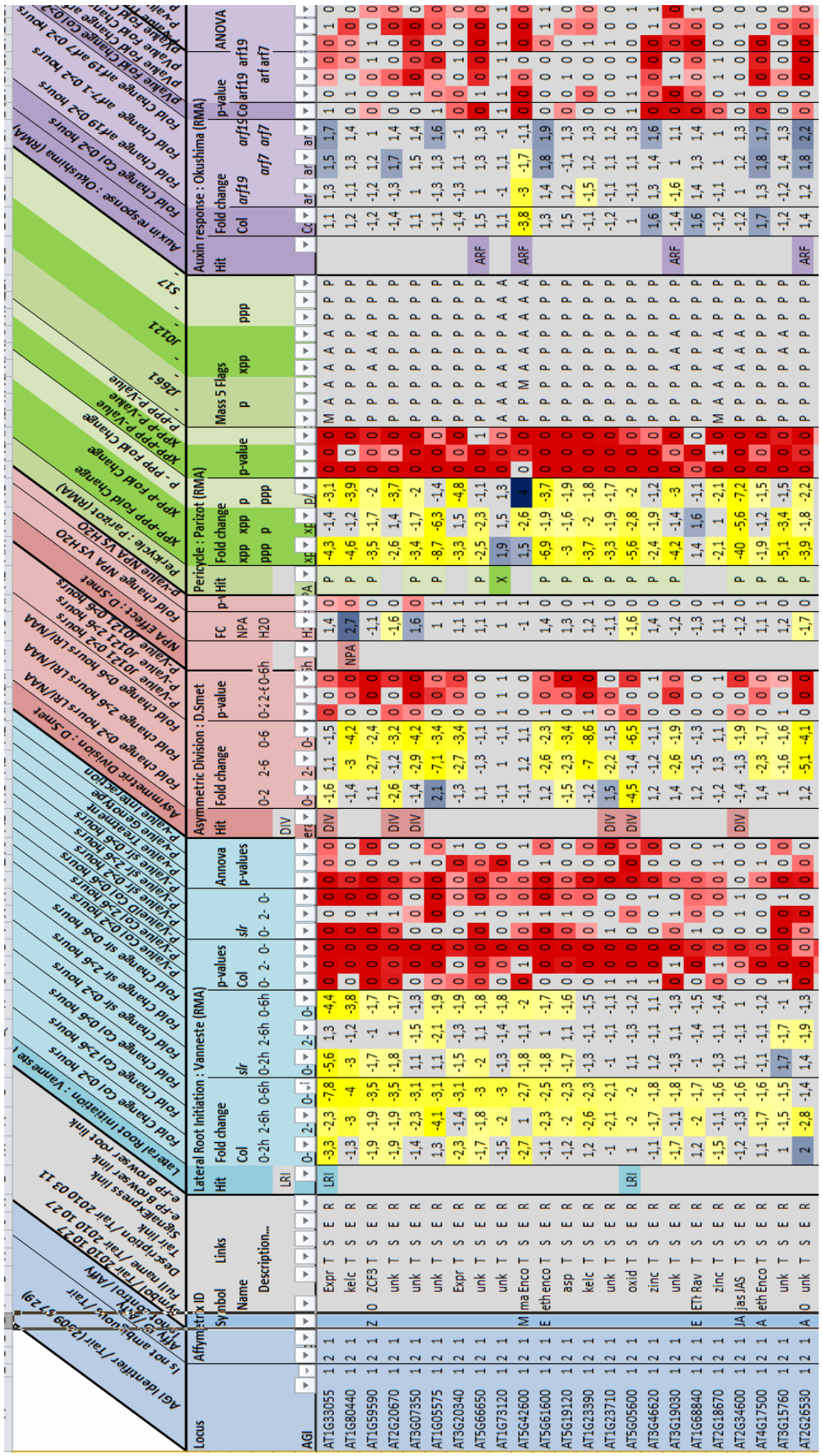


Figure S9 Action of NSRs in LR initiation in known signalling pathways.

Strikingly, according to the Visual Lateral Root Transcriptome Compendium (Visual LRTC, Parizot et al., 2009), A) out of 49 genes that are the most down-regulated in *nsra/nsrb* compared to WT plants in response to auxin, 7 are auxin-responsive genes. Out of these 7 genes, 5 are linked to lateral root initiation (dependent on *SOLITARY ROOT-1*, Fukaki et al., 2002) and 3 are dependent on ARF7 and/or ARF19, key regulators of lateral root development (Okushima et al., 2005). B) Among 100 most overexpressed transcripts in the *nsra/nsrb* compared to WT plants in response to auxin, we found 28 genes that are normally repressed by auxin, of which 3 are also linked to root initiation, whereas 6 are dependent on the ARF genes. These results integrate the action of NSRs in LR initiation in known signalling pathways.

Fig. S9B



auxin-inducible nc34 or the constitutive nc351, their differential behaviour was confirmed using RT-qPCR (Fig. S8A, S8B). These results link NSRs to the regulation of gene expression of both mRNAs and lncRNAs in response to auxin. Strikingly, according to the Visual Lateral Root Transcriptome Compendium (Visual LRTC, Parizot et al., 2009), out of 49 genes that are down-regulated in *nsra/nsrb* compared to WT plants in response to auxin, 7 are auxin-responsive genes. Out of these 7 genes, 5 are linked to lateral root initiation (dependent on *SOLITARY ROOT-1*, Fukaki et al., 2002) and 3 are dependent on ARF7 and/or ARF19, key regulators of lateral root development (Okushima et al., 2005, Fig. S9A). On the other hand, among the 100 overexpressed transcripts in the *nsra/nsrb* compared to WT plants in response to auxin, we found 28 genes that are normally repressed by auxin, of which 3 are also linked to root initiation, whereas 6 are dependent on the ARF genes (Fig. S9B). These results integrate the action of NSRs in LR initiation in known signalling pathways.

Considering that NSRs co-localise with spliceosome markers, we explored the impact of NSRs in splicing to address the mechanism by which NSRs might modulate gene expression. As no major defects in constitutive splicing were observed, we analysed 288 known alternatively spliced genes (Simpson et al., 2008). By comparing RNA from root tissues that were treated or not with auxin in WT and *nsra/nsrb* mutant plants, we found significant changes for 103 genes. In fact, auxin induces a major change in the relative amounts of alternatively spliced isoforms of these transcripts (180 genes of 288 known alternative genes; Fig. 3A), and the majority of these AS events (103) were dependent on NSRs. We then confirmed these variations using RT-PCR and a fluorescent gel assay for several genes, including the *ATPase1* (At1g27770) gene, which is normally transcribed as only one main isoform during both control and auxin-treated conditions but exhibits an intron retention event in the *nsra/nsrb* double-mutant treated by auxin (Fig. 3B, S10A and S10E). Other relevant examples of AS defects are the auxin-related gene (At2g33830) and the F-box gene (AT4G27050) transcripts, which are normally spliced into two isoforms during control conditions but exhibit differential splicing in response to auxin (Fig. 3C, 3D, S10B, S10C and S10E). This auxin-induced splicing variation is not detected in the *nsra/nsrb* mutant background (note in Fig. 3C the lack of variation of the second isoform in the *nsra/nsrb* mutant). These results strongly suggest that NSRs represent novel alternative splicing regulators in plants.

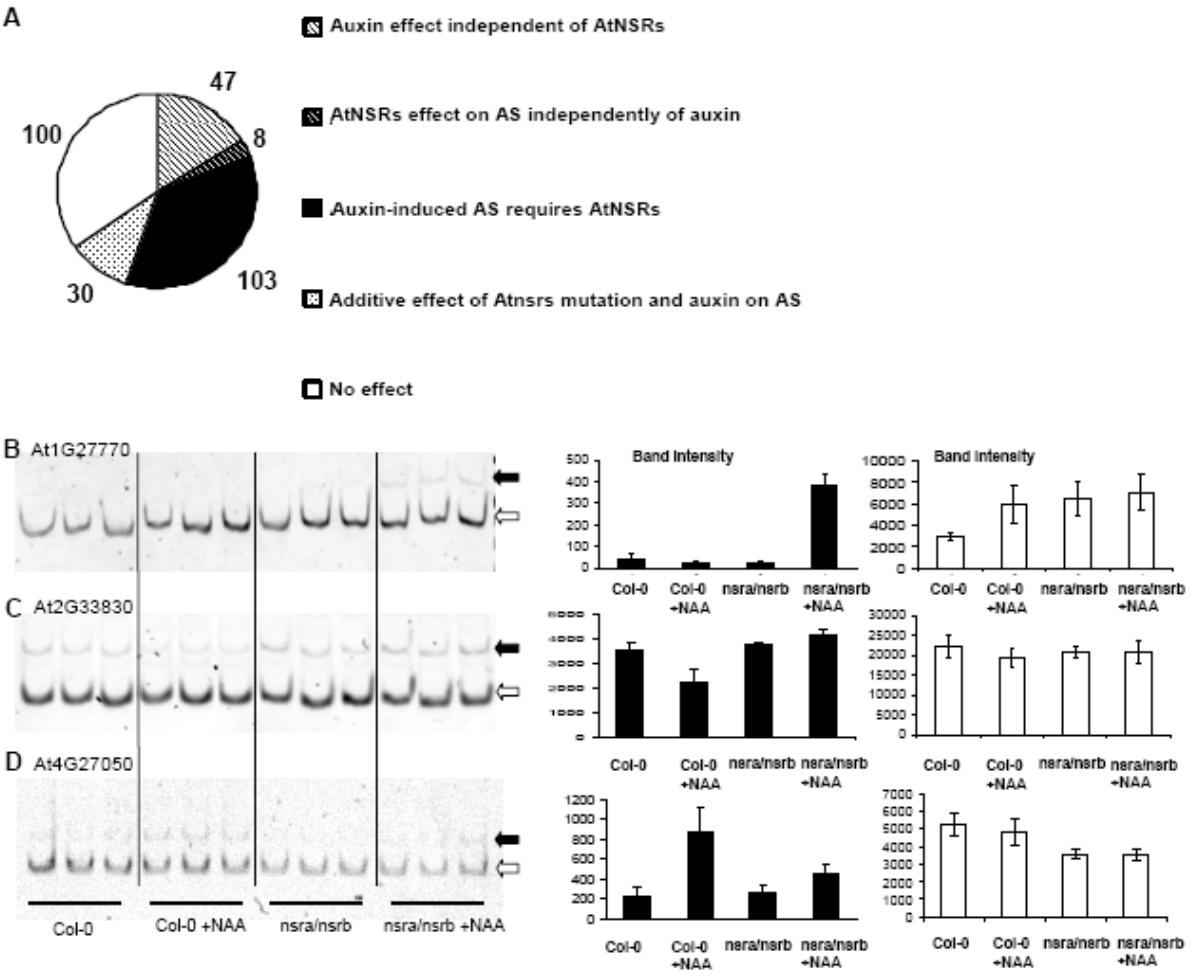


Fig. 3 Modification of AS in *tnsra/nsrb* mutant during auxin response.
(A) Proportion and gene number derived from the analysis of the High resolution RT-PCR panel coupled to HPLC in the *nsra/nsrb* double mutant treated or not with 1µM of auxin in comparison to WT. (B) The ATPase1 (At1G27770) is transcribed in a single mRNA in control conditions or during auxin treatment (1µM) whereas in the *nsra/nsrb* mutant treated by auxin a second isoform (an intron retention event; Fig. S10 A) is detectable. (C) The auxin related protein (At2G33830) expressed two isoforms which differ by an intron retention event (Fig. S10 B). The intron retained isoform is less detectable after auxin treatment, an auxin-induced variation lost in *nsra/nsrb* plants treated with auxin. (D) The F-box protein (At4G27050) is mainly transcribed in two isoforms, (Fig. S10 C) and, after auxin treatment, the intron-retained isoform is accumulated. This auxin-induced AS event is also lost in the *nsra/nsrb* mutant. Arrows indicate each AS isoform in PAGE gels and their shading correspond to its quantification in the accompanying bar graph panels.

As the *ENOD40* RNA interacts with AtNSRs to induce their cytoplasmic re-localisation, we wondered whether the heterologous expression of the *MtENOD40* RNA might affect alternative splicing in *Arabidopsis thaliana*. To this aim, we generated *Arabidopsis* lines expressing the *ENOD40* transcript and assayed the splicing of select NSR targets. Remarkably, the auxin-related (At2g33830) and *CCA1* (At2g46830) genes exhibited drastic changes in the ratio between isoforms in *ENOD40*-expressing plants (Fig. 4A, 4B, S10B, S10D and S10F), linking *ENOD40* expression to significant alterations of alternative splicing.

To further demonstrate a direct molecular link between NSRs and the alternatively spliced mRNAs, as well as lncRNAs, we performed an *in vivo* UV cross-linked RNA immunoprecipitation (RIP) experiment using NSRa or NSRb proteins that were fused to an HA-tail and expressed under the control of the endogenous NSRa promoter. These NSRa or NSRb fusion constructs were introgressed into the *nsra* or *nsrb* mutants, respectively. The RIP results revealed that NSRa and NSRb specifically bind to certain alternatively spliced mRNAs targets and to the *Arabidopsis* lncRNA nc351 *in planta* using either total or nuclear RNA (Fig. 5A and 5B).

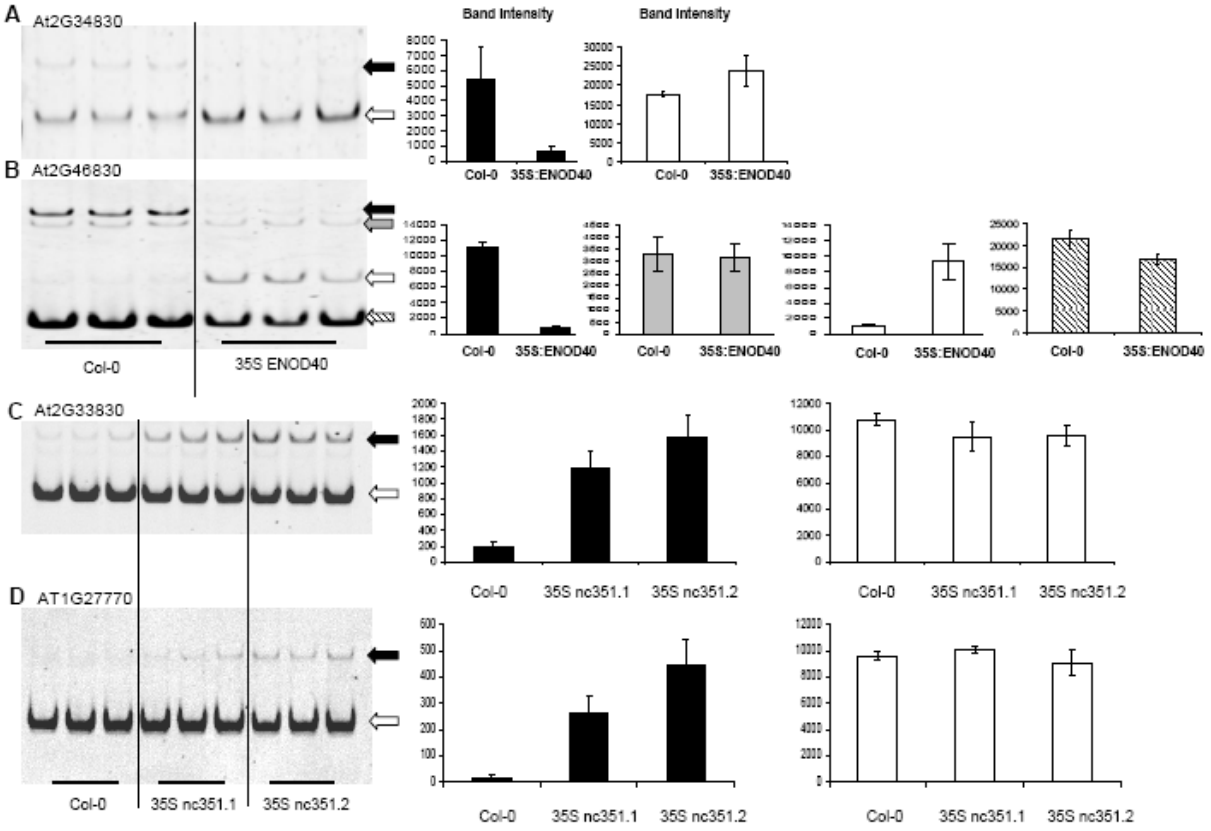


Fig. 4 Modification of AS in Arabidopsis plants expressing the *MtEnod40* or over expressing *Atnc351* lncRNAs.

(A) The auxin related protein (At2G33830) (shown in Fig.3C) also showed a change in isoform distribution in plants over-expressing the lncRNA *MtENOD40*. The intron retained isoform is less detectable in plant over expressing *MTENOD40* (Fig. S10 B). (B) The CCA1 gene (At2G46830) transcribed into at least 4 isoforms in control condition (Fig. S10D), presents a very significant switch of isoforms when *ENOD40* is expressed. (C) The auxin related protein (At2G33830) (shown in Fig.3C, 4A) showed a change in isoform distribution in plants over-expressing the lncRNA nc351. (D) The ATPase1 (At1G27770) is transcribed in a single mRNA in control conditions whereas in plants over expressing nc351 a second isoform (an intron retention event; Fig. S10A) is detectable, the same pattern observed in the *nsra/nsrb* double mutant treated with auxin (Fig. 3B). Arrows indicate each AS isoform in PAGE gels and their shading correspond to its quantification in the accompanying bar graph panels.

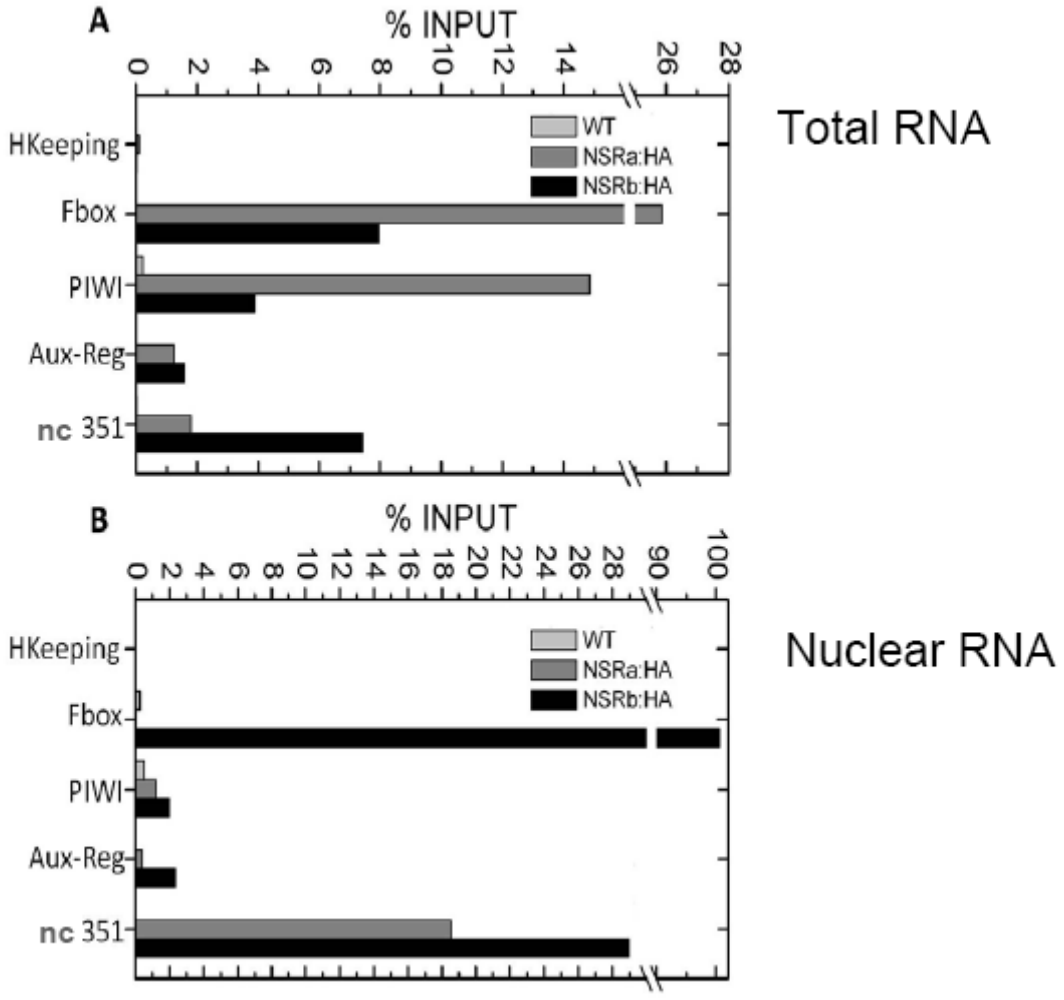


Fig. 5 NSRa and NSRb bind *in vivo* to alternatively spliced mRNAs and the lncRNA nc351.

RNA Immunoprecipitation (RIP) assays using HA-tagged NSRa or NSRb on (A) total cell lysates or (B) nuclei extracts of 10-day-old seedlings treated with 10 µM NAA for 24 h. Results of RT-qPCR are expressed as the percentage of the respective INPUT signal (total signal before RIP). Analyzed genes are HKeeping: AT1G13320 (Czechowski et al., 2005); Fbox: AT4G27050; PIWI factor: AT2G29210; Auxin-regulated gene: AT2G33830; ncRNA351: AT1G67105.

Globally, the auxin-inducible *NSRb* gene codes for a protein exhibiting increased affinity for its targets compared with NSRa. In the nuclear fraction, the enrichment of ncRNA351 in the NSRb::HA RIP is increased, supporting the action of lncRNA nc351 in nuclear-related processes (Fig. 5B). Remarkably, this ncRNA, which is a NSR direct target, corresponds to a lncRNA transcriptionally regulated in the *nsra/nsrb* double mutant (Fig. S8B). To further confirm the potential modulation of AS by NSR-interacting lncRNAs, we overexpressed the lncRNA nc351 in *Arabidopsis* plants. Two NSR-mRNA targets (encoding either the Auxin-Regulated Protein gene (At2g33830) or the ATPase1 gene (At1g27770), exhibited AS

changes in these plants (Fig. 4C, S10B and S10G). Moreover, the latter NSR target exhibited the same alteration in AS than the one observed in the *Atnsra/nsrb* double mutant. This intron retention occurs in plants overexpressing nc351 (Fig. 4D, S10A and S10G), suggesting that nc351 overexpression prevents NSR action. Hence, NSR-containing complexes bind to alternatively spliced NSR targets and lncRNAs, and the modulation of lncRNA expression can affect the AS of NSR-mRNA targets.

Our results demonstrate that NSRs are plant RNA-binding proteins that localise within nuclear speckles and regulate alternative splicing. The binding of the NSR-containing complexes to AS targets suggests that NSRs might modulate the action of splicing factors in a specific part of the nucleus where splicing takes place (Lorkovic et al., 2008). The growing importance assigned to AS in plants has been recently addressed using RNA sequencing techniques (Marquez et al., 2012), and the functional diversity of AS isoforms has been elegantly demonstrated for two isoforms of SR45 (Zhang and Mount, 2009). These isoforms differ by only eight amino acids, including a putative phosphorylation site, and can differentially complement petal or root developmental phenotypes in a *sr45* mutant (Zhang and Mount, 2009). Therefore, isoforms with very similar sequences can elicit substantially distinct morphological outcomes in different organs of the plant. The interaction of NSRs with ncRNAs might lead to drastic changes in developmental fates such as those occurring during nodule and lateral root formation, two processes where the *ENOD40* RNA is strongly expressed (Mathesius et al., 2000). Recent studies revealed a novel class of *sno-lncRNA* (a class of nuclear-enriched intron-derived lncRNAs processed on both ends by the snoRNA machinery) that may be strongly associated with the Fox family of splicing regulators and may alter patterns of splicing (Yin et al., 2012). As NSRs can also interact with lncRNAs to affect AS, we propose that lncRNAs can modulate the function of the splicing machinery through their interaction with NSRs. This hypothesis is supported by the molecular phenotype of the *nsra/nsrb* double mutant and the lncRNA *ENOD40* and nc351 overexpressing lines described in this work. In mammals, another lncRNA, *malate1*, can modulate alternative splicing in HeLa cells by interfering with protein phosphorylation (Tripathi et al., 2010). Here, we introduce a mechanism by which the interaction of lncRNA with NSR AS regulators can modulate alternative splicing during specific developmental transitions, such as the formation of new lateral roots by differentiated pericycle cells. We speculate that such lncRNAs mimic NMD intron-retained transcripts to hijack the splicing machinery and induce rapid changes in alternative splicing patterns.

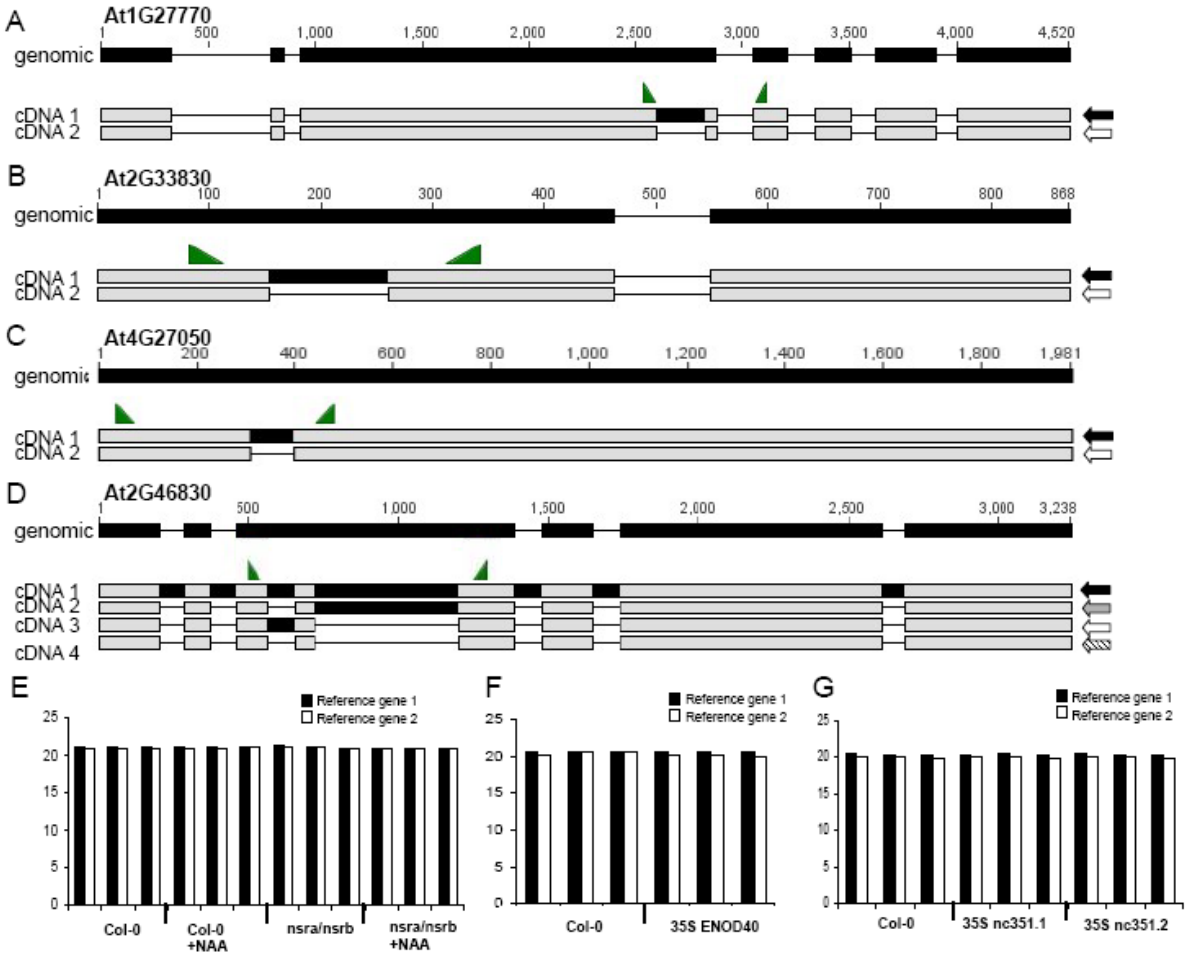


Fig. S10. Primer pairs used and AS gene models for the mRNA targets analysed in Figs. 3 and 4.

(A,B,C,D) gene models for alternatively spliced targets (according to www.arabidopsis.org): For each gene the different isoforms are represented. Primer pairs used in Fig. 3 are represented in green. Arrow shading correspond to shade arrows and quantification bars in Fig. 3 to design the specific isoform amplified (thick bars are included). (E,F,G) q-RT PCR of two constitutive genes (reference gene 1: AT1G13320 and reference gene 2: AT4G26410) on cDNA from biological triplicates of the experiments depicted in (Fig. 3B, 3C, 3D) ; (Fig. 4A,4B) and (Fig. 4C, 4D) respectively. These values showed the excellent calibration of the cDNAs used for the quantitative analysis of AS.

Materials and methods

Plant and bacteria material and growth conditions.

All mutants were in the Columbia-0 (Col-0) background. *Atnsra* (SALK_003214) and *Atnsrb* (Sail_717) were from the SALK and SAIL T-DNA collections, respectively. The plants were grown in long day (16-h light/8-h dark) or continuous light conditions at 23°C on soil or on solid half-strength MS medium. *E.coli* DH5α was used for subcloning.

Transient expression in Tobacco leaves and biolistics

All transient expression experiments were performed in *Nicotiana benthamiana,* infiltrated with 0.1 OD 600 nm *Agrobacterium thumefaciens* AGL-0 in infiltration medium containing half-strength MS medium, 0.01 M MgCL2, 0.01 M MES pH 5.2, 0.001 M acetosynringone. All constructs were co-expressed with HCpro (Anandalakshmi et al., 1998). For the co-expression of AtNSRb:GFP and MtNSR1:RFP, onion epidermis was bombarded with the two constructs. Biolistic introduction of plasmid DNA into plant cells was performed using the PDS-1000/He System (Bio-Rad). For each experiment, bombardments were performed at least five times, and an average of 200 transformed cells was obtained per bombardment. Equimolar quantities of plasmids were used in all cases. Epidermal tobacco leaves or onion cells were observed in the confocal microscope.

Confocal Microscopy and Image Processing

Cells were observed 48h after transient expression using a Leica DM RxA2 confocal microscope. The Leica confocal software was used for image acquisition and for the quantification of fluorescence profiles. Sequential scans were performed when necessary. Spectral profiles were calculated for five cells. Data processing was performed using ImageJ (http://rsbweb.nih.gov/ij/).

RT-PCR and RT-qPCR

Total RNA was prepared from roots and plantlets at different developmental stages using the Qiagen RNeasy plant mini kit. The DNAse treatment was performed according to the manufacturer's protocols. For reverse transcription with SuperScriptII (Invitrogen), 2.5 µg of total DNAse-treated RNA was used. One microliter of the resulting cDNA solution was used for RT-PCR or RT-qRT analyses using standard protocols. A complete list of PCR primers is available in (Table 1). Each cDNA sample was precisely calibrated and verified for two

constitutive genes (AT1G13320; AT4G26410; Czechowski & al., 2005). For RT-PCR, the amplification was as follows: one cycle of 4 min at 98°C, 26 cycles of 30 s at 98°C, 30 s at 59°C, and 1 min at 72°C. The products were separated in 7.5% acrylamide gel stained with SyBr green (Invitrogen) and revealed by Pharos Imager (Biorad). Band profiles were quantified using ImageJ (http://rsbweb.nih.gov/ij/). RT-qPCR was performed using a Roche Light Cycler 480 using standard protocols (40 cycles, 60°C annealing).

High resolution RT-PCR

The original panel (Simpson et al. 2008) was expanded to 289 primer pairs by identifying alternative splicing events that were either published or annotated in The Arabidopsis Information Resource (TAIR8 http://www.arabidopsis.org/) or in the Alternative Splicing in Plants database (ASIP, http://www.plantgdb.org/ASIP/). Primer pairs in which one primer was fluorescently labelled were designed as described previously (Simpson et al. 2008). Primer pairs used are listed in Supplementary Table S1. RT-PCR analysis was performed as described previously (Simpson et al. 2008). In brief, RT reactions were performed with total RNA using oligo-dT primers. The first-strand cDNA was aliquoted into microtitre plates, and PCR using the gene/alternative splicing event-specific primers was performed using 24 cycles. We have previously shown that 24 cycles still maintain the linear amplification range for various splicing substrates using [32P]-labelling (57) and for a number of AS primers that were used here to amplify transcripts of different quantities and sizes (Simpson et al. 2008). Our high-resolution RT-PCR system is capable of detecting multiple different AS transcripts from a single gene, distinguishing alternative splicing events involving small size differences in transcripts (as few as 2–3 nt) and identifying small but significant changes in the ratios of alternatively spliced variants. The AS variants for each of the genes were amplified simultaneously using the same primers within the same reaction. The different AS isoforms usually contain substantially conserved sequences, which reduce variation in amplification efficiency. In addition, if there were differences in amplification efficiency among particular AS isoforms, these differences occurred in the PCR reactions with the WT and mutants that were treated with or without auxin. Electropherograms produced by the ABI 3730 genotyping software identified the exact size of the RT-PCR products for each primer pair. The peak areas for each RT-PCR product were extracted from three replicates; the ratios of the different peaks were calculated to generate a mean value and standard error for each AS transcript as a percentage of the total transcript across the three replicates.

Transcriptomic Studies

Microarray analysis was performed at the Unité de Recherche en Génomique Végétale (Evry, France) using the CATMAv6.1 array, which is based on Roche-NimbleGen technology. A single high-density CATMAv6.1 microarray slide with twelve chambers was used; each chamber contained 135,000 primers representing all of the *Arabidopsis thaliana* genes: 30,834 probes corresponding to TAIRv8 annotation (including 476 probes of mitochondrial and chloroplast genes) + 1,289 probes corresponding to EUGENE software predictions. Moreover, the array included 5,352 probes corresponding to repeat elements, 658 probes for miRNA/MIR, 342 probes for other RNAs (rRNA, tRNA, snRNA, soRNA) and 36 controls. Each long primer represents a triplicate in each chamber for robust analysis. Two independent biological replicates were produced. For each biological repetition and each point, RNA samples were obtained by pooling RNAs from 10 plants. Plantlets were collected at 1.04 developmental growth stages (Boyes, 2001) and cultivated in ½ MS conditions. Total RNA was extracted using the Qiagen RNeasy kit according to the manufacturer's protocol. For each comparison, one technical replicate with fluorochrome reversal was performed for each biological replicate (i.e., four hybridisations per comparison). The labelling of cRNAs with Cy3-dUTP or Cy5-dUTP (Perkin-Elmer-NEN Life Science Products) and hybridisation to the slides were performed as described in Lurin *et al*., 2004. Two-micron scanning was performed using an InnoScan900 scanner (Innopsys^{R}, Carbonne, FRANCE), and the raw data were extracted using Mapix^{R} software (Innopsys^{R}, Carbonne, FRANCE).

Statistical Analysis of Microarray Data

Experiments were designed with the statistics group of the Unité de Recherche en Génomique Végétale. For each array, the raw data comprised the logarithm of median feature pixel intensity at wavelengths of 635 nm (red) and 532 nm (green). For each array, a global intensity-dependent normalisation was performed using the Loess procedure (Yang et al., 2002) to correct the dye bias. The differential analysis is based on the average log-ratios over duplicate probes and over technical replicates. Hence, the number of sets of available data for each gene equals the number of biological replicates, and that number was used to calculate the moderated t-test (Smyth, 2004).

Under the null hypothesis, no evidence for specific variance between probes is highlighted by Limma, and consequently, the moderated t-statistic is assumed to exhibit a standard normal distribution. To control the false discovery rate, adjusted p-values were calculated using the

optimised FDR approach (Storey et al, 2003). The probes with an adjusted p-value ≤ 0.05 were considered to be differentially expressed.

Statistical analysis was performed using the R software. The function SqueezeVar of the limma library was used to smooth the specific variances by computing empirical Bayes posterior mean values. The kerfdr library was used to calculate the adjusted p-values.

Data Deposition

Microarray data from this article were deposited at Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/), accession no. GSE 39659) and at CATdb (http://urgv.evry.inra.fr/CATdb/; Project: Blanc-08-01_2012_01_RNAPATHs_nsr) according to the "Minimum Information About a Microarray Experiment" standards.

RNA Immunoprecipitation

Atnsra and *Atnsrb* simple mutant lines were complemented with the constructs pNSRa:NSRa::HA::HA and pNSRa:NSRb::HA::HA, respectively. Ten-day-old plants were irradiated 3 times with UV using a UV cross-linker CL-508 (Uvitec) at 0.400 J/cm². Plants were grinded in liquid nitrogen, and total proteins were extracted in RIP extraction buffer (50 mM Tris-HCl pH 7.5; 150 mM NaCl; 5 mM MgCl2; Triton 0.1%; 10% Glycerol). The extracted suspension was filtered twice through Miracloth. After centrifugation for 15 min at 4°C 4000 RPM, the pellet was resuspended in 800 µl of nuclei lysis buffer + SDS (0,1% SDS; 10 mM EDTA; 50 mM Tris-HCL pH 7,4), as well as 20 µl of RNAse inhibitor and 20 µl of proteinase inhibitor. After a 1 h incubation at 4 °C in rotation, the sample was centrifuged at 1500 RPM for 5 min at 4°C. One hundred microliters of supernatant was used to prepare RNA for the input sample. To purify the NSR proteins, we used µMACs magnetic technology columns (Mitenyi). First, the columns were prepared following the manufacturer's protocols. The extracts were incubated for 1 h with the magnetic beads, which were coupled with HA antibodies and passed through the columns. Several wash steps were performed. First, the columns were washed twice with 200 µl of nuclei lysis buffer (0,1% SDS), twice with 200 µl of the Miltenyi washing buffer 1 (W1), twice with 200 µl of the Miltenyi washing buffer 2 (W2) and, finally, once with 200 µl of sterile water. The beads were collected and subjected to a native elution as indicated by the manufacturer's protocol. The resulting solutions were treated with proteinase K (RNAse-grade, Invitrogen) in 2 µl of RNAse inhibitor at 55°C for 1

h. Lastly, we passed the samples through new columns to exclude the beads. The total RNA was extracted from the template and from the input using Trizol reagent (Invitrogen) and then was resuspended in 60 µl of H_2O. The templates and the inputs were treated with DNAse, and subsequent RT-qPCR reactions were performed using specific primers.

Funding

This work was supported by the Agence Nationale de la Recherche ANR-08-BLAN-0082 (to ACM and AM); and ANR-10-LABX-40 (to MC and HV).

Acknowledgments:

This work has benefited from the facilities of the Imagif Cell Biology Unit of the Gif campus (www.imagif.cnrs.fr) which is supported by the Conseil Général de l'Essonne. Scanning electron microscopy (SEM) Pictures were obtained with the help of Séverine Domenichini (Institute de Biologie des Plantes). ncRNA34 expression was detected with the help of Olivier Vandepute (Université libre de bruxelles).

References:

Anandalakshmi, R., Pruss, G.J., Ge, X., Marathe, R., Mallory, A.C., Smith, T.H., and Vance, V.B. (1998). A viral suppressor of gene silencing in plants. Proceedings of the National Academy of Sciences of the United States of America **95**: 13079–84.

Barta, A., Kalyna, M., and Lorković, Z.J. (2008). Plant SR proteins and their functions. Current topics in microbiology and immunology **326**: 83–102.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., Kuehn, M.S., Taylor, C.M., Neph, S., Koch, C.M., Asthana, S., Malhotra, A., Adzhubei, I., Greenbaum, J.A., Andrews, R.M., Flicek, P., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature **447**: 799–816.

Blencowe, B.J. (2006). Alternative splicing: new insights from global analyses. Cell **126**: 37–47.

Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R., and Görlach, J. (2001). Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. The Plant cell **13**: 1499–510.

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

Charon, C., Sousa, C., Crespi, M., and Kondorosi, a (1999). Alteration of enod40 expression modifies medicago truncatula root nodule development induced by sinorhizobium meliloti. The Plant cell **11**: 1953–66.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.-R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant physiology **139**: 5–17.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G.K., Khatun, J., Williams, B.A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R.F., Alioto, T., et al. (2012). Landscape of transcription in human cells. Nature **489**: 101–8.

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.-K., and Mockler, T.C. (2010). Genome-wide mapping of alternative splicing in Arabidopsis thaliana. Genome research **20**: 45–58.

Fukaki, H., Tameda, S., Masuda, H., and Tasaka, M. (2002). Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of Arabidopsis. The Plant journal: for cell and molecular biology **29**: 153–68.

Guttman, M. and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature **482**: 339–46.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell **147**: 789–802.

James, A.B., Syed, N.H., Bordage, S., Marshall, J., Nimmo, G.A., Jenkins, G.I., Herzyk, P., Brown, J.W.S., and Nimmo, H.G. (2012). Alternative splicing mediates responses of the Arabidopsis circadian clock to temperature changes. The Plant cell **24**: 961–81.

Kalyna, M., Simpson, C.G., Syed, N.H., Lewandowska, D., Marquez, Y., Kusenda, B., Marshall, J., Fuller, J., Cardle, L., McNicol, J., Dinh, H.Q., Barta, A., and Brown, J.W.S. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. Nucleic acids research **40**: 2454–69.

Liu, Y., Pan, S., Liu, L., Zhai, X., Liu, J., Wen, J., Zhang, Y., Chen, J., Shen, H., and Hu, Z. (2012). A genetic variant in long non-coding RNA HULC contributes to risk of HBV-related hepatocellular carcinoma in a Chinese population. PloS one **7**: e35145.

Lorković, Z.J., Hilscher, J., and Barta, A. (2008). Co-localisation studies of Arabidopsis SR splicing factors reveal different types of speckles in plant cell nuclei. Experimental cell research **314**: 3175–86.

Lurin, C., Andrés, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyère, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., Lecharny, A., Le Ret, M., Martin-Magniette, M.-L., Mireau, H., Peeters, N., Renou, J.-P., Szurek, B., Taconnat, L., and Small, I. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. The Plant cell **16**: 2089–103.

Marquez, Y., Brown, J.W.S., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. Genome research **22**: 1184–95.

Mathesius, U., Charon, C., Rolfe, B.G., Kondorosi, A., and Crespi, M. (2000). Temporal and spatial order of events during the induction of cortical cell divisions in white clover by Rhizobium leguminosarum bv. trifolii inoculation or localized cytokinin addition. Molecular plant-microbe interactions: MPMI **13**: 617–28.

Matlin, A.J., Clark, F., and Smith, C.W.J. (2005). Understanding alternative splicing: towards a cellular code. Nature reviews. Molecular cell biology **6**: 386–98.

Okushima, Y., Overvoorde, P.J., Arima, K., Alonso, J.M., Chan, A., Chang, C., Ecker, J.R., Hughes, B., Lui, A., Nguyen, D., Onodera, C., Quach, H., Smith, A., Yu, G., and Theologis, A. (2005). Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in Arabidopsis thaliana: unique and overlapping functions of ARF7 and ARF19. The Plant cell **17**: 444–63.

Palusa, S.G., Ali, G.S., and Reddy, A.S.N. (2007). Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses. The Plant journal: for cell and molecular biology **49**: 1091–107.

Parizot, B., De Rybel, B., and Beeckman, T. (2010). VisuaLRTC: a new view on lateral root initiation by combining specific transcriptome data sets. Plant physiology **153**: 34–40.

Péret, B., De Rybel, B., Casimiro, I., Benková, E., Swarup, R., Laplaze, L., Beeckman, T., and Bennett, M.J. (2009). Arabidopsis lateral root development: an emerging story. Trends in plant science **14**: 399–408.

Rinn, J.L. and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annual review of biochemistry **81**: 145–66.

Rohrig, H., Schmidt, J., Miklashevichs, E., Schell, J., and John, M. (2002). Soybean ENOD40 encodes two peptides that bind to sucrose synthase. Proceedings of the National Academy of Sciences of the United States of America **99**: 1915–20.

Santi, C., Von Groll, U., Ribeiro, A., Chiurazzi, M., Auguy, F., Bogusz, D., Franche, C., and Pawlowski, K. (2003). Comparison of nodule induction in legume and actinorhizal symbioses: the induction of actinorhizal nodules does not involve ENOD40. Molecular plant-microbe interactions: MPMI **16**: 808–16.

Schindler, S., Szafranski, K., Hiller, M., Ali, G.S., Palusa, S.G., Backofen, R., Platzer, M., and Reddy, A.S.N. (2008). Alternative splicing at NAGNAG acceptors in Arabidopsis thaliana SR and SR-related protein-coding genes. BMC genomics **9**: 159.

Simpson, C.G., Lewandowska, D., Fuller, J., Maronova, M., Kalyna, M., Davidson, D., McNicol, J., Raczynska, D., Jarmolowski, A., Barta, A., and Brown, J.W.S. (2008). Alternative splicing in plants. Biochemical Society transactions **36**: 508–10.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology **3**: Article3.

Sousa, C., Johansson, C., Charon, C., Manyani, H., Sautter, C., Kondorosi, A., and Crespi, M. (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. Molecular and cellular biology **21**: 354–66.

Storey, J.D. and Tibshirani, R. (2003). Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America **100**: 9440–5.

Tanabe, N., Yoshimura, K., Kimura, A., Yabuta, Y., and Shigeoka, S. (2007). Differential expression of alternatively spliced mRNAs of Arabidopsis SR protein homologs, atSR30 and atSR45a, in response to environmental stress. Plant & cell physiology **48**: 1036–49.

Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., Blencowe, B.J., Prasanth, S.G., and Prasanth, K. V (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell **39**: 925–38.

Wang, G.-S. and Cooper, T.A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. Nature reviews. Genetics **8**: 749–61.

Wang, J., Ou, S.-W., Wang, Y.-J., Zong, Z.-H., Lin, L., Kameyama, M., and Kameyama, A. New variants of Nav1.5/SCN5A encode Na+ channels in the brain. Journal of neurogenetics **22**: 57–75.

Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., and Speed, T.P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic acids research **30**: e15.

Yin, Q.-F., Yang, L., Zhang, Y., Xiang, J.-F., Wu, Y.-W., Carmichael, G.G., and Chen, L.-L. (2012). Long Noncoding RNAs with snoRNA Ends. Molecular cell **48**: 219–30.

Zhang, X.-N. and Mount, S.M. (2009). Two alternatively spliced isoforms of the Arabidopsis SR45 protein have distinct roles during normal plant development. Plant physiology **150**: 1450–8.

	Accession	Gene	Oligo sequence	information
Alternative splicing RT-PCR	At4g27050	F-box family protein F	GGTTCTTTCTTTGTGAACAAC	
	At4g27050	F-box family protein R	GGGAAACTCCTTAGACAGCAG	
	At2g33830	Auxin related protein F	CCGGACCTAAACCGGAGCATGGCC	
	At2g33830	Auxin related protein R	CCGATCCTGGCGTCGTCGGAGTTCC	
	At2g46830	CCA1 Myb-like DNA binding F	GGACTGAGGAAGAACATAATAG	
	At2g46830	CCA1 Myb-like DNA binding R	GGTTTACGCTTAGGCCGTGG	
	At1g27770	calcium-transporting ATPase 1F	CGAAGCACTTCGAACTCTTTGCC	
	At1g27770	calcium-transporting ATPase 1R	CTCACCAAAGTATGCTTATCC	
RT-qPCR	AT1G13320	Hkeeping ref 1F	TAACGTGGCCAAAATGATGC	Czechowski et al., 2005
	AT1G13320	Hkeeping ref 1R	GTTCTCCACAACCGCTTGGT	Czechowski et al., 2005
	AT4G26410	Hkeeping ref 2F	GAGCTGAAGTGGCTTCCATGAC	Czechowski et al., 2005
	AT4G26410	Hkeeping ref 2R	GGTCCGACATACCCATGATCC	Czechowski et al., 2005
	AT1G67105	ncRNA 351F	CCCATCGCACTGATCGGCGG	
	AT1G67105	ncRNA 351R	TCGAGCGCTGCCGTCTTCAC	
	AT2G34655	ncRNA34F	CCTCTTGATCCGGTCATGTT	
	AT2G34655	ncRNA34R	ACAGCGGTGCCACCTATTAC	
	At1G76940	NSRa F	CTCAGGGACATTGTGGTACG	
	At1G76940	NSRa R	CTACACTCTCCGCGCATCTT	
	At1G21320	NSRb F	CACAGGATGAGGATCGAGGT	
	At1G21320	NSRb R	ACCTCCTTGTCGTACCATGC	
Cloning	At1G76940	NSRa F Cter furian	CACCATGGCGGATGGGTACTGGAACCAGC	For Pentr-D (Invitrogen)
	At1G76940	NSRa R Cter furian	CCTTCTTCCTCTTTGTCCTGGTCT	For Pentr-D (Invitrogen)
	At1G21320	NSRb F Cter furian	CACCTGTTTCCGTCGCCGTCAATG	For Pentr-D (Invitrogen)
	At1G21320	NSRb R Cter furian	CCTTCTTCCTCCACGCTGTCC	For Pentr-D (Invitrogen)
	AT4G02840	snRNP F Cter furian	CACCGCTGCTTTTCCTCCTCCCGCTC	For Pentr-D (Invitrogen)
	AT4G02840	snRNP R Cter furian	ACGACCACCACGGCCACGTCC	For Pentr-D (Invitrogen)
	TC177461 Version MtGI11	MtENOD40 expression		Charon et al., 1999 (35S promoter)
	TC177461 Version MtGI11	MtENOD40 expression		Charon et al., 1999 (35S promoter)
	At1G76940	NSRa promoter F	GGGGACAACTTTGTATAGAAAAGTTGGGGAGTCACAGACACTCGACGG	For P4P1R multisite Gateway
	At1G76940	NSRa promoter R	GGGGACTGCTTTTTTGTACAAACTTGTGGAAATCGAAGGATCAGAAG	For P4P1R multisite Gateway
	At1G76940	NSRa promoter GUS F	CACCCATTTGGTCTGGGTGCTTCT	For Pentr-D (Invitrogen)
	At1G76940	NSRa promoter GUS R	TTCTGCTCCGAAGCACTCTC	For Pentr-D (Invitrogen)
	At1G21320	NSRb1 promoter GUS F	CACCCTAAGTTCTAACTTTTGGGTTTG	For Pentr-D (Invitrogen)
	At1G21320	NSRb1 promoter GUS R	CAGTGGTGTTGGTTTCGGACC	For Pentr-D (Invitrogen)
	At1G21320	NSRb2 promoter GUS F	CACCCTAAGTTCTAACTTTTGGGTTTG	For Pentr-D (Invitrogen)
	At1G21320	NSRb2 promoter GUS R	GTGAGGGTTCGACGAAATTGG	For Pentr-D (Invitrogen)
	AT1G33980	UPF3 F	GGAGATAGAACCATGAAGGAACCTTTGCAGAAG	For pDONR221
	AT1G33980	UPF3 R	CCTCCGGATCMAGTACCGGATGATGGTTTTTGG	For pDONR221
	AT3G62800	DRB4 F	GGAGATAGAACCATGGATCATGTATACAAAGGTC	For pDONR221
	AT3G62800	DRB4 R	CCTCCGGATCMTGGCTTCACAAGACGATAGGC	For pDONR221
	At1G21320	NSRb F over expression	CACCTGTTTCCGTCGCCGTCAATG	For Pentr-D (Invitrogen)
	At1G21320	NSRb R over expression	CATTGTTCTTCACTAAATTCATCATCAC	For Pentr-D (Invitrogen)
	At1G76940	NSRa F over expression	CACCATGGCGGATGGGTACTGGAACCAGC	For Pentr-D (Invitrogen)
	At1G76940	NSRa R over expression	TCACCTTCTTCCTCTTTGTCCTGGTCT	For Pentr-D (Invitrogen)
	AT1G02840	SRP34 F Cter furian	CACCAATAAACCATGAGCAGTCGTTCGAGTAGA	For Pentr-D (Invitrogen)
	AT1G02840	SRP34 R Cter furian	CCTCGATGGACTCCTAGTGTG	For Pentr-D (Invitrogen)
RNA IP qPCR	At4g27050	F-box family protein F	GGTTCTTTCTTTGTGAACAAC	
	At4g27050	F-box family protein R	GGGAAACTCCTTAGACAGCAG	
	At2g33830	Auxin regulated protein F	CCGGACCTAAACCGGAGCATGGCC	
	At2g33830	Auxin regulated protein R	CCGATCCTGGCGTCGTCGGAGTTCC	
	AT1G67105	ncRNA 351F	TAGATGTAATCGGTGGTGGAGTC	
	AT1G67105	ncRNA 351R	ACCTATCCACGCGTCAGG	
	AT1G13320	Hkeeping ref 1F	TAACGTGGCCAAAATGATGC	Czechowski et al., 2005
	AT1G13320	Hkeeping ref 1R	GTTCTCCACAACCGCTTGGT	Czechowski et al., 2005
	At2g29210	Splicing factor PWI domain F	GCAAATGAAATAATGAAGAAGAGGG	
	At2g29210	Splicing factor PWI domain R	GTCAGCCCTTCTCCCATCCTCTGG	
Genotyping	At1G76940	nrra SALK F	GGATCTCGATGTTCCCGATA	
	At1G76940	nrra SALK R	CCTTCTTCCTCTTTGTCCTGGTCT	
	At1G21320	nrrb SAIL F	CACAGGATGAGGATCGAGGT	
	At1G21320	nrrb SAIL R	ACCTTCTTCCTCCACGCTGTCCTGG	

Table S1 Genetic constructions and oligonucleotides

All oligonucleotides used for cloning, RT-PCR and RT-qPCR are referenced.

1.3 Résultats complémentaires

Il n'existe pas de systeme *in vitro* recréant le spliceosome végétal afin d'étudier l'effet de l'ajout ou de la mutation d'un régulateur de l'épissage, nous avons donc imaginés une expérience de splicing exogène dans les feuilles de tabac grâce à des expériences d'agroinfiltration de feuilles de *N. tabacum* afin de confirmer l'action des ces protéines dans un système hétérologue.

Pour cela nous avons sur-exprimé dans ces feuilles des régulateurs de l'épissage (NSRa ou NSRb) ainsi qu'une cible des NSRs épissée alternativement chez Arabidopsis, le gène *Fbox* (At4G27050). De plus nous avons aussi pris des tabacs sur-exprimant la cible Fbox sans régulateur ou co-exprimé avec la mcherry comme contrôles négatifs. Après expression durant deux jours dans les feuilles de tabac, extraction des ARN et RT-PCR nous pouvons voir en gel d'acrylamide le profil d'épissage du gène *Fbox* (figure complémentaire). Nous avons d'abord vérifié la calibration des cDNA pour valider l'expérience, ainsi nous avons passé deux gènes constitutifs du tabac (référence 1 : UBi3 ; référence 2 : Eif1α).

Premièrement on remarque que le gène est épissé alternativement de la même façon que naturellement chez Arabidopsis. De plus en mesurant l'intensité des bandes et en réalisant le ratio de la bande haute sur la plus basse on remarque que les deux isoformes sont exprimé de façon similaire dans les contrôle et dans les extraits où NSRa est surexprimé indiquant que cette protéine ne serait pas impliquée dans la régulation de l'AS du gène *Fbox*. En revanche, la surexpression de NSRb entraînerait une diminution du ratio de moitié donc NSRb semble favoriser la production de l'isoforme totalement épissé plutôt que la variante non épissée.

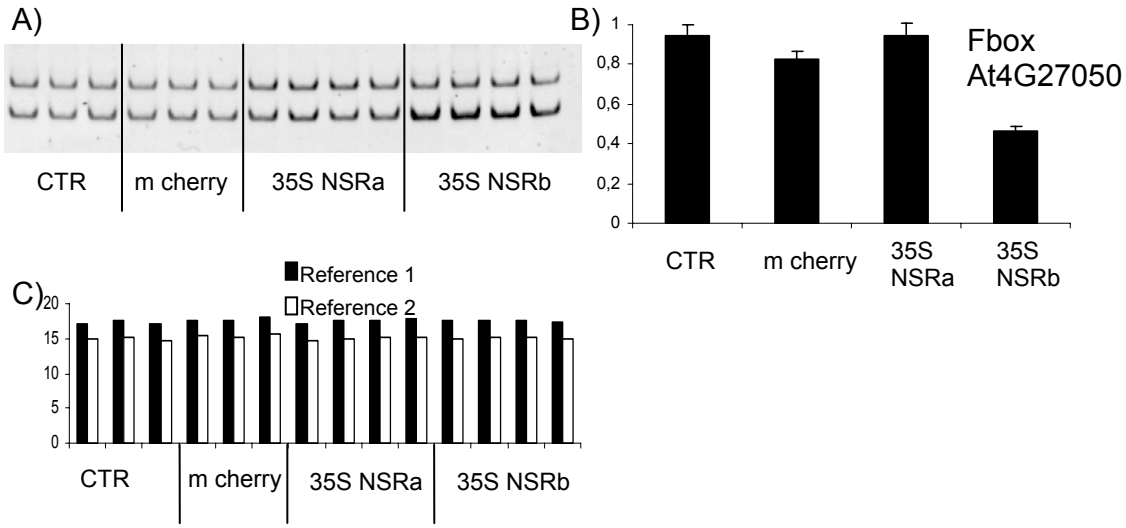


Figure complémentaire : Expérience d'épissage alternative en expression transitoire dans les feuilles de *N. tabacum*.

A) Gel d'acrylamide résultant de la RT-PCR dans les échantillons indiqués. B) Graphique représentant le ratio de la quantification entre l'intensité de la bande supérieure sur celle de la bande inférieure. C) Expression relative représentant l'expression de deux gènes constitutifs de N. tabacum indiquant la calibration des ADNc qui ont permis l'expérience. (ref1: UBi3: FW aggttgaatcttccgacacaat REV tcgactccttctggatgttgta ; ref2: Eif1α : FW GATTGGTGGTATTGGAACTGTC REV GAGCTTCGTGGTGCATCTC).

II.2- SGS14, SGS15, NSRa and NSRb encode alternative splicing-related factors that facilitate post-transcriptional gene silencing mediated by intron-containing transgenes.

Dans cette deuxième partie, je vais d'abord présenter une introduction sur les mécanismes de Silencing Post transcriptionnel des gènes ainsi que sur les lignées qui permettent l'étude du PTGS chez Arabidopsis. Puis je vais résumer les résultats obtenus lors de ma thèse sur ce sujet sous la forme d'un article, intitulé « SGS14, SGS15, NSRa and NSRb encode splicing-related factors that facilitate post-transcriptional gene silencing mediated by intron-containing transgenes ».

2.1 RBP et ncRNAs : Mécanismes du silencing

Cette deuxième partie est dévolue aux rôles des NSRs dans la régulation du silencing et en relation avec l'épissage. Je vais donc introduire les mécanismes qui sont mis en place lors de l'interférence par l'ARN.

Chez les plantes, l'interférence par l'ARN ou silencing par les ARN a été découverte à la suite d'expériences de transgenèse aboutissant à des résultats inattendus. Ainsi, la surexpression d'un gène ne conduisait pas toujours à une sur-accumulation de la protéine, mais parfois, au contraire, à un phénotype équivalent au mutant du gène étudié (Smith et al., 1990 ; van der krol et al., 1990 ; Napoli et al., 1990). L'analyse des plantes transgéniques inactivées pour des transgenes qui devraient être exprimés constitutivement a permis l'identification des deux principaux mécanismes de « RNA silencing » : le PTGS (Post Transcriptional Gene Silencing) et le TGS (Transcriptional Gene Silencing) aussi appelé depuis peu CTGS (Co-Transcriptional Gene Silencing). Dans cette partie, je ne reviendrais pas sur les protéines qui sont associées au mode d'action et à la biogenèse des miARN et des tasiARN traités précédemment (Ariel et al., 2012 chapitre 1.2 ; Charon et al., 2010 Cf annexe 1 et 3) ainsi qu'aux mécanismes moléculaires de silencing (Parent et al., 2012). Néanmoins, j'introduirai la description du silencing induit par un transgène dans la mesure où j'ai utilisé ce système dans cette partie. En effet, les transgènes peuvent être « silencés » par des mécanismes similaires aux mécanismes de régulation des gènes. Au niveau post-transcriptionnel, il est possible de distinguer le Sens PTGS ou (S-PTGS) des transgènes et l'Inverted Repeat PTGS (IR-PTGS).

2.2 Les mutants déficients en PTGS

L'obtention de lignées transgéniques, dont les transgènes déclenchent le silencing à des moments précis, a permis de mettre en place des cribles génétiques permettant l'identification de gènes impliqués dans les mécanismes moléculaires de silencing. Dans cette partie, j'introduirai les principales lignées déficientes en silencing que j'ai utilisées durant ma thèse.

2.2.1 La lignée *p35S::uidA* « L1 »

La lignée L1 est issue d'une transformation par *Agrobacterium tumefaciens* sur une plante Col-0, avec un ADN-T qui porte la construction *p35S::uidA* (GUS) (Elmayan et al., 1998 ; Figure 9a). La lignée L1 (pour *Low expressers1*) possède une simple insertion de ce transgène et présente une inactivation de GUS chez 100% des plantes à chaque génération. De plus, lors des transformations de Col-0 par cette construction, une lignée contrôle non silencée nommée « 6b4 » et présentant le signal GUS dans 100% des plantes à chaque génération a été obtenue. La lignée L1 produit des siARN spécifiques qui induisent la dégradation des ARNm du transgène GUS. Cette lignée est donc très utile pour caractériser des mutants déficients en S-PTGS. De plus, dans la mesure où il a été montré que le silencing pouvait être modifié par la présence d'un intron au sein du transgène une construction génétique 35S:GUS-int, nommée 159, a été réalisée. Elle ne diffère de la construction initiale L1 que par la présence d'un intron, d'origine végétale, dans le gène GUS. Cette lignée 159 présente comme pour L1 une forte expression du gène rapporteur GUS lors des stades précoces de développement ; cependant, le transgène est rapidement silencé.

Ces dernières années, la lignée L1 a permis de caractériser de nombreux gènes impliqués dans la biogenèse des siARN, leurs actions ou encore la régulation des mécanismes de silencing de type S-PTGS. Parmi ces gènes, on peut noter la découverte de : SGS1 (SUPRESSOR OF GENE SILENCING 1) *SGS2/RDR6* (RNA DEPENDANT RNA POLYMERASE), SGS3, SGS4/AGO1, SGS5/HEN1 (HUA ENHANCER 1) SGS6/MET1 (DNA METHYLTRANSFERASE 1), SGS7/SDE5 un facteur d'export de l'ARN, SGS8/JMJ14 une Deméthylase, SGS9/HPR1 un composant du complexe THO/TREX, SGS13/SDE3, SGS6/MET1 (Elmayan et al., 1998 ; Mourrain et al., 2000 ; Fagard et al., 2000 ; Morel et al., 2002; Boutet et al., 2003 ; Adenot et al., 2006 ; Elmayan et al., 2009 ; Jauvion et al., 2010 ; Le Masson et al., 2012).

2.2.2 La lignée *p35S::NIA2* « *2a3* »

La nitrate réductase codée par les gènes *NIA1* et *NIA*2 est nécessaire à la croissance des plantes en conditions naturelles. L'introduction dans le génome de plantes Col-0 du transgène *p35S::NIA2* devait conduire à la sur-accumulation de la protéine NIA2 ; elle a en fait conduit a la co-supression des gènes *NIA1* et *NIA2* endogènes ainsi que du transgène *p35S::NIA2* (Figure 9b). Dans cette expérience, 297 des 298 transformants primaires obtenus sont morts ; seul le transformant *2a3* a produit des graines à partir desquelles une lignée homozygote a été générée. Chez la lignée homozygote *2a3,* en condition de culture standard, seulement 20% des plantes atteignent le stade de fructification les 80% restant meurent avant de former leurs graines (Adenot et al., 2006). Cette lignée *2a3* comme la lignée L1 permet l'étude des mécanismes impliqués dans le S-PTGS et contient 2 introns.

2.2.3 La lignée pSUC2::antisensPDS « JAP3 »

Afin d'étudier la propagation du signal de silencing, une lignée transgénique dont l'initiation et l'action de silencing se déroule dans un tissu précis a été préparée (Smith et al., 2007). Pour cela une construction RNAi du gène codant une phytoène désaturase (PDS) a été placé sous le contrôle d'un promoteur spécifique du phloème : Suc2 (codant pour un transporteur du sucrose spécifique du phloème ; Figure 9c). Le gène PDS est nécessaire à la synthèse de chlorophylle dans les feuilles des plantes. Lorsque la construction RNAi entraîne le silencing du gène PDS endogène les cellules subissent une décoloration très facilement observable. SUC2 ne s'exprimant que dans le phloème la propagation du silencing dans les cellules voisines peut être suivie grâce à la décoloration de celles-ci. Des siARN de 21 et de 24 nucléotides sont produits dans la lignée pSuc2::PDSRNAi, indiquent que ce transgène active plusieurs voies du silencing. Cette construction possède un intron entre les brin sens et antisens qui contiennent la partie de l'ARNm *PDS.*

Figure 9 Constructions génétiques utilisées pour étudier le silencing chez les plantes.
(a) Le promoteur 35S et le terminateur 3' *Rubisco* (*Rbs*) sont représentés en blanc, la partie noire représente la partie codante *uidA.* (b) Le promoteur *35S* et le terminateur endogène de *NIA2* ainsi que le terminateur *NOS* sont représentés en blanc. Les parties noires représentent les exons du gène *NIA2,* les traits noirs représentent les introns du gène *NIA2.* (c) Le promoteur *pAtSuc2* et le terminateur pA35S sont représentés en blanc. Les parties noires représentent le gène PDS et les flèches représentent le gène PDS en sens et en antisens. L'intron représenté par un trait noir correspond à l'intron 1 du gène WRKY33.

2.3 Le S-PTGS des transgènes et la propagation du silencing

A ce jour diverses questions concernant le S-PTGS (Sens-PTGS) restent floues comme : de quelle façon un dsARN primaire est produit par un transcrit sens pendant le S-PTGS ? A quelle fréquence le S-PTGS varie d'une lignée transgénique à l'autre ? (Vaucheret et al., 1995). Plusieurs hypothèses ont été proposées pour expliquer le déclenchement du PTGS dans le S-PTGS : la première repose sur l'idée d'une transcription involontaire et locus-dépendant d'un antisens ARN qui formerait un dsRNA avec les transcrits du transgène ; la seconde soutient que la surexpression du S-transgène donnerait lieu à un ARN aberrant sens qui serviraient de matrice à RDR6. Quoi qu'il en soit, une fois produit, le dsARN peut être traité par une ou plusieurs protéines DCL. Dans le cas du locus L1 d'Arabidopsis, la disponibilité de DCL2 semble être le premier déterminant qui initie le S-PTGS (Mlotshwa et al., 2008). Probablement car DCL2 produit des siARN de 22-nt qui, après association avec AGO1, clive la cible ARNm pour générer des produits de clivages (Figure 10). Ces produits sont protégés de la dégradation par SGS3 et transformés en dsARN secondaires par RDR6 assisté par SDE5 (Jauvion et al., 2011). Le locus L1 produit un niveau semblable de siARN de 21-nt et de 22-nt du transgène GUS (Mlotshwa et al., 2008), ce qui semble indiquer que les dsARN secondaires sont donc traités de la même façon par DCL4 et DCL2 produisant des dsARN 21-nt et de 22-nt respectivement. En outre, les siARN de 22-nt dépendant de DCL2, peuvent générer une boucle d'amplification de

siRNAs, ce qui permet d'amplifier le silencing et donc de diminuer le niveau d'expression de l'ARNm cible. Cette boucle est essentielle pour la diffusion du S-PTGS. En effet, le S-PTGS semble se déclencher localement pour ensuite diffuser de cellule à cellule grâce aux plasmodesmes ou même à plus longue distance à travers les tissus vasculaires comme l'a montré des expériences de greffes de plantes transgéniques silencées ou pas (Palauqui et al., 1996, Dunoyer et al., 2010). Ces résultats suggèrent qu'un signal de silencing séquence spécifique permettrait de propager le S-PTGS dans la plante (Voinnet et al., 1998). Plus récemment, il a été montré que les siARN sont mobiles (Dusnoyer et al., 2010 ; Molnar et al., 2010). ; les siARN primaires ou secondaires produits dans la cellule ou le S-PTGS a été initié peuvent se déplacer de cellule à cellule et dans certains cas dans la plante entière via les tissus vasculaires. Dans la cellule qui reçoit les sRNA, ces derniers peuvent déclencher le S-PTGS à leur tour sur leur séquence homologue (figure 10).

La propagation du silencing est séquence-spécifique, ce qui indique que le signal comprend probablement un ARN, e.g. un ARN double brin (dsARN) ou le small interfering ARN (siARN) généré lors du silencing des ARN (Dunoyer et al, 2005). Les enzymes Dicer-like (DCL), de la famille RNase de type III, clivent l'ARNdb en siARN, et dans un système expérimental, la propagation du silencing est perdu dans le mutant dcl4, en équation avec l'idée que la production de siARN est nécessaire à la production du signal (Dunoyer et al., 2005). Les siARN, mesurant seulement 21 à 24 nucléotides de long peuvent passer à travers les plasmodesmes. De plus les échantillons de sève de phloème contiennent des molécules apparentées aux siARN (Yoo et al., 2004). Cependant, des transcrits plus long en orientation sens et antisens sont également présents dans la sève du phloème, il est donc possible que les long ARN soient également mobiles (Haywood et al, 2005). Le mouvement de cellule à cellule par les plasmodesmes et la translocation longue distance du signal de silencing à travers le phloème sont parfois considérés comme des mécanismes distincts, car ils peuvent être inhibés de façon différentiel par un traitement au cadmium, par des protéines virales, ou par mutation de gènes végétaux (Himber et al, 2003 ; Schwach et al, 2005).

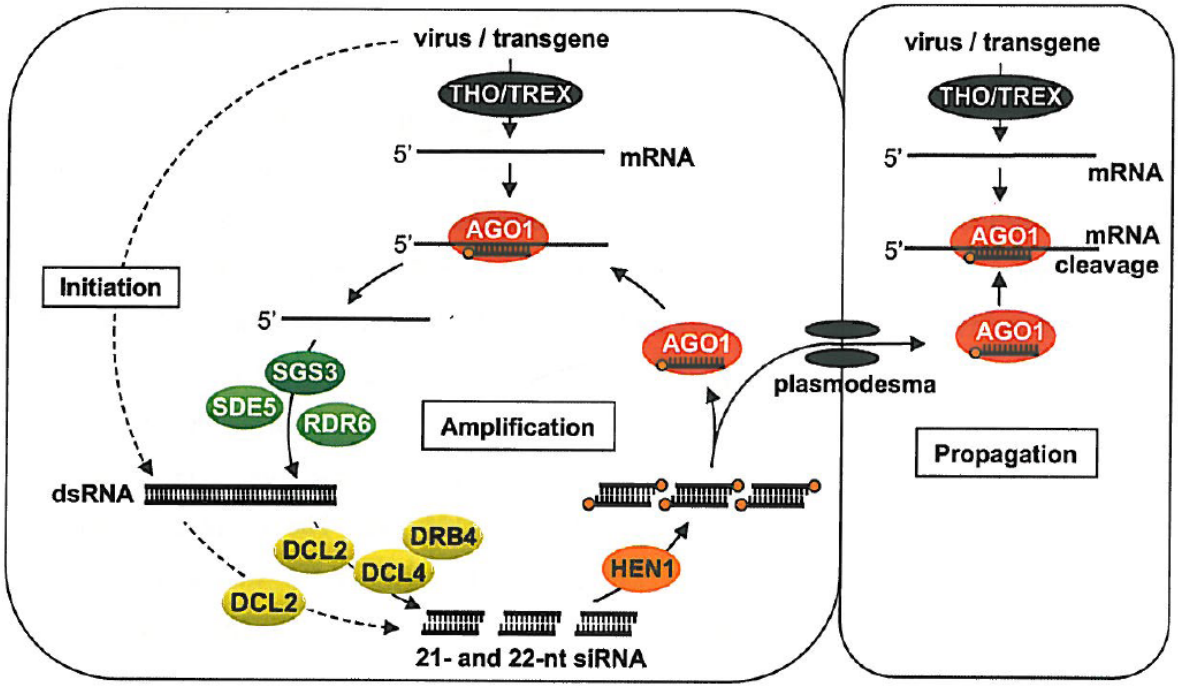


Figure 10 Composants générale du S-PTGS

Les composants protéiques représentés de la même couleur sont impliqués dans les mêmes voies métaboliques. Le mécanisme est décrit dans le texte. (Jauvion et al., 2011).

2.4 la voie du IR-PTGS des transgènes

Les trangènes sens qui produisent un niveau limité de dsARN font intervenir RDR6, SDE3, SDE5 et SGS3 pour amplifier le signal de silencing. Par contre, les transgènes en IR (inverted repeat) produisent directement des dsARN et ne nécessitent pas ces protéines pour produire des siARN et silencer leurs cibles par PTGS (IR-PTGS ou Inverted Repeat PTGS ; Beclin et al., 2002 ; Himber et al., 2003 ; Dunoyer et al., 2005 ; Smith et al ., 2007 ; Dunoyer et al., 2010). Par conséquent, il serait logique de penser que les transgènes IR utilisent les mêmes voies que les loci en IR endogènes, comme IR71 et IR2039. Ces loci endogènes produisent via DCL4, DCL2 et DCL3 respectivement des siARN de 21-nt et 22-nt ainsi que de 24-nt, et sont capables d'éteindre les cibles endogènes complémentaires grâce à AGO1 (Henderson et al., 2006 ; Dunoyer et al., 2010). l'IR-PTGS basé sur les lignées pSUC2-SUL et pSUC2-PDS a servi à identifier des nouveaux allèles de DCL4 et AGO1 (Dunoyer et al., 2005 ; Smith et al., 2007) mais aussi de RDR2, les sous unités NRPD1 et NRPD2 de Pol_IV ainsi qu'une nouvelle protéine de remodelage de la chromatine nommée CLSY1 pour la diffusion à courte distance du PTGS (figure 11). (Dunoyer et al., 2005 ;

Dunoyer et al., 2007 ; Smith et al., 2007 ; Dunoyer et al., 2010). Puisque le niveau de siARN de PDS dans la lignée pSUC2-PDS est réduit dans le fond génétique clsy1, nrpd1 et rdr2 mutants (des mutants perturbés dans la production des siRNAs de 24nt mais pas d'autres siRNAs) il faut donc considérer que l'IR-PTGS du locus pSUC2-PDS subit un control transcriptionnel dépendant (au moins en partie) des siARN de 24-nt produits par ce locus. D'autres données soutiennent cette hypothèse, en effet, des mutations dans différentes protéines en relation avec le remodelage de la chromatine, dont le facteur d'assemblage de la chromatine/nucléosome FVE ainsi que l'histone deméthylase FLD et JMJ14, affectent l'IR-PTGS dans la lignée pSUC2-PDS (Baurle et al., 2007 ; Searle et al., 2010).

Les mécanismes de silencing sont donc principalement mis en place et agissent grâce à différentes RBPs. Ces protéines utilisent une variété de npcARN dans le but moduler l'expression des gènes afin d'optimiser la réponse physiologique de la plante aux variations environnementales. Ces mécanismes impliquant des ribonucléoprotéines agissent de façon coordonnée avec les RBPs impliquées dans la maturation, l'export, le contrôle qualité, ainsi que la séquestration et la dégradation des ARN.

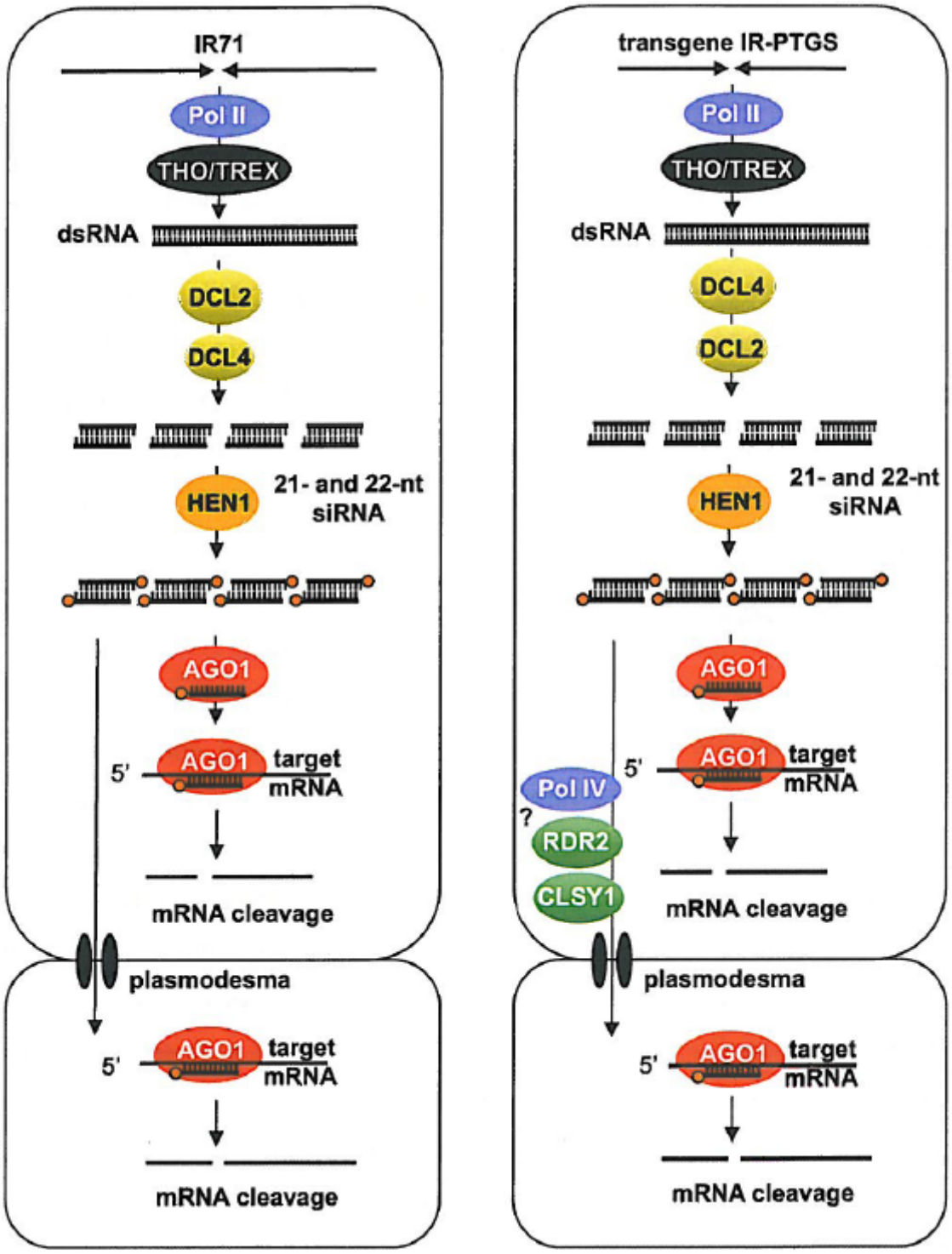


Figure 11 Composants générale du S-PTGS

Les composants protéiques représentés de la même couleur sont impliqués dans les mêmes voies métaboliques. La taille du rond autour de DCL2 et DCL4 dans L'IR-PTGS endogène et des transgènes représente leur action hiérarchique. Le point d'interrogation à coté de RDR2 et NRPD1 indique que cette étape n'est pas connue. Le mécanisme est décrit dans le texte. (Jauvion et al., 2011).

II.3- Publication: SGS14, SGS15, NSRa and NSRb encode alternative splicing-related factors that facilitate post-transcriptional gene silencing mediated by intron-containing transgènes

Dans cette partie, je décris les principaux résultats obtenus (sous la forme d'un article) sur le rôle des protéines NSRs en relation avec le silencing, mais aussi ceux obtenus sur deux autres protéines issu d'un crible « Supressor of Gene Silencing » (sgs14 et sgs15) réalisé par Hervé Vaucheret (INRA, Versaille). Ces deux protéines SGS étant deux homologues avec des protéines en relation avec la régulation de l'épissage, nous nous sommes interrogés sur un potentiel lien entre l'AS et mécanismes du silencing. Certains résultats restants en suspens, le papier comporte encore quelques zones d'ombre auquel il faudra répondre avant sa soumission. Les deux mutants SGS ont été cartographiés par Vincent Jauvion lors de sa thèse dans le laboratoire d'Hervé Vaucheret. Il a aussi réalisé les northern blots qui comparent les différents mutants issus du crible SGS (tasiARN, GUS, NIA2) avec l'aide d'Emilie Elvira-Matelot et de Nathalie Bouteiller. De plus Hervé Vaucheret a réalisé les croisements des différents mutants, lignées reportrices, et les transformations des lignées pour complémenter les phénotypes et ou l'effet sur le silencing de lignées reportrices en rapport avec les gènes SGS. J'ai quand à moi réalisé les différents clonages qui ont permis de complémenter les différents mutants, et j'ai étudié la localisation de ces protéines. De plus j'ai réalisé les mutants nsra/nsrb/JAP3 ainsi que toutes les manipulations en relation avec les lignées JAP3 dans les différents fonds mutants (nothern, RT-PCR etc…). De plus j'ai réalisé les testes d'AS sur les différentes cibles et sur les différents mutants. Enfin avec l'aide d'Hervé Vaucheret et Martin Crespi nous avons rédigé l'article.

Title: SGS14, SGS15, NSRa and NSRb encode alternative splicing-related factors that facilitate post-transcriptional gene silencing mediated by intron-containing transgenes.

Authors: Florian Bardou[1], Vincent Jauvion[2], Emilie Elvira-Matelot[2], Nathalie Bouteiller[2], Martin Crespi[1] and Hervé Vaucheret[2]

[1] Institut des Sciences du Végétal, Centre National de la Recherche Scientifique, F-91198 Gif-sur-Yvette cedex, France

[2] Institut Jean-Pierre Bourgin, UMR1318, INRA, Versailles, France

Corresponding author: Hervé Vaucheret and Martin Crespi

Keywords: PTGS, splicing regulator, Alternative splicing

Abstract:

Post-transcriptional gene silencing (PTGS) is a molecular mechanism that down-regulates exogenous RNAs such as transgenes or viruses through small 21-22nt siRNAs. PTGS can spread cell-to-cell and systemically. Diverse transgenic lines showing constitutive or spreading PTGS were used to identify molecular components of PTGS such as proteins involved in siRNA production, transport or action as well as chromatin status. Here, we identified that the *SGS14* and *SGS15* genes controlling PTGS of intron-containing transgenes encode RNA processing-related nuclear proteins homologous to splicing regulators in other species. SGS14/LSM9a and SGS15/PRP39 localize in nuclear "speckles" showing partial co-localisation with the splicing-related factor SRP34 and the alternative splicing regulators NSRa and NSRb. Although no major splicing defect was found in these plants, mutations in *SGS14*, *SGS15* and *NSRa/NSRb* affect the alternative splicing of common endogenous mRNAs in different manners. Furthermore, the nrsa/nsrb double mutant is affected in the spreading of PTGS. These results revealed a new link between alternative splicing regulation and PTGS.

INTRODUCTION

Post-transcriptional gene silencing is a process that leads to mRNA degradation through the production of small 21nt siRNAs. PTGS can be triggered by invading RNAs such as viruses or transgenes that form double-stranded RNA which are processed into siRNAs. These small siRNAs triggered the sequence-specific cleavage of mRNAs containing complementary sequences including the invading RNA but also related endogenous mRNAs. Using transgenic lines where transgenes are silenced by PTGS, several genes involved in PTGS have been identified (Jauvion et al., 2011). A forward genetic screen based on the line *L1,* which carries a post-transcriptionally silent *p35S-GUS* sense transgene, identified a series of S-PTGS-deficient mutants that defined 12 independent *SUPPRESSOR OF GENE SILENCING (SGS)* loci. Mutations in these *SGS* loci also impaired silencing of line *2a3*, which carries a *p35S-NIA2* sense intron-containing transgene that triggers co-suppression of the endogenous genes *NIA1* and *NIA2.* A novel forward genetic screen based on line *2a3* identified three additional loci (*SGS13, SGS14* and *SGS15)* required for *2a3* but not *L1* silencing.

Previous analyses revealed the functions of *SGS2/RDR6*, *SGS3*, *SGS4/AGO1*, *SGS5/HEN1*, *SGS6/MET1*, *SGS7/SDE5*, *SGS8/JMJ14*, *SGS9/HPR1* and *SGS13/SDE3* (Elmayan et al, 1998; Mourrain et al, 2000; Fagard et al, 2000; Morel et al, 2002; Boutet et al, 2003; Jauvion et al, 2010; Le Masson et al, 2012). These genes are involved in siRNA biogenesis, transport or action or in the modulation of chromatin status. In addition, components of RNA processing complexes not linked to siRNAs (biogenesis, transport or action) were also shown to modulate gene silencing. Known endogenous PTGS suppressors include the XRN2, XRN3, XRN4 ribonucleases and FRY1 (Gy et al., 2007), P-body and other RQC components (Motomura et al., 2012), NMD components (Arciga-Reyes et al., 2006), the 3' end processing factor ESP1, ESP4 and ESP5 and the putative splicing factor ESP3/PRP2 (Herr et al, 2006).

Here, we show that *SGS14* and *SGS15* encode RNA processing-related nuclear proteins homologous to known splicing regulators in other species. Co-localisation studies suggest that SGS14 and to a lesser extent SGS15 interact in nuclear "speckles" with the splicing-related factor SRP34 and the alternative splicing regulators NSRa and NSRb. Consistently, mutations in *SGS14*, *SGS15* and *NSRa/NSRb* affect the alternative splicing of overlapping sets of endogenous mRNAs as well as the efficiency of PTGS triggered by intron-containing transgenes. These results revealed a novel link between alternative splicing regulation and gene silencing.

RESULTS

SGS14 encodes a Like-Sm Ribonucleoprotein

The SGS14 locus is defined by a unique allele: *sgs14-1* that was identified using the 2a3 line expressing the intron-containing transgene *NIA2* (Elmayan et al., 1998). This mutant exhibits developmental defects (Fig. 1A) and PTGS-deficiency. The *sgs14* mutation was mapped to a 164 kb interval on chromosome 4. Whole-genome sequencing revealed that fast-neutron mutagenesis had induced a deletion in this interval, which removed six protein-coding genes (At4g02800, At4g02810, At4g02820, At4g02830, At4g02840 and At4g02850). Mutant lines harboring T-DNA insertions in the ORF of At4g02800, At4g02810, At4g02820, At4g02830 and At4g02850 did not exhibit developmental defects, suggesting that the deletion of At4g02840 was responsible for the developmental defects. However, the only available mutant in this gene had an insertion upstream of the ORF and did not exhibit developmental defects. Therefore, we transformed the *sgs14-1* mutant with a 6 kb genomic fragment carrying the At4g02840 gene. At first, we transformed *sgs14-1* mutant plants from which the 2a3 locus has been segregated away. Among 40 primary transformants, 39 exhibited a wild-type phenotype, indicating that the deletion of At4g02840 is responsible for the developmental defects of the sgs14-1 deletion mutant. However, because the deletion removed six adjacent genes, PTGS-deficiency could be due to the deletion of any of the six genes. Thus, we also transformed mutant plants carrying the 2a3 locus. Among 86 primary transformants, 84 exhibited the typical 2a3 co-suppresion phenotype, i.e. they died after two weeks, indicating that the deletion of At4g02840 is responsible for both developmental defects and PTGS-deficiency in the *sgs14-1* mutant. At4g02840 is predicted to encode a Like-Sm Ribonucleoprotein (LSM), hereafter referred to as LSM9a.

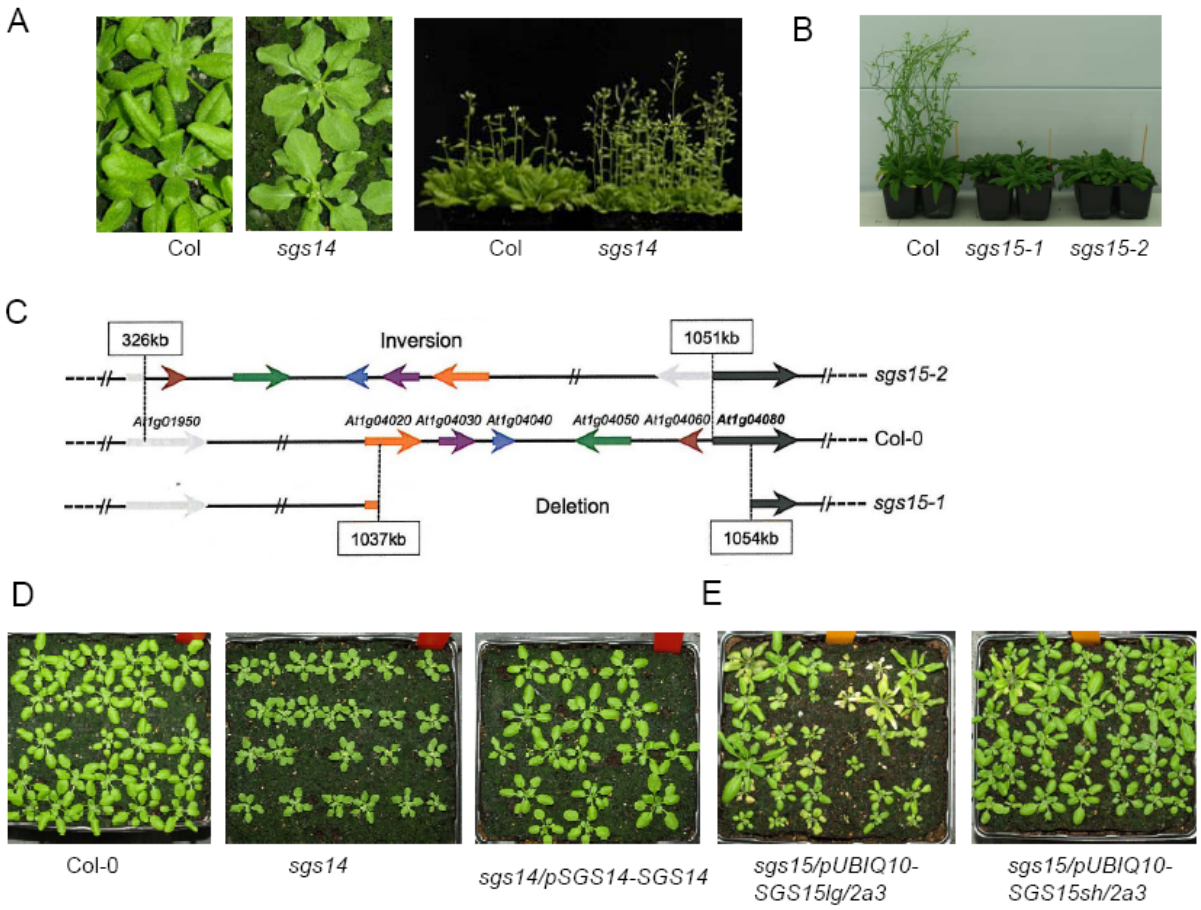


Fig. 1 Developmental phenotypes of *sgs14/lsm9a* and *sgs15/prp39* plants
A) *sgs14* mutant show a leaf serration developmental phenotype (upper panel) and accelerated flowering time compared to Col-0. B) *sgs15.1* (deletion mutant) and *sgs15.2* (inversion mutant) exhibit a late flowering time phenotype compared to WT. C) Schematic representation of the top of chromosome 1 in WT and *sgs15* mutants. The genes *ARGK2* (At1G01950, in light gray) and *PRP39* (At1G04080, in black) are indicated. The physical position of the inversion and deletion and the corresponding Kb from the beginning of the chromosome are indicated. A 725Kb genomic region that undergoes an inversion in *sgs15.2* and a deletion of 16Kpb in *sgs15.1* is represented. D) Picture showing the complementation of developmental defects in sgs14.1. The pictures show Col-0 WT phenotype and *sgs14.1* mutant phenotype and his complementation by *pSGS14-SGS14* (15/16 plants are complemented). E) co-supression deficiency in the line *sgs15/pUBIQ10-SGS15lg/2a3* is shown and his complementation could be observed only when the long annotation of SGS15 is used (*sgs15/pUBIQ10-SGS15lg* ; 35/37 complemented). The short anatation can't complements sgs15.1 mutation (0/39 plants are complemented).

SGS15 encodes an HAT-helix protein

The SGS15 locus is defined by two alleles: *sgs15-1* and *sgs15-2*. These two mutants exhibit no obvious developmental defects, with the exception of a late flowering phenotype (Fig. 1B), which co-segregates with PTGS-deficiency. The *sgs15* mutations were mapped to a 354 kb interval on chromosome 1. Whole-genome sequencing of *sgs15-1* revealed that fast-neutron mutagenesis had induced a 16 kb deletion in this interval, which removed six protein-coding

genes (At1g04020, At1g04030, At1g04040, At1g04050, At1g04070 and At1g04080). Sequencing of these genes in *sgs15-2* did not reveal any mutation in At1g04020, At1g04030, At1g04040, At1g04050 or At1g04070. However, part of the At1g04080 gene could not be amplified. Inverse PCR (iPCR) analysis revealed a 725kb inversion involving these loci. One end of the inversion is in the 5'UTR of At1g04080, while the other end is in the third intron of At1g01950/ARK2 (ARMADILLO REPEAT KINESIN 2). This inversion disrupts the junction between the At1g04080 ORF and its promoter, thus preventing transcription (Fig. 1C). Because At1g04080 is the only gene affected in both alleles (deleted in *sgs15-1* and not transcribed in *sgs15-2*), we assumed that it could correspond to SGS15. Consistent with this hypothesis, three mutant lines harboring T-DNA insertions in the ORF of At1g04080 exhibited the *sgs15* late flowering phenotype, indicating that the impairment of At1g04080 is responsible for the developmental defects of the *sgs15* mutants. Moreover, crosses between one of these T-DNA mutants and line *2a3* generated double homozygous plants that are impaired in cosuppression as the *sgs15-1* and *sgs15-2* mutants, indicating that At1g04080 is required for efficient PTGS. At last, transformation of *sgs15-1* mutant plants that do not carry the 2a3 locus with a pUBQ10-At1g04080-GFP fusion construct (see below) restored a wild-type developmental phenotype, while transformation of *sgs15-1* mutant plants that carry the 2a3 locus with the pUBQ10-At1g04080-GFP fusion construct restored cosuppression, indicating that the deletion of At1g04080 is responsible for both developmental defects and PTGS-deficiency in the *sgs15-1* mutant (Fig. 1D showing complementation of phenotype and NIA2 co-suppression). At1g04080 is predicted to encode a HAT-helix (Half A Tetra tricopeptide repeats (TPR) protein, also known as PRP39. Interestingly, the flowering time regulator (FLC) is up-regulated in a *prp39-1* mutant (Wang et al., 2007) explaining the change in flowering behavior of this mutant.

SGS14/LSM9a and SGS15/PRP39 are required for post-transcriptional gene silencing mediated by intron-containing transgenes.

Like SGS13/SDE3, SGS14/LSM9a and SGS15/PRP39 have been identified in a screen for PTGS-deficient mutants based on the 2a3 line. In fact, 80% of the control 2a3 plants show co-suppression of the 35-*NIA2* transgene and *NIA2* endogenous gene causing their death (Fig. 2A). Moreover, analysis on mix of plants from sgs14/lsm9a and sgs15/prp39 11 days after germination shows the accumulation of the *NIA2* mRNA with no *NIA2* siRNA like it was observed for rdr6, sgs3, hen1 (Fig. 2B). In contrast these genes are not required for PTGS

triggered by the L1 locus (Fig. 2C). Indeed, L1/sgs14-1 and L1/sgs15-1 exhibit low GUS activity and low GUS mRNA levels and accumulate GUS siRNA like the control line L1 (Fig. 2C). Moreover, SGS14/LSM9a and SGS15/PRP39 are not required for the endogenous tasiRNA pathway. Indeed, unlike in *sgs2/rdr6, sgs3, sgs4/ago1, sgs5/hen1, sgs7/sde5* and *sgs8/hpr1* mutants, no change in *TAS2* precursor and cleavage products or mature tasiRNA level was observed in *sgs14* and *sgs15*, indicating that *SGS14 and SGS15* do not encode general components of the PTGS and tasiRNA pathways (Fig. 2E).

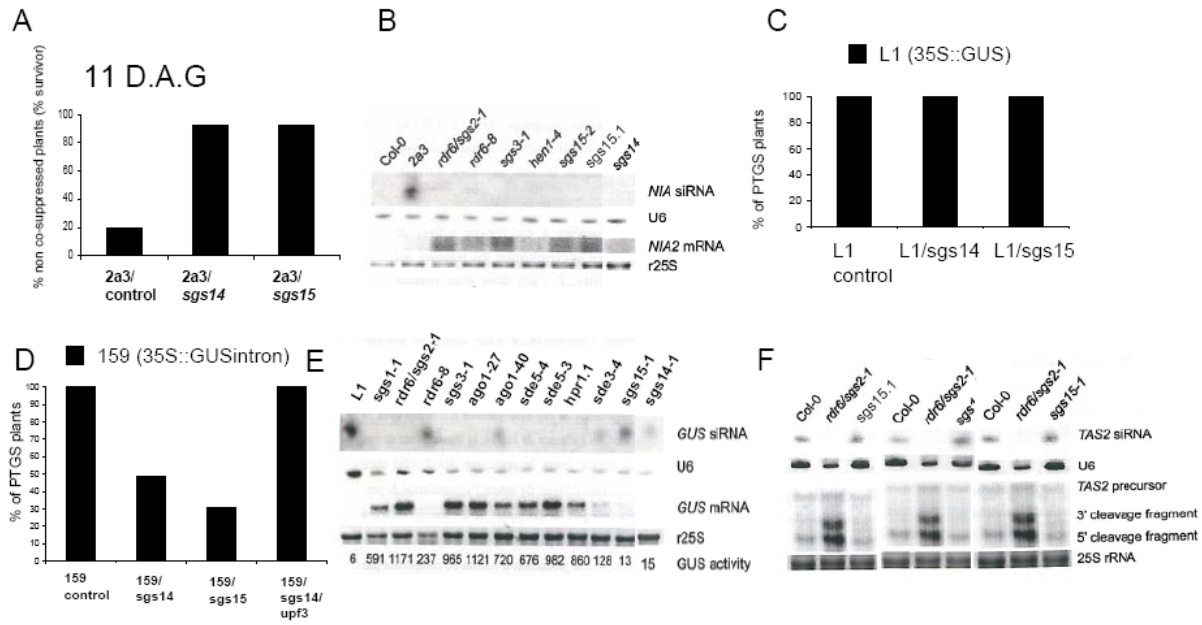


Fig. 2 SGS14/LSM9a and SGS15/PRP39 are required for post-transcriptional gene silencing mediated by intron-containing transgenes.
A) Graph representing the percentage of non co-suppressed plants (% survivor) 11 days after germination in line 2a3 (expressing an intron-containing Nia2 transgene undergoing PTGS) in WT, *sgs14* or *sgs15* backgrounds. B) Northern blot analysis showing the accumulation of *NIA2* mRNA and *NIA2* siRNA in the indicated genotype. Note the absence of siRNAs in *sgs15* and *sgs14* genotypes. C) Graph indicating the percentage of PTGS plants in line L1 (transgene without an intron) in WT and different genetic backgrounds. D) Graph indicating the percentage of PTGS plants in line 159 (same transgene than L1 with an intron) in WT and different genetic backgrounds E) Northern blot analysis showing the accumulation of GUS mRNA and siRNA and GUS activity measurement (UF/min/µg prot.) in the indicated genotype. F) Northern blot analysis showing the accumulation of tasiRNA and *TAS2* precursors (1Kb) in indicated genotype. 5' (0.4Kb) and 3' (0.6Kb) cleavage products are indicated.

Lines 2a3 and L1 differ in many aspects, including the presence of introns in the 2a3 transgene but not in the L1 transgene. To test if the specific effect of sgs14 and sgs15 on 2a3 PTGS could be due to the presence of introns in the 2a3 transgene but not in the L1 transgene, we modified the L1 transgene by inserting a plant intron into the GUS sequence. The 35S-GUS-int transgene was introduced into wild type Arabidopsis and we selected lines showing

efficient PTGS, i.e. lines exhibiting high GUS activity at early stages of development and low GUS activity at later stages in 100% of the population. One line, hereafter referred to as 159, exhibited this pattern at each generation. Consistent with PTGS characteristics, this line accumulated GUS siRNA at late stages of development. Line 159 was crossed to sgs14-1 and sgs15-1, and double homozygous plants were selected in the F2. Unlike control 159 populations, in which no GUS positive plant could be found at late stages of development, 159/sgs14-1 and 159/sgs15-1 populations exhibited 61% and 69% of GUS positive plants, respectively, at late stages of development, indicating that the presence of an intron in the 35S-GUS transgene makes PTGS sensitive to mutations in *SGS14/LSM9a* and *SGS15/PRP39* (Fig. 2D). Moreover the PTGS phenotype obtains in sgs14/159 can be complemented by introducing upf3 mutation (Fig2D), indicating that aberrant RNAs are required for PTGS triggered by introns containing transgene in sgs14 mutant. Finally, we verified that the intron contained in the transgenes from the different PTGS deficient lines 2a3 and 159 are correctly spliced when sgs14 is mutated as well as the splicing of the endogenous *NIA2* mRNA (Fig. S1A, S1B). These results indicate that the silencing phenotype observed is not linked with perturbed splicing of the transgene or the endogenous mRNA.

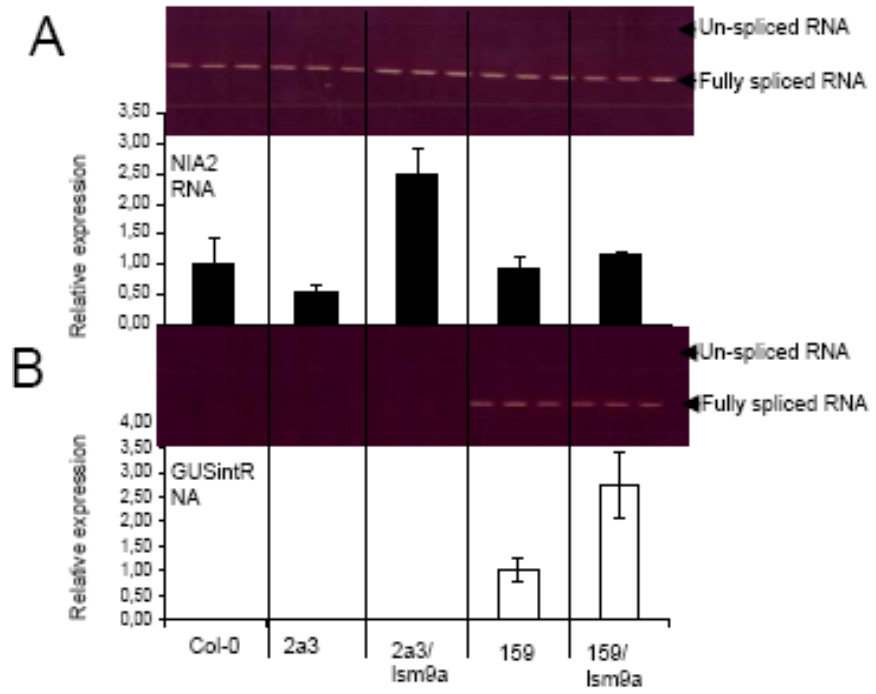


Fig. S1 transgene intron in line 2a3 and 159 are correctly expressed and spliced when *sgs14/lsm9a* is introgressed.
A) The NIA2 endogenous mRNA and the transgene NIA2 are expressed in all conditions and the intron is totally spliced in sgs14 compare to WT or to 2a3 control line.
B) GUS transgene is expressed in 159 and 159/*lsm9a* and the transgene intron is always spliced.

NSRa and NSRb encode two nuclear alternative splicing regulators required for post-transcriptional gene silencing mediated by intron-containing inverted-repeat transgenes.

We recently identified Nuclear Speckles RNA binding Protein a (NSRa) and NSRb, which regulate alternative splicing in *Arabidopsis* (Bardou et al., submitted). These plant conserved proteins localize in nuclear speckles linked to the splicing machinery. Interestingly, *nsra/nsrb* double mutants exhibit an early flowering phenotype, similar to *sgs14* mutant whereas nsra and nsrb simple mutants show only smaller variations (Figure 1A, 3A, 3B). Although nsra/nsrb do not show any variation in rosette diameter but present a reduced leaf number at flowering when compared to WT (Fig. 3B). This phenotype can be linked to the fact that the RNA binding protein FCA (which negatively controls FLC expression) is up-regulated in the *nsra/nsrb* mutant background (Fig. 3E). Furthermore, no significant changes in *FPA* expression in sgs14 and sgs15 mutants were detectable (Fig. 3E).

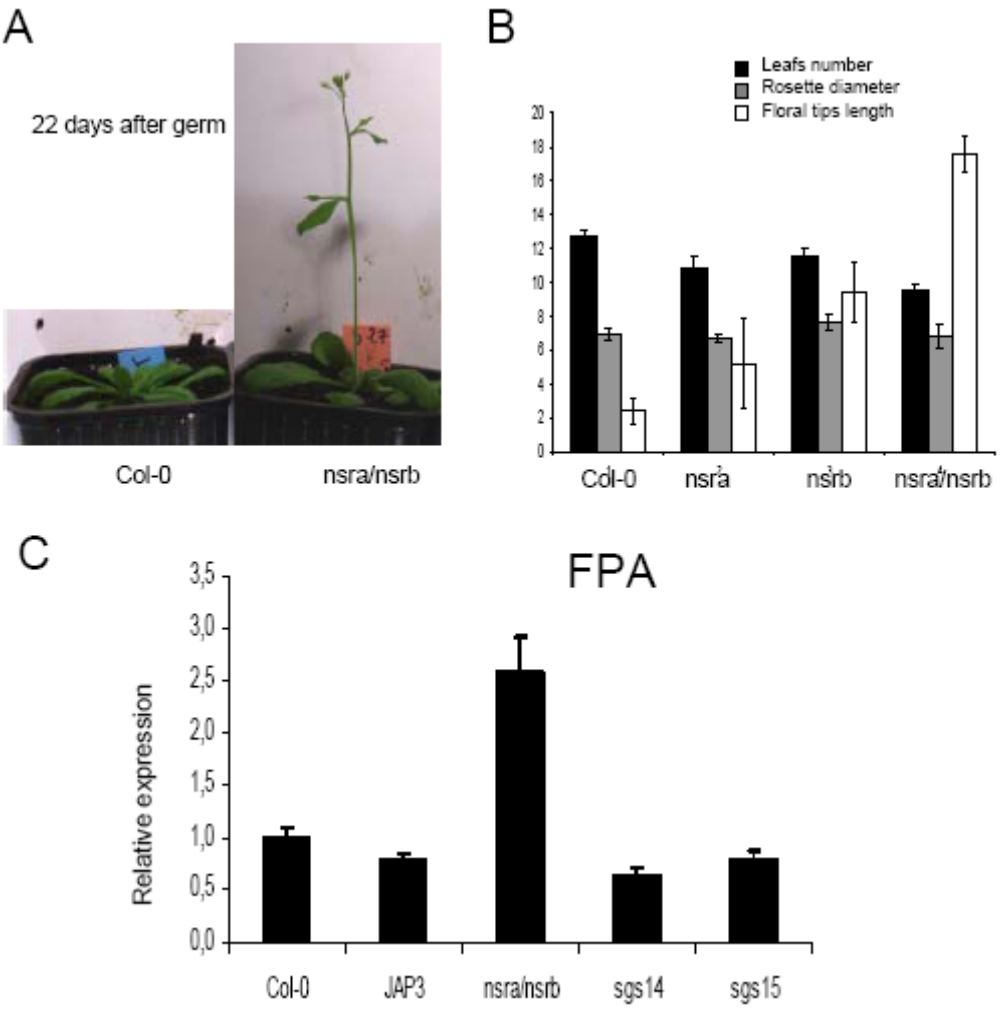


Fig. 3 The *nsra/nsrb* mutant exhibits an early floral phenotype.
A) The *nsra/nsrb* double mutant shows an accelerated floral development compare to Col-0. B) Measurement of leaf number, rosette diameter and floral tips length in *nsr* simple or double mutants compared to WT. C) qRT-PCR showing the de-regulation of the floral regulator FPA in the *nsra/nsrb, sgs14 and sgs15* compared to WT. (22 days after germination)

Given that *sgs14* and *sgs15* mutations affect PTGS triggered by intron-containing transgenes, we aimed at determining if the *nsra/nsrb* double mutation also affects PTGS. Unfortunately, the *nsra/nsrb* mutant relies on T-DNA insertions from the SALK collection, which cannot be used for analyzing silencing effects on lines 2a3, L1 and 159 due to interference of the 35S promoter in the transgene (Daxinger et al., 2008). Indeed, SALK lines produce 35S siRNA that cause TGS on transgenes expressed under the control of the 35S promoter, such as those carried by lines 2a3, L1 and 159. As there is no 35S-free *nsra* mutant available, we compared the effect of *sgs14*, *sgs15* and *nsra/nsrb* on a 35S-free transgene silenced by PTGS. The JAP transgene consists in two inverted repeat PDS segments separated by an intron, expressed under the control of the phloem companion cells-specific SUC2 promoter. Thus, JAP is a

35S-free intron-containing transgene. The transgenic JAP3 locus produces a PDS double-stranded RNA (containing an intron) in vascular tissues (due to the SUC2 promoter) and targets PDS transcripts, required for chlorophyll synthesis. When silencing occurs, a progressive bleaching starts in the vascular tissues of the leaves and spreads into other mesophyl cells (Smith et al., 2007). The JAP3 line was crossed to *sgs14, sgs15* and *nsra/nsrb* mutants, and double or triple homozygous plants were identified in the F2. The *sgs14/lsm9a* and the *nsra/nsrb* double mutations impaired JAP3 silencing, whereas *sgs15/prp39* mutation seemed to slightly increase JAP3 silencing (Fig 4A). As *SGS14/LSM9a, SGS15/PRP39, NSRa* and *NSRb* encode putative splicing regulators, we analyzed the processing of the JAP3 transgene that contains an intron. No significant change in splicing of this transgenic intron was observed in any of the mutant lines (*nsra/nsrb*, *sgs14/Lsm9a* and *sgs15/prp39*) (Fig. 4B). We then tested PDS siRNA production to characterize more in detail the silencing phenotype (Fig. 4C). Moreover, in all mutants line (*nsra/nsrb*, *sgs14/Lsm9a* and *sgs15/prp39*) JAP3 transgene is expressed and PDS expression is down regulated compared to WT (Fig. 4D).
Altogether, these data indicate that the *sgs14, sgs15* and *nsra/nsrb* mutations affect PTGS efficiency in distinct ways, which are not linked to the perturbation of transgene splicing.

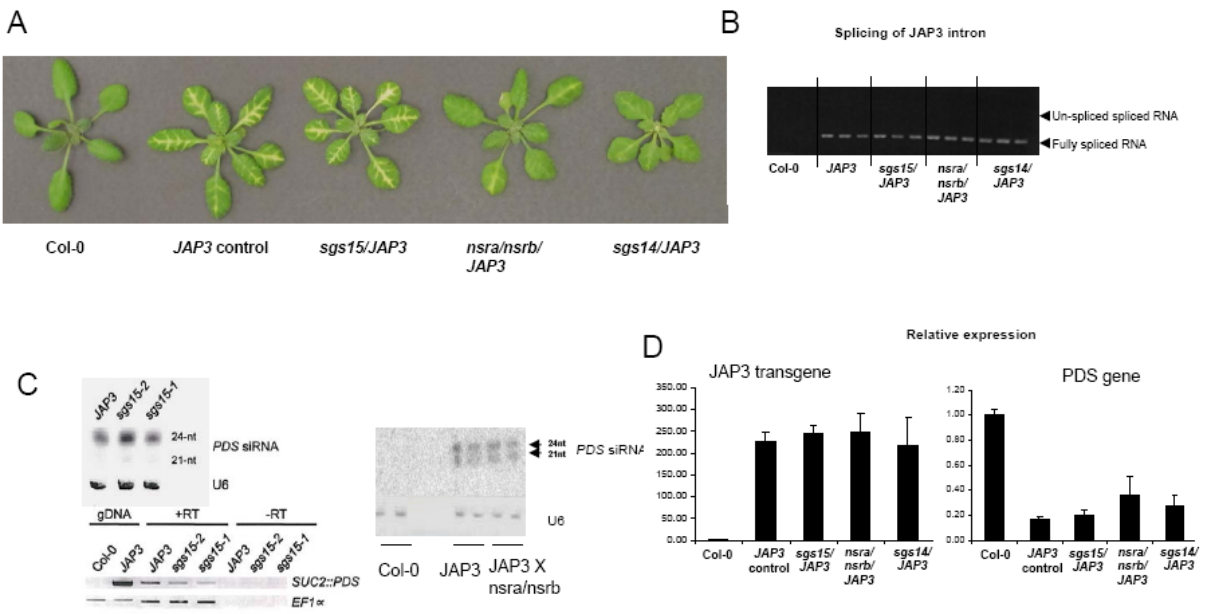


Fig. 4 SGS14/LSM9a, SGS15/PRP39 and NSRa/NSRb contribute to the silencing phenotypes induced by an IR transgene expressed in vascular tissues (JAP3).
A) PDS RNA silencing spreading from phloem companion cells is detected by a chlorotic phenotype. This phenotype is clearly observable in JAP3 transgenic plants (compared to on-transgenic Col 0) and in JAP3/*sgs15* plants. The chlorotic leaf phenotype is lost or restricted to the vasculature in *sgs14* or *nsra/nsrb* mutants introgressed into JAP3 line. B) RT-PCR showing JAP3 intron transgene splicing in indicated accession. C) Northern blot analysis showing the accumulation of *PDS* siRNA in the

indicated genotype. D) Relative expression of JAP3 trangene and of *PDS* mRNA in indicated mutants compare to WT obtained by RT-qPCR.

SGS14/LSM9a colocalizes with splicing-related factors NSRa and SRP34 in nuclear speckles

We then explored where these proteins are expressed in the cell, notably in relation to known components of the splicing or silencing machineries. Indeed, the NSRA and NSRB proteins co-localised with SRP34 and snRNP protein fusions in nuclear splicing-related "speckles" (Bardou et al., submitted). Fusion proteins were generated between GFP and SGS14/LSM9a or SGS15/PRP39 and expressed under the control of the UBQ10 promoter. At first, the functionality of these constructs was verified by transformation of *sgs14-1* and *sgs15-1* mutant plants carrying or not the 2a3 locus. As a single gene model is predicted for SGS14/LSM9a, a single construct was generated, carrying the SGS14/LSM9a genomic sequences comprised between the ATG and the last amino acid before the stop codon fused to the GFP in 3'. The pUBQ10-SGS14-GFP construct restored wild type development when introduced in *sgs14-1* plants (Fig. 1D). In contrast, two gene models are predicted for SGS15/PRP39, which have similar 3' ends but distinct 5' ends. Therefore, two constructs (a long and a short) were generated, carrying SGS15/PRP39 genomic sequences comprised between the ATG and the last amino acid before the stop codon of each model. The long pUBQ10-SGS15-GFP construct restored a wild type development when introduced into *sgs15-1* plants and efficient co-suppression of Nia2 when introduced in 2a3/*sgs15-1* plants (Fig. 1E). In contrast, the short pUBQ10-SGS15-GFP construct did not complement the developmental and co-suppression defects of sgs15-1 mutant plants.

Transient expression of pUBQ10-SGS14-GFP and long pUBQ10-SGS15-GFP constructs in *Nicotiana benthamiana* leaves revealed nuclear localization of both fusion proteins (Fig. 5A, 5B). Interestingly, they showed specific sub-cellular localization patterns inside the nucleus. The SGS14-GFP fusion localized in nuclear dots, which strongly resembling the splicing-related "speckles" (Fig. 5B, 6B, 6D, 6F). Co-infiltration of SGS14-GFP with NSRa-RFP or SRP34-RFP confirmed the co-localization of these proteins in the nuclear "speckles" (Fig. 6A, 6B, 6D). In contrast the SGS15-GFP fusion showed a nuclear signal that was more spread on the nucleoplasm and showing much less particles (Fig. 5B, 6C, 6E). Moreover, only few dots co-localised with NSRa-RFP (Fig. 6C). Finally, analysis of the localization of NSRa, NSRb, LSM9a and PRP39 fluorescent protein fusions suggest that these protein are excluded from nucleolus (Fig. 5), as also confirmed using the nucleolar markers XRN2-GFP or XRN2-

RFP (Fig. 6E to 6G). However, LSM9a-GFP shows also a peri-nucleolar signal that can not be found in the PRP39 and NSRs (Fig. 5A, 6B, 6F). Therefore, these proteins show specific subcellular localization in the cell nucleus that may suggest a potential link with splicing regulation.

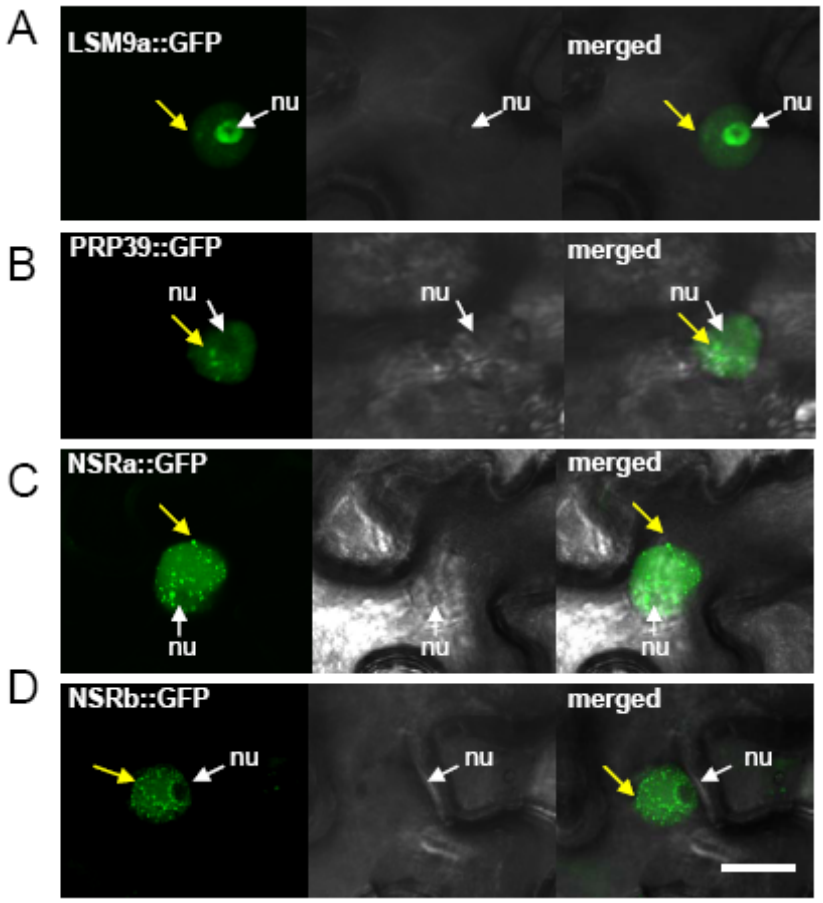


Fig. 5 SGS14/LSM9a, SGS15/PRP39, NSRa and NSRb localised in nuclear speckles.
Confocal images of transformed tobacco leaf cells with the indicated constructs 2 days after transformation (left panels) and direct light transmission (middle panels) or the merged images (right panel). White and yellow arrows indicate nucleolus and nuclear dots, respectively. Scale=10µm, A) The pUBQ10-LSM9a-GFP fusion is localized in nuclear dots and presents a peri-nucleolar signal. B) C) and D) pUBQ10-PRP39-GFP, pUBQ10-NSRa-GFP, pUBQ10-NSRb-GFP localized in nuclear dots and are not present in the nucleolus.

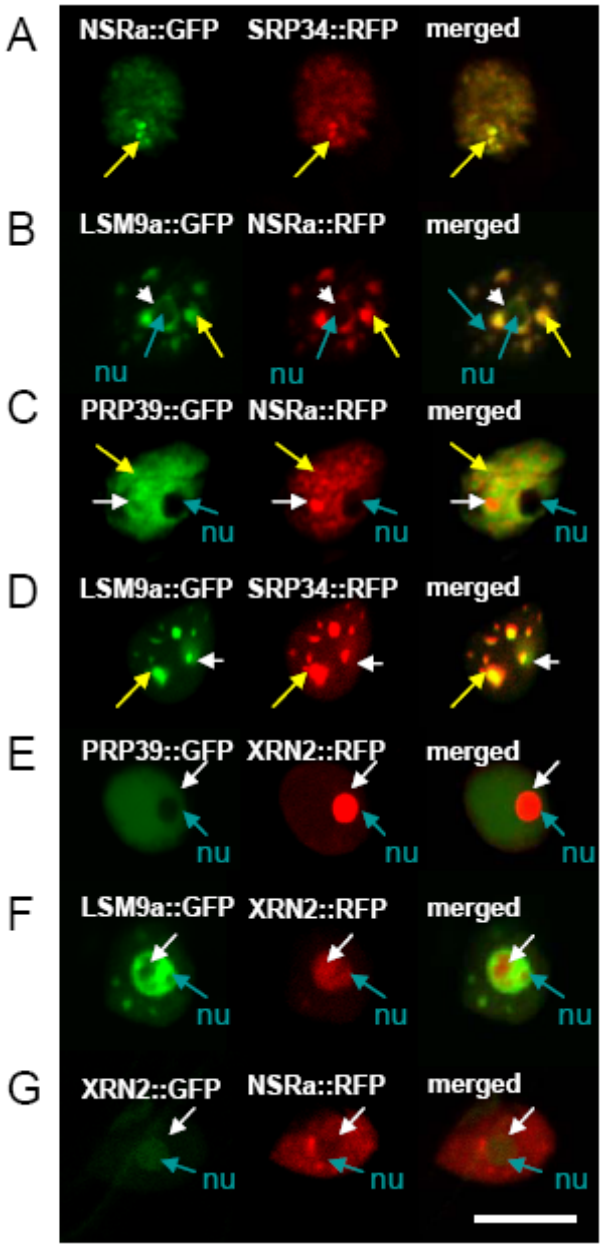


Fig. 6 SGS14/LSM9a, SGS15/PRP39, NSRa and SRP34 partially co-localised in nuclear speckles.

Confocal imaging of co-transformed tobacco leaves. White and yellow arrows indicate non co-localised or co-localised particles, respectively. nu=nucleolus. Scale=10µm. The constructs are indicated on each panel and the merged image with both signals is shown in the left panel.

A) 35S-NSRa-GFP and pUBQ10-SRP34-RFP. B) pUBQ10-LSM9a-GFP and 35S-NSRa-RFP. C) pUBQ10-PRP39-GFP. D) pUBQ10-LSM9a-GFP and pUBQ10-SRP34-RFP. E) pUBQ10-PRP39-GFP and 35S-XRN2-RFP. F) pUBQ10-LSM9a-GFP and 35S-XRN2-RFP. G) 35S-XRN2-GFP and 35S-NSRa-RFP.

SGS14/LSM9a, SGS15/PRP39, NSRa and NSRb have differential roles in splicing

We previously reported that the nsra/nsrb double mutant exhibit splicing defects in relation to auxin in several genes known to be alternatively spliced (Bardou et al., submitted). Because NSRs, LSM9a, and PRP39 are splicing-related proteins showing nuclear signals that co-localize with the splicing machinery, we tested if these proteins are able to modulate alternative splicing using three genes known to produce alternatively spliced transcripts.

The CCA1 (At2g48830) locus produces four alternatively spliced transcripts, CCA1.1 producing wildtype CCA-α and CCA1.2/3/4 producing truncated CCA- β due to intron retentions (Fig. S2A). Our results show that the nsra/nsrb double mutations up-regulate the formation of CCA- β (Fig. 7A, S2A, S2E). The sgs14/lsm9a and sgs15/prp39 mutants give an opposite results, i.e. they both decrease isoforms producing CCA1- β compare to WT.

The *Fbox* (At4g27050) locus produces four alternatively spliced transcripts (Fig. 7B, S2B, S2E). Intron retention increases in sgs14/lsm9a and sgs15/prp39 mutants compare to WT, as shown by the decrease in the accumulation of the fully spliced transcript. In contrast, no changes are observed in the nsra/nsrb double mutant.

The Transformer Serine/Arginine Rich protein (At4g35785) locus produces at least four alternatively spliced transcripts, One of them is down-regulated in the prp39 mutant, but not in the other mutants (Fig. 7C, S2C, S2E).

Altogether, these results indicate that mutations in SGS14/LSM9a, SGS15/PRP39 and NSRa/b distinctly affect alternative splicing.

Mutations in SGS14/LSM9a, SGS15/PRP39, NSRa and NSRb have marginal effects on the PTGS machinery

To test if the silencing defects observed in sgs14/lsm9a, sgs15/prp39 and nsra/nsrb mutants could be due to changes in the expression of key silencing genes, we performed q-PCR on 18 genes involved in silencing (SDE5.1, SDE5.2, RDR6, DCL4, DRB4, HEN1.1, HEN1.2, AGO1.1, AGO1.3, SQN, HSP90.2, DCL2, MAD3, CLSY1, MAD4, NRPD1 a, NRPD1 b, RDR2, DCL3, AGO4 and JMJ14). The majority of these genes did not show any major changes in expression or splicing. Only one of the two annotations of HEN1 and AGO1 were differentially expressed in *nsra/nsrb* or *sgs14, sgs15* mutants (Fig. S3A, S3B).

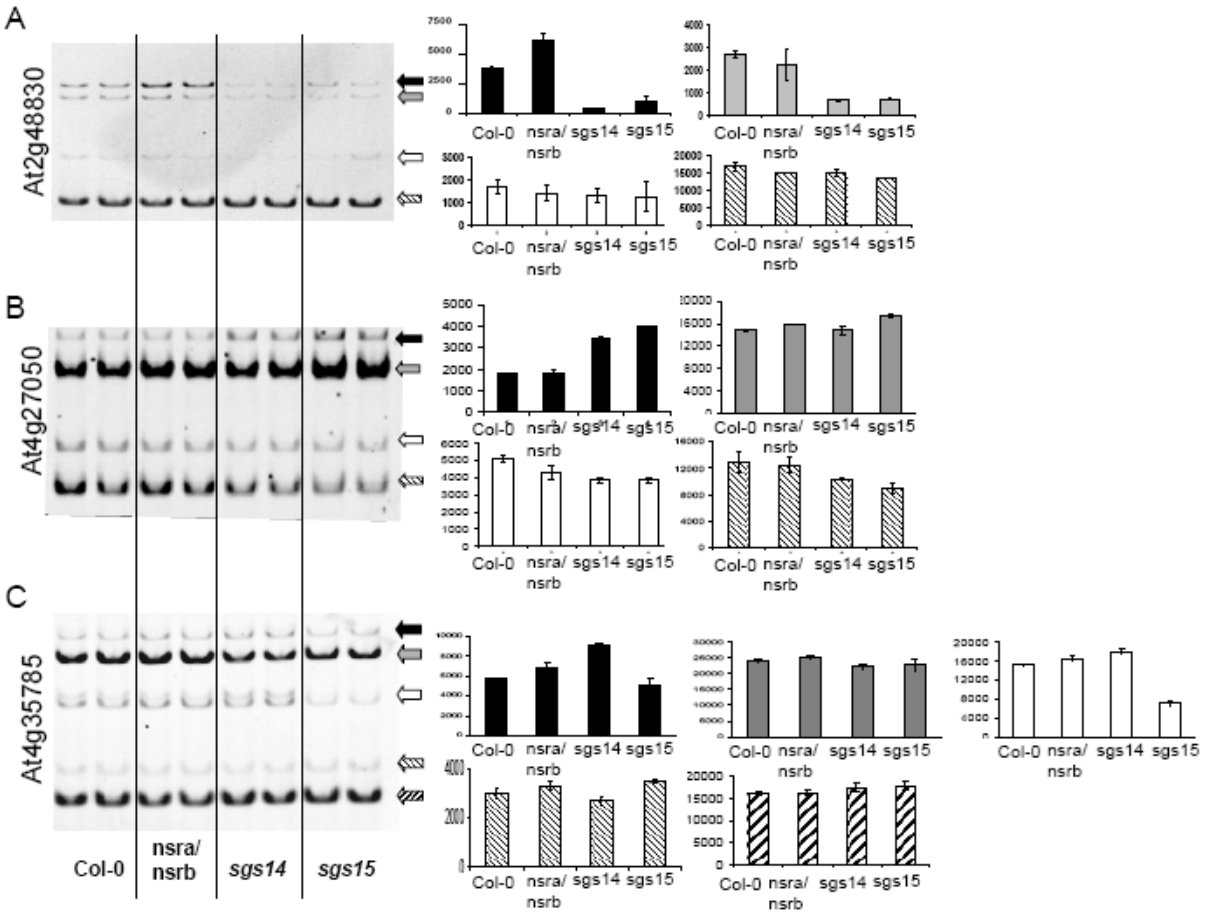


Fig. 7 SGS14/LSM9a, SGS15/PRP39, NSRa and NSRb have differential roles in alternative splicing of common targets.
Quantitative changes in isoforms distribution for selected mRNA targets undergoing alternative splicing. Right panels show gels depicting the different isoforms using oligos spanning alternatively spliced exons (indicated in Fig. S3). Left graphs showed the quantification (based on triplicate experiments) of the fluorescent signal observed for each isoforms (according to the color code of the arrow). Note that for the different targets, distinct accumulations of isoforms are observed between *nsra/nsb*, *sgs14 or sgs15* mutants. Qualibration of cDNA used for this experiment are shown in Fig. S3D).

Moreover, we performed RT-PCR gel assay on several genes that were predicted to be alternatively spliced, to visualize their AS profiles in nsra/nsrb, sgs14/lsm9a or sgs15/prp39 mutants lines. Only the HSP90.2 AT5G56030 molecular chaperon protein shows variations in sgs14/Lsm9a, sgs15/prp39, nsra/nsrb. (Fig. S3C, S2D, S2E)

Altogether our results indicate that the effect of mutations in *SGS14/LSM9a, SGS15/PRP39, NSRa and NSRb* on PTGS is not due to an indirect effect on the PTGS machinery.

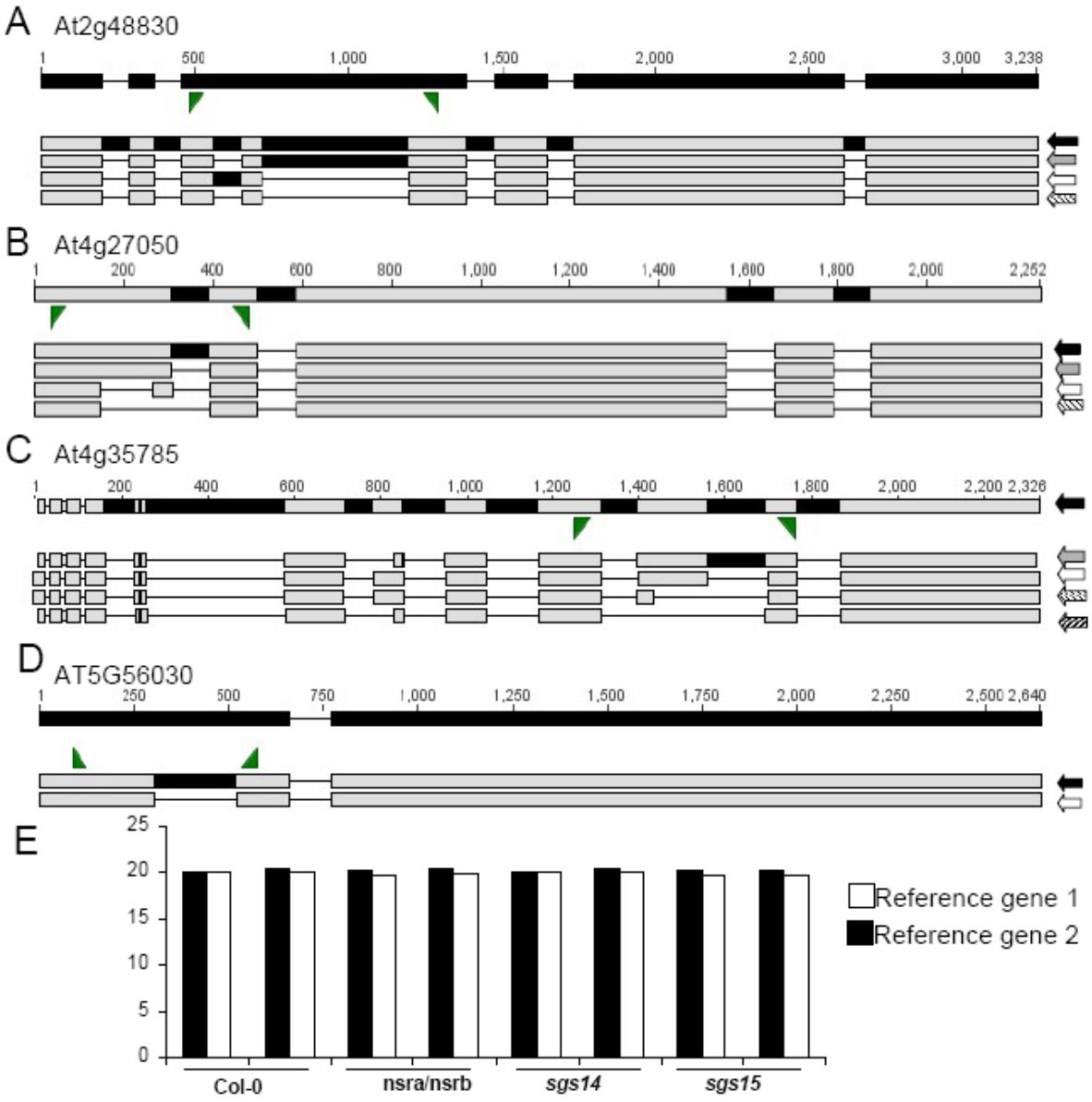


Fig. S2 AS targets genes model.
Green arrows indicate primer used for RT-PCR AS assays and all band quantified in Fig.8 are represented by arrows in corresponding model. A) *CCA1* gene (At2g48830). B) *Fbox* gene (At4g27050). C) *Transformer Serine/Arginine Rich protein* coding gene (At4g35785). D) HSP90.2 (AT5G56030) E) RT-qPCR representing the qualibration of the cDNA used for the AS experiments for each replicats in each mutant background with the two constitutive genes (AT1G13320; AT4G26410; Czechowski & al., 2005).

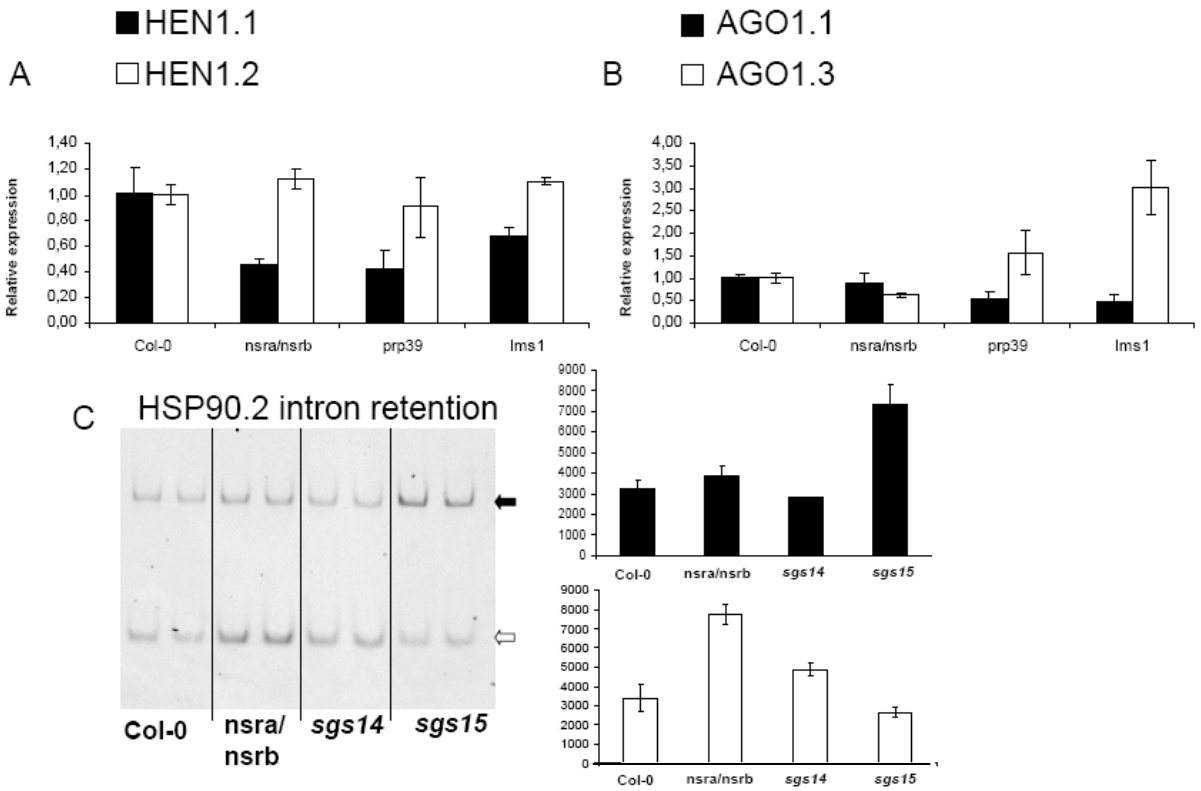


Fig. S3 Expression and splicing of key components of silencing mechanisms
RT-qPCR analyses reveal the relative expression of A) *HEN1.1*, *HEN1.2* and B) *AGO1.1*, *AGO1.3* in indicated mutants. C) RT-PCR showing the AS of *HSP90.2* and band quantifications.

DISCUSSION:

This work demonstrates a new link between splicing and PTGS and identified components of the splicing machinery that play a role in the PTGS of intron-containing transgenes.

Components of RNA processing complexes other than those involved in siRNAs biogenesis or action were shown to modulate gene silencing including exoribonucleases XRN2, XRN3, XRN4 and their regulator FRY1, P-bodies components DCP1, DCP2 and VCS, NMD components UPF1 and UPF3, and 3' end processing factor FCA, FPA, FLK, FY, ESP1, ESP4 and ESP5 (Herr et al, 2006; Gy et al., 2007; Motomura et al., 2012; Arciga-Reyes et al., 2006; Bäurle et al., 2007; Simpson et al., 2007). However, only the putative splicing factor ESP3/PRP2 has linked splicing to PTGS (Herr et al, 2006). Nevertheless, one paper has suggested that the presence of introns decreased transgene susceptibility to PTGS (Christie et al, 2011).

Alternative splicing is a process that is regulated to increase protein diversity and may lead to the formation of NMD substrates. In Arabidopsis AS in introns in 5' or 3'UTRs modulate NMD-sensitivity of mRNA transcripts (Kalyna et al., 2012) and suggest that AS can generate potential aberrant RNAs like those generated by the silencing machinery revealing the extensive cross-talk between these different mRNA processing activities. We have shown that alternative splicing is regulated by AtNSRs (previous chapter). We show here that the nsra/nsrb double mutant exhibits defects in PTGS of the JAP3 transgene, suggesting another link between transgene PTGS and splicing. It is noteworthy, that splicing regulators are also involved in the control of miRNA processing (e.g. Serrate) (Laubinger et al. 2008) whose precursors are long non-coding RNAs, sometimes showing complex splicing patterns such as miR162a (Hirsch et al., 2006), and as such, potential NMD substrates.

Splicing factors have been previously described to be localized in different nuclear speckles or interchromatin granules under the name speckles and housing the various components of the splicing machinery in Arabidopsis (Lorković et al., 2008). These different splicing speckles may explain the variable localization patterns found for PRP39 and SRP34. In contrast, SRP34 co-localizes with the SGS15 and NRSs proteins in the large majority of the particles observed suggesting a more coordinated function. Interestingly localization of NSRs in speckles depends on their RRM domains (Campalans et al., 2004) suggesting that these ribonucleoprotein particles aggregate around their RNA substrates to perform splicing reactions.

RT-PCR has shown several defects in alternative splicing in the sgs14, sgs15 and nsra/nsb mutants but not in the same way. In fact, The CIRCADIAN CLOCK-ASSOCIATED1 (CCA1; At2g48830) is one of the core clock components in Arabidopsis and is a common target of these regulators. This transcription factor has previously be described to be alternatively spliced. Interestingly Seo et al., 2012 demonstrate the self-regulation of Arabidopsis thaliana CCA1 activity by a splice variant, CCA1 β, linking the clock to cold acclimation. CCA1 alternative splicing produces two isoforms, the full-size CCA1α and the truncated CCA1β. The CCA1β isoform has the protein-protein interaction domain that mediates dimer formation. However, it lacks the N-terminal MYB DNA binding domain, unlike the CCA1α isoform. CCA1β competitively inhibits CCA1α activity by forming non-functional heterodimers CCA1α-CCA1β and LHY-CCA1β, which have reduced DNA-binding affinities (Park et al., 2012). These results demonstrate the link between AS and developmental variations.

The Fbox AS target was previously shown to be differentially spliced in response to auxin treatment in roots (Bardou et al., submitted, previous chapter). The AS splicing variation induced by auxin was lost in the *nsra/nsrb* double mutant treated with auxin. This event occurs on the 3'UTR and have no consequences on the coding region however, alternative splicing occurring in the 5' and 3' untranslated regions (UTRs) of mRNAs can have other consequences on mRNA stability or translational regulation. The understanding of the significance and the regulation of these variations is rather limited. The mouse Gli1 oncogene, which is the terminal transcriptional effector of the Hedgehog (HH) signaling pathway is alternatively spliced in the 5'UTR and the Gli1 5' UTRs had an impact on translational capacity, with the shorter and the exon 1B-skipped mRNA variants being most effective (Palaniswamy et al., 2010). In Arabidopsis recent data has shown the importance of AS in introns in 5' or 3'UTRs also in modulating NMD-sensitivity of these mRNA transcripts (Kalyna et al., 2012).

The link between PTGS and alternative splicing can be further linked to chromatin interactions. Indeed, several genes involved in the spreading of gene silencing (Dunoyer and Voinnet, 2009; Dunoyer et al., 2007) encode for genes linked to chromatin-related pathways such as the DCL3/RDR2/AGO4 pathway. Other chromatin regulators have been identified to play roles in gene silencing by modifying the chromatin status of the transgene (Le Masson et al., 2012). In our work, we did not detect major changes in splicing patterns of the transgenes triggering PTGS. Alternatively, in human cell, a link between alternative splicing and epigenetic is emerging (Luco et al., 2010). Distinctive histone modification signatures

correlate with the alternative splicing outcome in a set of human genes, and the modulation of histone modifications in exon regions can lead to switches in splice sites. In fact, the authors propose that histone marks affect splicing outcome by influencing the recruitment of splicing regulators to specific mRNA targets. We can propose that the NSRa, NSRb, SGS14 and SGS15 recruit chromatin-binding proteins and other epigenetic regulators to the mRNA from intron containing transgenes. This may change the nature of the RNAs produced by this locus, although not his splicing pattern, and enhance the amount of aberrant RNAs required to make transgenes prone to trigger PTGS.

EXPERIMENTAL PROCEDURES

Plant material and growth conditions.

All Arabidopsis plants are in the Columbia accession. The JAP3 line (Smith et al, 2007) was a kind gift of D. Baulcombe. Lines L1 and 2a3 have been described before (Jauvion et al, 2010). Line 159 exhibits 100% PTGS at each generation, like line L1, and carries a 35S-GUS transgene similar to L1, expect for the presence of an intron. sgs14 and sgs15 mutants were identified in a PTGS forward genetic screen based on the 2a3 locus (Jauvion et al, 2010). nsra and nsrb correspond to SALK_003214 and SAIL_717_F03.v1, respectively. Plants were grown in long day (16-h light/8-h dark) or continuous light conditions at 23°C on soil or on solid half-strength MS medium containing 0,5% sucrose.

Cloning procedures

All the clones were made using the Gateway technology (Invitrogen) and planned using Geneious (http://www.geneious.com). A list of the oligonucleotides used for cloning is provided in Table 1.

Nicotiana benthamiana agro-infiltration

Agrobacterium (ASE or Agl0 strains) carrying plasmids of interest were grown overnight at 30°C in 3ml LB medium containing the appropriate antibiotics to a final OD600 of between 1.0 and 2.0. The bacteria were pelleted and resuspended in 1ml of infiltration medium (10mM MgCl2, 10mM MES pH5.2, 150 mM acetosyringone) to a final OD600 of 0.1. The bacterial solution was infiltrated into the abaxial side of leaves using a 1ml syringe and samples were assayed 3 days after infiltration.

Confocal imaging

For confocal imaging, agro-infiltrated tobacco leaves (mounted in water) were directly imaged on a Leica TCS SP2 (Leica Microsystems). The CFP was imaged with 458nm excitation using the dichroic mirror DD458/514, and detection window of 465-505nm; the GFP was imaged with 488nm excitation using the dichroic mirror DD488/543, and a detection window of 500-580nm; and the RFP was imaged with 543nm excitation using the dichroic mirror DD488/543, and a detection window of 580-670nm. For the colocalisations, all of the images were taken by sequential acquisition. Image analysis was performed using the National Institute of Health ImageJ (http://rsb.info.nih.gov/ij/) software.

RT-PCR and RT-qPCR

Total RNA was prepared from roots and plantlets at different developmental stages using the Qiagen RNeasy plant mini kit. The DNAse treatment was performed according to the manufacturer's protocols. For reverse transcription with SuperScriptII (Invitrogen), 2.5 µg of total DNAse-treated RNA was used. One microliter of the resulting cDNA solution was used for RT-PCR or RT-qRT analyses using standard protocols. A complete list of PCR primers is available in Table 1. Each cDNA sample was precisely calibrated and verified for two constitutive genes (AT1G13320; AT4G26410; Czechowski & al., 2005). For RT-PCR, the amplification was as follows: one cycle of 4 min at 98°C, 26 cycles of 30 s at 98°C, 30 s at 59°C, and 1 min at 72°C. The products were separated in 7.5% acryl gel stained with SyBr green (Invitrogen) and revealed by Pharos Imager (Biorad). Band profiles were quantified using ImageJ (http://rsbweb.nih.gov/ij/). Mutants in Figure 4C were all in JAP3 background. RT-qPCR was performed using a Roche Light Cycler 480 using standard protocols (40 cycles, 60°C annealing).

Acknowledgments

This work has benefited from the facilities and expertise of the Imagif Cell Biology Unit of the Gif campus (www.imagif.cnrs.fr) which is supported by the Conseil Général de l'Essonne.

References

Arciga-Reyes, L., Wootton, L., Kieffer, M., and Davies, B. (2006). UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in Arabidopsis. The Plant journal: for cell and molecular biology **47**: 480–9.

Boutet, S., Vazquez, F., Liu, J., Béclin, C., Fagard, M., Gratias, A., Morel, J.B., Crété, P., Chen, X., and Vaucheret, H. (2003). Arabidopsis HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. Current biology **13**: 843–8.

Bäurle, I., Smith, L., Baulcombe, D.C., and Dean, C. (2007). Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. Science (New York, N.Y.) **318**: 109–12.

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

Christie, M., Croft, L.J., and Carroll, B.J. (2011). Intron splicing suppresses RNA silencing in Arabidopsis. The Plant journal: for cell and molecular biology **68**: 159–67.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.-R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant physiology **139**: 5–17.

Daxinger, L., Hunter, B., Sheikh, M., Jauvion, V., Gasciolli, V., Vaucheret, H., Matzke, M., and Furner, I. (2008). Unexpected silencing effects from T-DNA tags in Arabidopsis. Trends in plant science **13**: 4–6.

Dunoyer, P., Himber, C., Ruiz-Ferrer, V., Alioua, A., and Voinnet, O. (2007). Intra- and intercellular RNA interference in Arabidopsis thaliana requires components of the microRNA and heterochromatic silencing pathways. Nature genetics **39**: 848–56.

Dunoyer, P. and Voinnet, O. (2009). Movement of RNA silencing between plant cells: is the question now behind us? Trends in plant science **14**: 643–4.

Elmayan, T., Balzergue, S., Béon, F., Bourdon, V., Daubremet, J., Guénet, Y., Mourrain, P., Palauqui, J.C., Vernhettes, S., Vialle, T., Wostrikoff, K., and Vaucheret, H. (1998). Arabidopsis mutants impaired in cosuppression. The Plant cell **10**: 1747–58.

Fagard, M., Boutet, S., Morel, J.B., Bellini, C., and Vaucheret, H. (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. Proceedings of the National Academy of Sciences of the United States of America **97**: 11650–4.

Gy, I., Gasciolli, V., Lauressergues, D., Morel, J.-B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C. (2007). Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. The Plant cell **19**: 3451–61.

Herr, A.J., Molnàr, A., Jones, A., and Baulcombe, D.C. (2006). Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **103**: 14994–5001.

Jauvion, V., Elmayan, T., and Vaucheret, H. (2010). The conserved RNA trafficking proteins HPR1 and TEX1 are involved in the production of endogenous and exogenous small interfering RNA in Arabidopsis. The Plant cell **22**: 2697–709.

Jauvion, V., Mallory, A.C., and Vaucheret, H. (2011). Deciphering post-transcriptional Gene Silencing Pathways Through Genetic Screens. In (RNAi Technology), pp. 17 –46.

Kalyna, M., Simpson, C.G., Syed, N.H., Lewandowska, D., Marquez, Y., Kusenda, B., Marshall, J., Fuller, J., Cardle, L., McNicol, J., Dinh, H.Q., Barta, A., and Brown, J.W.S. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. Nucleic acids research **40**: 2454–69.

Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Rätsch, G., and Weigel, D. (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America **105**: 8795–800.

Lorković, Z.J., Hilscher, J., and Barta, A. (2008). Co-localisation studies of Arabidopsis SR splicing factors reveal different types of speckles in plant cell nuclei. Experimental cell research **314**: 3175–86.

Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. Science (New York, N.Y.) **327**: 996–1000.

Le Masson, I., Jauvion, V., Bouteiller, N., Rivard, M., Elmayan, T., and Vaucheret, H. (2012). Mutations in the Arabidopsis H3K4me2/3 Demethylase JMJ14 Suppress Posttranscriptional Gene Silencing by Decreasing Transgene Transcription. The Plant cell **24**: 3603–12.

Morel, J.-B., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., Proux, F., and Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. The Plant cell **14**: 629–39.

Motomura, K., Le, Q.T., Kumakura, N., Fukaya, T., Takeda, A., and Watanabe, Y. (2012). The role of decapping proteins in the miRNA accumulation in Arabidopsis thaliana. RNA biology **9**: 644–52.

Mourrain, P., Béclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N., Rémoué, K., Sanial, M., Vo, T.A., and Vaucheret, H. (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell **101**: 533–42.

Palaniswamy, R., Teglund, S., Lauth, M., Zaphiropoulos, P.G., and Shimokawa, T. (2010). Genetic variations regulate alternative splicing in the 5' untranslated regions of the mouse glioma-associated oncogene 1, Gli1. BMC molecular biology **11**: 32.

Seo, P.J., Park, M.-J., Lim, M.-H., Kim, S.-G., Lee, M., Baldwin, I.T., and Park, C.-M. (2012). A self-regulatory circuit of CIRCADIAN CLOCK-ASSOCIATED1 underlies the circadian clock regulation of temperature responses in Arabidopsis. The Plant cell **24**: 2427–42.

Simpson, G.G., Quesada, V., Henderson, I.R., Dijkwel, P.P., Macknight, R., and Dean, C. (2004). RNA processing and Arabidopsis flowering time control. Biochemical Society transactions **32**: 565–6.

Smith, L.M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F.K., Herr, A.J., Pikaard, C.S., and Baulcombe, D.C. (2007). An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in Arabidopsis. The Plant cell **19**: 1507–21.

Wang, C., Tian, Q., Hou, Z., Mucha, M., Aukerman, M., and Olsen, O.-A. (2007). The Arabidopsis thaliana AT PRP39-1 gene, encoding a tetratricopeptide repeat protein with similarity to the yeast pre-mRNA processing protein PRP39, affects flowering time. Plant cell reports **26**: 1357–66.

Genetic constructions

Accession	gene	native stop yes/no	entry vector	donor vector	resistance coli	resistance Agro	agro	promoter 35S/UBI	tag
At1G76940	NSRa	no	pENTR/D	pUBQ10-GFP	KM	KM + Chloramph	ASE	UBI	GFP
At1G21320	NSRb	no	pENTR/D	pUBQ10-GFP	KM	KM + Chloramph	ASE	UBI	GFP
At1G76940	NSRa	no	pENTR/D	pB7RWG2	KM	KM + Chloramph	ASE	35S	RFP
AT4G02840	LSM9a	no	pENTR/D	pUBQ10-GFP	KM	KM + Chloramph	ASE	UBI	GFP
At1g04080	PRP39sh	no	pENTR/D	pUBQ10-GFP	KM	KM + Chloramph	ASE	UBI	GFP
At1g04080	PRP39lg	no	pENTR/D	pUBQ10-GFP	KM	KM + Chloramph	ASE	UBI	GFP
AT1G02840	SRP34	no	pENTR/D	pB7RWG2	KM	KM + Chloramph	ASE	35S	RFP
AT5G42540	XRN2	no	pDONR221	pB7RWG2	KM	KM + Chloramph	ASE	35S	RFP

oligo for cloning

Accession	gene	oligo FW	oligo REV
At1G76940	NSRa:GFP	caccATGGCGGATGGGTACTGGAACCAGC	CCTTCTTCCTCTTTGTCCTGGTCT
At1G21320	NSRb	caccTGTTTCCGTCGCCGTCAATG	CCTTCTTCCTCCACGCTGTCC
At1G76940	NSRa:RFP	caccATGGCGGATGGGTACTGGAACCAGC	CCTTCTTCCTCTTTGTCCTGGTCT
AT4G02840	LSM9a	caccGCTGCTTTTCCTCCTCCCGCTC	ACGACCACCACGGCCACGTCC
At1g04080	PRP39sh	caccGGCTTAGGTAGCCCTTGACTG	GTAGTACGTGTTGTAGTACG
At1g04080	PRP39lg	caccGCGTCGCCTTTCGTTTTGGTTC	GTAGTACGTGTTGTAGTACG
AT1G02840	SRP34	caccAATAAACCATGAGCAGTCGTTCGAGTAGA	CCTCGATGGACTCCTAGTGTG
AT5G42540	XRN2	GGAGATAGAACCATGGGAGTTCCGTCGTTCTACAG	CCTCCGGATCMAGCTGTTTTGGGAGGAGCTCC

oligo for RT-PCR (Alternative splicing and/or RT-qPCR)

Accession	gene	oligo FW	oligo REV
At2g43410	FPA	AGCCCGGGATTCACCTGCCA	CCCAGCTTGAGGCGCTGCAT
JAP3 transgene	PDS	TCAGCGGCCGCTTTGTATGCCAGTAGTGGATCATA	CTAGTCTAGAATCGATCGGAAG
PDS mRNA	PDS	GAGCGTGTATGCCGACATGTCC	CATTTACTGAAGAGAGTAAGG
JAP3 splicing	PDS	CCGTGCCATCGTCATTGAGCTCAA	CCGTGCCATCGTCATTGAGCTCAA
At2g33830	Auxin rlt	CCGGACCTAAACCGGAGCATGGCC	CCGATCCTGGCGTCGTCGGAGTTCC
AT5G56030	HSP90.2	GGCGGACGCTGAAACCTTTGC	GCTTGCTCTTGTCCGTCAAGGACTC
At2g48830	CCA1	ggactgaggaagaacataatag	ggtttacgcttaggccgtgg
At4g35785	ser/arg rich	ggtgcattaaatatctcaaccag	ggagcttttcaagccaagatagtgg
AT1G13320	HK1	TAACGTGGCCAAAATGATGC	GTTCTCCACAACCGCTTGGT
AT4G26410	HK2	GAGCTGAAGTGGCTTCCATGAC	GGTCCGACATACCCATGATCC
AT4G20910	HEN1.1	CGTTGACAATATTTATGTGTACC	TCAGGTAGCTGCAAGTGGCAACG
AT4G20910	HEN1.2	TCTGTTGCTCTCTCCTGAGTTCTTCT	TCAGGTAGCTGCAAGTGGCAACG
AT1G48410	AGO1.1	GCCAGAGACATCAGACAGTGGC	GTGATGAAATATCCAAACACACG
AT1G48410	AGO1.2	GCCAGAGACATCAGACAGTGGC	TGCTGGTTAAGAGATGGAAGAG
2a3 Transgene	NIA2splicing	CTATCCTTCGCAAGACCCTTCCTC	GCACATACGTTCCATGTCTCTCC
159 Transgene	GUSsplicing	CGATGCGGTCACTCATTACGGC	GCCACCACCTGCCAGTCAACAG
AT1G37130	NIA2	GCTCTCAACGTCTGCTTCGAAGGG	GCACATACGTTCCATGTCTCTCC
159 transgene	GUS	CGATGCGGTCACTCATTACGGC	GCCCCAATCCAGTCCATTAA

Table 1 Genetic constructions and oligonucleotides

All genetic constructions done in this study and the oligonucleotides used for cloning are referenced. Oligos used for RT-PCR and RT-qPCR are referenced.

II.4- Résultats complémentaires

En plus d'étudier la localisation des gènes qui codent les SGS LSMa9 et PRP39, j'ai aussi étudié la localisation d'une nouvelle protéine SGS qui correspond à la protéine AAR2. La localisation de AAR2-GFP étant situé dans des particules cytoplasmiques, j'ai donc tenté d'identifier la nature de ces corps cytoplasmiques. Pour cela, j'ai co-localisés cette protéine avec les protéines SGS3, RBP47 et DCP1, des protéines qui marquent respectivement les siRNA bodies, les granules de stress et les P-bodies. Pour cela j'ai donc réalisé le clonage UBQ10-AAR2-GFP et nous avons utilisé l'expression transitoire pour réaliser ces co-localisations. Nos résultats suggèrent que AAR2 co-localise bien avec les stress granules lors d'un stress à la chaleur (Figure complémentaire 1B) et avec les silencing bodies (Figure complémentaire 1C) indiquant que ces deux bodies peuvent donc parfois co-localiser. De plus les spots de DCP1 (P-bodies) semblent mis à part quelques exceptions ne pas co-localiser avec AAR2 (Figure complémentaire 1A). Les marqueurs cytoplasmiques ont été fournis par Ana Moreno et Alexis Maizel.

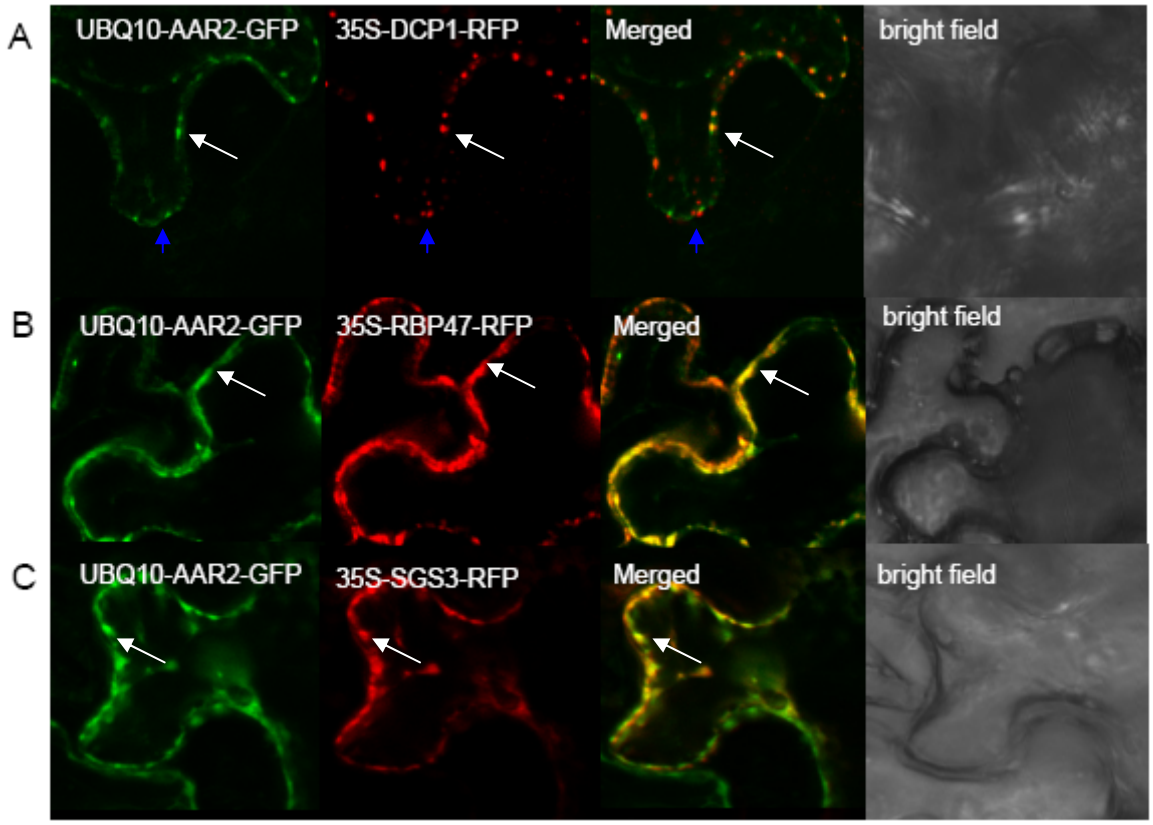


Figure complémentaire 1 Etude de la localisation de AAR2-GFP
A) Les protéines AAR2-GFP co-localisent avec DCP1 (P-bodies), B) avec RBP47 lors d'un stress à la chaleur 10 minutes à 42°C (stress granules) et C) mais très peut SGS3 (silencing bodies) (les flèches blanches indiquent les points co-localisés, les flèches bleues les points non co-localisés).

Enfin nous avons entrepris l'étude du phénotype de la racine en mesurant la racine principale des différents mutants SGS étudiés lors de cette thèse (lsm9a, aar2mutant ems= AAR2mut, aar2 salk, prp39sail, prp39salk) et en condition contrôle ou lors de la réponse à l'auxine (Figure complémentaire 2). Ces différents mutants semblent tous affectés lors de la croissance de la racine principale. En revanche le traitement à l'auxine ne semble pas affecter différentielement la croissance de ces racines en comparaison à celles du sauvage.

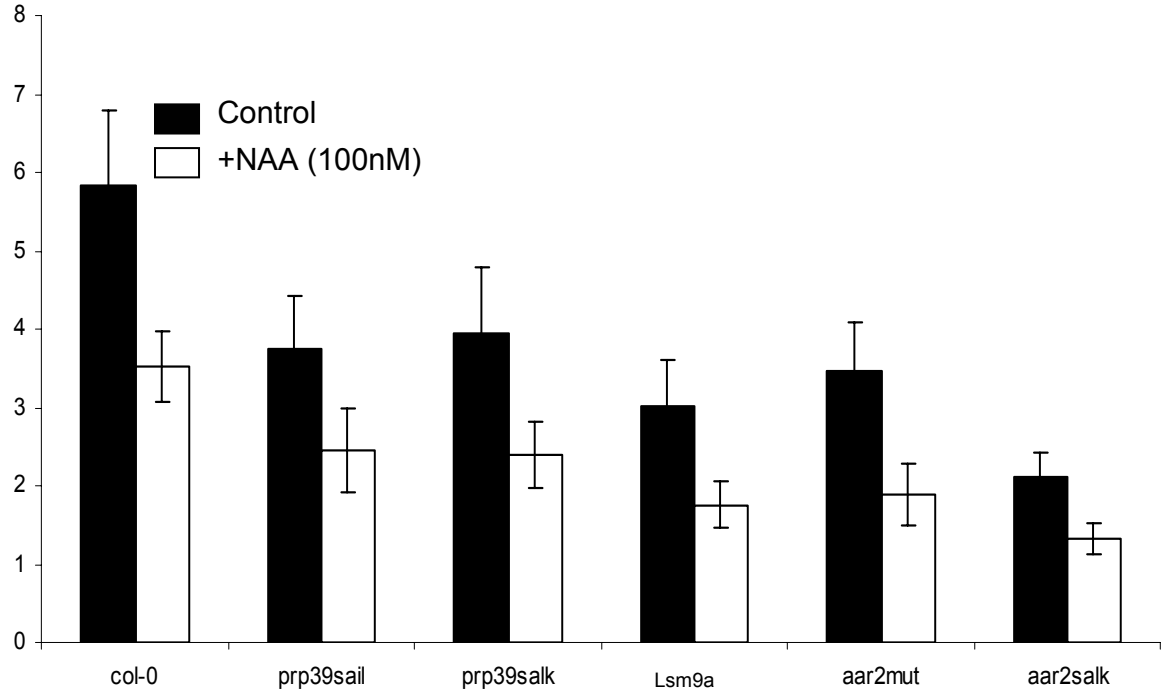


Figure complémentaire 2 Etude de la croissance de la racine principale
La longueur de la racine principale est représentée graphiquement dans les différents mutants indiqués traités ou non avec 100nM NAA. Graduation (en cm ; 12 jours après germination).

III- Discussion

III- Discussion

Lors de cette thèse, j'ai étudié le rôle des RBPs Nuclear Speckles RNA binding proteins *d'Arabidopsis thaliana* (AtNSRs) dans l'épissage alternatif et le silencing ainsi que leurs interactions avec les lncARNs. Tout d'abord, nous avons étudié la localisation des NSRs dans les speckles nucléaires (publication 1 Figures 1A, 1B, 2B) et montré que ces protéines contrôlent l'épissage alternatif chez Arabidopsis (publication 1 Figure 3). Elles peuvent se lier à des ARNm cible qui subissent l'AS mais aussi à des ARN non codants (publication 1 Figure 5). L'interaction avec le lncARN exogène *MtENOD40* provoquerait un changement de localisation des NSRs depuis les speckles nucléaires vers des particules cytoplasmiques (publication 1 Figure 1A). De plus, la surexpression de deux ARN non codants (lncARN351 et *MTENOD40* semble aussi pouvoir réguler le profil d'AS de certain des gènes contrôlés par les NSRs (publication 1 Figure 4). Cette connexion entre ncRNA et épissage via les NSRs nous a mené à étudier l'effet des NSRs sur le PTGS, un processus impliquant des ARN aberrants non-codants dérivés de trangènes. Nous avons découvert que les NSRs pouvaient réguler la diffusion de cellule à cellule du PTGS induit par un transgene RNAi qui contient un intron (publication 2 Figure 4). Nous avons aussi étudié le rôle des NSRs ainsi que de deux autres protéines issu d'un crible « suppressor of gene silencing » (SGS14 et SGS15) et qui sont deux protéines reliées à l'épissage. Nous avons montré que ces protéines peuvent elles aussi réguler le PTGS induit par des transgènes contenant un intron, induire des variations dans le profil d'épissage de certains gènes et réguler aussi le temps de floraison (publication 2 Figure 1). Ces différents résultats obtenus au niveau moléculaire provoquent différents effets sur la physionomie de la plante. En effet les NSRs peuvent moduler la formation de racines latérales induites par un traitement auxine (publication 1 Figure 2), elles modifient aussi le temps de floraison (publication 2 Figure 3).

1- Localisation des NSR et interaction avec les lncARN

Nous avions identifiés les protéines NSRs grâce à leur interaction avec l'ARN *ENOD40*. Elles seraient relocalisées du noyau vers le cytoplasme par l'action de ce lncARN (Campalans et al., 2004). Les études de localisation des protéines NSRs fusionnées à la GFP ont montré un pattern de localisation sub-cellulaire dans les speckles nucléaires. Ces granules inter-chromatiniens sont donc principalement constitués de complexes RNP et ne sont pas délimités par une membrane (Spector and Lamond 2011). L'association spécifique entre les différents partenaires ARN et/ou protéines permet le maintien de ce « compartiment » nucléaire. Les facteurs qui sont localisés dans les speckles nucléaires, tels

que des protéines SR ou encore les petits ARN nucléaires (snARN) U1, U2, U4, U5, font la navette entre le noyau et le cytoplasme dans les cellules HeLa. Cette translocation est importante pour leur maturation structurale et leur fonction (Patel et Bellini 2008 ; Sapra et al, 2009). Il est donc important de noter que le profil de localisation de ces speckles peut être modifié notamment lors de variations environnementales ainsi qu'en différents tissus ou stades développementaux. En effet, la quantité des composants protéiques et ARN présent dans une cellule à un moment donné peut influer sur la nature des speckles et donc sur leur nombre et leur activité. De même, chez les végétaux il y aurait pas un type de speckles mais plusieurs de ces granules. Ainsi les protéines de type SR ne co-localiseraient pas parfaitement voire pas de tout pour certaines suggérant la présence des speckles avec des compositions différentes (Lorković et al., 2008). Dans notre cas, on a observé que les protéines PRP39-GFP et NSRa-RFP ne co-localisent pas et même serait localisées dans des speckles différents (publication 2 Figure 6C). Ceci peut être lié au fait que ces deux protéines ont aussi des effets antagonistes sur l'AS de certaines cibles (publication 2 Figure 7A).

Les lncARN peuvent avoir différents modes d'action. De récentes études semblent indiquer que les mécanismes des lncARN sont classés en 4 archétypes (Wang et Chang 2011). Premièrement les lncARN peuvent agir comme des « molécules signales » quand leur expression reflète fidèlement l'action de facteurs de transcription contrôlant l'expression de gènes dans l'espace et le temps. Deuxièmement, les lncARN peuvent aussi agir comme des « leurres » en séquestrant des facteurs de transcription ou d'autres protéines régulatrices de la chromatine, en modifiant leur localisation sub-cellulaire ou en les empêchant d'agir. Autre exemple de lncARN leurre sont les lncARN qui se lient aux miARN (ARN mimicry ; Franco-Zorrilla et al., 2007) et ainsi dérégulent leur ARNm cibles. Troisièmement, les lncARN peuvent aussi servir de guides en recrutant des enzymes de modification de la chromatine sur des gènes cibles, que ce soit en cis (à proximité du site de production du lncARN) ou en -trans sur d'autres loci. Le lncARN-RNP peut ainsi agir sur la chromatine en ayant une incidence sur les modifications des histones ou de la méthylation de l'ADN. Enfin, ils peuvent servir comme échafaudages en rassemblant de multiples protéines pour former des complexes RNPs ayant un rôle structural (stabilisation des structures nucléaires ou des complexes de signalisation). Plusieurs études récentes convergent pour indiquer que les lncARN sont capables de structurer un compartiment de type RNP (Shevtsov and Dundr 2010; Caudron-Herger and Rippe 2012). Tel est le cas du lncARN MALATE1, qui a été montré comme un acteur majeur de l'intégrité spatiale des speckles dans le noyau (Triphati et al., 2010 ; Miyagawa et al., 2012). Chez les humains, les speckles ont comme composés de base les protéines SRm300 (une protéine co-activatrice de l'épissage de type Ser/Arg) et SRSF2 (SC35 : Un facteur d'épissage) (Hall et al. 2006 ;

Hutchinson et al. 2007). La répression par ARN interférence de la protéine SRm300 perturbe la localisation de type speckles de SRSF2 (SC35), ainsi que celle du lncRNAMALAT-1 (Miyagawa et al., 2012). Ces résultats suggèrent que la structure et donc l'activité du spliceosome pourrait donc être étroitement liée avec les partenaires présents dans ces corps, qu'ils soient ribonucléiques ou protéiques. Le lncARN MALAT-1 pourrait ainsi moduler l'épissage alternatif d'un pre-ARNm en régulant l'état de phosphorylation de protéines SR du spliceosome. Nos résultats soutiennent que deux partenaires des NSRs, ncARN351 et *ENOD40* semblent capables de modifier le profil d'épissage alternatif de gène cibles (publication 1 Figure 4). Dans notre cas, cette interaction avec les lncRNAs modifierait la localisation des NSR, notamment pour l'ARN lncENOD40, et ainsi perturberait l'AS, à la différence du MALAT-1 qui module l'état de phosphorylation des protéines SR liés à la machinerie de base de l'épissage. Ces travaux mettent donc en lumière, l'importance que peuvent porter les ncARN grâce à leur conformation secondaire hautement structurée quand ils interagissent avec différents ribonucléoprotéines.

Mis à part sa localisation dans des speckles contenant des enzymes de l'épissage comme SRP34, les NSRs ont été montrées comme parfois légèrement co-localisée avec UPF3 (une protéine clé du NMD) ainsi qu'avec les protéines en relation avec le silencing DRB4 (DOUBLE-STRANDED ARN BINDING) (publication 1 Figure S3B). Or, des études récentes montrent que la transcription, les mécanismes de maturation des ARN ainsi que la régulation des ARN ne se déroulent pas de façon séquentiel mais plutôt de manière simultanée. Les ARNs en cours de transcription seraient donc épissés durant leur transcription. Ainsi la co-localisation entre UPF3 et NSR pourrait indiquer que les produits de rétention d'intron observé dans le double mutant nsra/nsrb pourraient être des cibles potentiel pour le NMD, entraînant donc la dégradation de certains de ces transcrits et ajoutant un niveau supplémentaire dans régulation des ARN. De plus, la co-localisation partielle entre les NSRs et la protéine DRB4 (une protéine qui lie les ARN double brin), semble indiquer que les NSRs pourraient donc interagir avec des protéines clés pour la mise en place des mécanismes de silencing. Ceci peut être lié au fait que les doubles mutant nsra/nsrb modifient le profil de silencing des lignées reportrices du silencing comme JAP3 (publication 2 Figure 4).

2- Les AtNSRs, des régulateurs nucléaires de l'épissage alternatif

Nos résultats montrent que les NSRs jouerait un rôle redondant comme régulateurs de l'épissage alternatif. En effet, les simples mutants nsr ne présentent pas de phénotype majeur en comparaison avec le double mutant nsra/nsrb qui présente un phénotype altéré

de réponse à l'auxine (publication 1 Figure 2) ainsi qu'un temps de floraison accéléré (publication 2 Figure 3). En outre, on peut remarquer que les deux simples mutants nsra ou nsrb présentent un léger phénotype de transition florale qui s'additionne dans le double mutant. Ces résultats, en accord avec les observations microscopiques de co-localisation entre les NSRs semblent indiquer un rôle redondant pour ces protéines. De plus, les lignées qui sur-expriment les NSRs présentent toutes un phénotype nain ainsi que des nécroses sur les feuilles (publication 1 Figure S6). Finalement, un des arguments qui soutient encore un rôle d'action commun pour ces deux protéines NSRs est lié au fait qu'elles ont une affinité pour les mêmes cibles D'AS et pour le même ncARN351 et toutes les deux sont relocalisées dans le cytoplasme par l'ARN *ENOD40* (publication 1 Figure 1A). Néanmoins, les variations observées au niveau moléculaire dans le fond double mutant nsra/nsrb ne seraient pas toutes imputables aux deux protéines NSRs. L'expérience de RIP (publication 1 Figure 5) nous indique une capacité d'association parfois différente entre NSRa et NSRb ce qui pourrait entraîner des effets divergents entre un fond mutant nsra et nsrb.

Les NSRs jouerait un rôle dans l'AS tout comme certaines protéines liés au PTGS tel que SGS14 ou SGS15. Une des cibles communes aux NSRs et à SGS14 et SGS15 est l'ARNm codant pour le facteur de transcription CIRCADIAN CLOCK-ASSOCIATED1 (*CCA1*; At2g48830) qui est un des composants principaux de l'horloge circadienne chez Arabidopsis. Ce gène est aussi associé avec l'acclimatation au froid et la régulation du temps de floraison et il a été précédemment décrit comme étant alternativement épissé. De façon intéressante, Seo et al., 2012 ont démontré l'autorégulation de l'activité de CCA1 par une variante de son épissage, CCA1β reliant l'horloge circadienne à l'acclimatation au froid. L'épissage alternatif de CCA1 produit dans leurs études 2 transcrits, un de pleine longueur nommé CCA1α et la version tronquée CCA1β. L'isoforme CCA1β possède le domaine d'interaction protéine-protéine qui permet la formation de dimère mais ne possède pas le domaine MYB de liaison à l'ADN à l'inverse de l'isoforme CCA1α. CCA1β inhibe par compétition l'activité de CCA1α en formant des hétéro-dimères CCA1α-CCA1β et LHY-CCA1β qui ont affinité de liaison à l'ADN diminuée (Seo et al., 2012). Ainsi les variations d'AS du gène CCA1 pourraient être une explication du phénotype de floraison précoce des NSRs (publication 2 Figure 3A, 7A).

Une autre cible des NSRs est l'ARNm codant une protéine Fbox et qui subit l'épissage alternatif (publication 2 Figure 7B). Cet événement d'AS se déroule au niveau de la région 3'UTR de cet ARNm, L'AS étant un phénomène souvent détecté dans les régions non transcrites 5' ou 3' (UTRs) des ARNm. Un exemple démontrant l'importance d'un tel événement est l'oncogene Gli1 chez la souris, qui est l'effecteur terminal de la voie de signalisation Hedgehog (HH). Ce gène subit l'AS au niveau de la partie 5' UTR de son ARNm et ceci a un impact sur sa propre capacité de traduction. La variante de l'ARNm la plus courte ainsi que la version ne possédant pas l'exon 1B sont favorisées (Palaniswamy et

al., 2010). Chez Arabidopsis, l'AS d'introns qui se situent dans des parties 5' ou 3' UTR peut modifier la sensibilité des transcrits d'ARNm pour le NMD (Kalyna et al., 2012).

3- Les AtNSRs et le silencing

En général, les transgènes ne contiennent pas d'introns, une caractéristique partagée avec les transposons, l'extinction de l'expression de ces derniers étant le premier objectif du silencing. Étant donné que les introns sont très fréquents dans les gènes endogènes, et sont souvent absents dans les transgènes et les transposons, une hypothèse soutient que les introns peuvent supprimer le silencing des gènes. Une étude récente sur le génome entier a permis de calculer les densités des petits d'ARN dans les exons de gènes sans introns en comparaison avec des gènes qui contiennent des introns chez *Arabidopsis thaliana* (Christie et al., 2011). La densité de petits ARN dans les exons de gènes sans intron était comparable aux exons d'éléments transposables, cette densité étant bien inférieure dans des gènes qui possède un intron. Pour étudier l'effet des introns sur le silencing *in vivo*, un système rapporteur transgénique utilisant la green fluorescent protéine (GFP) (Brosnan et al., 2007) a été analysé. Lors de l'utilisation d'une construction sans intron, 85% des Plantes T1 montrent une GFP « silencée ». Ce silencing est dépendant de RDR6, ce qui démontre un mécanisme PTGS. Ces résultats ont été comparés avec ceux obtenus avec la même construction GFP mais qui contient cette fois-ci un intron. Les lignées contenant un intron ont réussi à réduire de quatre fois le « silencing » des gènes des lignées transgéniques d'Arabidopsis. Cette suppression du silencing de l'ARN médiée par les introns sur l'expression du transgène GFP dépend de l'efficacité de l'épissage de l'intron introduit, et de ABH1, l'orthologue humain de la cap-binding protéine 80 (CBP80) d'Arabidopsis. En effet, la suppression du « silencing » médiée par les introns est perdue dans le fond génétique mutant abh1 (Christie et al., 2011). En outre, il a été montré que la biogenèse des siARN secondaires RDR6-dépendant est atténuée le long des transgènes épissés par rapport aux transcrits issus d'un transgène sans intron. L'épissage des introns fournirait donc un mécanisme permettant d'éviter aux gènes endogènes le silencing RDR6-dépendant chez Arabidopsis.

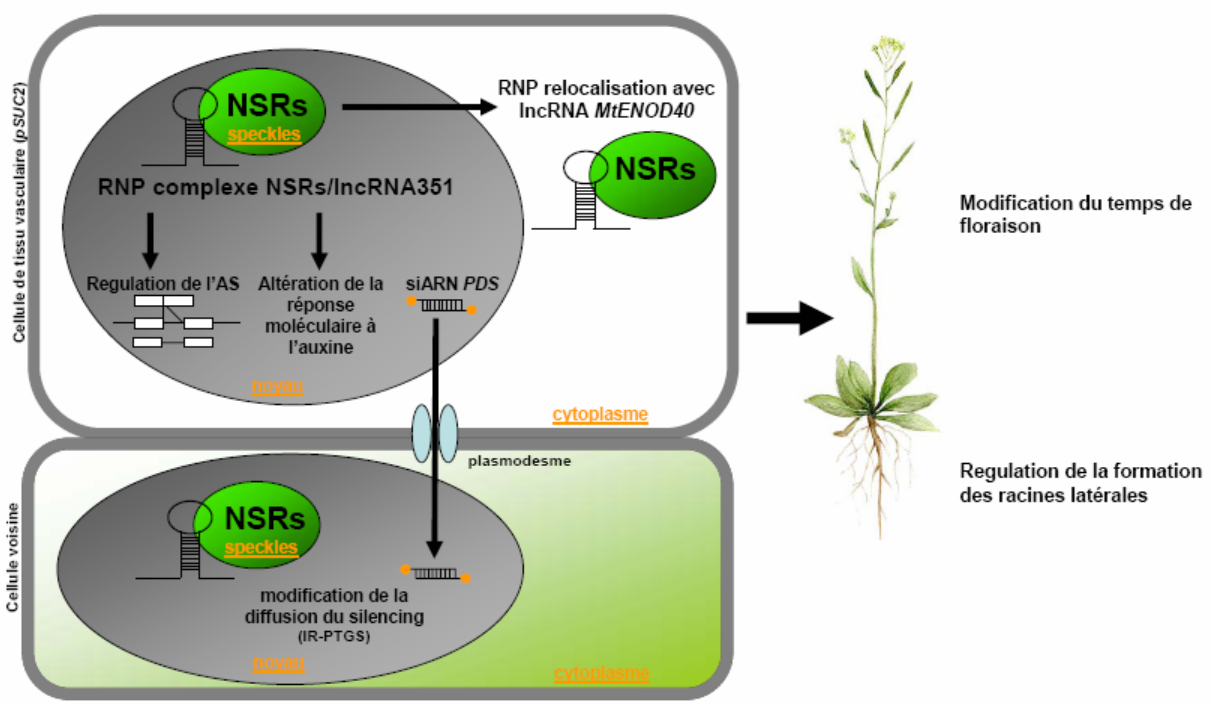
D'autres résultats semblent indiquer aussi un lien entre le PTGS et l'épissage alternatif (Laubinger et al., 2008). La biogenèse des microARN d'*Arabidopsis thaliana* (miARN) à partir d'un transcrit primaires (pri-miARN) nécessite l'activité de plusieurs protéines, y compris DICER-like1 (DCL1), la protéine de liaison au dsARN HYPONASTIC leaves1 (HYL1), et une protéine à doigt de zinc nommée SERRATE (SE). Tant SE comme le nuclear cap-binding complex (CBP20/CBP80) sont nécessaires pour la biogenèse/production des pri-miARN. L'inactivation de l'une des protéines ABH1/CBP80 ou

CBP20 entraîne une baisse des niveaux de miARN matures accompagnés d'une stabilisation apparente du pri-miARN. Un « tilling array » sur l'ensemble du génome révèlent que les mutants se, abh1/cbp80 et cbp20 présentent des défauts d'épissage similaires, conduisant à l'accumulation d'un grand nombre de transcrits partiellement épissés. Finalement, de récents travaux montrent que ABH1 et SE sont nécessaires pour la suppression du silencing des ARN mis en place par les introns (Christie et carroll, 2011). Ainsi, on peut spéculer que les NSRs ainsi que les protéines SGS14 et SGS15 sont capables de moduler l'AS et le PTGS des siARN de la même façon que ABH1 ou SE régulent l'AS et le PTGS. Dans notre travail, nous n'avons pas détecté des changements majeurs dans l'épissage des transgènes qui entraînent le déclenchement du PTGS et donc le rôle des AtNSRs dans le PTGS et sa propagation reste à comprendre.

Le lien entre PTGS et l'épissage alternatif peut être aussi reliée aux interactions entre épissage et l'état de la chromatine. En effet, plusieurs gènes impliqués dans la propagation du silencing codent pour des gènes liés à des voies de régulation en relation avec la chromatine telles que la voie de régulation par DCL3/RDR2/AGO4 (Dunoyer et Voinnet, 2009 ; Dunoyer et al, 2007). D'autres régulateurs de la chromatine ont été identifiés comme jouant un rôle dans le silencing génique en modifiant le statut de la chromatine du transgène (Le Masson et al., 2012).

Dans les cellules humaines, un lien entre l'épissage alternatif et l'épigénétique semble émerger (Luco et al., 2010). En effet, un rôle direct des modifications d'histone dans la régulation de l'AS a été identifié. Plusieurs signatures de modifications d'histones sont en corrélation avec le profil d'épissage dans un ensemble de gènes humains et cette modulation des modifications d'histone causerait des variations dans les sites d'épissage. Les marques d'histone affecteraient le profil d'épissage de certains gènes en recrutant des régulateurs de l'épissage via une protéine de liaison à la chromatine. On peut donc imaginer que les NSR ou SGS14 en modifiant l'AS pourrait introduire des modifications de l'état épigénétique du transgène et ainsi contribuer à la propagation du silencing. Nous pensons que SGS14 et les NSRs recrutent des protéines de liaison à la chromatine et d'autres régulateurs épigénétiques sur l'ARNm des transgènes contenant un intron. En absence de SGS14 ou des NSRs, la localisation ou le marquage épigénétique du locus (159, L1 ou JAP3) pourrait être modifié. Dans ce cas, les ARN produits par les locus 2a3 et 159 ou encore JAP3 seraient détournés vers le NMD au lieu d'entrer dans la voie du PTGS, expliquant la diminution de PTGS dans les lignées reportrices du PTGS contenant un intron. Une hypothèse serait que ces gènes permettent le recrutement de protéines lié à la chromatine sur le site d'insertion du transgène ce qui changerait la nature des ARN produit par ce locus mais pas le profil d'épissage ce qui permettrait d'augmenter la production d'ARN aberrant de

ce locus permettant de déclancher le PTGS. Ainsi la mutation des NSRs et SGS entraînerait bien une diminution de l'efficacité du PTGS.



Modèle représentant le rôle des protéines NSRs chez *Arabidopsis thaliana.*

Modèle résumant les résultats obtenus durant cette thèse. Les NSRs sont donc localisées dans des speckles, elles peuvent moduler la réponse à l'auxine ainsi que l'AS. Les NSRs peuvent être relocalisées en présence du lncARn *ENOD40*. Les NSRs interagissent aussi avec le ncARN351 (endogène) qui est lui aussi capable de réguler l'AS. Enfin les NSRs peuvent moduler le PTGS dans la lignée JAP3. Ces différents effets à l'échelle moléculaire entraînent des variations morphologiques notamment sur le développement floral et la formation des racines latérales induite par un traitement auxine.

Cette hypothèse est appuyée par les résultats obtenu en croisant 159 (100% PTGS) et sgs14 (159/sgs14 31% PTGS) avec upf3 (mutant UP-FRAMESHIFT 3) (159/sgs14/upf3 100% PTGS) (publication 2 Figure 2D). Ces résultats indiquent que lorsqu'on bloque upf3 et donc une partie du NMD on augmente la formation d'ARN aberrant et ainsi on permet de restaurer le PTGS de la lignée 159/sgs14. Ce résultat semble donc bien indiqué que les NSRs pourraient permettrent la formation d'ARN qui entraînerait le déclanchement du PTGS

Nos résultats nous permettent donc de d'établir le modèle ci-dessus. Les protéines NSRs peuvent interagirent avec des lncARN dont le ncARN exogène MtENOD40 ce qui entraîne un changement de localisation dans les tissus vasculaires (là où sont naturellement exprimé les NSRs chez Arabidopsis et ENOD40 chez Medicago). Les NSRs interagissent aussi avec le lncARN351 endogène de Arabidopsis. Les NSRs ont été montrées comme des régulateurs de l'épissage alternatif au même titre que le lncARN351 et ENOD40. Enfin, de façon similaire aux protéines SGS14 et SGS15 qui régulent elles aussi l'AS, les NSRs sont capables de permettre la diffusion du signal PTGS induit par un transgène contenant un intron jusqu'aux cellules voisines. La cellule qui initie le PTGS (où s'exprime le promoteur SUC2 dans la lignée JAP3) est représenté blanche pour symbolisé la chlorose et donc le PTGS. La cellule dite voisine dans le schéma est représentée en dégradé de blanc et vert afin de symboliser une cellule en cour de chlorose à cause de la diffusion du PTGS du transgène JAP3 de cellule à cellule grâce aux NSRs (la construction JAP3 étant une construction RNAi ciblant le gène PDS qui est impliqué dans la synthèse de chlorophylle ; Smith et al., 2007). Ainsi les résultats obtenus dans cette thèse permettent de supposer que les NSRs fonctionnent à l'interphase entre la régulation de L'AS et la propagation du silencing.

IV- Perspectives

Cette thèse ouvre plusieurs perspectives :

Afin d'étudier si, comme chez les cellules humaines un lncARN pourrait structurer les speckles, Il serait intéressant d'étudier la localisation de protéines marqueurs de speckles (NSRs, protéines SR) en microscopie en utilisant les constructions ARNi ou qui surexpriment le ncARN351 ce qui nous donnerai plus d'indications sur un hypothétique rôle du ncRNA351 dans le maintient de la structure des speckles. De même, il faudrait étudié la capacité de relocalisation des NSRs en présence du npcARN351.

Un deuxième point serait de comprendre la fonction du lncRNA351 en utilisant des approches de génétique reverse (RNAi ou lignées surexprimmant) afin d'étudier l'impact de ces modifications au niveau de l'AS « genome-wide » et de la morphologie, notamment de la racine latérale L'introduction de la construction surexprimmant le lnc351 dans le mutant Atnsra/Atnsrb permettra de confirmer l'action de cette ARN sur les NSRs.

A plus long terme, Il serait intéressant de regarder les cibles ARN des NSRs d'une manière globale en réalisant des expériences d'immunoprécipitation d'ARN comme celles décrites dans la thèse mais cette fois-ci en faisant du séquençage globale. Ces résultats nous permettrons de voir l'ensemble des gènes impliqués dans les mécanismes de régulation des NSRs. L'utilisation des racines latérales sera un plus pour décortiquer finement l'action des NSRs.

Enfin, afin de valider notre hypothèse selon laquelle les ARN produits par les locus JAP3 seraient détournés vers le NMD par les NSRs au lieu d'entrer dans la voie du PTGS, nous devrions introgresser le mutant upf3 dans la lignée nsra/nsrb/JAP3. Ceci nous permettra de préciser la nature des RNA aberrants accumulés (transcrits partiellement épissés) et voir si on peut restaurer le phénotype de PTGS déclenché par JAP3. Enfin, les différents mutants identifiés comme reliés à l'épissage et le PTGS ouvrent des larges perspectives pour analyser leurs interactions au niveau génétique et subcellulaire.

Enfin il serait intéressant d'analyser l'état chromatinien du transgène et son épissage en détail dans tous les mutants SGS et NSRs.

V- Matériels et méthodes

V- Matériels et méthodes

La majeure partie des matériels et méthodes des expériences qui m'ont permis d'obtenir les résultats de cette thèse sont expliqués dans les parties matériels et méthodes des articles présentés dans la partie résultats. Dans cette partie je vais donc uniquement apporter des précisions sur l'expérience d' immunoprécipitation (RIP) des ARN, je proposerais donc un protocole de RIP détaillé.

RIP for nuclear or cytoplasmic proteins

Nuclei extract preparation

1. Prepare between 2g and 5g of fresh plant material.
2. Crosslink by UV exposure (3 times 400 mJ/cm^2).
3. Freeze in N_2liq.
4. Grind to fine dust.
5. Resuspend the dust in **Lysis Buffer I** (5ml/g).
6. Filter the suspension through a 200 µm mesh.
7. Centrifuge 1500g 10min 4°C.
8. Wash in **NRBT Buffer** (5ml/g).
9. Centrifuge 1500g 10min 4°C and discard the supernatant.
10. Repeat steps 8 and 9.
11. Resuspend the pellet in **Nuclei Lysis Buffer I** (200µl/g). Transfer to a 1.5ml tube.
12. Add 10 µl of RNAse-FREE DNAse and incubate at 37°C, 10min.
13. Add SDS to a final concentration of 0.1% (2.5µl SDS 20%/500µl). Invert several times to mix.
14. Incubate 1h at 4°C.
15. Centrifuge 13000 rpm 30 min 4°C.
16. Transfer the supernatant to a new tube and discard the pellet.
17. Take 10% of the volume **(INPUT SAMPLE)** and conserve it at -20°C.

Total lysate instead of nuclei extract

Please, change from step 5 on:

5bis. Resuspend the dust in **Lysis Buffer II** (5ml/g).
6bis. Filter the suspension through a 200 µm mesh.
7bis. Incubate in the ice for 10min.
8bis. Centrifuge 1500g 10min 4°C.
9bis. Resuspend the pellet in **Nuclei Lysis Buffer II** (200µl/g). Transfer to a 1.5ml tube.
10bis. Incubate 1h at 4°C.
11bis. Centrifuge 13000 rpm 10 min 4°C.
12bis. Transfer the supernatant to a new tube and discard the pellet.
13bis. Take 10% of the volume **(INPUT SAMPLE)** and conserve it at -20°C.
Continue in step 18.

Immunoprecipitation

18. Follow the instructions of Miltenyi µMACS Epitope Tag (e.g. HA) Kit, as follows (points 2.4 and 3.2):
19. Add 50 µl of anti-tag microbeads to the nuclei lysate.
20. Mix well at 4°C during 30min.
21. Place a µ Column in the magnetic field of the µMACS Separator.

22. Prepare the µ Column by applying 200 µl of Miltenyi Lysis Buffer.
23. Once the incubation has finished, apply the lysate onto the Column and let it run through.
24. Rinse the Column 4 times with 200 µl of Wash Buffer 1.
25. Rinse the Column 3 times with 100 µl of Wash Buffer 2.
26. **Extra**: Rinse the Column 3 times with 200 µl of RNAse-FREE water.
27. Take out the Column from the magnetic field.
28. Apply 50 µl of water and collect the brown eluate.
29. Add 150 µl of **Proteinase K Buffer** and collect the rest of the eluate in the same tube.

Reverse Crosslink and RNA purification

30. Add 40 µl of 20mg/ml Proteinase K.
31. Incubate 30min at 55°C.
32. Add SDS at a final concentration of 1%.
33. Incubate 30min at 55°C.
34. Place a new µ Column in the magnetic field and calibrate it with 200 µl of the adequate Buffer.
35. Wash the Column with 500 µl of water.
36. Apply the Proteinase K treated sample and collet the **flow-through**, eliminating the magnetic beads. Repeat the procedure until obtaining a clear sample.
37. Use TRIZOL followed by isopropanol precipitation for RNA purification, both of the RIP sample and the INPUT sample.
38. After qPCR, the results are expressed as a % of the INPUT, considering the relationship between the volumes.

Plant RIP-UV solutions

Lysis Buffer I	**For 100ml**	**For 200ml**
20mM Tris-HCl pH7,4	2ml of 1M	4ml of 1M
25% Glycerol	25ml	50ml
20mM KCl	2ml of 1M	4ml of 1M
2mM EDTA	400µl of 0,5M	800µl of 0,5M
2,5mM $MgCl_2$	250µl of 1M	500µl of 1M
250mM Sucrose	12,5ml of 2M	25ml of 2M
Protease Inhibitor	2ml	4ml
H_20 to volume		

Lysis Buffer II	**For 100ml**	**For 200ml**
50mM Tris-HCl pH7,4	5ml of 1M	10ml
10% Glycerol	10ml	20ml
5mM $MgCl_2$	500µl of 1M	1ml
150mM NaCl	3ml of 5M	6ml
0.1% Triton	0,5ml of 20%	1ml
Protease Inhibitor	2ml	4ml
H_20 to volume		

NRBT	**For 100ml**	**For 600ml**
20mM Tris-HCl pH7,4	2ml of 1M	12ml of 1M
25% Glycerol	25ml	150ml
2,5mM $MgCl_2$	250µl of 1M	1,5ml of 1M
Protease Inhibitor	2ml	12ml
0,5% Triton	2,5ml of 20%	
H_20 to volume		

Nuclei Lysis Buffer I	**For 5ml**	**For 10ml**
50mM Tris-HCl pH7,4	250µl of 1M	500µl
1% Triton	250µl of 20%	0.5ml
100mM NaCl	100µl of 5M	200µl
1mM $MgCl_2$	5µl of 1M	10µl
0,1mM $CaCl_2$	5µl of 100mM	10µl
Protease Inhibitor	200µl	400µl
RNAase-in (20U/ml)	12,5µl	25µl
H_20 to volume		

Nuclei Lysis Buffer II	**For 5ml**	**For 10ml**
50mM Tris-HCl pH7,4	250µl of 1M	500µl
0,1% SDS	25µl 20%	50µl 20%
10mM EDTA	100µl 0.5M	200µl
RNAase-in (20U/ml)	12,5µl	25µl
H_20 to volume		

PK Buffer	**For 1ml**
100mM Tris-HCl pH7,4	100µl of 1M
50mM NaCl	10µl of 5M
10mM EDTA	20µl of 0,5M
Proteinase K 4mg/ml	200ul of 20mg/ml
RNAase-in (20U/ml)	5µl
H_20 to volume	

VI- Bibliographie

VI- Bibliographie

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouché, N., Gasciolli, V., and Vaucheret, H. (2006a). DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. Current biology: CB **16**: 927–32.

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouché, N., Gasciolli, V., and Vaucheret, H. (2006b). DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. Current biology: CB **16**: 927–32.

Aguilera, A. (2005a). Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. Current opinion in cell biology **17**: 242–50.

Aguilera, A. (2005b). mRNA processing and genomic instability. Nature structural & molecular biology **12**: 737–8.

Ali, G.S., Golovkin, M., and Reddy, A.S.N. (2003). Nuclear localization and in vivo dynamics of a plant-specific serine/arginine-rich protein. The Plant journal: for cell and molecular biology **36**: 883–93.

Allen, E., Xie, Z., Gustafson, A.M., and Carrington, J.C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell **121**: 207–21.

Ambrosone, A., Costa, A., Leone, A., and Grillo, S. (2012). Beyond transcription: RNA-binding proteins as emerging regulators of plant response to environmental constraints. Plant science: an international journal of experimental plant biology **182**: 12–8.

Ben Amor, B., Wirth, S., Merchan, F., Laporte, P., d'Aubenton-Carafa, Y., Hirsch, J., Maizel, A., Mallory, A., Lucas, A., Deragon, J.M., Vaucheret, H., Thermes, C., and Crespi, M. (2009). Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. Genome research **19**: 57–69.

Amrani, N., Sachs, M.S., and Jacobson, A. (2006). Early nonsense: mRNA decay solves a translational problem. Nature reviews. Molecular cell biology **7**: 415–25.

Anderson, P. and Kedersha, N. (2006). RNA granules. The Journal of cell biology **172**: 803–8.

Andrei, M.A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R., and Lührmann, R. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. RNA (New York, N.Y.) **11**: 717–27.

Ariel, F., Diet, A., Verdenaud, M., Gruber, V., Frugier, F., Chan, R., and Crespi, M. (2010). Environmental regulation of lateral root emergence in Medicago truncatula requires the HD-Zip I transcription factor HB1. The Plant cell **22**: 2171–83.

Ariel, F., Moreno, A.B., Bardou, F., and Crespi, M. (2012). non-protein-coding RNAs and root développement. **2** 29-48

Arrigo, A.P. and Welch, W.J. (1987). Characterization and purification of the small 28,000-dalton mammalian heat shock protein. The Journal of biological chemistry **262**: 15359–69.

Barbazuk, W.B., Fu, Y., and McGinnis, K.M. (2008). Genome-wide analyses of alternative splicing in plants: opportunities and challenges. Genome research **18**: 1381–92.

Barta, A., Kalyna, M., and Lorković, Z.J. (2008). Plant SR proteins and their functions. Current topics in microbiology and immunology **326**: 83–102.

Bezerra, I.C., Michaels, S.D., Schomburg, F.M., and Amasino, R.M. (2004). Lesions in the mRNA cap-binding gene ABA HYPERSENSITIVE 1 suppress FRIGIDA-mediated delayed flowering in Arabidopsis. The Plant journal: for cell and molecular biology **40**: 112–9.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annual review of biochemistry **72**: 291–336.

Blencowe, B.J. (2006). Alternative splicing: new insights from global analyses. Cell **126**: 37–47.

Blilou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Friml, J., Heidstra, R., Aida, M., Palme, K., and Scheres, B. (2005). The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature **433**: 39–44.

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.-K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. Cell **123**: 1279–91.

Boutet, S., Vazquez, F., Liu, J., Béclin, C., Fagard, M., Gratias, A., Morel, J.B., Crété, P., Chen, X., and Vaucheret, H. (2003). Arabidopsis HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. Current biology **13**: 843–8.

Brosnan, C.A., Mitter, N., Christie, M., Smith, N.A., Waterhouse, P.M., and Carroll, B.J. (2007). Nuclear gene silencing directs reception of long-distance mRNA silencing in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **104**: 14741–6.

Brown, J.M., Leach, J., Reittie, J.E., Atzberger, A., Lee-Prudhoe, J., Wood, W.G., Higgs, D.R., Iborra, F.J., and Buckle, V.J. (2006). Coregulated human globin genes are frequently in spatial proximity when active. The Journal of cell biology **172**: 177–87.

Brown, J.W.S., Echeverria, M., and Qu, L.H. (2003). Plant snoRNAs: functional evolution and new modes of gene expression. Trends in plant science **8**: 42–9.

Bäurle, I. and Dean, C. (2008). Differential interactions of the autonomous pathway RRM proteins and chromatin regulators in the silencing of Arabidopsis targets. PloS one **3**: e2733.

Bäurle, I., Smith, L., Baulcombe, D.C., and Dean, C. (2007). Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. Science (New York, N.Y.) **318**: 109–12.

Béclin, C., Boutet, S., Waterhouse, P., and Vaucheret, H. (2002). A branched pathway for transgene-induced RNA silencing in plants. Current biology: CB **12**: 684–8.

Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. Genes & Development **1**: 1183–1200.

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

Cao, S., Jiang, L., Song, S., Jing, R., and Xu, G. (2006). AtGRP7 is involved in the regulation of abscisic acid and stress responses in Arabidopsis. Cellular & molecular biology letters **11**: 526–35.

Carmo-Fonseca, M., Mendes-Soares, L., and Campos, I. (2000). To be or not to be in the nucleolus. Nature cell biology **2**: E107–12.

Casimiro, I., Marchant, a, Bhalerao, R.P., Beeckman, T., Dhooge, S., Swarup, R., Graham, N., Inzé, D., Sandberg, G., Casero, P.J., and Bennett, M. (2001). Auxin transport promotes Arabidopsis lateral root initiation. The Plant cell **13**: 843–52.

Casson, S.A., Chilley, P.M., Topping, J.F., Evans, I.M., Souter, M.A., and Lindsey, K. (2002). The POLARIS gene of Arabidopsis encodes a predicted peptide required for correct root growth and leaf vascular patterning. The Plant Cell Online **14**: 1705.

Caudron-Herger, M., Müller-Ott, K., Mallm, J.-P., Marth, C., Schmidt, U., Fejes-Tóth, K., and Rippe, K. (2011). Coding RNAs with a non-coding function: maintenance of open chromatin structure. Nucleus (Austin, Tex.) **2**: 410–24.

Caudron-Herger, M. and Rippe, K. (2012). Nuclear architecture by RNA. Current opinion in genetics & development **22**: 179–87.

Chapman, E.J. and Estelle, M. (2009). Mechanism of auxin-regulated gene expression in plants. Annual review of genetics **43**: 265–85.

Charon, C., Moreno, A.B., Bardou, F., and Crespi, M. (2010a). Non-protein-coding RNAs and their interacting RNA-binding proteins in the plant cell nucleus. Molecular plant **3**: 729–39.

Charon, C., Moreno, A.B., Bardou, F., and Crespi, M. (2010b). Non-protein-coding RNAs and their interacting RNA-binding proteins in the plant cell nucleus. Molecular plant **3**: 729–39.

Charon, C., Sousa, C., Crespi, M., and Kondorosi, a (1999). Alteration of enod40 expression modifies medicago truncatula root nodule development induced by sinorhizobium meliloti. The Plant cell **11**: 1953–66.

Chekanova, J. a, Gregory, B.D., Reverdatto, S. V, Chen, H., Kumar, R., Hooker, T., Yazaki, J., Li, P., Skiba, N., Peng, Q., Alonso, J., Brukhin, V., Grossniklaus, U., Ecker, J.R., and Belostotsky, D. a (2007). Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. Cell **131**: 1340–53.

Chiba, Y. and Green, P.J. (2009). mRNA Degradation Machinery in Plants. Journal of Plant Biology **52**: 114–124.

Chilley, P.M., Casson, S. a, Tarkowski, P., Hawkins, N., Wang, K.L.-C., Hussey, P.J., Beale, M., Ecker, J.R., Sandberg, G.K., and Lindsey, K. (2006). The POLARIS peptide of Arabidopsis regulates auxin transport and root growth via effects on ethylene signaling. The Plant cell **18**: 3058–72.

Christie, M., Croft, L.J., and Carroll, B.J. (2011). Intron splicing suppresses RNA silencing in Arabidopsis. The Plant journal: for cell and molecular biology **68**: 159–67.

Cioce, M. and Lamond, A.I. (2005). Cajal bodies: a long history of discovery. Annual review of cell and developmental biology **21**: 105–31.

Collier, N.C. and Schlesinger, M.J. (1986). The dynamic state of heat shock proteins in chicken embryo fibroblasts. The Journal of cell biology **103**: 1495–507.

Collier, S., Pendle, A., Boudonck, K., Van Rij, T., Dolan, L., and Shaw, P. (2006). A distant coilin homologue is required for the formation of cajal bodies in Arabidopsis. Molecular biology of the cell **17**: 2942–51.

Conti, E. and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Current opinion in cell biology **17**: 316–25.

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

Crespi, M. and Frugier, F. (2008). De novo organ formation from differentiated cells: root nodule organogenesis. Science signaling **1**: re11.

Crespil, M.D., Jurkevitchl, E., Poiret, M., Aubenton-carafa, Y., Petrovics, G., Kondorosi, E., and Kondorosi, A. (1994). expressed during nodule codes for a non-translatable RNA involved in plant growth. EMBO Journal **13**: 5099–5112.

Decker, C.J. and Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harbor perspectives in biology **4**: a012286.

Docquier, S., Tillemans, V., Deltour, R., and Motte, P. (2004). Nuclear bodies and compartmentalization of pre-mRNA splicing factors in higher plants. Chromosoma **112**: 255–66.

Dolan, L., Janmaat, K., Willemsen, V., Linstead, P., Poethig, S., Roberts, K., and Scheres, B. (1993). Cellular organisation of the Arabidopsis thaliana root. Development (Cambridge, England) **119**: 71–84.

Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. Nature reviews. Molecular cell biology **3**: 195–205.

Dubrovsky, J.G., Gambetta, G. a, Hernández-Barrera, a, Shishkova, S., and González, I. (2006). Lateral root initiation in Arabidopsis: developmental window, spatial patterning, density and predictability. Annals of botany **97**: 903–15.

Dubrovsky, J.G., Rost, T.L., Colón-Carmona, A., and Doerner, P. (2001). Early primordium morphogenesis during lateral root initiation in Arabidopsis thaliana. Planta **214**: 30–6.

Dubrovsky, J.G., Sauer, M., Napsucialy-Mendivil, S., Ivanchenko, M.G., Friml, J., Shishkova, S., Celenza, J., and Benková, E. (2008). Auxin acts as a local morphogenetic trigger to specify lateral root founder cells. Proceedings of the National Academy of Sciences of the United States of America **105**: 8790–4.

Dundr, M. and Misteli, T. (2010). Biogenesis of nuclear bodies. Cold Spring Harbor perspectives in biology **2**: a000711.

Dunoyer, P., Himber, C., Ruiz-Ferrer, V., Alioua, A., and Voinnet, O. (2007). Intra- and intercellular RNA interference in Arabidopsis thaliana requires components of the microRNA and heterochromatic silencing pathways. Nature genetics **39**: 848–56.

Dunoyer, P., Himber, C., and Voinnet, O. (2005). DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. Nature genetics **37**: 1356–60.

Dunoyer, P., Schott, G., Himber, C., Meyer, D., Takeda, A., Carrington, J.C., and Voinnet, O. (2010). Small RNA duplexes function as mobile silencing signals between plant cells. Science (New York, N.Y.) **328**: 912–6.

Dunoyer, P. and Voinnet, O. (2009). Movement of RNA silencing between plant cells: is the question now behind us? Trends in plant science **14**: 643–4.

Duque, P. (2011). A role for SR proteins in plant stress responses. Plant signaling & behavior **6**: 49–54.

Elmayan, T., Adenot, X., Gissot, L., Lauressergues, D., Gy, I., and Vaucheret, H. (2009). A neomorphic sgs3 allele stabilizing miRNA cleavage products reveals that SGS3 acts as a homodimer. The FEBS journal **276**: 835–44.

Elmayan, T., Balzergue, S., Béon, F., Bourdon, V., Daubremet, J., Guénet, Y., Mourrain, P., Palauqui, J.C., Vernhettes, S., Vialle, T., Wostrikoff, K., and Vaucheret, H. (1998). Arabidopsis mutants impaired in cosuppression. The Plant cell **10**: 1747–58.

Estévez, A.M., Lehner, B., Sanderson, C.M., Ruppert, T., and Clayton, C. (2003). The roles of intersubunit interactions in exosome stability. The Journal of biological chemistry **278**: 34943–51.

Fagard, M., Boutet, S., Morel, J.B., Bellini, C., and Vaucheret, H. (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. Proceedings of the National Academy of Sciences of the United States of America **97**: 11650–4.

Fang, Y., Hearn, S., and Spector, D.L. (2004). Tissue-specific expression and dynamic organization of SR splicing factors in Arabidopsis. Molecular biology of the cell **15**: 2664–73.

Fang, Y. and Spector, D.L. (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. Current biology: CB **17**: 818–23.

Feng, Y., Chen, M., and Manley, J.L. (2008). Phosphorylation switches the general splicing repressor SRp38 to a sequence-specific activator. Nature structural & molecular biology **15**: 1040–8.

Fenger-Grøn, M., Fillman, C., Norrild, B., and Lykke-Andersen, J. (2005). Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. Molecular cell **20**: 905–15.

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.-K., and Mockler, T.C. (2010). Genome-wide mapping of alternative splicing in Arabidopsis thaliana. Genome research **20**: 45–58.

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., García, J.A., and Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. Nature genetics **39**: 1033–7.

Friml, J. Subcellular trafficking of PIN auxin efflux carriers in auxin transport. European journal of cell biology **89**: 231–5.

Fukaki, H., Nakao, Y., Okushima, Y., Theologis, A., and Tasaka, M. (2005). Tissue-specific expression of stabilized SOLITARY-ROOT/IAA14 alters lateral root development in Arabidopsis. The Plant journal: for cell and molecular biology **44**: 382–95.

Fukaki, H., Tameda, S., Masuda, H., and Tasaka, M. (2002). Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of Arabidopsis. The Plant journal: for cell and molecular biology **29**: 153–68.

Fukuhara, N., Ebert, J., Unterholzner, L., Lindner, D., Izaurralde, E., and Conti, E. (2005). SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. Molecular cell **17**: 537–47.

Fukunaga, R. and Doudna, J.A. (2009). dsRNA with 5' overhangs contributes to endogenous and antiviral RNA silencing pathways in plants. The EMBO journal **28**: 545–55.

Gall, J.G., Bellini, M., Wu, Z., and Murphy, C. (1999). Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. Molecular biology of the cell **10**: 4385–402.

Gasciolli, V., Mallory, A.C., Bartel, D.P., and Vaucheret, H. (2005). Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Current biology: CB **15**: 1494–500.

Gazzani, S., Lawrenson, T., Woodward, C., Headon, D., and Sablowski, R. (2004). A link between mRNA turnover and RNA interference in Arabidopsis. Science (New York, N.Y.) **306**: 1046–8.

Gerbi, S.A., Borovjagin, A. V, and Lange, T.S. (2003). The nucleolus: a site of ribonucleoprotein maturation. Current opinion in cell biology **15**: 318–25.

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P. (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. Molecular biology of the cell **15**: 5383–98.

Girard, G., Roussis, A., Gultyaev, A.P., Pleij, C.W. a, and Spaink, H.P. (2003). Structural motifs in the RNA encoded by the early nodulation gene enod40 of soybean. Nucleic acids research **31**: 5003–15.

Gläser, H.U., Thomas, D., Gaxiola, R., Montrichard, F., Surdin-Kerjan, Y., and Serrano, R. (1993). Salt tolerance and methionine biosynthesis in Saccharomyces cerevisiae involve a putative phosphatase gene. The EMBO journal **12**: 3105–10.

Goeres, D.C., Van Norman, J.M., Zhang, W., Fauver, N.A., Spencer, M. Lou, and Sieburth, L.E. (2007). Components of the Arabidopsis mRNA decapping complex are required for early seedling development. The Plant cell **19**: 1549–64.

Golovkin, M. and Reddy, A.S. (1999). An SC35-like protein and a novel serine/arginine-rich protein interact with Arabidopsis U1-70K protein. The Journal of biological chemistry **274**: 36428–38.

Golovkin, M. and Reddy, A.S. (1996). Structure and expression of a plant U1 snRNP 70K gene: alternative splicing of U1 snRNP 70K pre-mRNAs produces two different transcripts. The Plant cell **8**: 1421–35.

Golovkin, M. and Reddy, A.S. (1998). The plant U1 small nuclear ribonucleoprotein particle 70K protein interacts with two novel serine/arginine-rich proteins. The Plant cell **10**: 1637–48.

González-Carranza, Z.H., Elliott, K.A., and Roberts, J.A. (2007). Expression of polygalacturonases and evidence to support their role during cell separation processes in Arabidopsis thaliana. Journal of experimental botany **58**: 3719–30.

Graveley, B.R. (2001). Alternative splicing: increasing diversity in the proteomic world. Trends in genetics: TIG **17**: 100–7.

Gregory, B.D., O'Malley, R.C., Lister, R., Urich, M. a, Tonti-Filippini, J., Chen, H., Millar, a H., and Ecker, J.R. (2008). A link between RNA metabolism and silencing affecting Arabidopsis development. Developmental cell **14**: 854–66.

Gultyaev, A.P. and Roussis, A. (2007). Identification of conserved secondary structures and expansion segments in enod40 RNAs reveals new enod40 homologues in plants. Nucleic acids research **35**: 3144–52.

Guthrie, C. and Steitz, J. (2005). Nucleus and gene expression Coordinated nuclear events regulate mRNA synthesis, processing, export and turnover. Current Opinion in Cell Biology **17**: 239–241.

Gutierrez, L., Bussell, J.D., Pacurar, D.I., Schwambach, J., Pacurar, M., and Bellini, C. (2009). Phenotypic plasticity of adventitious rooting in Arabidopsis is controlled by complex regulation of AUXIN RESPONSE FACTOR transcripts and microRNA abundance. The Plant cell **21**: 3119–32.

Gy, I., Gasciolli, V., Lauressergues, D., Morel, J.-B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C. (2007). Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. The Plant cell **19**: 3451–61.

Hall, L.L., Smith, K.P., Byron, M., and Lawrence, J.B. (2006). Molecular anatomy of a speckle. The anatomical record. Part A, Discoveries in molecular, cellular, and evolutionary biology **288**: 664–75.

Han, M.-H., Goud, S., Song, L., and Fedoroff, N. (2004). The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. Proceedings of the National Academy of Sciences of the United States of America **101**: 1093–8.

Hanada, K., Zhang, X., Borevitz, J.O., Li, W.-H., and Shiu, S.-H. (2007). A large number of novel coding small open reading frames in the intergenic regions of the Arabidopsis thaliana genome are transcribed and/or under purifying selection. Genome research **17**: 632–40.

Handwerger, K.E. and Gall, J.G. (2006). Subnuclear organelles: new insights into form and function. Trends in cell biology **16**: 19–26.

Haywood, V., Yu, T.-S., Huang, N.-C., and Lucas, W.J. (2005). Phloem long-distance trafficking of GIBBERELLIC ACID-INSENSITIVE RNA regulates leaf development. The Plant journal: for cell and molecular biology **42**: 49–68.

He, G., Elling, A.A., and Deng, X.W. (2011). The epigenome and plant development. Annual review of plant biology **62**: 411–35.

Heidi M. Cederholm, Anjali S. Iyer ‐ Pascuzzi, P.N.B. (2012). Patterning the primary root in Arabidopsis. WIREs Developmental Biology: 1: 675–691.

Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J., and Jacobsen, S.E. (2006). Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nature genetics **38**: 721–5.

Henz, S.R., Cumbie, J.S., Kasschau, K.D., Lohmann, J.U., Carrington, J.C., Weigel, D., and Schmid, M. (2007). Distinct expression patterns of natural antisense transcripts in Arabidopsis. Plant physiology **144**: 1247–55.

Himanen, K., Boucheron, E., Vanneste, S., Engler, J.D.A., Inzé, D., and Beeckman, T. (2002). Auxin-Mediated Cell Cycle Activation during Early Lateral Root Initiation. **14**: 2339–2351.

Himber, C., Dunoyer, P., Moissiard, G., Ritzenthaler, C., and Voinnet, O. (2003). Transitivity-dependent and -independent cell-to-cell movement of RNA silencing. The EMBO journal **22**: 4523–33.

Hiraguri, A., Itoh, R., Kondo, N., Nomura, Y., Aizawa, D., Murai, Y., Koiwa, H., Seki, M., Shinozaki, K., and Fukuhara, T. (2005). Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in Arabidopsis thaliana. Plant molecular biology **57**: 173–88.

Hofacker, I.L., Fekete, M., and Stadler, P.F. (2002). Secondary structure prediction for aligned RNA sequences. Journal of molecular biology **319**: 1059–66.

Hofmann, I., Casella, M., Schnölzer, M., Schlechter, T., Spring, H., and Franke, W.W. (2006). Identification of the junctional plaque protein plakophilin 3 in cytoplasmic particles containing RNA-binding proteins and the recruitment of plakophilins 1 and 3 to stress granules. Molecular biology of the cell **17**: 1388–98.

Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. Nature reviews. Molecular cell biology **7**: 529–39.

Hugouvieux, V., Kwak, J.M., and Schroeder, J.I. (2001). An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in Arabidopsis. Cell **106**: 477–87.

Hutchinson, J.N., Ensminger, A.W., Clemson, C.M., Lynch, C.R., Lawrence, J.B., and Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC genomics **8**: 39.

Ingelfinger, D., Arndt-Jovin, D.J., Lührmann, R., and Achsel, T. (2002). The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. RNA (New York, N.Y.) **8**: 1489–501.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell **78**: 657–68.

Jakab, G., Mougin, A., Kis, M., Pollák, T., Antal, M., Branlant, C., and Solymosy, F. (1997). Chlamydomonas U2, U4 and U6 snRNAs. An evolutionary conserved putative

third interaction between U4 and U6 snRNAs which has a counterpart in the U4atac-U6atac snRNA duplex. Biochimie **79**: 387–95.

James, A.B., Syed, N.H., Bordage, S., Marshall, J., Nimmo, G.A., Jenkins, G.I., Herzyk, P., Brown, J.W.S., and Nimmo, H.G. (2012). Alternative splicing mediates responses of the Arabidopsis circadian clock to temperature changes. The Plant cell **24**: 961–81.

Jauvion, V., Elmayan, T., and Vaucheret, H. (2010). The conserved RNA trafficking proteins HPR1 and TEX1 are involved in the production of endogenous and exogenous small interfering RNA in Arabidopsis. The Plant cell **22**: 2697–709.

Jauvion, V., Mallory, A.C., and Vaucheret, H. (2011). Deciphering post-transcriptional Gene Silencing Pathways Through Genetic Screens. In (RNAi Technology), pp. 17 –46.

Jouannet, V., Moreno, A.B., Elmayan, T., Vaucheret, H., Crespi, M.D., and Maizel, A. (2012). Cytoplasmic Arabidopsis AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. The EMBO journal **31**: 1704–13.

Kaiser, T.E., Intine, R. V, and Dundr, M. (2008). De novo formation of a subnuclear body. Science (New York, N.Y.) **322**: 1713–7.

Kalyna, M. and Barta, A. (2004). A plethora of plant serine/arginine-rich proteins: redundancy or evolution of novel gene functions? Biochemical Society transactions **32**: 561–4.

Kalyna, M., Simpson, C.G., Syed, N.H., Lewandowska, D., Marquez, Y., Kusenda, B., Marshall, J., Fuller, J., Cardle, L., McNicol, J., Dinh, H.Q., Barta, A., and Brown, J.W.S. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. Nucleic acids research **40**: 2454–69.

Kedersha, N., Cho, M.R., Li, W., Yacono, P.W., Chen, S., Gilks, N., Golan, D.E., and Anderson, P. (2000). Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. The Journal of cell biology **151**: 1257–68.

Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. The Journal of cell biology **169**: 871–84.

Kertész, S., Kerényi, Z., Mérai, Z., Bartos, I., Pálfy, T., Barta, E., and Silhavy, D. (2006). Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucleic acids research **34**: 6147–57.

Khan, G. a, Declerck, M., Sorin, C., Hartmann, C., Crespi, M., and Lelandais-Brière, C. (2011). MicroRNAs as regulators of root development and architecture. Plant molecular biology **77**: 47–58.

Kim, J.Y., Park, S.J., Jang, B., Jung, C.-H., Ahn, S.J., Goh, C.-H., Cho, K., Han, O., and Kang, H. (2007). Functional characterization of a glycine-rich RNA-binding protein 2 in

Arabidopsis thaliana under abiotic stress conditions. The Plant journal: for cell and molecular biology **50**: 439–51.

Kim, W.J., Back, S.H., Kim, V., Ryu, I., and Jang, S.K. (2005). Sequestration of TRAF2 into stress granules interrupts tumor necrosis factor signaling under stress conditions. Molecular and cellular biology **25**: 2450–62.

Kim, W.Y., Kim, J.Y., Jung, H.J., Oh, S.H., Han, Y.S., and Kang, H. Comparative analysis of Arabidopsis zinc finger-containing glycine-rich RNA-binding proteins during cold adaptation. Plant physiology and biochemistry: PPB / Société française de physiologie végétale **48**: 866–72.

Kim, Y.-O. and Kang, H. (2006). The role of a zinc finger-containing glycine-rich RNA-binding protein during the cold adaptation process in Arabidopsis thaliana. Plant & cell physiology **47**: 793–8.

Koritzinsky, M., Magagnin, M.G., Van den Beucken, T., Seigneuric, R., Savelkouls, K., Dostie, J., Pyronnet, S., Kaufman, R.J., Weppler, S.A., Voncken, J.W., Lambin, P., Koumenis, C., Sonenberg, N., and Wouters, B.G. (2006). Gene expression during acute and prolonged hypoxia is regulated by distinct mechanisms of translational control. The EMBO journal **25**: 1114–25.

Van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N., and Stuitje, A.R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. The Plant cell **2**: 291–9.

Kuhn, J.M., Breton, G., and Schroeder, J.I. (2007). mRNA metabolism of flowering-time regulators in wild-type Arabidopsis revealed by a nuclear cap binding protein mutant, abh1. The Plant journal: for cell and molecular biology **50**: 1049–62.

Kumakura, N., Takeda, A., Fujioka, Y., Motose, H., Takano, R., and Watanabe, Y. (2009). SGS3 and RDR6 interact and colocalize in cytoplasmic SGS3/RDR6-bodies. FEBS letters **583**: 1261–6.

Kurihara, Y., Matsui, A., Hanada, K., Kawashima, M., Ishida, J., Morosawa, T., Tanaka, M., Kaminuma, E., Mochizuki, Y., Matsushima, A., Toyoda, T., Shinozaki, K., and Seki, M. (2009). Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **106**: 2453–8.

Kurihara, Y., Takashi, Y., and Watanabe, Y. (2006). The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA (New York, N.Y.) **12**: 206–12.

Kutschera, U. and Niklas, K.J. (2009). Evolutionary plant physiology: Charles Darwin's forgotten synthesis. Die Naturwissenschaften **96**: 1339–54.

Kwak, K.J., Jung, H.J., Lee, K.H., Kim, Y.S., Kim, W.Y., Ahn, S.J., and Kang, H. (2012). The minor spliceosomal protein U11/U12-31K is an RNA chaperone crucial for

U12 intron splicing and the development of dicot and monocot plants. PloS one **7**: e43707.

Kwak, K.J., Kim, Y.O., and Kang, H. (2005). Characterization of transgenic Arabidopsis plants overexpressing GR-RBP4 under high salinity, dehydration, or cold stress. Journal of experimental botany **56**: 3007–16.

De la Fuente van Bentem, S., Anrather, D., Roitinger, E., Djamei, A., Hufnagl, T., Barta, A., Csaszar, E., Dohnal, I., Lecourieux, D., and Hirt, H. (2006). Phosphoproteomics reveals extensive in vivo phosphorylation of Arabidopsis proteins involved in RNA metabolism. Nucleic acids research **34**: 3267–78.

Lareau, L.F., Green, R.E., Bhatnagar, R.S., and Brenner, S.E. (2004). The evolving roles of alternative splicing. Current opinion in structural biology **14**: 273–82.

Laskowski, M.J., Williams, M.E., Nusbaum, H.C., and Sussex, I.M. (1995). Formation of lateral root meristems is a two-stage process. Development (Cambridge, England) **121**: 3303–10.

Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Rätsch, G., and Weigel, D. (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America **105**: 8795–800.

Lawrence, J.B. and Clemson, C.M. (2008). Gene associations: true romance or chance meeting in a nuclear neighborhood? The Journal of cell biology **182**: 1035–8.

Lazar, G., Schaal, T., Maniatis, T., and Goodman, H.M. (1995). Identification of a plant serine-arginine-rich protein similar to the mammalian splicing factor SF2/ASF. Proceedings of the National Academy of Sciences of the United States of America **92**: 7672–6.

Leicht, B.G., Biessmann, H., Palter, K.B., and Bonner, J.J. (1986). Small heat shock proteins of Drosophila associate with the cytoskeleton. Proceedings of the National Academy of Sciences of the United States of America **83**: 90–4.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Molecular cell **12**: 675–87.

Lejeune, F. and Maquat, L.E. (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Current opinion in cell biology **17**: 309–15.

Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proceedings of the National Academy of Sciences of the United States of America **100**: 189–92.

Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y. V, Chan, S.W.-L., Lagrange, T., Pikaard, C.S., and Jacobsen, S.E. (2006). An ARGONAUTE4-

containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. Cell **126**: 93–106.

Licatalosi, D.D. and Darnell, R.B. (2006). Splicing regulation in neurologic disease. Neuron **52**: 93–101.

Lim, M.-H., Kim, J., Kim, Y.-S., Chung, K.-S., Seo, Y.-H., Lee, I., Kim, J., Hong, C.B., Kim, H.-J., and Park, C.-M. (2004). A new Arabidopsis gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. The Plant cell **16**: 731–40.

Liu, F., Quesada, V., Crevillén, P., Bäurle, I., Swiezewski, S., and Dean, C. (2007). The Arabidopsis RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate FLC. Molecular cell **28**: 398–407.

Liu, J.-L., Wu, Z., Nizami, Z., Deryusheva, S., Rajendra, T.K., Beumer, K.J., Gao, H., Matera, A.G., Carroll, D., and Gall, J.G. (2009). Coilin is essential for Cajal body organization in Drosophila melanogaster. Molecular biology of the cell **20**: 1661–70.

Liu, Q., Shi, L., and Fang, Y. (2012). Dicing bodies. Plant physiology **158**: 61–6.

Ljung, K., Hull, A.K., Celenza, J., Yamada, M., Estelle, M., Normanly, J., and Sandberg, G. (2005). Sites and regulation of auxin biosynthesis in Arabidopsis roots. The Plant cell **17**: 1090–104.

Lopato, S., Waigmann, E., and Barta, A. (1996). Characterization of a novel arginine/serine-rich splicing factor in Arabidopsis. The Plant cell **8**: 2255–64.

Lorkovic, Z.J., Lopato, S., Pexa, M., Lehner, R., and Barta, A. (2004). Interactions of Arabidopsis RS domain containing cyclophilins with SR proteins and U1 and U11 small nuclear ribonucleoprotein-specific proteins suggest their involvement in pre-mRNA Splicing. The Journal of biological chemistry **279**: 33890–8.

Lorković, Z.J. (2009). Role of plant RNA-binding proteins in development, stress response and genome organization. Trends in plant science **14**: 229–36.

Lorković, Z.J. and Barta, A. (2002). Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. Nucleic acids research **30**: 623–35.

Lorković, Z.J., Hilscher, J., and Barta, A. (2008). Co-localisation studies of Arabidopsis SR splicing factors reveal different types of speckles in plant cell nuclei. Experimental cell research **314**: 3175–86.

Lorković, Z.J., Wieczorek Kirk, D.A., Klahre, U., Hemmings-Mieszczak, M., and Filipowicz, W. (2000). RBP45 and RBP47, two oligouridylate-specific hnRNP-like proteins interacting with poly(A)+ RNA in nuclei of plant cells. RNA (New York, N.Y.) **6**: 1610–24.

Lorković, Z.J., Wieczorek Kirk, D.A., Lambermon, M.H., and Filipowicz, W. (2000). Pre-mRNA splicing in higher plants. Trends in plant science **5**: 160–7.

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

Lucas, M., Guédon, Y., Jay-Allemand, C., Godin, C., and Laplaze, L. (2008). An auxin transport-based model of root branching in Arabidopsis thaliana. PloS one **3**: e3673.

Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. Science (New York, N.Y.) **327**: 996–1000.

Ludwig-Müller, J. (2011). Auxin conjugates: their role for plant development and in the evolution of land plants. Journal of experimental botany **62**: 1757–73.

López-Bucio, J., Cruz-Ramírez, A., and Herrera-Estrella, L. (2003). The role of nutrient availability in regulating root architecture. Current Opinion in Plant Biology **6**: 280–287.

Macknight, R., Bancroft, I., Page, T., Lister, C., Schmidt, R., Love, K., Westphal, L., Murphy, G., Sherson, S., Cobbett, C., and Dean, C. (1997). FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. Cell **89**: 737–45.

Malamy, J.E. and Benfey, P.N. (1997). Organization and cell differentiation in lateral roots of Arabidopsis thaliana. Development (Cambridge, England) **124**: 33–44.

Mao, Y.S., Zhang, B., and Spector, D.L. (2011). Biogenesis and function of nuclear bodies. Trends in genetics: TIG **27**: 295–306.

Maquat, L.E. (2004). Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nature reviews. Molecular cell biology **5**: 89–99.

Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V., and Gleba, Y. (2005). Systemic Agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants. Nature biotechnology **23**: 718–23.

Marin, E., Jouannet, V., Herz, A., Lokerse, A.S., Weijers, D., Vaucheret, H., Nussaume, L., Crespi, M.D., and Maizel, A. (2010). miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. The Plant cell **22**: 1104–17.

Marquez, Y., Brown, J.W.S., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. Genome research **22**: 1184–95.

Martinez-Contreras, R., Cloutier, P., Shkreta, L., Fisette, J.-F., Revil, T., and Chabot, B. (2007). hnRNP proteins and splicing control. Advances in experimental medicine and biology **623**: 123–47.

Mashiguchi, K., Tanaka, K., Sakai, T., Sugawara, S., Kawaide, H., and Natsume, M. (2011). The main auxin biosynthesis pathway in Arabidopsis. 1–6.

Le Masson, I., Jauvion, V., Bouteiller, N., Rivard, M., Elmayan, T., and Vaucheret, H. (2012). Mutations in the Arabidopsis H3K4me2/3 Demethylase JMJ14 Suppress Posttranscriptional Gene Silencing by Decreasing Transgene Transcription. The Plant cell **24**: 3603–12.

Matera, A.G., Terns, R.M., and Terns, M.P. (2007). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nature reviews. Molecular cell biology **8**: 209–20.

Matlin, A.J., Clark, F., and Smith, C.W.J. (2005). Understanding alternative splicing: towards a cellular code. Nature reviews. Molecular cell biology **6**: 386–98.

Mattick, J.S. and Makunin, I. V (2006). Non-coding RNA. Human molecular genetics **15 Spec No**: R17–29.

Mintz, P.J. and Spector, D.L. (2000). Compartmentalization of RNA processing factors within nuclear speckles. Journal of structural biology **129**: 241–51.

Mitchell, P., Petfalski, E., Shevchenko, a, Mann, M., and Tollervey, D. (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. Cell **91**: 457–66.

Miyagawa, R., Tano, K., Mizuno, R., Nakamura, Y., Ijiri, K., Rakwal, R., Shibato, J., Masuo, Y., Mayeda, A., Hirose, T., and Akimitsu, N. (2012). Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. RNA (New York, N.Y.) **18**: 738–51.

Mlotshwa, S., Pruss, G.J., Peragine, A., Endres, M.W., Li, J., Chen, X., Poethig, R.S., Bowman, L.H., and Vance, V. (2008). DICER-LIKE2 plays a primary role in transitive silencing of transgenes in Arabidopsis. PloS one **3**: e1755.

Mockler, T.C., Yu, X., Shalitin, D., Parikh, D., Michael, T.P., Liou, J., Huang, J., Smith, Z., Alonso, J.M., Ecker, J.R., Chory, J., and Lin, C. (2004). Regulation of flowering time in Arabidopsis by K homology domain proteins. Proceedings of the National Academy of Sciences of the United States of America **101**: 12759–64.

Molnar, A., Melnyk, C.W., Bassett, A., Hardcastle, T.J., Dunn, R., and Baulcombe, D.C. (2010). Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. Science (New York, N.Y.) **328**: 872–5.

Montgomery, T.A., Howell, M.D., Cuperus, J.T., Li, D., Hansen, J.E., Alexander, A.L., Chapman, E.J., Fahlgren, N., Allen, E., and Carrington, J.C. (2008). Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in TAS3 Trans-Acting siRNA Formation. Cell: 128–141.

Moore, M.J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. Science (New York, N.Y.) **309**: 1514–8.

Morel, J.-B., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., Proux, F., and Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. The Plant cell **14**: 629–39.

Mourrain, P., Béclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N., Rémoué, K., Sanial, M., Vo, T.A., and Vaucheret, H. (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell **101**: 533–42.

Muench, D.G., Zhang, C., and Dahodwala, M. Control of cytoplasmic translation in plants. Wiley interdisciplinary reviews. RNA **3**: 178–94.

Munns, R., James, R. a, Sirault, X.R.R., Furbank, R.T., and Jones, H.G. (2010). New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. Journal of experimental botany **61**: 3499–507.

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

Narsai, R., Howell, K. a, Millar, a H., O'Toole, N., Small, I., and Whelan, J. (2007). Genome-wide analysis of mRNA decay rates and their determinants in Arabidopsis thaliana. The Plant cell **19**: 3418–36.

Neuwald, A.F., Krishnan, B.R., Brikun, I., Kulakauskas, S., Tomcsanyi, T., Leyh, T.S., and Berg, D.E. (1992). cysQ , a Gene Needed for Cysteine Synthesis in Escherichia coli K-12 Only during Aerobic Growth. Synthesis.

Niedojadło, J., Mikulski, Z., Dełeńko, K., Szmidt-Jaworska, A., Smoliński, D.J., and Epstein, A.L. (2012). The perichromatin region of the plant cell nucleus is the area with the strongest co-localisation of snRNA and SR proteins. Planta **236**: 715–26.

Nizami, Z., Deryusheva, S., and Gall, J.G. (2010a). The Cajal body and histone locus body. Cold Spring Harbor perspectives in biology **2**: a000653.

Nizami, Z.F., Deryusheva, S., and Gall, J.G. (2010b). Cajal bodies and histone locus bodies in Drosophila and Xenopus. Cold Spring Harbor symposia on quantitative biology **75**: 313–20.

Nover, L. and Scharf, K.D. (1984). Synthesis, modification and structural binding of heat-shock proteins in tomato cell cultures. European journal of biochemistry / FEBS **139**: 303–13.

Nover, L., Scharf, K.D., and Neumann, D. (1983). Formation of cytoplasmic heat shock granules in tomato cell cultures and leaves. Molecular and cellular biology **3**: 1648–55.

Okushima, Y., Overvoorde, P.J., Arima, K., Alonso, J.M., Chan, A., Chang, C., Ecker, J.R., Hughes, B., Lui, A., Nguyen, D., Onodera, C., Quach, H., Smith, A., Yu, G., and Theologis, A. (2005). Functional genomic analysis of the AUXIN RESPONSE

FACTOR gene family members in Arabidopsis thaliana: unique and overlapping functions of ARF7 and ARF19. The Plant cell **17**: 444–63.

Oldroyd, G.E.D. and Downie, J.A. (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. Annual review of plant biology **59**: 519–46.

Olmedo, G., Guo, H., Gregory, B.D., Nourizadeh, S.D., Aguilar-henonin, L., Li, H., An, F., Guzman, P., and Ecker, J.R. (2006). exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1 □ 2. Gene Expression.

Olson, M.O.J., Hingorani, K., and Szebeni, A. (2002). Conventional and nonconventional roles of the nucleolus. International review of cytology **219**: 199–266.

Osmont, K.S., Sibout, R., and Hardtke, C.S. (2007). Hidden branches: developments in root system architecture. Annual review of plant biology **58**: 93–113.

Palaniswamy, R., Teglund, S., Lauth, M., Zaphiropoulos, P.G., and Shimokawa, T. (2010). Genetic variations regulate alternative splicing in the 5' untranslated regions of the mouse glioma-associated oncogene 1, Gli1. BMC molecular biology **11**: 32.

Palusa, S.G., Ali, G.S., and Reddy, A.S.N. (2007). Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses. The Plant journal: for cell and molecular biology **49**: 1091–107.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature genetics **40**: 1413–5.

Papp, I., Mette, M.F., Aufsatz, W., Daxinger, L., Schauer, S.E., Ray, A., Van der Winden, J., Matzke, M., and Matzke, A.J.M. (2003). Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. Plant physiology **132**: 1382–90.

Papp, I., Mur, L. a, Dalmadi, A., Dulai, S., and Koncz, C. (2004). A mutation in the Cap Binding Protein 20 gene confers drought tolerance to Arabidopsis. Plant molecular biology **55**: 679–86.

Parent, J.-S., Martínez de Alba, A.E., and Vaucheret, H. (2012). The origin and effect of small RNA signaling in plants. Frontiers in plant science **3**: 179.

Parker, R. and Sheth, U. (2007). P bodies and the control of mRNA translation and degradation. Molecular cell **25**: 635–46.

Patel, S.B. and Bellini, M. (2008). The assembly of a spliceosomal small nuclear ribonucleoprotein particle. Nucleic acids research **36**: 6482–93.

Pederson, T. (1998). The plurifunctional nucleolus. Nucleic acids research **26**: 3871–6.

Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (2004a). and the production of trans -acting siRNAs in Arabidopsis for juvenile development and the production of trans-acting siRNAs in Arabidopsis. Genes & Development: 2368–2379.

Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (2004b). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. Genes & development **18**: 2368–79.

Petrásek, J. and Friml, J. (2009). Auxin transport routes in plant development. Development (Cambridge, England) **136**: 2675–88.

Pontes, O., Li, C.F., Costa Nunes, P., Haag, J., Ream, T., Vitins, A., Jacobsen, S.E., and Pikaard, C.S. (2006). The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell **126**: 79–92.

Pontes, O. and Pikaard, C.S. (2008). siRNA and miRNA processing: new functions for Cajal bodies. Current opinion in genetics & development **18**: 197–203.

Raczynska, K.D., Simpson, C.G., Ciesiolka, A., Szewc, L., Lewandowska, D., McNicol, J., Szweykowska-Kulinska, Z., Brown, J.W.S., and Jarmolowski, A. (2010). Involvement of the nuclear cap-binding protein complex in alternative splicing in Arabidopsis thaliana. Nucleic acids research **38**: 265–78.

Reddy, A.S.N. (2007). Alternative splicing of pre-messenger RNAs in plants in the genomic era. Annual review of plant biology **58**: 267–94.

Reddy, A.S.N. and Shad Ali, G. Plant serine/arginine-rich proteins: roles in precursor messenger RNA splicing, plant development, and stress responses. Wiley interdisciplinary reviews. RNA **2**: 875–89.

Richardson, D.N., Rogers, M.F., Labadorf, A., Ben-Hur, A., Guo, H., Paterson, A.H., and Reddy, A.S.N. (2011). Comparative analysis of serine/arginine-rich proteins across 27 eukaryotes: insights into sub-family classification and extent of alternative splicing. PloS one **6**: e24542.

Riera, M., Redko, Y., and Leung, J. (2006). Arabidopsis RNA-binding protein UBA2a relocalizes into nuclear speckles in response to abscisic acid. FEBS letters **580**: 4160–5.

Rohrig, H., Schmidt, J., Miklashevichs, E., Schell, J., and John, M. (2002). Soybean ENOD40 encodes two peptides that bind to sucrose synthase. Proceedings of the National Academy of Sciences of the United States of America **99**: 1915–20.

Roman, G., Lubarsky, B., Kieber, J.J., Rothenberg, M., and Ecker, R. (1995). Genetic Analysis of Ethylene Signal Transduction i n. October.

De Rybel, B., Vassileva, V., Parizot, B., Demeulenaere, M., Grunewald, W., Audenaert, D., Van Campenhout, J., Overvoorde, P., Jansen, L., Vanneste, S., Möller, B., Wilson, M., Holman, T., Van Isterdael, G., Brunoud, G., Vuylsteke, M., Vernoux, T., De Veylder, L., Inzé, D., Weijers, D., et al. (2010). A novel aux/IAA28 signaling

cascade activates GATA23-dependent specification of lateral root founder cell identity. Current biology: CB **20**: 1697–706.

Rymarquis, L. a, Souret, F.F., and Green, P.J. (2011). Evidence that XRN4, an Arabidopsis homolog of exoribonuclease XRN1, preferentially impacts transcripts with certain sequences or in particular functional categories. RNA (New York, N.Y.) **17**: 501–11.

Rühl, C., Stauffer, E., Kahles, A., Wagner, G., Drechsel, G., Rätsch, G., and Wachter, A. (2012). Polypyrimidine tract binding protein homologs from Arabidopsis are key regulators of alternative splicing with implications in fundamental developmental processes. The Plant cell **24**: 4360–75.

Sapra, A.K., Ankö, M.-L., Grishina, I., Lorenz, M., Pabis, M., Poser, I., Rollins, J., Weiland, E.-M., and Neugebauer, K.M. (2009). SR protein family members display diverse activities in the formation of nascent and mature mRNPs in vivo. Molecular cell **34**: 179–90.

Schomburg, F.M., Patton, D.A., Meinke, D.W., and Amasino, R.M. (2001). FPA, a gene involved in floral induction in Arabidopsis, encodes a protein containing RNA-recognition motifs. The Plant cell **13**: 1427–36.

Schwach, F., Vaistij, F.E., Jones, L., and Baulcombe, D.C. (2005). An RNA-dependent RNA polymerase prevents meristem invasion by potato virus X and is required for the activity but not the production of a systemic silencing signal. Plant physiology **138**: 1842–52.

Schwartz, S., Gal-Mark, N., Kfir, N., Oren, R., Kim, E., and Ast, G. (2009). Alu exonization events reveal features required for precise recognition of exons by the splicing machinery. PLoS computational biology **5**: e1000300.

Schwartz, S.H., Silva, J., Burstein, D., Pupko, T., Eyras, E., and Ast, G. (2008). Large-scale comparative analysis of splicing signals and their corresponding splicing factors in eukaryotes. Genome research **18**: 88–103.

Schöning, J.C., Streitner, C., Meyer, I.M., Gao, Y., and Staiger, D. (2008). Reciprocal regulation of glycine-rich RNA-binding proteins via an interlocked feedback loop coupling alternative splicing to nonsense-mediated decay in Arabidopsis. Nucleic acids research **36**: 6977–87.

Schüttpelz, M., Schöning, J.C., Doose, S., Neuweiler, H., Peters, E., Staiger, D., and Sauer, M. (2008). Changes in conformational dynamics of mRNA upon AtGRP7 binding studied by fluorescence correlation spectroscopy. Journal of the American Chemical Society **130**: 9507–13.

Searle, I.R., Pontes, O., Melnyk, C.W., Smith, L.M., and Baulcombe, D.C. (2010). JMJ14, a JmjC domain protein, is required for RNA silencing and cell-to-cell movement of an RNA silencing signal in Arabidopsis. Genes & development **24**: 986–91.

Secco, D., Baumann, A., and Poirier, Y. (2010). Characterization of the rice PHO1 gene family reveals a key role for OsPHO1;2 in phosphate homeostasis and the evolution of a distinct clade in dicotyledons. Plant physiology **152**: 1693–704.

Seo, P.J., Park, M.-J., Lim, M.-H., Kim, S.-G., Lee, M., Baldwin, I.T., and Park, C.-M. (2012). A self-regulatory circuit of CIRCADIAN CLOCK-ASSOCIATED1 underlies the circadian clock regulation of temperature responses in Arabidopsis. The Plant cell **24**: 2427–42.

Shao, W., Kim, H.-S., Cao, Y., Xu, Y.-Z., and Query, C.C. (2012). A U1-U2 snRNP interaction network during intron definition. Molecular and cellular biology **32**: 470–8.

Shaw, P.J. and Brown, J.W.S. (2004). Plant nuclear bodies. Current opinion in plant biology **7**: 614–20.

Sheth, U. and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science (New York, N.Y.) **300**: 805–8.

Shevtsov, S.P. and Dundr, M. (2011). Nucleation of nuclear bodies by RNA. Nature cell biology **13**: 167–73.

Shopland, L.S., Johnson, C. V, Byron, M., McNeil, J., and Lawrence, J.B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. The Journal of cell biology **162**: 981–90.

Shukla, G.C. and Padgett, R.A. (1999). Conservation of functional features of U6atac and U12 snRNAs between vertebrates and higher plants. RNA (New York, N.Y.) **5**: 525–38.

Simpson, C.G., Manthri, S., Raczynska, K.D., Kalyna, M., Lewandowska, D., Kusenda, B., Maronova, M., Szweykowska-Kulinska, Z., Jarmolowski, A., Barta, A., and Brown, J.W.S. (2010). Regulation of plant gene expression by alternative splicing. Biochemical Society transactions **38**: 667–71.

Simpson, G.G., Clark, G.P., Rothnie, H.M., Boelens, W., Van Venrooij, W., and Brown, J.W. (1995). Molecular characterization of the spliceosomal proteins U1A and U2B" from higher plants. The EMBO journal **14**: 4540–50.

Sleeman, J.E. and Lamond, A.I. (1999). Nuclear organization of pre-mRNA splicing factors. Current opinion in cell biology **11**: 372–7.

De Smet, I., Lau, S., Voss, U., Vanneste, S., Benjamins, R., Rademacher, E.H., Schlereth, A., De Rybel, B., Vassileva, V., Grunewald, W., Naudts, M., Levesque, M.P., Ehrismann, J.S., Inzé, D., Luschnig, C., Benfey, P.N., Weijers, D., Van Montagu, M.C.E., Bennett, M.J., Jürgens, G., et al. (2010). Bimodular auxin response controls organogenesis in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **107**: 2705–10.

De Smet, I., Tetsumura, T., De Rybel, B., Frey, N.F.D., Laplaze, L., Casimiro, I., Swarup, R., Naudts, M., Vanneste, S., Audenaert, D., Inzé, D., Bennett, M.J., and

Beeckman, T. (2007). Auxin-dependent regulation of lateral root positioning in the basal meristem of Arabidopsis. Development (Cambridge, England) **134**: 681–90.

De Smet, I., Vassileva, V., De Rybel, B., Levesque, M.P., Grunewald, W., Van Damme, D., Van Noorden, G., Naudts, M., Van Isterdael, G., De Clercq, R., Wang, J.Y., Meuli, N., Vanneste, S., Friml, J., Hilson, P., Jürgens, G., Ingram, G.C., Inzé, D., Benfey, P.N., and Beeckman, T. (2008). Receptor-like kinase ACR4 restricts formative cell divisions in the Arabidopsis root. Science (New York, N.Y.) **322**: 594–7.

Smith, C.J., Watson, C.F., Bird, C.R., Ray, J., Schuch, W., and Grierson, D. (1990). Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. Molecular & general genetics: MGG **224**: 477–81.

Smith, L.M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F.K., Herr, A.J., Pikaard, C.S., and Baulcombe, D.C. (2007). An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in Arabidopsis. The Plant cell **19**: 1507–21.

Sommerville, J. (1992). RNA-binding proteins: masking proteins revealed. BioEssays: news and reviews in molecular, cellular and developmental biology **14**: 337–9.

Song, L., Han, M.-H., Lesicka, J., and Fedoroff, N. (2007). Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. Proceedings of the National Academy of Sciences of the United States of America **104**: 5437–42.

Souret, F.F., Kastenmayer, J.P., and Green, P.J. (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. Molecular cell **15**: 173–83.

Sousa, C., Johansson, C., Charon, C., Manyani, H., Sautter, C., and Kondorosi, A. (2001). Translational and Structural Requirements of the Early Nodulin Gene enod40 , a RNA , for Elicitation of a Cell-Specific Growth Response in the Alfalfa Root Cortex Translational and Structural Requirements of the Early Nodulin Gene enod40 , a Short-Open Re. Society.

Spector, D.L. (2011). Immunofluorescence localization of nuclear proteins. Cold Spring Harbor protocols **2011**: 1276–80.

Spector, D.L. (2001). Nuclear domains. Journal of cell science **114**: 2891–3.

Spector, D.L. and Lamond, A.I. (2011). Nuclear speckles. Cold Spring Harbor perspectives in biology **3**.

Staiger, D. (2002). Circadian rhythms in Arabidopsis: time for nuclear proteins. Planta **214**: 334–44.

Stamm, S. (2008). Regulation of alternative splicing by reversible protein phosphorylation. The Journal of biological chemistry **283**: 1223–7.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A., and Soreq, H. (2005). Function of alternative splicing. Gene **344**: 1–20.

Stewart, C.L., Roux, K.J., and Burke, B. (2007). Blurring the boundary: the nuclear envelope extends its reach. Science (New York, N.Y.) **318**: 1408–12.

Streitner, C., Danisman, S., Wehrle, F., Schöning, J.C., Alfano, J.R., and Staiger, D. (2008). The small glycine-rich RNA binding protein AtGRP7 promotes floral transition in Arabidopsis thaliana. The Plant journal: for cell and molecular biology **56**: 239–50.

Swarup, K., Benková, E., Swarup, R., Casimiro, I., Péret, B., Yang, Y., Parry, G., Nielsen, E., De Smet, I., Vanneste, S., Levesque, M.P., Carrier, D., James, N., Calvo, V., Ljung, K., Kramer, E., Roberts, R., Graham, N., Marillonnet, S., Patel, K., et al. (2008). The auxin influx carrier LAX3 promotes lateral root emergence. Nature cell biology **10**: 946–54.

Syed, N.H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J.W.S. (2012). Alternative splicing in plants--coming of age. Trends in plant science **17**: 616–23.

Tacke, R. and Manley, J.L. (1999). Determinants of SR protein specificity. Current opinion in cell biology **11**: 358–62.

Tanabe, N., Yoshimura, K., Kimura, A., Yabuta, Y., and Shigeoka, S. (2007). Differential expression of alternatively spliced mRNAs of Arabidopsis SR protein homologs, atSR30 and atSR45a, in response to environmental stress. Plant & cell physiology **48**: 1036–49.

Teixeira, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M., and Parker, R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. RNA (New York, N.Y.) **11**: 371–82.

Terzi, L.C. and Simpson, G.G. (2008). Regulation of flowering time by RNA processing. Current topics in microbiology and immunology **326**: 201–18.

Tian, C., Muto, H., Higuchi, K., Matamura, T., Tatematsu, K., and Koshiba, T. (2004). Disruption and overexpression of auxin response factor 8 gene of Arabidopsis affect hypocotyl elongation and root growth habit , indicating its possible involvement in auxin homeostasis in light condition. **3**: 333–343.

Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J., and Guigó, R. (2009). Nucleosome positioning as a determinant of exon recognition. Nature structural & molecular biology **16**: 996–1001.

Topping, J.F., May, V.J., Muskett, P.R., and Lindsey, K. (1997). Mutations in the HYDRA1 gene of Arabidopsis perturb cell shape and disrupt embryonic and seedling morphogenesis. Development (Cambridge, England) **124**: 4415–24.

Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., Blencowe, B.J., Prasanth, S.G., and Prasanth, K. V (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell **39**: 925–38.

Ule, J. and Darnell, R.B. (2006). RNA binding proteins and the regulation of neuronal synaptic plasticity. Current opinion in neurobiology **16**: 102–10.

Unterholzner, L. and Izaurralde, E. (2004). SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. Molecular cell **16**: 587–96.

Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. Genes & development **20**: 515–24.

Vaucheret, H. (2005). MicroRNA-dependent trans-acting siRNA production. Science's STKE: signal transduction knowledge environment **2005**: pe43.

Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes & development **20**: 759–71.

Vaucheret, H., Palauqui, J.C., Elmayan, T., and Moffatt, B. (1995). Molecular and genetic analysis of nitrite reductase co-suppression in transgenic tobacco plants. Molecular & general genetics: MGG **248**: 311–7.

Vazquez, F., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A.C., Hilbert, J., Bartel, D.P., and Cre, P. (2004). Endogenous trans-Acting siRNAs Regulate the Accumulation of Arabidopsis mRNAs. October **16**: 69–79.

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

Veley, K.M. and Michaels, S.D. (2008). Functional redundancy and new roles for genes of the autonomous floral-promotion pathway. Plant physiology **147**: 682–95.

Voinnet, O., Vain, P., Angell, S., and Baulcombe, D.C. (1998). Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. Cell **95**: 177–87.

Walker, M.P., Tian, L., and Matera, A.G. (2009). Reduced viability, fertility and fecundity in mice lacking the cajal body marker protein, coilin. PloS one **4**: e6171.

Wang, G.-S. and Cooper, T.A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. Nature reviews. Genetics **8**: 749–61.

Wang, K.C. and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Molecular cell **43**: 904–14.

Weber, C., Nover, L., and Fauth, M. (2008). Plant stress granules and mRNA processing bodies are distinct from heat stress granules. The Plant journal: for cell and molecular biology **56**: 517–30.

Wilczynska, A., Aigueperse, C., Kress, M., Dautry, F., and Weil, D. (2005). The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. Journal of cell science **118**: 981–92.

Woodward, A.W. and Bartel, B. (2005). Auxin: regulation, action, and interaction. Annals of botany **95**: 707–35.

Workman, P. (2004). Inhibiting the phosphoinositide 3-kinase pathway for cancer treatment. Biochemical Society transactions **32**: 393–6.

Xie, Z., Allen, E., Wilken, A., and Carrington, J.C. (2005). DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in Arabidopsis thaliana. PNAS.

Xie, Z. and Qi, X. (2008). Diverse small RNA-directed silencing pathways in plants. Biochimica et biophysica acta **1779**: 720–4.

Xiong, L., Lee Bh, Ishitani, M., Lee, H., Zhang, C., and Zhu, J.K. (2001). FIERY1 encoding an inositol polyphosphate 1-phosphatase is a negative regulator of abscisic acid and stress signaling in Arabidopsis. Genes & development **15**: 1971–84.

Xu, J., Yang, J.-Y., Niu, Q.-W., and Chua, N.-H. (2006). Arabidopsis DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. The Plant cell **18**: 3386–98.

Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.-S., Zhang, C., Yeo, G., Black, D.L., Sun, H., Fu, X.-D., and Zhang, Y. (2009). Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. Molecular cell **36**: 996–1006.

Yan, K., Liu, P., Wu, C.-A., Yang, G.-D., Xu, R., Guo, Q.-H., Huang, J.-G., and Zheng, C.-C. (2012). Stress-induced alternative splicing provides a mechanism for the regulation of microRNA processing in Arabidopsis thaliana. Molecular cell **48**: 521–31.

Yang, W.C., Katinakis, P., Hendriks, P., Smolders, A., Vries, F., Spee, J., Kammen, A., Bisseling, T., and Franssen, H. (1993). Characterization of Gmenod40, a gene showing novel patterns of cell-specific expression during soybean nodule development. The Plant Journal **3**: 573–585.

Yasuda, J. and Hayashizaki, Y. (2008). The RNA continent. Advances in cancer research **99**: 77–112.

Yoo, B.-C., Kragler, F., Varkonyi-Gasic, E., Haywood, V., Archer-Evans, S., Lee, Y.M., Lough, T.J., and Lucas, W.J. (2004). A systemic small RNA signaling system in plants. The Plant cell **16**: 1979–2000.

Yoshikawa, M., Peragine, A., Park, M.Y., and Poethig, R.S. (2005). A pathway for the biogenesis of trans -acting siRNAs in Arabidopsis A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. Genes & Development: 2164–2175.

Zakrzewska-Placzek, M., Souret, F.F., Sobczyk, G.J., Green, P.J., and Kufel, J. (2010). Arabidopsis thaliana XRN2 is required for primary cleavage in the pre-ribosomal RNA. Nucleic acids research **38**: 4487–502.

Zhang, X.-N. and Mount, S.M. (2009). Two alternatively spliced isoforms of the Arabidopsis SR45 protein have distinct roles during normal plant development. Plant physiology **150**: 1450–8.

Zhao, B., Schneid, C., Iliev, D., Schmidt, E.-M., Wagner, V., Wollnik, F., and Mittag, M. (2004). The circadian RNA-binding protein CHLAMY 1 represents a novel type heteromer of RNA recognition motif and lysine homology domain-containing subunits. Eukaryotic cell **3**: 815–25.

VII- annexes

VII- annexes

Annexe 1:

Ariel, F., Moreno, A.B., Bardou, F., and Crespi, M. (2012). non-protein-coding RNAs and root développement. Chapter 2 From the book « Root genetic and soil interactions »

Chapter 2

The complex eukaryotic transcriptome: non-protein-coding RNAs and root development

F. Ariel, A.B. Moreno, F. Bardou and M. Crespi*

Institut des Sciences du Végétal (ISV), Centre National de la Recherche Scientifique (CNRS)

1, Av de la Terrasse, F91198

Gif sur Yvette cedex, France

F. Ariel: fariel@isv.cnrs-gif.fr
A. Moreno: moreno@isv.cnrs-gif.fr
F. Bardou: bardou@isv.cnrs-gif.fr
* Corresponding author: M. Crespi: e-mail: crespi@isv.cnrs-gif.fr

Abstract

In recent years, in addition to mRNAs, the non-protein-coding RNAs (or npcRNAs) have emerged as a major part of the eukaryotic transcriptome. New genomic approaches allowed the discovery of many new regulatory npcRNAs and their characterization suggests a role for specific regulatory npcRNAs in the generation of evolutionary complexity in multicellular organisms. These regulatory npcRNAs range from small RNAs to the so-called long npcRNAs and were shown to play diverse functions in plant development. Small RNAs act in the regulation of gene expression at transcriptional or posttranscriptional level and several among them have been linked to root growth and development. On the other hand, only few long npcRNAs have been implicated in root responses such as abiotic stress responses or symbiotic nodule formation. RNA metabolism is globally determined by a variety of RNA-binding proteins, with which RNAs form ribonucleoprotein complexes (RNPs) and regulatory RNAs may interfere and/or modulate the action of these different RNPs. The understanding of these RNA networks may reveal novel mechanisms involved in root development and its adaptation to changing environmental conditions.

1. Genomic approaches reveal novel aspects of the eukaryotic transcriptome

Although the term "transcriptome" usually refers to the whole set of messenger RNA (mRNA) molecules in the cell, it is worth noting that mRNA only represents a fraction of all the transcripts that are actually transcribed. In fact, many transcripts are not translated into proteins, the so-called non-protein-coding RNAs (npcRNAs), but still play a role in relevant cellular functions and complexes (e.g. ribonucleoproteins). NpcRNAs are a class of RNAs with very poor protein coding potential but their function is associated to the RNA molecule itself. They form a heterogeneous group of RNAs that could be divided into three different classes according to their length and function. Generally, npcRNAs can range from 18 to 25 nucleotides for the families of microRNAs and small interfering RNAs (siRNAs) linked to post-transcriptional and transcriptional gene silencing, 20 to 300 nucleotides for small RNAs commonly found as transcriptional and translational regulators (e.g. small nuclear RNAs), or up to and beyond 10,000 nucleotides for medium and large npcRNAs mainly linked to the chromatin environment and epigenetic mechanisms. It is now known that there exist transcripts spanning chromosomal regions larger than 2 Mb (Carninci 2010). Indeed, transcripts originated and terminated in chromosomal regions that are as long as several megabases have been identified by RACE (Djebali et al. 2008). In the last years, bioinformatics and experimental strategies have revealed a remarkable number of novel npcRNA candidates in various model organisms from yeast or plants to *Homo sapiens* (Mattick and Makunin 2006; Yasuda and Hayashizaki 2008; Mercer et al. 2009). NpcRNAs are generally encoded in intergenic regions, and there exist a high number of antisense RNA transcripts, pseudogenes and truncated transcripts in eukaryotes.

The transcriptome complexity involving npcRNAs have been hypothesized to exert a regulatory role that is required for the development and function of higher organisms (Carninci et al. 2010). Indeed, certain npcRNAs have been implicated in different regulatory mechanisms in plant development (Brown et al. 2008; Voinnet 2009), in biotic interactions and the response to environmental abiotic stress (BenAmor et al. 2009; Jay et al. 2010; Sunkar 2010). Generally, npcRNAs are produced by RNA polymerase II and are capped and polyadenylated. Recent exploration of in the poly(A)-/+ transcriptomes of HeLa and H9 cells revealed that while the majority of protein coding transcripts are poly(A)+, a significant portion of them are either poly(A)- or

bimorphic, being found in both the poly(A)+ and poly(A)- populations. Only stable excised introns were identified as a class of poly(A)- long npcRNAs (Yang et al. 2011). In plants, certain npcRNAs are produced by specific RNA polymerases IV and V (Voinnet 2009) and then processed into smaller RNA molecules (see below). In fact, a transcriptome may be surprisingly complex with long npcRNAs often overlapping with or interspersed between coding transcripts. This dynamic molecular picture significantly changed the understanding of gene expression in eukaryotes, if we consider that a single DNA sequence can be transcribed in multiple sense and antisense transcripts, intronic npcRNAs, intergenic or promoter associated RNAs (Mercer et al. 2009). In *Arabidopsis thaliana*, whole-genome mapping based on the use of tiling arrays revealed that over 30% of observed transcription was intergenic and that many antisense RNA transcripts actually exist (Yamada et al. 2003).

Long and small npcRNAs represent an emerging class of riboregulators, which either act directly in this long form or are processed to shorter molecules (Figure 1). Several long npcRNAs are processed into small RNAs due to their folding as double stranded RNA (dsRNA) loops derived from endogenous loci (as the miRNAs) or due to the action of RNA-dependent RNA polymerases that generate long dsRNAs. These double stranded RNA structures are processed into siRNAs by member(s) of the Dicer family (DCL, Dicer-like, Vaucheret 2006). Small si/miRNAs induce mRNA cleavage and translational inhibition in the so-called RISC complexes (RNA-induced silencing complex) through pairing with specific mRNA targets, mainly in the cytoplasm, or lead to transcriptional gene silencing, heterochromatin formation and *de novo* DNA methylation in the nucleus (Vaucheret 2006; Jamalkandi and Masoudi-Nejad 2009; Verdel et al. 2009). Another class of endogenous siRNAs deriving from pairs of natural *cis*-antisense transcripts was discovered in plants. Natural Antisense Transcripts (NATs) can be defined as endogenous RNA molecules that are transcribed from the opposite DNA strands, resulting in partially or entirely complement transcripts. As in most species, the majority of *Arabidopsis* NATs pairs (72%) overlapped at their 3' end (Wang and Metzlaff 2005) and, for 99% of them, the overlapping region included exon sequences. Both sense and antisense RNAs can encode proteins or be npcRNAs, although in most cases antisense transcription consists of a protein coding RNA overlapping with an npcRNA (Faghihi and Wahlestedt 2009). NATs can give place to active siRNAs formed from their overlapped region (Borsani et al. 2005). Small RNAs have the fascinating property to move from cell to cell and act in recipient tissues. The recent discovery that microRNAs are also able to move between root tissues (Martienssen 2010) raises the intriguing possibility that different cells from a tissue signals other cells to coordinate growth. Although there are very few examples of miRNA movement, it is well established that siRNAs can move between plant tissues even at long distances (Dunoyer et al. 2010) and be biologically active in the recipient cells. Indeed, specific miRNAs (such as miR390) target long npcRNAs (e.g. TAS genes) to produce secondary tasiRNAs (for trans-acting siRNAs) that control mRNA expression patterns in a diffusible manner in meristems (Chitwood and Timmermans 2010) and during lateral root growth (Marin et al. 2010). These mobile siRNAs can repress gene expression at post-transcriptional level and even transmit new epigenetic patterns into root tissues (e.g. changes in DNA methylation) as elegantly shown using grafting experiments with different silencing mutants (Melnyk et al. 2011). If members of these classes of mobile signals (si/miRNAs) function in this manner in the root, their production and distribution in primary and lateral meristems during growth of the root system promises to be a key research goal for the coming years. Interestingly, recent data showed that plant miRNAs, like the rice miR168a, may be orally acquired by animals through food intake. Strikingly, miR168a was detected in the animal sera and tissues (Zhang et al. 2011). Moreover, it has been shown that miR168a could bind to the human/mouse low-density lipoprotein receptor adapter protein 1 (LDLRAP1) mRNA, inhibiting LDLRAP1 expression in liver, and consequently decreasing LDL removal from mouse plasma. These data indicate clearly an evidence of cross-kingdom regulation by miRNA.

Apart from the known described small RNAs (si/miRNAs) other small RNAs may play different roles in development. Deep sequencing analysis has recently revealed that under some growth or stress conditions tRNA cleaved pathways could be activated in different organisms like human yeast or plants (e.g. during oxidative stress in *Arabidopsis* seedling, Thompson et al. 2008). Indeed, an abundant population of small RNAs of 30 to 40 nt aligns with tRNA sequences which can derive either from 5' or 3' end of mature tRNAs or from tRNA precursors by a cleavage near or in the anticodon loop (Pederson et al. 2010). The two types of tRNA-derived small RNAs (tsRNAs) are non-random cleavage products: type I (5'-tsRNAs) and II (3'-tsRNAs) (Haussecker et al. 2010). Both types are produced from tRNA precursors and respectively generated by DICER (Babiarz et al. 2008) and endoribonuclease (Thompson and Parker, 2009). In plants, it has been shown that phloem sap tsRNA

interfere with wheat germ extract translation (Zhang et al. 2009). It has been shown that specific tRNA fragments may interfere with ribosomal activity, although these data were only obtained with artificially produced tRNA fragments in an *in vitro* experiment. However, the tsRNAs ability to cause the assembly of stress granules suggests that they could affect the initiation of translation. Over ten years ago, Asano and co-workers (2001) have shown that specific tsRNAs were able to displace eIF4F from capped mRNA. Moreover, pool-down analysis reveals that tsRNAs bind to a complex containing AGO2, which is involved in miRNAs/siRNAs silencing in animals. tsRNAs represent 34% of small RNA reads in phosphate starved roots and among this the Gly-tRNATCC represented over 80% of the population (Hsieh et al. 2009). Moreover, these tsRNAs were more accumulated in roots. The precise role of these small npcRNAs remains unknown but recently, tsRNAs have been observed with a specific pattern in embryogenic and non embryogenic rice calli (Chen et al. 2011). As certain aspects of the induction of calli formation may resemble an ectopic lateral root development program (Sugimoto et al. 2010) a role for tsRNAs also in root development may be considered.

Lately, bioinformatic approaches on full-length cDNA databases resulted in the identification of 76 *Arabidopsis* npcRNAs (Hirsch et al. 2006; Ben Amor et al. 2009), including several precursors of known miRNAs. Eleven of these npcRNAs turned out to be antisense to protein-coding mRNAs, suggesting that they act as *cis*-regulatory molecules, whereas 5 others are likely to be siRNA precursors. Expression analyses of the 76 npcRNAs in mutant plants of several RNA binding proteins involved in npcRNA processing, shed light on the molecular pathways in which some of them are involved. Abiotic stress, such as phosphate starvation, drought or salt stress altered the accumulation of 22 npcRNAs. In the same work, over-expression analyses in *Arabidopsis* identified 2 npcRNAs as regulators of root growth during salt stress and leaf morphology, respectively. These novel results supported the idea that together with small RNAs, long npcRNAs can also exert diverse roles during growth and differentiation.

Long npcRNAs have been shown to have specific localisations at tissular, cellular and sub-cellular levels (Campalans et al. 2004; Zhan and Lukens 2010) however their mechanisms of action remain largely unknown. It has been demonstrated in several organisms that long npcRNAs interact with histone modifiers (Yu et al. 2009; Swiezewski et al. 2009; Yap et al. 2010). Among the best studied npcRNAs in animals, several were identified in the nucleus and act either as *cis*- or *trans*-acting epigenetic regulators of chromatin. This is the case for Xist that triggers X-chromosome inactivation (Chaumeil et al. 2006; Clemson et al. 1996), Kcnq1ot1 and Air that control genomic imprinting (Mancini-Dinardo et al. 2006; Nagano et al. 2008; Pandey et al. 2008; Redrup et al. 2009; Sleutels et al. 2002; Terranova et al. 2008) and *HOTAIR* that regulates homeobox (HOX) genes (Rinn et al. 2007). It has been shown that *HOTAIR* participates in epigenetic changes by recruiting chromatin remodelling complexes to specific genomic loci (Rinn et al. 2007). In the same context, the long npcRNA MIRA activates the target genes *HOXa6* and *HOXa7* by recruiting epigenetic activators, implicating long npcRNAs in gene regulation during vertebrate cell-fate determination (Bertani et al. 2011). Recently, another long intergenic npcRNA called *HOTTIP* transcribed from the 5' tip of the *HOXa* locus was shown to coordinate the activation of several *HOXa* genes *in vivo*. A chromosomal looping brings *HOTTIP* into close proximity to its target genes. *HOTTIP* RNA binds to adaptor proteins and target remodeling complexes across HOXA, driving histone H3 Lys 4 trimethylation and gene transcription (Wang et al. 2011). In *Arabidopsis*, the study of the regulation of the MADS box transcriptional regulator *FLOWERING LOCUS C* (*FLC*) revealed a set of antagonistic pathways comprising evolutionary conserved activities that link chromatin regulation, transcription level and co-transcriptional RNA metabolism (Crevillen and Dean 2011). Cold triggers enrichment of tri-methylated Histone H3 Lys 27 at the chromatin of the floral repressor *FLC* and results in its epigenetically stable repression. It has been demonstrated that a long intronic npcRNA (termed as *COLDAIR*) physically associates with a component of Polycomb Repressive Complex 2 (PRC2) and targets it to *FLC*, allowing the vernalization-mediated epigenetic repression of the locus (Heo and Sung, 2010). In addition, a group of related antisense npcRNAs (termed *COOLAIR*) from *FLC* have been proposed to be involved in vernalization-mediated *FLC* repression (Swiezewski et al. 2009). The RNA-binding proteins FCA and FPA participate in the repression of *FLC* but they may play a more global role in plant development and epigenetic control. Indeed, in addition to FLC, the double mutant *fca-fpa* misexpressed a large number of previously un-annotated genomic segments corresponding to intergenic regions in which DNA methylation was altered. FCA and FPA would play important roles in the *Arabidopsis* genome in RNA 3' processing and transcription termination, thus limiting intergenic transcription (Sonmez et al. 2011).

Long npcRNAs including antisense npcRNAs have to bypass several RNA-quality control (RQC) mechanisms occurring in the cell that actively degrade "aberrant" mRNAs such as incompletely spliced products, transcripts without coding capacity or pseudogenes (Lykke-Andersen et al. 2009). For example, a genome wide analysis of exosome substrates in *Arabidopsis* revealed, in addition to mRNA and miRNA processing intermediates, hundreds of npcRNAs and antisense RNAs not previously described (Chekanova et al. 2007). The exosome is a macromolecular complex that mediates RNA processing and degradation and is generally essential for viability in eukaryotes. It was shown that the plant exosome exhibits a remarkable functional plasticity in comparison to yeast and metazoans. The npcRNAs only detected in exosome mutants include large numbers of antisense RNAs as they are rapidly and actively degraded in wild-type plants (Chekanova et al. 2007). Similarly, the non-sense mediated mRNA decay (or NMD) is an RQC mechanism related to cytoplasmic foci known as P-bodies, that recognizes premature nonsense or stop codons within an mRNA (Conti and Izaurralde 2005). After recognition of an incorrectly positioned stop codon, the NMD system signals the elimination of the mRNAs through decapping, deadenylation, and exonucleolytic degradation. The UP-frameshift proteins are essential for NMD and 3 *UPF* genes exist in *Arabidopsis*. The analysis of these mutants revealed that, in addition to the expected NMD substrates, most npcRNAs including large numbers of antisense RNAs are degraded by this pathway, suggesting that one of the most important roles of NMD is the genome-wide suppression of aberrant or antisense RNAs (Kurihara et al. 2009). In *Saccharomyces cerevisiae*, regulatory npcRNAs are destabilized by the XRN1 5'→3' RNA exonuclease, and the use of strand-specific RNA sequencing (RNA-seq), helped to identify a novel class of over 1,000 XRN1-sensitive unstable transcripts (XUTs) in which the majority are antisense to open reading frames. These transcripts are polyadenylated and RNA pol II-dependent. The accumulation of most of XUTs in response to stress suggests that they might have a role in adaptive responses to changes in growth conditions (van Dijk et al. 2011). Another major class of npcRNAs recently described consists of the Cryptic Unstable Transcripts (CUTs), which are very short-lived RNA pol II transcripts only detectable when stabilized by mutations impairing the nuclear exosome (Neil and Jacquier 2011). On the other hand, the yeast Stable Uncharacterized Transcripts (SUTs) are detectable also in wild type cells. These transcripts are partially susceptible to the nuclear exosome, and primarily degraded by the NMD (Marguardt et al. 2011). The role of various RNA Binding Proteins in RQC mechanisms will be discussed below. However, it is clear that long npcRNAs, including small RNA precursors, regulating development must bypass these control mechanisms in order to reach the different ribonucleoproteic complexes where they act on their mRNA targets.

It is worth noting that the majority of the transcriptome studies usually measure the actual amount of transcripts present in a cell, resulting from the balance between transcription and degradation rates. This means that the complete set of sequences actually transcribed is significantly wider than the originally annotated genomic features. The steady-state of npcRNAs should be considered in relation to their influence on gene expression and their own regulation may also affect the stability of their mRNA targets. Although several of these long npcRNAs have been experimentally identified in plants (Charon et al. 1999; Hirsch et al. 2006; Ma et al. 2008; Ben Amor et al. 2009; Franco-Zorrilla et al. 2009), their activity, subcellular localisation or molecular roles remain largely unknown. Globally, npcRNAs have been far less studied in plants than in animals, where diverse mechanisms involving npcRNAs in the regulation of gene expression have been discovered (Prasanth and Spector 2007; Voinnet 2009; Wilusz et al. 2009).

2. The role of RNA binding proteins in npcRNA metabolism and activity

The post-transcriptional regulation of gene expression mediated by npcRNAs as well as all aspects of RNA metabolism is globally determined by a variety of RNA-binding proteins (RBPs; Lorkovic 2009). Most RBPs contain one or more conserved domains, as the RNA-recognition motif (RRM) and the Khomology (KH) motif. Based on the characteristics of those conserved domains, the analysis of the *A. thaliana* genome revealed that plants express a complex set of RBPs, with 196 RRM- and 26 KH-containing proteins, the majority of them being plant-specific (Barta et al. 2008; Lorkovic and Barta 2002). Most of these proteins have not been characterized experimentally, and it is largely unclear how their action may control gene expression and development, primarily due to the difficulty in the identification of their RNA partners (Lorkovic 2009). Nevertheless, following forward and reverse genetic approaches, some studies begin to reveal the requirement of specific RBPs that have crucial roles in RNA metabolism during plant development.

Most RBPs are likely to have multiple RNA partners including mRNAs and npcRNAs (e.g. antisense RNAs, intergenic npcRNAs or small RNAs) that may integrate

different RNPs (Ribonucleoproteins) to generate RNA networks in which npcRNAs can determine their localization, interfere or modulate their action (acting as competitors or activators against other substrates). Identification of the RBPs with which npcRNAs associate will contribute to understand their role in RNP networks in the cell. For example, several RBPs are involved in the biogenesis and action of small RNAs (e.g. DCLs, RDRs or AGOs, Valencia-Sanchez et al. 2006). The RNAi pathways have been largely diversified and several steps occur in the nucleus and/or the cytoplasm. The different small RNA silencing pathways differ mainly in the way of generation of the small RNA (Vaucheret et al. 2006). One RNA silencing mechanism is initiated by endogenous loci able to form double-stranded stem-loops that are processed by DICER (double-stranded RNAses) into small RNAs called miRNAs (miRNA pathway). In the siRNA pathway, a single-stranded RNA is targeted by RNA-dependent RNA polymerases to form long dsRNAs that then is cut by DICERs into small RNAs called siRNAs. In addition, long dsRNAs can be precursors of siRNAs when an npcRNA (a trans acting or TAS long non-coding RNAs) is targeted by a specific miRNA and cleaved, becoming a substrate of RDR polymerases to form a long dsRNA qui sera processes en small siRNAs, called tasiRNAs (derived from TAS genes). Finally, two independent transcripts may form a complementary dsRNA (natural antisense genes) and this region of the dsRNA molecule is processed by DICER into dsRNA small molecules or natsiRNAs. In all cases, one strand of the processed si/miRNAs duplex is incorporated into the RISC complex containing AGO proteins. The presence of this small RNA provides the RISC complex a sequence specificity to target a particular mRNA molecule. Gene Silencing can occur at both PostTranscriptional (PTGS, through mRNA cleavage and/or translation inhibition) and/or Transcriptional levels (TGS, through DNA methylation) (Vaucheret 2006).

In the cell there exist several sources of long npcRNAs that may lack of coding capacity because they are aberrant RNAs, e.g. mRNAs without either the 5' cap or the 3' polyA tail. Those aberrant RNAs can be converted in double stranded RNA by RDRs triggering silencing mechanisms through the action of DCLs or be eliminated by the normal degradation machinery, involving the exosome, XRN enzymes or the NMD pathway (Conti and Izaurralde 2005; Chekanova et al. 2007; Kurihara et al. 2009). These survey mechanisms conform what we previously defined as RQC mechinery and aim to ensure the correct expression of the genes (Houseley and Tollervey 2009).

As mentioned before, one of those control survey mechanisms is the NMD, for non-sense-mediated mRNA decay, whose core RBPs are the UP-frameshift proteins (UPFs). This pathway recognizes premature nonsense or stop codons (PTC) within an mRNA by the action of the exon-junction complex (EJC) that marks correctly fused exons (Pontes and Pikaard 2008). Such PTC containing transcripts can arise as a result of genomic frameshifts, nonsense mutations or inefficiently spliced pre-mRNAs for example (Maquat 2004). In *Arabidopsis* there are three UPF genes (UPF1, UPF2 and UPF3), and a genome-wide analysis using a tiling array of the mutants *upf1-1* and *upf3-1* revealed that in addition to the expected NMD substrates, that is coding mRNAs, most npcRNAs, including large numbers of antisense RNAs, are degraded by this pathway (the tiling-array data can be viewed at http://omicspace.riken.jp/gps/group/psca3). In these mutants the percentage of up-regulated messenger-like-npcRNAs (mlnpcRNAs) to all expressed mlnpcRNAs was much higher than the percentage of up-regulated mRNAs to all expressed mRNAs. This suggests that in fact, besides the recognition of nonsense mRNAs, another very important role of NMD is the genome-wide suppression of the mlnpc-RNAs that are recognized as aberrant transcripts by this machinery (Kurihara et al. 2009). After recognition of an incorrectly positioned stop codon, the NMD system through the action of other RBPs, signals the elimination of this RNA by recruiting decapping enzymes followed by 5'→3' exonuclease activities (XRN enzymes), and/or deadenylation enzymes followed by 3'→5' exonucleolytic degradation (Amrani et al. 2006; Conti and Izaurralde 2005; Lejeune and Maquat 2005). This 3'→5' exonucleolytic degradation is known to take place in the exosome complex. Hence, npcRNAs may affect the stability of other aberrant transcripts by interfering with this pathway.

The exosome is an evolutionary conserved macromolecular complex that mediates numerous reactions of 3'– 5' RNA processing and degradation, being essential for eukaryotic cell viability (Estevez et al. 2003; Mitchell et al. 1997). Some of its activities relay in the homeostatic mRNA turnover, decay of unstable mRNAs, nonsense-mediated mRNA decay and degradation of the mRNA fragments derived from endonucleolytic cleavage mediated by mi/siRNAs in RISC complexes (RNA-induced silencing complex) (Houseley et al. 2006). Combining genetic, proteomic, and whole-transcriptome analyses to investigate the function of the exosome complex Chekanova and co-workers (2007) found that individual subunits of the exosome are

functionally specialized, ranging from being dispensable for growth and development (CSL4) to being essential for the development of female gametophytes (RRP41) or embryogenesis (RRP4). These findings demonstrate an unexpected degree of functional plasticity in the plant exosome core. Concerning the exosome targets they found multiple classes of stable structural RNAs, a select subset of mRNAs, primary miRNA (primiRNA) processing intermediates, tandem repeat-associated siRNA precursor species, as well as numerous long npcRNAs, as mentioned before, such as npcRNAs associated with heterochromatic regions in plants. In addition, the results revealed a novel layer of the transcriptome composed of intergenic npcRNAs that are tightly downregulated by constitutive exosome activity. This elegant work gave rise to a publicly available exosome-regulated transcriptome database (http://signal.salk.edu/cgi-bin/exosome) that will certainly help future work to elucidate regulatory mechanisms in complex eukaryotic transcriptomes.

Taking into account the roles of the NMD and the exosome, one could expect that the transcripts up-regulated in the *upf* mutants ought to overlap with the exosome substrates, as aberrant transcripts recognized by the NMD complex should be degraded from the 3' end by deadenylation and subsequent 3'→5' exonuclease activity in the exosome (Mitchell and Tollervey 2003; Lejeune et al. 2003). However it was not the case and only slight overlap was found between transcripts up-regulated in the *upf* mutants and exosome substrates. In addition, in the population identified as exosome substrates there were no natural antisense RNAs, or NAT-RNAs. It is possible that these differences are probably due to the experimental variations, such as age of plants, difference in growth conditions and statistical analysis (Kurihara et al. 2009) and certainly future work will need to address these discrepancies.

As mentioned before, after recognition of an abnormal RNA by one of the cellular survey mechanisms, the RNA molecules enters the decay mechanisms usually starting by deadenylation of the 3' poly(A) tail and then by decapping of the 5' cap structure, followed by degradation in the 3'→5' and/or 5'→3' directions by the exosome or XRN exoribonucleases, respectively (Chiba and Green 2009). In addition to aberrant transcripts derived from incorrect splicing or other cellular mechanisms, RNA substrates for this step also arrive from the silencing pathways after the cleavage of mRNA targets by si/miRNAs (Souret et al. 2004; Gy et al. 2007; Gregory et al. 2008; Rymarquis et al. 2011).

Despite that several components of the RQC machinery are required early on for plant growth and their mutations are lethal, inactivation of certain other components may lead to specific phenotypes. The *Arabidopsis* genome contains three XRN proteins (XRN2, XRN3 and XRN4) with different subcellular localizations and cellular functions. XRN2 and XRN3 are localized in the nucleus, are partially redundant and are required for primary cleavage and processing of pre-ribosomal RNAs (Zakrzewska-Placzek et al. 2010). On the other hand, XRN4, also called ETHYLENE INSENSITIVE 5 (EIN5), is cytoplasmic and necessary for a correct ethylene signalling in the plant and for the degradation of the 3' products resulting from the miRNA-mediated cleavage of target mRNAs (Roman et al. 1995; Olmedo et al. 2006; Gregory et al. 2008). Recent work pointed the association of XRN4 with transcripts encoding specific sequence motifs and select functional groups suggesting that these RNAs may be specifically targeted to the 5'→3' decay pathway for degradation in *Arabidopsis*. The RNA sequence thus would play a major role in this targeting (Rymarquis et al. 2011). In addition to these functions on survey mechanisms, the three XRN proteins are involved in PTGS acting as endogenous RNA silencing suppressors. The exoribonuclease XRN4 was shown to act as a PTGS suppressor, possibly through the degradation of RNA-dependent RNA polymerase (RdRp) templates. In this way, mutations in XRN4 lead to the accumulation of aberrant, uncapped RNAs derived from transgenes, which could enhance PTGS (Gazzani et al. 2004). Likewise, in another work XRN2 and XRN3, as well as FIERY1 (FRY1), were also shown to be suppressors of PTGS as the XRN activity is inhibited in a *fry1* background (Gy et al. 2007). FRY1, also known as SAL1 and HOS2, is a dual function 3'(2'),5'-bisphosphate nucleotidase/inositol polyphosphate 1-phosphatase orthologous to Hal2 and CysQ in yeast and *Escherichia coli*, respectively (Neuwald et al. 1992; Glaser et al. 1993). It was isolated in a genetic screen based on ABA- and stress-inducible gene transcription, and the *Arabidopsis* mutation *fry1* results in super-induction of ABA- and stress-responsive genes. Accordingly, *fry1* mutants are more sensitive to ABA or stress inhibition, and present less tolerance to cold, drought, and salt stresses (Xiong et al. 2001). Accordingly to their activity as suppressors of PTGS, both *fry1* and the *xrn* double and triple mutants accumulate RNA intermediates of miRNA-directed post-transcriptional regulation which are templates of XRN4 (Souret et al. 2004), and miRNA loops, which are templates of both XRN2 and XRN3 (Gy et al. 2007). To further analyze the role of XRN in the *fry1* phenotype, they generated an *xrn2 xrn3 xrn4* triple mutant that was

fertile, unlike the sterile *xrn2 xrn3* double mutant. Although the mechanism for the partial phenotypic rescue is unclear, it suggests that *xrn4* mutations act to partially suppress the *xrn2 xrn3* phenotypic effects. Neither XRN2 nor XRN4 is critical for proper plant development. However, the embryo-lethality of null *xrn3* alleles and the developmental defects exhibited by hypomorphic *xrn3-3* mutants suggest an essential role of XRN3. The *xrn2 xrn3 xrn4* triple mutant displays the *fry1* lateral root and drought tolerance phenotypes but does not affect primary root. Microscopical observations revealed that the altered root architecture in *fry1* mutants was due to reduced meristem activity in the primary root and to a lateral root initiation defect. Altogether, these results suggest that the pleiotropic phenotype of the *fry1* mutants results, at least in part, from a general perturbation in XRN activities.

Recently, in a genetic screen for *Arabidopsis* mutants deregulated in the expression of Phosphate High affinity Transporter 1;4 (PHT1;4), a novel *fry1* allele was described. The authors identified a novel FRY1 function modulating the transcription of several Pi starvation markers in the root stele, however this *fry1* mutant phenotype is independent of XRN activities. A transcriptomic analysis confirmed that the phenotype observed corresponded to a point mutation in the transcript corresponding to the *fry* gene.

The two periphery marks of the extreme borders of a eukaryotic mRNA are defined by the 5' m7G-cap structure and the 3' poly(A) tail, and in the nuclei of eukaryotic organisms the 5' cap is recognized by the cap-binding complex (CBC). In *Arabidopsis* there are two single genes for both the large and small subunits, ABH1 (ABA Hypersensitive 1) and CBP20 respectively, that form the CBC heterodimer and play numerous roles in RNA metabolism (Aguilera 2005; Hugouvieux et al. 2001 and 2002). Mutant plants for the ABH1 RBP present ABA hypersensitivity in seed germination, stomata closure, reduced wilting during drought and ABA-induced guard cell calcium increases (Hugouvieux et al. 2001). *ABH1* activity is necessary for the correct expression level of a subset of genes in the *Arabidopsis* genome (Hugouvieux et al. 2001), suggesting a specific connection between mRNA metabolism and ABA signalling. Consistent with their intimately linked activities, inactivation of CBP20 causes a similar serrated leaf phenotype and increased drought resistance as seen in *abh1* mutants (Hugouvieux et al. 2001; Papp et al. 2004). Although the morphological and physiological effects of ABH1 and CBP20 have been quite well described (Hugouvieux et al. 2001; Bezerra et al. 2004; Papp et al. 2004), it is largely unclear how these proteins relate to the biochemical functions of the CBC. Nevertheless, through the analysis of the developmental defects on *xrn4-abh1* double mutant plants, Gregory and co-workers (2008) demonstrated surprising roles in RNA silencing pathways for these two proteins, XRN4 and ABH1, involved in general RNA metabolism. They found that the loss of ABH1 decreases the levels of mature miRNAs, suggesting that this protein functions in the miRNA-mediated RNA silencing pathway. Indeed the *Arabidopsis* CBC would be important for proper pri-miRNA processing eventually providing a platform for recruitment of miRNA maturation factors (Laubinger et al. 2008). On the other hand, XRN4 affected the abundance of a distinct class of mainly 21 nt small RNAs, processed from both sense and antisense strands of some endogenous transcripts. Apparently those transcripts are converted to double-stranded RNA (dsRNA) and subsequently processed, and regularly accumulate in an uncapped form in *xrn4* mutant plants. Taken together, these results suggest that an additional fate for endogenous uncapped transcripts is shuttling into an RNA silencing pathway where they become small RNA-biogenesis substrates.

Accordingly, a very recent work showed that there is a link between the *Arabidopsis* cap-binding protein ABH1 and the suppression of silencing (Christie et al. 2011). Their results indicate that genes containing introns are less susceptible to PTGS and that this intron suppression of gene silencing requires an efficient splicing that is dependent on ABH1. At the end of the 80s, it was been already published that endogenous genes generate much higher levels of gene expression than their cDNA counterparts (Callis et al. 1987). As well, in transgene-expressed viral RNA genomes, the addition of introns has also been shown to significantly enhance their accumulation (Marillonnet et al. 2005). Finally, genome-wide mRNA decay rates show that transcripts from intronless genes are significantly more unstable than those from intron-containing genes (Narsai et al. 2007). Based on a GFP-transgenic reporter system and varying the number of introns, this recent study provides a molecular basis to elucidate those evidences. They show that efficiently spliced introns may reduce RDR6 activity along spliced GFP transcripts via a mechanism requiring the cap-binding protein ABH1. Indeed the ABH1 protein has been previously correlated to pre-mRNA splicing in plants (Kuhn et al. 2007; Gregory et al. 2008; Laubinger et al. 2008) and various studies in yeast and animals have shown that the CBC is functionally and physically coupled to pre-mRNA splicing (Izaurralde et al. 1994). To investigate this hypothesis, they transformed *abh1* mutants with the same GFP-transgenic reporter system and

unexpectedly concluded that intron suppression of transgene silencing was lost in those mutants. They propose a model where an efficient intron splicing could decrease the quantity of aberrant RNA by-products produced by transcription (as uncapped or improperly terminated transcripts), or alternatively, facilitate recruitment of enzymes that degrade aberrant RNA as they are formed. As discussed before, various RQC pathways exist in the cell and compete against endogenous RdR polymerases activities to prevent extensive amplification of silencing. In other words, there is a constant fight between RQC and the silencing pathways competing for the same aberrant or endogenous npcRNA substrates produced in the cell. Thus it is possible that the large varieties of npcRNAs found in the cell may interfere with the RNA surveillance and degradation pathways and affect the processes of RNA maturation.

As mentioned before, the total number of protein coding genes in diverse organisms varies much less than the number of different transcripts along evolutionary scales (Mattick and Makunin 2006; Yasuda and Hayashizaki 2008). This supports the idea that npcRNAs are essential to understand the huge complexity of multi-cellular organisms. Indeed, the large diversity of npcRNAs identified up to now in eukaryotes, and their increasing number, may reflect the importance of riboregulation, mediated by npcRNA–RBPs interactions, in the determination of differentiation and adaptability in eukaryotes. Concretely, plants display a notable flexibility in their architecture and growing patterns in response to external stimuli, characterized by a great developmental plasticity. This quality allows higher plants to adapt to different environmental conditions, with individuals with the same genotype giving rise to different phenotypes. Thus, future challenges lie in understanding the implication of the various RNP networks to determine growth and developmental outcomes under different environmental conditions.

3. Non protein coding RNAs in root development

In this section, we will discuss more specific npcRNAs linked to root development (Table 1). The root architecture of the plant constitutes an important model to study how developmental plasticity is translated into growth responses under stress conditions. Indeed, primary root development and the formation of *de novo* meristems to generate lateral roots (LRs) are conditioned by the soil environment (Osmont et al. 2007). The remarkable developmental plasticity called the attention of Charles Darwin and his son Francis. In their monograph on the *Power of Movements* they referred to the behavior (i.e. gravitropism) of the growing root, postulating that the root tip acts like a plant brain: "We believe that there is no structure in plants more wonderful, as far as its functions are concerned, than the tip of the radicle... It is hardly an exaggeration to say that the tip of the radicle thus endowed, and having the power of directing the movements of the adjoining parts, acts like the brain of the lower animals" (cited by Kutschera and Niklas 2009).

Several of the best described riboregulators in plant biology, the miRNAs, have been linked to root development (Table 1), such as the control of primary root growth or the formation of organs from *de novo* meristems, such as lateral and adventitious roots, or the legume-specific nitrogen-fixing nodules, through symbiotic interactions with soil bacteria (Khan et al. 2011).

External cues influence plant growth by modulating hormone levels and signaling. Auxin is one of the main phytohormones regulating root growth and architecture. Auxin activity is mediated by the *AUXIN RESPONSE FACTOR* (*ARF*) genes, a plant-specific family of transcriptional regulators (Okushima et al. 2005). TAS3 is a trans-acting siRNA, whose biogenesis requires the initial miR390-mediated cleavage of the TAS3 precursor (Marin et al. 2010). The cleavage product is then converted to double-stranded RNA through the RdR6 activity and sequential DCL4-mediated cleavage events (Peragine et al. 2004; Vazquez et al. 2004; Allen et al. 2005; Gasciolli et al. 2005; Xie et al. 2005; Yoshikawa et al. 2005; Adenot et al. 2006). Of the four tasiRNAs precursors identified (TAS1-4) in *Arabidopsis*, cleavage of TAS3 is unique since it requires the specific action of the miR390/AGO7 complex for ta-siRNA production (Montgomery et al. 2008). These tasiRNAs inhibit ARF2, ARF3, and ARF4, thus releasing repression of lateral root growth (Marin et al. 2010). In addition, ARF2, ARF3 and ARF4 affect auxin-induced miR390 accumulation. Positive and negative feedback regulation of miR390 by ARF2, ARF3, and ARF4 thus ensures the proper definition of the miR390 expression pattern resulting in the adaptation of the root system architecture (RSA) by auxin. In *Arabidopsis*, it was shown that auxin-associated miRNAs tightly control adventitious root formation through a complex regulation that involved various ARFs (Gutierrez et al. 2009). Whereas miR160 positively regulates adventitious root formation by controlling ARF17, the auxin-related miR167, which targets ARF6 and 8, is a negative regulator of adventitious root development. In

Arabidopsis, ARF8 and ARF17 play antagonist roles in auxin homeostasis (Tian et al. 2004; Sorin et al. 2005).

Other molecular mechanisms involved in root plasticity in response to the environment and endogenous signals are the natural antisense RNA (NATs), which are transcripts complementary to *cis* or *trans*-mRNAs and exert a repressive activity on them. Considering that NATs may encode proteins, some of them can be classified as dual RNAs because of their double function: both NAT and mRNA (Bardou et al. 2011). NATs have been described in several organisms like yeast, human, mice and plants. We can distinguish two different classes of NATs: *cis*-NATs, which are generated by antisense transcription at the same genomic locus, and *trans*-NATs, which are generated from different loci. Interestingly, a large number of transcripts were predicted to have both *cis*- and *trans*-NATs, suggesting that antisense transcripts can form a complex regulatory network (Henz et al. 2007). In *Arabidopsis*, a *cis*-NAT pair encoding SRO5 and P5CDH, was shown to repress translation, and improving *Arabidopsis* salt tolerance (Borsani et al. 2005). High salt concentration is toxic for plants when up-taken from the soil by the root (Ariel et al. 2010), strongly affecting root metabolism and development. In response, plants have developed biological mechanisms that prevent NaCl accumulation or absorption (Munns et al. 2010). P5CDH is constitutively expressed and encodes the D1-pyrroline-5-carboxylate dehydrogenase, which prevents proline accumulation, whereas SRO5 is induced by salt stress and encodes an unknown protein. Under high levels of NaCl, both genes form a natural double-stranded pair of transcripts proved to be cleaved by DCL2 and DCL1 generating 24- and 21-nucleotide nat-siRNAs that will repress the constitutively expressed P5CDH mRNAand lead to increased salt tolerance (Borsani et al. 2005).

In many cases, the function or coding capacity of NAT-RNAs remain largely unknown. *PHO* genes participate in phosphate (Pi) transport in the cellular response to Pi starvation in plants. Phosphate is normally up-taken from the soil and is essential for plant development. In rice, three *PHO1* genes can form a *cis*-NAT pair with npcRNAs. Surprisingly, the *cis*-NAT associated with the OsPHO1-3 gene does not appear to be regulated either developmentally or in response to Pi deficiency, in contrast to the *cis*-NATs of OsPHO1-1 and OsPHO1-2, which are strongly up-regulated by Pi starvation, while the expression of the complementary sense transcript remains relatively stable (Secco et al. 2010). However, considering that the analysis was performed using whole roots, a hypothesis is that the expression of the sense and antisense OsPHO1 transcripts may not occur in the same cells or tissues in all cases. In *Arabidopsis*, a search for npcRNAs (Ben Amor et al. 2009) allowed to identify 13 antisense npcRNAs complementary to protein-coding transcripts. One of these (npc536) forms a *cis*-NAT with AT1G67930, and its over-expression allowed plants to grow under salt stress without modifying AT1G67930 mRNA accumulation. Furthermore, *npc536* mutants do not show any misregulation of the antisense transcript. As this NAT contains a short open reading frame (ORF) conserved in rice, npc536 may act through this encoded peptide. Alternatively, npc536 may regulate translation of the AT1G67930 mRNA or act as a *trans*-NAT, with an unidentified complementary target that plays a role in the salt stress response.

Apart from npcRNAs that lead to small RNAs such as the TAS or the NAT genes, only few npcRNAs have been implicated in root developmental processes (Charon et al. 2010). The *ENOD40* genes code for highly structured plant mRNAs that contain a series of short ORFs without any long ORF (Charon et al. 1999; Gultyaev et al. 2007) and are involved in legume-specific root nodule organogenesis. Root nodules are nitrogen-fixing symbiotic plant organs that result from the interaction of soil bacteria of the genus *Rhizobium* with the root cells of host legumes (Oldroyd and Downie 2008; Crespi et al. 2008). This process initiates with cell-specific division in the roots, where *ENOD40* is strongly expressed (Yang et al. 1993; Crespi et al. 1994). The *ENOD40* gene is characterized by specific conserved nucleotide sequences that can be also found in some non-leguminous plants (Gultyaev et al. 2007). Furthermore, transgenic lines with increased or decreased levels of *ENOD40* exhibit accelerated or reduced nodulation, respectively (Charon et al. 1999). It was first proposed that *ENOD40* was an npcRNA due to its highly stable RNA secondary structure, a characteristic of known npcRNAs (Crespi et al. 1994; Hofacker et al. 2002) however other authors proposed that this transcript may encode a small primary oligopeptide of around 13 aminoacids (Charon et al. 2010). Translational analysis identified two short ORFs (sORF I and II; 13 and 27 amino acids long, respectively) that could be translated from this transcript in *Medicago truncatula* (Sousa et al. 2001) and sORF I contains a conserved nucleotide region across legumes, but not other plants, in contrast to the highly conserved stem-loops of the *ENOD40* RNA throughout all known plant species (Girard et al. 2003). A cell-specific assay for the action of *ENOD40* in *Medicago sativa* was developed using a biolistic process and suggested that translation of these sORF may be biologically relevant. Interestingly, mutations in the predicted structured RNA region also strongly

inhibited this biological activity (Sousa et al. 2001). These results confirm the importance of both the sORF peptides and the RNA secondary structure of *ENOD40* in its activity and suggest that *ENOD40* encodes a bi-functional or dual RNA. To gain further insight into the action of *ENOD40*, molecules that interact with the peptides or RNA were identified. A novel RNA-binding protein MtRBP1 (for *M. truncatula* RNA-Binding Protein 1), which interacts with the *ENOD40* RNA, was identified using a yeast three-hybrid screen. Immunolocalization studies and the use of an MtRBP1-DsRed2 fusion construct showed that MtRBP1 localized to nuclear "speckles," which are nuclear ribonucleoprotein complexes known to house the splicing machinery in plant cells (Cioce and Lamond, 2005; Handwerger and Gall, 2006; Li et al. 2006). These nuclear speckles (or inter- chromatin granule clusters) are spotted shapeless structures containing elevated concentrations of splicing snRNPs and other splicing-related proteins that participate in the co-transcriptional splicing of mRNAs at the chromosomes (Shaw and Brown, 2004). Interestingly, MtRBP1 was located in the cytoplasm of *ENOD40* expressing cells in *M. truncatula* nodules. The direct involvement of the *ENOD40* RNA in MtRBP1 relocalization into cytoplasmic granules was confirmed using a transient expression assay and an MS2 bacteriophage system to tag the *ENOD40* RNA (Campalans et al. 2004). This *in vivo* approach to monitor RNA-protein interactions demonstrated that the cytoplasmic relocalization of MtRBP1 was mediated by *ENOD40* and suggested that the relocalization of nuclear RNA-binding proteins during specific developmental processes could be a new function mediated by npcRNAs (Campalans et al. 2004). On the other hand, the *ENOD40* peptides expressed in soybeans were shown to bind to sucrose synthase (SUC1) suggesting a potential role of these peptides in the regulation of sucrose utilization in the nodules (Rohrig et al. 2002). These results further highlight that npcRNA genes may act as bi-functional RNAs in plants as many genes contain potentially active sORF-encoded peptides. In *Arabidopsis*, more than 3000 sORFs are transcribed, suggesting that huge numbers of sORF-encoded peptides are still hidden in genomic regions that have not been annotated yet (Hanada et al. 2007). For example, the *POLARIS* (*PLS*) gene in *Arabidopsis* was identified experimentally using a promoter-trap approach and shown to have a root-specific expression pattern (Topping et al. 1997). *pls* mutant plants have short roots with radially expanded cells and reduced leaf vascularization (Casson et al. 2002). The *PLS* gene is transcribed as a relatively short, 500-nucleotide mRNA, which contains three short ORFs that encode putative peptides of 8, 9 and 36 amino acids. Over-expression of the ORF encoding the 36-amino acid peptide partially rescues the short-root phenotype. Although the function of *PLS* has not been fully elucidated, a role in hormonal homeostasis, including ethylene signaling and auxin transport, and in the regulation of microtubule cytoskeletal dynamics was proposed (Chilley et al. 2006). These dual RNA further highlights the large variety of mechanisms that npcRNAs may trigger in root tissues.

4. Future perspectives

In recent years, growing evidence has demonstrated that RNAs can not only encode proteins, but can also exert a wide range of molecular functions, including the modulation of mRNA expression and RNA processing or localization; the regulation of protein activity and structure; and can act as precursors to small RNAs or sORF-encoded peptides. The transcriptome complexity and the remarkable portion of npcRNAs indicate that many of the roles played by these RNAs remain largely unknown. Plant developmental plasticity in response to environmental changes depends entirely on molecular flexibility in gene expression. Therefore, future work will help to unravel the crucial role of npcRNAs in root development.

5. References

Adenot, X. et al., 2006. DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. *Current biology: CB*, 16(9), 927-932.

Aguilera, A., 2005. Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. *Current opinion in cell biology*, 17(3), 242-250.

Allen, E. et al., 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, 121(2), 207-221.

Amrani, N., Sachs, M.S. & Jacobson, A., 2006. Early nonsense: mRNA decay solves a translational problem. *Nature reviews Molecular Cell Biology*, 7(6), 415-425.

Ariel, F. et al., 2010. Environmental regulation of lateral root emergence in Medicago truncatula requires the HD-Zip I transcription factor HB1. *The Plant Cell*, 22(7), 2171-2183.

Asano, K. et al., 2011 A Multifactor Complex of eIF1, eIF2, eIF3, eIF5, and $tRNA_i^{Met}$ Promotes Initiation Complex Assembly and Couples GTP Hydrolysis to AUG Recognition Cold Spring Harbor Symposia on Quantitative Biology, 66, 403-416.

Babiarz, J.E. et al., 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes & development*, 22(20), 2773-2785.

Bardou, F. et al., 2011. Dual RNAs in plants. *Biochimie*, 93, 1950-1954.

Bari, R. et al., 2006. Phosphate-Signaling Pathway in Plants. *Plant Physiology*, 141, 988-999.

Barta, A., Kalyna, M., & Lorkovic, Z.J., 2008. Plant SR proteins and their functions. *Current Topics in Microbiology and Immunology*, 326, 83-102.

Ben Amor, B. et al., 2009. Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. *Genome research*, 19(1), 57-69.

Bertani S., Sauer S., Bolotin E., & Sauer F., 2011. The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Molecular Cell*, 43(6), 1040-6.

Bezerra, I.C. et al., 2004. Lesions in the mRNA cap-binding gene ABA HYPERSENSITIVE 1 suppress FRIGIDA-mediated delayed flowering in Arabidopsis. *The Plant journal*, 40(1), 112-119.

Borsani, O. et al., 2005. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell*, 123(7), 1279-1291.

Boualem, A. et al., 2008. MicroRNA166 controls root and nodule development in Medicago truncatula. *The Plant journal*, 54(5), 876-887.

Brown, J.W.S., Marshall, D.F. & Echeverria, M., 2008. Intronic noncoding RNAs and splicing. *Trends in plant science*, 13(7), 335-342.

Callis, J., Fromm, M. & Walbot, V., 1987. Introns increase gene expression in cultured maize cells. *Genes & Development*, 1(10), 1183-1200.

Campalans, A., Kondorosi, A. & Crespi, Martin, 2004. Enod40 , a Short Open Reading Frame - Containing mRNA , Induces Cytoplasmic Localization of a Nuclear RNA Binding Protein in Medicago truncatula. *The Plant Cell*, 16, 1047-1059.

Carninci, P., 2010. RNA dust: where are the genes? *DNA research*, 17(2), 51-59.

Casson, S.A. et al., 2002. The POLARIS gene of Arabidopsis encodes a predicted peptide required for correct root growth and leaf vascular patterning. *The Plant Cell*, 14(8), 1705.

Charon, C et al., 1999. Alteration of enod40 expression modifies medicago truncatula root nodule development induced by Sinorhizobium meliloti. *The Plant Cell*, 11(10), 1953-1966.

Charon, Celine et al., 2010. Non-protein-coding RNAs and their interacting RNA-binding proteins in The Plant Cell nucleus. *Molecular plant*, 3(4), 729-739.

Chaumeil, J. et al., 2006. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes & development*, 20(16), 2223-2237.

Chekanova, J. a et al., 2007. Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell*, 131(7), 1340-1353.

Chen C.J. et al., 2011. Genome-wide discovery and analysis of microRNAs and other small RNAs from rice embryogenic callus. *RNA Biology*, 8(3), 538-547.

Chiba, Y. & Green, P.J., 2009. mRNA Degradation Machinery in Plants. *Journal of Plant Biology*, 52(2), 114-124.

Chilley, Paul M et al., 2006. The POLARIS peptide of Arabidopsis regulates auxin transport and root growth via effects on ethylene signaling. *The Plant Cell*, 18(11), 3058-3072.

Chitwood, D.H. & Timmermans, M.C.P., 2010. Small RNAs are on the move. *Nature*, 467(7314), 415-419.

Christie, M., Croft, L.J. & Carroll, B.J., 2011. Intron splicing suppresses RNA silencing in Arabidopsis. *The Plant journal*, 68(1), 4676-4676.

Cioce, M. & Lamond, A.I., 2005. Cajal bodies: a long history of discovery. *Annual review of cell and developmental biology*, 21, 105-131.

Clemson, C.M. et al., 1996. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *The Journal of cell biology*, 132(3), 259-275.

Combier, J.P. et al., 2006. MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. *Genes & development*, 20(22), 3084-3088.

Conti, E. & Izaurralde, Elisa, 2005. Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Current opinion in cell biology*, 17(3), 316-325.

Crespi, M. & Frugier, F., 2008. De novo organ formation from differentiated cells: root nodule organogenesis. *Science signaling*, 1(49), re11.

Crespil, M.D. et al., 1994. expressed during nodule codes for a non-translatable RNA involved in plant growth. *EMBO Journal*, 13(21), 5099-5112.

Crevillén, P. & Dean, C., 2011. Regulation of the floral repressor gene FLC: the complexity of transcription in a chromatin context. *Current opinion in plant biology*, 14(1), 38-44.

Djebali, S. et al., 2008. Efficient targeted transcript discovery via array-based normalization of RACE libraries. *Nature Methods*, 5(7), 629-635.

Dunoyer, P. et al., 2010. Small RNA duplexes function as mobile silencing signals between plant cells. *Science*, 328(5980), 912-916.

Estévez, A.M. et al., 2003. The roles of intersubunit interactions in exosome stability. *The Journal of biological chemistry*, 278(37), 34943-34951.

Faghihi, M.A. et al., 2010. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome biology*, 11(5), R56.

Franco-Zorrilla, J.M. et al., 2007. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature genetics*, 39(8), 1033-1037.

Franco-Zorrilla, J.M. et al., 2009. Genome-wide identification of small RNA targets based on target enrichment and microarray hybridizations. *The Plant journal*, 59(5), 840-850.

Jiao, Y. et al., 2010. Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nature genetics*, 42(6), 541-544.

Gasciolli, V. et al., 2005. Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Current Biology*, 15(16), pp.1494-500.

Gazzani, S. et al., 2004. A link between mRNA turnover and RNA interference in Arabidopsis. *Science*, 306(5698), 1046-1048.

Gifford, M.L. et al., 2008. Cell-specific nitrogen responses mediate developmental plasticity. *Proceedings of the National Academy of Sciences USA*, 105(2), 803-808.

Girard, G. et al., 2003. Structural motifs in the RNA encoded by the early nodulation gene enod40 of soybean. *Nucleic acids research*, 31(17), 5003-5015.

Gläser, H.U. et al., 1993. Salt tolerance and methionine biosynthesis in Saccharomyces cerevisiae involve a putative phosphatase gene. *The EMBO journal*, 12(8), 3105-3110.

Gregory, B.D. et al., 2008. A link between RNA metabolism and silencing affecting Arabidopsis development. *Developmental cell*, 14(6), 854-866.

Grigg, S.P. et al., 2009. Repression of apical homeobox genes is required for embryonic root development in Arabidopsis. *Current Biology*, 19(17), 1485-1490.

Gultyaev, A.P. & Roussis, A., 2007. Identification of conserved secondary structures and expansion segments in enod40 RNAs reveals new enod40 homologues in plants. *Nucleic acids research*, 35(9), 3144-3152.

Guo, H.-shan et al., 2005. MicroRNA Directs mRNA Cleavage of the Transcription Factor NAC1 to Downregulate Auxin Signals for Arabidopsis Lateral Root Development. *Development*, 17, 1376-1386.

Gutierrez, L. et al., 2009. Phenotypic plasticity of adventitious rooting in Arabidopsis is controlled by complex regulation of AUXIN RESPONSE FACTOR transcripts and microRNA abundance. *The Plant Cell*, 21(10), 3119-3132.

Gy, I. et al., 2007. Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *The Plant Cell*, 19(11), 3451-3461.

Hanada, K. et al., 2007. A large number of novel coding small open reading frames in the intergenic regions of the Arabidopsis thaliana genome are transcribed and/or under purifying selection. *Genome research*, 17(5), 632-640.

Handwerger, K.E. & Gall, J.G., 2006. Subnuclear organelles: new insights into form and function. *Trends in cell biology*, 16(1), 19-26.

Haussecker, D. et al., 2010. Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* , 16(4), 673-695.

Henz, S.R. et al., 2007. Distinct expression patterns of natural antisense transcripts in Arabidopsis. *Plant physiology*, 144(3), 1247-1255.

Heo, J.B. & Sung, S., 2011. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* , 331(6013), 76-79.

Hirsch, J. et al., 2011. A novel fry1 allele reveals the existence of a mutant phenotype unrelated to 5'-3' exoribonuclease (XRN) activities in Arabidopsis thaliana roots. *PloS one*, 6(2), e16724.

Hirsch, J. et al., 2006. Characterization of 43 Non-Protein-Coding mRNA Genes in Arabidopsis, Including the *MIR162a*-Derived Transcripts. *Plant Physiology*, 140: 1192-1204.

Hofacker, I.L., Fekete, M. & Stadler, P.F., 2002. Secondary structure prediction for aligned RNA sequences. *Journal of molecular biology*, 319(5), 1059-1066.

Houseley, J. & Tollervey, D., 2009. The many pathways of RNA degradation. *Cell*, 136(4), 763-776.

Houseley, J., LaCava, J. & Tollervey, D., 2006. RNA-quality control by the exosome. *Nature reviews. Molecular cell biology*, 7(7), 529-539.

Hsieh, L.C. et al., 2009. Uncovering small RNA-mediated responses to phosphate deficiency in Arabidopsis by deep sequencing. *Plant physiology*, 151(4), 2120-2132.

Hugouvieux, V., Kwak, J.M. & Schroeder, J.I., 2001. An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in Arabidopsis. *Cell*, 106(4), 477-487.

Hugouvieux, V. et al., 2002. Localization, Ion Channel Regulation , and Genetic Interactions during Abscisic Acid Signaling of the Nuclear mRNA Cap-Binding Protein, ABH1 1. *Plant Physiology*, 130, 1276-1287.

Izaurralde, E. et al., 1994. A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell*, 78(4), 657-668.

Jamalkandi, S.A. & Masoudi-Nejad, A., 2009. Reconstruction of Arabidopsis thaliana fully integrated small RNA pathway. *Functional & integrative genomics*, 9(4), 419-432.

Jay, F., Renou, J.P., Voinnet O., Navarro L., 2010. Biotic stress-associated microRNAs: identification, detection, regulation, and functional analysis. *Methods Molecular Biology*, 592, 183-202.

Khan, G. a et al., 2011. MicroRNAs as regulators of root development and architecture. *Plant molecular biology*, 77(1-2), 47-58.

Kuhn, J.M., Breton, G. & Schroeder, J.I, 2007. mRNA metabolism of flowering-time regulators in wild-type Arabidopsis revealed by a nuclear cap binding protein mutant, abh1. *The Plant journal*, 50(6), 1049-1062.

Kurihara, Y. et al., 2009. Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in Arabidopsis. *Proceedings of the National Academy of Sciences USA*, 106(7), 2453-2458.

Kutschera, U. & Niklas, K.J., 2009. Evolutionary plant physiology: Charles Darwin's forgotten synthesis. *Die Naturwissenschaften*, 96(11), 1339-54.

Laubinger, S. et al., 2008. Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences USA*, 105(25), 8795-8800.

Lejeune, F. & Maquat, L.E., 2005. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Current opinion in cell biology*, 17(3), 309-315.

Lejeune, F., Li, X. & Maquat, L.E., 2003. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Molecular cell*, 12(3), 675-687.

Li, C.F. et al., 2006. An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. *Cell*, 126(1), 93-106. Li, H. et al., 2010. Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant physiology*, 153(4), 1759-1770.

Li, W.-X. et al., 2008. The Arabidopsis NFYA5 transcription factor is regulated transcriptionally and posttranscriptionally to promote drought resistance. *The Plant Cell*, 20(8), 2238-2251.

Liang, G., Yang, F. & Yu, D., 2010. MicroRNA395 mediates regulation of sulfate accumulation and allocation in Arabidopsis thaliana. *The Plant journal*, 62(6), 1046-1057. Lorković, Z.J., 2009. Role of plant RNA-binding proteins in development, stress response and genome organization. *Trends in plant science*, 14(4), 229-236.

Lorković, Z.J. & Barta, A., 2002. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. *Nucleic acids research*, 30(3), 623-635.

Lykke-Andersen, S., Brodersen, D.E. & Jensen, T.H., 2009. Origins and activities of the eukaryotic exosome. *Journal of cell science*, 122(Pt 10), 1487-1494.

Ma, L. et al., 2008. A novel RNA-binding protein associated with cell plate formation. *Plant physiology*, 148(1), 223-234.

Mancini-Dinardo, D. et al., 2006. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes & development*, 20(10), 1268-1282.

Maquat, L.E., 2004. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nature reviews. Molecular cell biology*, 5(2), 89-99.

Marillonnet, S. et al., 2005. Systemic Agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants. *Nature biotechnology*, 23(6), 718-723.

Marin, E. et al., 2010. miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *The Plant Cell*, 22(4), 1104-1117.

Marquardt, S., Hazelbaker, D.Z. & Buratowski, S., 2011. Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. *Transcription*, 2(3), 145-154.

Martienssen, R., 2010. Small RNA makes its move. *Science*, 328(5980), 834-835.

Mattick, J.S. & Makunin, I.V., 2006. Non-coding RNA. *Human molecular genetics*, 15 (1), 17-29.

Melnyk, C.W., Molnar, A. & Baulcombe, D.C., 2011. Intercellular and systemic movement of RNA silencing signals. *The EMBO Journal*, 30(17), 3553-3563.

Mercer, T.R. et al., Long non-coding RNAs: insights into functions. *Nature Reviews Genetics*, 10, 155-159.

Mitchell, P. et al., 1997. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-5' exoribonucleases. *Cell*, 91(4), 457-466.

Mitchell, P. & Tollervey, D., 2003. An NMD Pathway in Yeast Short Article Involving Accelerated Deadenylation and Exosome-Mediated 3 → 5 Degradation. *Yeast*, 11, 1405-1413.

Montgomery, T.A. et al., 2008. Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in TAS3 Trans-Acting siRNA Formation. *Cell*, 128-141.

Munns, R. et al., 2010. New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. *Journal of experimental botany*, 61(13), 3499-3507.

Nagano, T. et al., 2008. The air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science*, 322, 1717-1720.

Narsai, R. et al., 2007. Genome-wide analysis of mRNA decay rates and their determinants in Arabidopsis thaliana. *The Plant Cell*, 19(11), 3418-3436.

Naya, L. et al., 2010. Cleavage of a non-conserved target by a specific miR156 isoform in root apexes of Medicago truncatula. *Plant Signaling Behavior*, 5(3), 328-331.

Neil, H. & Jacquier, A., 2011. Enrichment of Unstable Non-coding RNAs and Their Genome-Wide Identification. *Methods in Molecular Biology*, 759, Part 2, 87-106.

Neuwald, A.F. et al., 1992. cysQ , a Gene Needed for Cysteine Synthesis in Escherichia coli K-12 Only during Aerobic Growth. *Journal of Bacteriology*, 174(2), 415-425.

Okushima, Y. et al., 2005. Functional Genomic Analysis of the AUXIN RESPONSE FACTOR Gene Family Members in Arabidopsis thaliana: Unique and Overlapping Functions of ARF7 and ARF19. *The Plant Cell*, 17, 444-463.

Oldroyd, G.E.D. & Downie, J.A., 2008. Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annual review of plant biology*, 59, 519-46.

Olmedo, G. et al., 2006. *ETHYLENE-INSENSITIVE5* encodes a 5'-3' exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. *Proceedings of the National Academy of Sciences USA*, 103(36), 13286-13293.

Osmont, K.S., Sibout, R. & Hardtke, C.S., 2007. Hidden branches: developments in root system architecture. *Annual review of plant biology*, 58, 93-113.

Pandey, R.R. et al., 2008. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Molecular cell*, 32(2), 232-246.

Papp, I. et al., 2004. A mutation in the Cap Binding Protein 20 gene confers drought tolerance to Arabidopsis. *Plant molecular biology*, 55(5), 679-686.

Pederson, T., 2010. Regulatory RNAs derived from transfer RNA? Regulatory RNAs derived from transfer RNA? *RNA*, 16, 1865-1869.

Peragine, A. et al., 2004. *SGS3* and *SGS2/SDE1/RDR6* are required for juvenile development and the production of *trans*-acting siRNAs in *Arabidopsis Genes & Development*, 18, 2368-2379.

Pontes, O. & Pikaard, C.S., 2008. siRNA and miRNA processing: new functions for Cajal bodies. *Current opinion in genetics & development*, 18(2), 197-203.

Prasanth, K.V. & Spector, D.L., 2007. Eukaryotic regulatory RNAs: an answer to the "genome complexity" conundrum. *Genes & development*, 21(1), 11-42.

Raguz, S. et al., 2010. Molecular Interplay of the Noncoding RNA ANRIL and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in Transcriptional Silencing of INK4a. *Molecular Cell*, 38, 662-674.

Redrup, L. et al., 2009. The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development (Cambridge, England)*, 136(4), 525-530.

Rinn, J.L. et al., 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, 129(7), 1311-1323.

Rohrig, H. et al., 2002. Soybean ENOD40 encodes two peptides that bind to sucrose synthase. *Proceedings of the National Academy of Sciences USA*, 99(4), 1915-1920.

Roman, G. et al., 1995. Genetic Analysis of Ethylene Signal Transduction in Arabidqsis thaliana: Five Novel Mutant Loci Integrated into a Stress Response Pathway. *Genetics*, 139, 1393-1409.

Rymarquis, L. a, Souret, Frederic F & Green, P.J., 2011. Evidence that XRN4, an Arabidopsis homolog of exoribonuclease XRN1, preferentially impacts transcripts with certain sequences or in particular functional categories. *RNA*, 17(3), 501-11.

Secco, D., Baumann, A. & Poirier, Y., 2010. Characterization of the rice PHO1 gene family reveals a key role for OsPHO1;2 in phosphate homeostasis and the evolution of a distinct clade in dicotyledons. *Plant physiology*, 152(3), 1693-1704.

Shaw, P.J. & Brown, J.W.S., 2004. Plant nuclear bodies. *Current opinion in plant biology*, 7(6), 614-20.

Sleutels, F., Zwart, R. & Barlow, D.P., 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*, 415, 810-813

Sonmez, C. et al., 2011. RNA 3' processing functions of Arabidopsis FCA and FPA limit intergenic transcription. *Proceedings of the National Academy of Sciences USA*, 108(20), 8508-8513.

Sorin et al., 2005. Auxin and Light Control of Adventitious Rooting in Arabidopsis Require ARGONAUTE1. *The Plant Cell*, 17, 1343-1359.

Souret, Frédéric F, Kastenmayer, J.P. & Green, P.J., 2004. AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. *Molecular cell*, 15(2), 173-183.

Sousa, Carolina 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. *Molecular and Cellular Biology*, 21(1), 354-366.

Sugimoto, K., Jiao, Y. & Meyerowitz, E.M., 2010. Article Arabidopsis Regeneration from Multiple Tissues Occurs via a Root Development Pathway. *Developmental Cell*, 18(3), 463-471.

Sunkar, R., 2010. MicroRNAs with macro-effects on plant stress responses. *Seminars in cell & developmental biology*, 21(8), 805-811.

Swiezewski, S. et al., 2009. Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature*, 462(7274), 799-802.

Terranova, R. et al., 2008. Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Developmental cell*, 15(5), 668-679.

Thompson, D.M. & Parker, R., 2009. Stressing out over tRNA cleavage. *Cell*, 138(2), 215-219.

Thompson, D.M. et al., 2008. tRNA cleavage is a conserved response to oxidative stress in eukaryotes tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA*, 2095-2103.

Tian, C. et al., 2004. Disruption and overexpression of auxin response factor 8 gene of Arabidopsis affect hypocotyl elongation and root growth habit, indicating its possible involvement in auxin homeostasis in light condition. , 3, 333-343.

Topping, J.F. et al., 1997. Mutations in the HYDRA1 gene of Arabidopsis perturb cell shape and disrupt embryonic and seedling morphogenesis. *Development*, 124(21), 4415-4424.

Valencia-Sanchez, M.A. et al., 2006. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & development*, 20(5), 515-524.

van Dijk, E.L. et al., 2011. XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature*, 475(7354), 114-117.

Vaucheret, Hervé, 2006. Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes & Development*, 759-771.

Vazquez, F. et al., 2004. Endogenous trans-Acting siRNAs Regulate the Accumulation of Arabidopsis mRNAs. *Molecular Cell*, 16, 69-79.

Verdel, A. et al., 2009. Common themes in siRNA-mediated epigenetic silencing pathways. *Intenational Journal of Development Biology*, 53, 245-257.

Vidal, E. a et al., 2010. Nitrate-responsive miR393/AFB3 regulatory module controls root system architecture in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences USA*, 107(9), 4477-4482.

Voinnet, O., 2009. Origin, biogenesis, and activity of plant microRNAs. *Cell*, 136(4), 669-687.

Wang, K.C. et al., 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, 472(7341), 120-124.

Wang, M.-B. & Metzlaff, M., 2005. RNA silencing and antiviral defense in plants. *Current opinion in plant biology*, 8(2), 216-222.

Williams, L. et al., 2005. A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the Arabidopsis ARF2, ARF3, and ARF4 genes. *Proceedings of the National Academy of Sciences USA*, 102(27), 9703-9708.

Wilusz, J.E., Sunwoo, H. & Spector, D.L., 2009. Long noncoding RNAs: functional surprises from the RNA world. *Genes & development*, 23(13), 1494-1504.

Xie, Z. et al., 2005. DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 102, 12985-12989.

Xiong, L. et al., 2001. FIERY1 encoding an inositol polyphosphate 1-phosphatase is a negative regulator of abscisic acid and stress signaling in Arabidopsis. *Genes & development*, 15(15), 1971-84.

Yamada, K. et al., 2003. Empirical analysis of transcriptional activity in the Arabidopsis genome. *Science* , 302(5646), 842-6.

Yang et al., 1993. Characterization of *GmENOD40* , a gene showing novel patterns of cell-specific expression during soybean nodule development. *The Plant Journal*, 3(4), 573-585.

Yang, J.H. et al., 2006. "Evidence of an auxin signal pathway, microRNA167-ARF8-GH3, and its response to exogenous auxin in cultured rice cells". *Nucleic acids research*, 34(6), 1892-1899.

Yang, Z. et al., 2011. Ends of Vaccinia Virus Early mRNAs Delineates Regulatory Sequences of Annotated and Anomalous Transcripts Genome-Wide Analysis of the 5' and 3' Ends of Vaccinia Virus Early mRNAs Delineates Regulatory Sequences of Annotated and Anomalous Transcripts. *Journal of Virology*, 85(12), 5896-5908.

Yap K.L., et al., 2010. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Molecular Cell*, 38(5), 662-674.

Yasuda, J. & Hayashizaki, Y., 2008. The RNA continent. *Advances in cancer research*, 99(07), 77-112.

Yoshikawa, M. et al., 2005. A pathway for the biogenesis of trans -acting siRNAs in Arabidopsis A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes & Development*, 2164-2175.

Yu, Q. et al., 2009. Saccharomyces cerevisiae linker histone Hho1p functionally interacts with core histone H4 and negatively regulates the establishment of transcriptionally silent chromatin. Journal of Biological Chemistry. 284, 740-750.

Zakrzewska-Placzek, M. et al., 2010. Arabidopsis thaliana XRN2 is required for primary cleavage in the pre-ribosomal RNA. *Nucleic acids research*, 38(13), 4487-4502.

Zhan, S. & Lukens, L., 2010. Identification of novel miRNAs and miRNA dependent developmental shifts of gene expression in Arabidopsis thaliana. *PloS one*, 5(4), p.e10157.

Zhang S., Sun L., & Kragler F., 2009. The Phloem-Delivered RNA Pool Contains Small Noncoding RNAs and Interferes with Translation. *Plant Physiology*; 150, 378-387.

Zhang, L. et al., 2011. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. *Cell Research*, 1-20.

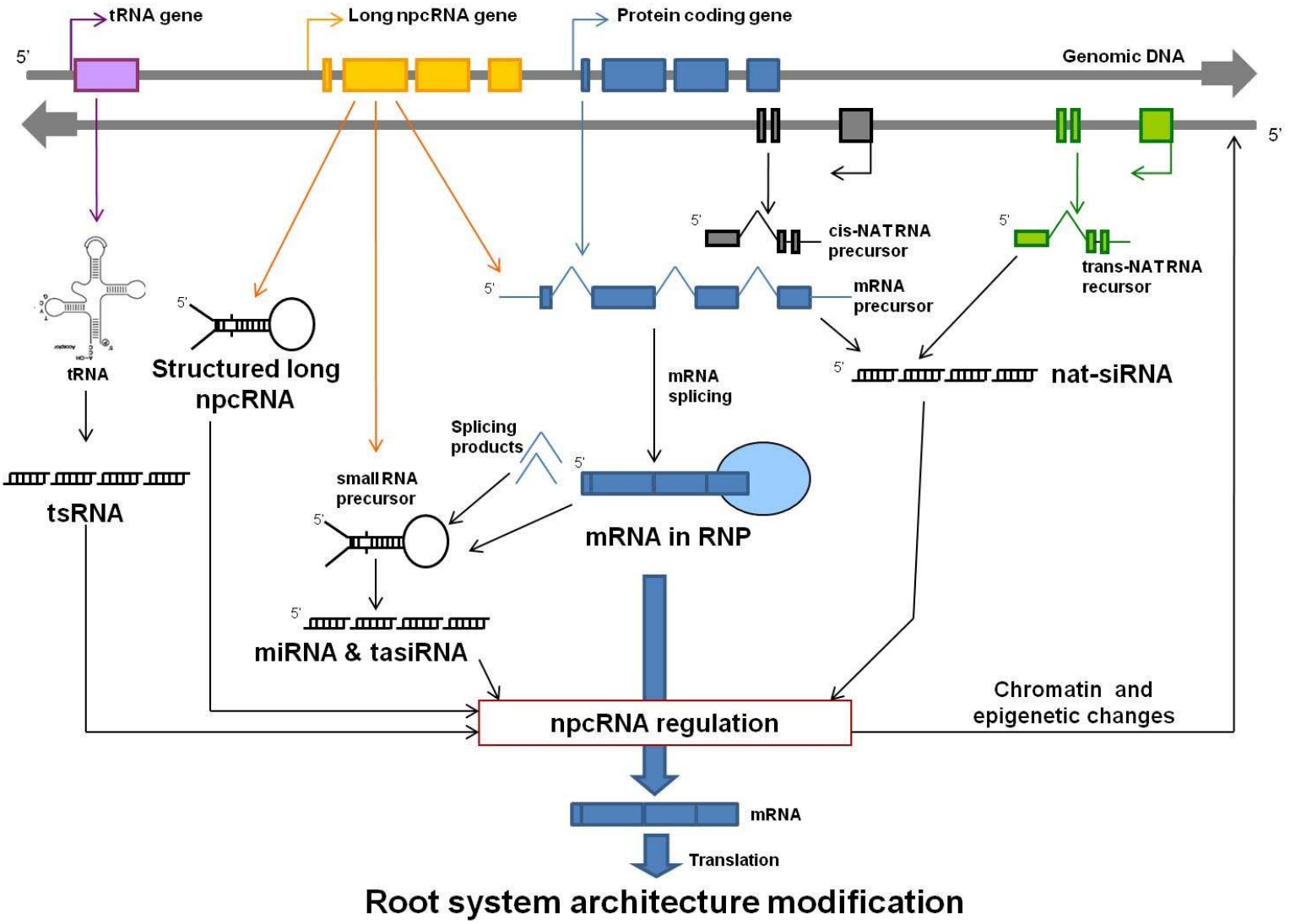
Tables

anism	Targets	Role	Reference
uncatula	HD-ZIP III	root vascular tissues diferenciation, lateral root and nodule formation	Boualem et al. 2008
aliana	PHB/PHV (HD-ZIP III)	embryonic root development	Grigg et al. 2009
aliana	ARF17	adventicious root formation	Gutierrez et al. 2009
aliana	ARF16	lateral root initiation	Wang et al. 2005
aliana	ARF10/ARF16	auxin signaling and root growth and gravitropism	Wang et al. 2005
aliana	GRF	root growth	Rodriguez et al. 2010
ativa	SPL	global growth	Jiao et al. 2010
miR156	*M.truncatula*	WD40	root growth	Naya et al. 2010
miR164	*A. thaliana*	NAC1	lateral root initiation and development	Guo et al. 2005
miR390	*O. sativa*	TAS3/ARF2/ARF3/ARF4	potential role in auxin signaling and root development	Williams et al. 200
miR390	*A. thaliana*	TAS3/ARF2/ARF3/ARF4	auxin signaling and lateral root growth	Marin et al. 2010
miR167	*O. sativa*	ARF6/ARF8	auxin signaling and adventicious root formation	Yang et al. 2006
miR167	*A. thaliana*	ARF6/ARF8	auxin signaling and adventicious root formation	Gutierrez et al. 20
miR169	*M.truncatula*	Mt-HAP2-1	nodule differentiation	Combier et al. 200
miR482	*G. max*	R gene receptor kinases	nodule formation	Li et al. 2010
miR1511	*G. max*	different	nodule formation	Li et al. 2010
miR1512	*G. max*	different	nodule formation	Li et al. 2010
miR1515	*G. max*	DCL	nodule formation	Li et al. 2010
miR167a	*A. thaliana*	ARF8	root response to N and lateral root emergence	Gifford et al. 2008
miR393	*A. thaliana*	AFB1/AFB2/AFB3/TIR1	root response to N, auxin and potential role in lateral root formation	Vidal et al. 2010
miR399	*O. sativa*	PHO2	phosphate signaling, potential role in lateral root formation	Bari et al. 2006
miR399	*A. thaliana*	PHO2	phosphate signaling and lateral root formation	Franco-Zorrilla et a
miR395	*A. thaliana*	SULTR2;1/APS1/APS3/APS4	sulphate metabolism and potential role in laterale root formation	Liang et al. 2010
miR169	*A. thaliana*	NF-YA5	drought stress, potential role in root development	Li et al. 2008

Table 1. List of several miRNAs involved in root development, stress response and hormone signaling of various plant species.

Figure Legends

Fig. 1. The non-protein-coding RNA network in plants. Long npcRNAs can be precursors of small RNAs or cis/trans-antisense RNAs (NATs) of other transcripts. Once incorporated into the RISC effector complex, small RNAs from different pathways of RNA silencing (miRNA, siRNA or nat-siRNA) can act directly on the target mRNA or other npcRNAs. These regulatory pathways can lead to epigenetic changes and DNA methylation (TGS). Hence, long and short npcRNAs form a network of ribonucleoproteins within the cell that may affect the expression patterns of coding mRNAs, affecting root development.


tRNA gene
Long npcRNA gene
Protein coding gene
5'
Genomic DNA
5'
cis-NAT RNA precursor
trans-NAT RNA recursor
mRNA precursor
tRNA
Structured long npcRNA
mRNA splicing
nat-siRNA
Splicing products
small RNA precursor
mRNA in RNP
tsRNA
miRNA & tasiRNA
Chromatin and epigenetic changes
npcRNA regulation
mRNA
Translation
Root system architecture modification

Annexe 2:

Ana Beatriz Moreno, Angel Emilio Martínez de Alba, Florian Bardou, Martin D. Crespi, Hervé Vaucheret, Alexis Maizel, Allison C. Mallory (2013), Cytoplasmic and nuclear quality control and turnover of single stranded RNA modulate post-transcriptional gene silencing in plants

Published online 12 March 2013

Nucleic Acids Research, 2013, Vol. 41, No. 8 ***4699–4708***
doi:10.1093/nar/gkt152

Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-transcriptional gene silencing in plants

Ana Beatriz Moreno[1], Angel Emilio Martínez de Alba[2], Florian Bardou[1], Martin D. Crespi[1], Hervé Vaucheret[2], Alexis Maizel[1,3,*] and Allison C. Mallory[2,*]

[1]Institut des Sciences du Végétal, CNRS UPR 2355, SPS Saclay Plant Sciences, 91198 Gif-sur-Yvette, France, [2]Institut Jean-Pierre Bourgin, UMR1318 INRA, SPS Saclay Plant Sciences, 78026 Versailles, France and [3]Center for Organismal Studies, University of Heidelberg, Im Neuenheimer Feld 230, 69120 Heidelberg, Germany

Received January 6, 2013; Revised February 13, 2013; Accepted February 15, 2013

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

ABSTRACT

Eukaryotic RNA quality control (RQC) uses both endonucleolytic and exonucleolytic degradation to eliminate dysfunctional RNAs. In addition, endogenous and exogenous RNAs are degraded through post-transcriptional gene silencing (PTGS), which is triggered by the production of double-stranded (ds)RNAs and proceeds through short-interfering (si)RNA-directed ARGONAUTE-mediated endonucleolytic cleavage. Compromising cytoplasmic or nuclear 5′–3′ exoribonuclease function enhances sense-transgene (S)-PTGS in *Arabidopsis*, suggesting that these pathways compete for similar RNA substrates. Here, we show that impairing nonsense-mediated decay, deadenylation or exosome activity enhanced S-PTGS, which requires host RNA-dependent RNA polymerase 6 (RDR6/SGS2/SDE1) and SUPPRESSOR OF GENE SILENCING 3 (SGS3) for the transformation of single-stranded RNA into dsRNA to trigger PTGS. However, these RQC mutations had no effect on inverted-repeat–PTGS, which directly produces hairpin dsRNA through transcription. Moreover, we show that these RQC factors are nuclear and cytoplasmic and are found in two RNA degradation foci in the cytoplasm: siRNA-bodies and processing-bodies. We propose a model of single-stranded RNA tug-of-war between RQC and S-PTGS that ensures the correct partitioning of RNA substrates among these RNA degradation pathways.

INTRODUCTION

Eukaryotic gene expression produces large amounts of both protein-coding and non-coding RNA species. To ensure proper cellular function and viability, a high level of fidelity must be sustained. To tackle this challenge, RNA surveillance and decay serve three main purposes: first, to ensure RNA quality control (RQC) mechanisms that scrutinize RNA integrity and eliminate defective messenger RNA (mRNA), thus dampening the production of potentially toxic proteins, second, to regulate mRNA turnover to control protein abundance and third, to detect invading RNAs, to defend the cell against them (1–4) and to regulate selected endogenous mRNAs through an endonucleolytic cleavage process called post-transcriptional gene silencing (PTGS) (5–8). How RQC and PTGS pathways interact and the processes that regulate the partitioning of RNA substrates into these pathways are not well understood.

Nonsense-mediated decay (NMD) is an extensively studied RQC pathway involved in the genome-wide suppression of transcripts (9–11) in which translation is arrested either owing to the presence of a premature termination codon or owing to excessive 3′untranslated region (UTR) length (12–16). Although there are several different mechanisms by which NMD can be triggered, once instigated, NMD generally involves the recruitment and activation of conserved UPFRAMESHIFT 1 (UPF1),

*To whom correspondence should be addressed. Tel: +33 1 30 83 31 82; Fax: +33 1 30 83 3319; Email: allison.mallory@versailles.inra.fr
Correspondence may also be addressed to Alexis Maizel. Tel: +49 6221 546 456; Fax: +49 6221 546 424; Email: alexis.maizel@cos.uni-heidelberg.de

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2013. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

UPF2 and UPF3 proteins to defective transcripts that are translationally stalled. However, the presence of an exon junction complex (EJC) is not always required to evoke NMD because it can target intronless transcripts in yeast, mammals, flies and plants (17–21). This recruitment, either by invoking decapping and deadenylation pathways or via endonucleolytic cleavage, as is the case in *Drosophila* and humans, generates aberrant RNAs [RNAs lacking a 5′-cap structure or a 3′-poly(A) tail] that are subsequently degraded through exonucleolytic cleavage [for reviews see (2,22,23)].

Exonucleolytic RNA degradation in *Arabidopsis* exploits a suite of processes including, but not limited to, the shortening of the 3′-poly(A) tail (deadenylation), which is catalysed by the conserved 3′–5′ POLY(A)-SPECIFIC RIBONUCLEASE (PARN) as well as by the conserved CARBON CATABOLITE REPRESSOR 4 (CCR4) complex (24–27). It also involves the removal of the 5′-cap structure, which is accomplished by a set of conserved decapping proteins: DCP1, DCP2 (TDT), DCP5, VARICOSE (VCS) and possibly DEA(D/H)-box RNA HELICASE 1 (DHH1) (28–30). Decapping and deadenylation are a prerequisite for most RNA to be degraded by 5′–3′ XRN exoribonucleases and the multimeric 3′–5′ exoribonuclease exosome complex. *Arabidopsis* expresses three XRN proteins, the nuclear XRN2 and XRN3 and the cytoplasmic XRN4 (31). Biochemical and molecular characterization of the *Arabidopsis* exosome core complex revealed the subunits RRP4, RRP40, RRP41, RRP42, RRP43, RRP45 (CER7), RRP46, CSL4 and MTR3 (32). Additional components likely involved in exosome function include RRP44, RRP6L1, RRP6L2, RRP6L3 and MTR4 (32–35).

In addition to these RNA degradation mechanisms, plants and other eukaryotes use PTGS to defend against foreign invading RNAs, such as viruses and high levels of transgenic mRNAs (36–40). PTGS also is required to modulate the abundance or expression of cellular mRNAs important during developmental transitions, such as the mRNAs targets of the trans-acting small interfering (ta-si)RNA pathway (41,42). Double-stranded (ds)RNA is the priming trigger of PTGS and is generated though several processes such as viral replication, sense-antisense transcription or transcription of inverted-repeat (IR) sequences, whose transcripts are self-complementary and thus fold-back on themselves to form dsRNA. It can also be produced by the cellular RNA-DEPENDENT RNA POLYMERASE 6 (RDR6/SGS2/SDE1), which is coupled to the RNA stabilizing protein SUPPRESSOR OF GENE SILENCING 3 (SGS3). Once the dsRNA is produced, it is processed by DICER-LIKE (DCL) enzymes into 21–22-nt siRNAs, which serve as sequence-specific guides for ARGONAUTE 1 (AGO1)-dependent endonucleolytic cleavage of complementary transcripts (6,43,44). AGO1-mediated cleavage generates RNAs that are, in most cases, subjected to XRN- and exosome-mediated degradation (45). In the case of viruses, once PTGS is instigated, amplification of the siRNAs ensures that tissues are primed against subsequent infection by the same virus or expression of a transgene bearing virus sequences (46,47).

Previous data suggested that defects in RNA processing and degradation that lead to the accumulation of decapped and deadenylated RNA, including mutations in RNA splicing, 3′-end formation and 5′–3′ exoribonuclease XRN-mediated degradation, promote PTGS (48–50). Moreover, removing transgene 3′-terminator sequences enhanced PTGS, while having multiple terminators reduced PTGS (51). Here, we explore the ways in which an array of nuclear and cytoplasmic RQC factors and PTGS interact mechanistically and spatially in plants. Impairing either nuclear or cytoplasmic NMD UPF1 and UPF3, deadenylation PARN and CCR4a and exosome RRP4, RRP6L1, RRP41 and RRP44A components enhanced sense (S)-PTGS but had no effect on an IR-PTGS system. In the cytoplasm, RQC factors localized in siRNA-body and processing (P)-body RNA degradation foci. These findings show that nuclear and cytoplasmic aberrant RNAs are instrumental during this type of RNA silencing process, as opposed to IR-PTGS, which produces dsRNA, a direct template for the DCLs. The correct partitioning of aberrant RNA substrates among these RNA degradation mechanisms ensures the discrimination of dysfunctional self-RNA and invading non–self-RNA from functional self-RNA and acts as a barrier to prevent the undesired triggering of PTGS of self-RNA.

MATERIALS AND METHODS

Plant material

All *Arabidopsis thaliana* are in the Columbia accession (52). The *JAP3* line was the kind gift of D. Baulcombe and the inducible RNA interference (iRNAi) lines *rrp41*^{iRNAi} and *rrp4*^{iRNAi} (32) were the kind gift of J. Ecker. The parn [fast neutron mutant *ahg2-1*; (53)] was kindly provided by T. Hirayama. The *upf1-5* (*SALK_112922*, insertion located in the 3′UTR) was obtained from NASC. Homozygous *ccr4a* (*SAIL_784_A07*, insertion located in intron 9/10), *ccr4b* (*SAIL_635_B07*, insertion located in exon 2/11), *upf1-6* (*SAIL_1295_E07*, insertion located 148 bp upstream of the ATG), *upf3-3* (*SAIL_122_G02*, insertion located 183 bp upstream of the ATG), *upf3-1* (*SALK_025175*, insertion located in exon 5/12) and *rrp6L1* (*rrp6A*; *SAIL_1306_C10* insertion located in intron 12/13) mutants were generated during this study (see Supplementary Figure S1 for molecular characterization). Seeds were obtained from NASC.

Generation of artificial miRNA lines

The artificial miRNA *amiR-RRP44Aa* (5′-UAUGAGUA UACAGGCGUGCUG-3′) was generated using the WMD3 microRNA designer (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi) and expressed under the ubiquitin promoter in the context of the *MIR319a* backbone. PTGS reporter lines were transformed using the floral dip methods (54) and transformed plants were selected on 15 μg/ml of glufosinate. PTGS was analysed in the

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

progeny of 3 T2 lines harbouring a single *UB::amiR-RRP44a* insertion.

RNA extraction and RNA gel blot analysis

For RNA gel blot analyses, frozen tissue was homogenized in a buffer containing 0.1 M NaCl, 2% sodium dodecyl sulphate (SDS), 50 mM Tris–HCl (pH 9.0), 10 mM ethylenediaminetetraacetic acid (pH 8.0) and 20 mM β-mercaptoethanol, and RNAs were extracted two times with phenol and recovered by ethanol precipitation. To obtain high molecular weight (HMW) RNA, total RNA was precipitated overnight in 2 M LiCl at 4°C and recovered by centrifugation. For low molecular weight (LMW) RNA analysis, total RNA was separated on a 15% denaturing polyacrylamide gel electrophoresis gel, stained with ethidium bromide and transferred to nylon membrane (HybondNX, Amersham). LMW RNA and *U6* hybridizations were at 50°C with hybridization buffer containing 5× saline-sodium citrate (SSC), 20 mM Na_2HPO_4, pH 7.2, 7% SDS, 2× Denhardt's solution and denatured sheared salmon sperm DNA (Invitrogen). HMW RNA hybridization was at 37°C in SigmaPerfectHyb buffer (Sigma). Blots were hybridized with a radioactively labelled random-primed DNA probes for beta-glucuronidase (*GUS*) mRNA and *GUS* siRNAs, and end-labelled oligonucleotide probes for *TAS1* ta-sRNA, *TAS2* tasRNA and *U6* detection.

GUS activity quantification

With the exception of *amiR-RRP44A*, $rrp41^{iRNAi}$ and $rrp4^{iRNAi}$ lines, plants were grown on Bouturage 2 medium (Duchefa Biochemie) in standard long-day conditions (16 h light, 8 h dark at 20–22°C), transferred to soil after 2 weeks and grown in controlled growth chambers in standard long-day conditions. To induce expression of the RNAi lines, $rrp41^{iRNAi}$ and $rrp4^{iRNAi}$ plants were grown on Bouturage media containing 8 μM estradiol for 12 days in standard long-day conditions, and then transferred to soil and grown in controlled growth chambers in standard long-day conditions. Total protein was extracted from cauline leaves of flowering plants and GUS activity was quantified as in (49) by measuring (Fluoroscan II; Thermo Scientific) the quantity of 4-methylumbelliferone produced from the substrate 4-methylumbelliferyl-b-D-glucuronide (Duchefa Biochemie).

Semi-quantitative reverse transcriptase-polymerase chain reaction

RNAs were extracted using the RNeasy plant mini kit (Qiagen), and 1 μg of RNA was reverse transcribed using oligo dT and Super ScriptII reverse transcriptase (Invitrogen). Twenty-seven cycles of polymerase chain reaction (PCR) were used to amplify *RRP44A*, *CCR4a*, *CCR4b* and *EF1-alpha*, and 28 cycles of PCR were used to amplify *UPF1* and *UPF3* to non-saturation. The number of cycles used to amplify *RRP4* and *RRP6L1* to non-saturation is indicated above each lane in Supplementary Figure S1. *EF1-alpha* amplification was used as a control.

Nicotiana benthamiana agro-infiltration

Agrobacterium (ASE or Agl0 strains) carrying plasmids of interest were grown overnight at 30°C in 3 ml Lysogeny Broth (LB) medium containing the appropriate antibiotics to a final OD600 of between 1.0 and 2.0. The bacteria were pelleted and resuspended in 1 ml of infiltration medium (10 mM $MgCl_2$, 10 mM 2-(N-morpholino)ethanesulfonic acid (MES), pH 5.2, 150 mM acetosyringone) to a final OD600 of 0.1. The bacterial solution containing the plasmid(s) of interest was coinfiltrated with a bacterial solution expressing HELPER COMPONENT-PROTEINASE (HC-Pro), a viral suppressor of silencing, into the abaxial side of leaves using a 1 ml syringe, and samples were assayed 3 days after infiltration. HC-Pro was used to better visualize the fluorescent signals and did not have an observable impact on the localization pattern of the tested RQC and PTGS components.

Confocal imaging

For confocal imaging, agro-infiltrated tobacco leaves (mounted in water) were directly imaged on a Leica Confocal TCS SP2 (Leica Microsystems). The CFP was imaged with 458 nm excitation using the dichroic mirror DD458/514 and detection window of 465–505 nm; the GFP was imaged with 488 nm excitation using the dichroic mirror DD488/543 and a detection window of 500–580 nm; the RFP was imaged with 543 nm excitation using the dichroic mirror DD488/543 and a detection window of 580–670 nm. For the co-localizations, all of the images were taken by sequential acquisition. Image analysis was performed using the National Institute of Health ImageJ (http://rsb.info.nih.gov/ij/) software.

Cloning procedures

All the clones were made using the Gateway technology (Invitrogen) and planned using Geneious (http://www.geneious.com). A list of the oligonucleotides used for cloning is provided in Supplementary Table S2. *UPF3* (*AT1G33980*), *SGS3* (*AT5G23570*) and *RRP41* (*AT3G61620*) were PCR amplified from complementary DNA (cDNA) and cloned into the vector *pDONR221* to generate entry clones, whereas *PARN* (*AT1G55870*), *CCR4a* (*AT3G58560*) and *RRP4* (*AT1G03360*) were PCR amplified from genomic DNA and cloned into the vector *pENTR-D* to generate entry clones. To obtain the GFP fusions under the control of the *35S* promoter, *SGS3*, *CCR4a*, *PARN* and *RRP41* entry clones were recombined in the expression vector *pH7WGF2*, whereas the *UPF3* entry clone was recombined in the expression vector *pH7FWG2*. To obtain the RFP fusion proteins under the control of the *35S* promoter, the entry clone containing *PARN* was recombined in the expression vector *pB7WGR2*, and the one containing *RRP4* in the expression vector *pB7RWG2*. For the GFP fusion under the control of the *Ubiquitin10* promoter, the *RRP4* entry clone was recombined in the expression vector *pUBN-GFP* (55). The *35S:RFP:DCP1* and *35S:CFP:DCP1* constructs were made by recombination of an entry clone containing *DCP1* (*AT1G08370*, gift from C. Antonelli) in the

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

expression vectors *pB7WGR2* and *pB7WGC2*, respectively (56). The *35S:RDR6:GFP* construct was made by PCR amplifying *RDR6* (*AT3G49500*) from cDNA and adding the restriction sites SalI and NotI to each terminus to generate a Gateway entry clone in the plasmid *pENTR1A* that was then recombined in the expression vector *pH7FWG2*. The construct *pGFP-N-Bin:UPF1* (*AT5g47010*) was generously provided by A. Pendle and J. Brown. The construct *35S:RFP:UPF1* was obtained by recombining the entry clone *UPF1* cDNA *pDONR207* (kindly provided by A. Pendle and J. Brown) into the expression vector *pB7WGR2*. The *35S::HC-Pro* plasmid was the kind gift of J. Carrington.

RESULTS

To investigate the possible crosstalk between PTGS and other RNA degradation pathways, we isolated loss-of-function *Arabidopsis* mutants in many key components of RQC and RNA turnover pathways and characterized their impact on S-PTGS. In the cases where loss-of-function caused lethality, we examined the impact of partial-loss-of-function mutants when possible. The effect of RQC and RNA turnover mutants on S-PTGS was determined using the well-characterized *Arabidopsis* reporter lines *Hc1* and *6b4*. Both lines carry a *35S::GUS* transgene, but whereas 6b4 stably produces GUS, silencing of the *GUS* transgene is spontaneously triggered in 20% of *Hc1* plants at each generation add (57,58). These reporter systems allowed us to reveal both positive and negative effects of the RQC mutations on S-PTGS. To avoid the *35S* interference phenomenon reported to occur when introducing the *35S::GUS* transgene carried by the *6b4* and *Hc1* into mutants already carrying a *35S* T-DNA insertion (59), we analysed S-PTGS uniquely in mutants containing either *35S*-free T-DNA insertions or fast neutron-generated mutations.

Mutations in NMD, deadenylation and exosome factors enhance S-PTGS

To examine the contribution of NMD to PTGS, we searched publicly available mutant collections and identified *upf1* (*SAIL_1295_E07*, hereafter referred to as *upf1-6*), and *upf3* (*SAIL_122_G02*, hereafter referred to as *upf3-3*) partial-loss-of-function mutants (Supplementary Figure S1A), and these mutants were crossed with *Hc1* and *6b4* lines. Quantitative GUS assays performed on the progeny of plants homozygous for both the *Hc1* locus and either the *upf1* or *upf3* mutation indicated that *Hc1* silencing was enhanced from 20% in line *Hc1* to 44% in *Hc1*/*upf1-6* and to 78% in *Hc1*/*upf3-3* (Figure 1A). To determine the strength of the silencing enhancement, we also analysed the effect of these mutations on line *6b4*. The *upf3-3* mutation triggered silencing in 13% of the *6b4* plants analysed (Figure 1B), whereas *upf1-6* did not appear to have an effect on *6b4* silencing (0/32 plants were silenced at the *6b4* locus). Characteristic of PTGS, *GUS* siRNAs accumulated and *GUS* mRNA levels were reduced to nearly undetectable levels in the silenced *Hc1*/*upf1-6*, *Hc1*/*upf3-3* and *6b4*/*upf3-3* lines (Figure 1C and D), indicating that both UPF1 and UPF3 are endogenous PTGS suppressors.

Arabidopsis PARN has poly(A) RNA degradation activity and complete loss-of-function parn mutants are lethal, indicating that it is an essential ribonuclease (13). Nevertheless, a fast neutron-generated partial-loss-of-function alternative splicing parn mutant ahg2-1 has been described (60). Quantitative GUS assays on plants homozygous for both the *Hc1* transgene and the *ahg2-1* (*parn*) mutation indicated that silencing of *Hc1* was increased from 20% to nearly 50% (Figure 1A). In addition, *ahg2-1* triggered silencing in nearly 40% of *6b4* plants (Figure 1B), indicating that PARN is a suppressor of PTGS. Like the *parn* mutant that negatively impacts deadenylation, a mutation in the putative deadenylation factor CCR4a enhanced *Hc1* silencing from 20% to nearly 60% (Figure 1A); however, unlike the *parn* mutant, the *ccr4a* mutation did not trigger silencing of line *6b4* (0/30 of *6b4*/*ccr4a* plants were silenced). Silencing triggered by both *parn* and *ccr4a* deadenylation mutants led to the accumulation of *GUS* siRNAs and a reduction in *GUS* mRNA levels (Figure 1C and D). In contrast to *ccr4a*, a mutation in the related *CCR4b* gene, which is located adjacent to the *CCR4a* gene, did not impact *Hc1* or *6b4* silencing (18%, 10/56 *Hc1*/*ccr4b* plants and 0%, 0/39 *6b4*/*ccr4b* plants were silenced), suggesting that CCR4b could be partially redundant with CCR4a. Both *ccr4a* and *ccr4b* mutants appeared to be full-loss-of-function mutants, as they did not produce detectable *CCR4a* and *CCR4b* transcripts, respectively (Supplementary Figure S1B), but additional work will be required to determine whether these proteins are partially redundant.

The multimeric exosome complex contains 3′–5′ exoribonucleases that degrade RNA with unprotected 3′-ends. To determine if perturbations in exosome function could influence PTGS, we characterized the impact on S-PTGS in the *Hc1* and *6b4* reporter lines of mutants defective in the *Arabidopsis* core exosome subunits RRP4 and RRP41, the latter of which exhibits catalytic 3′–5′ RNAse activity, unlike the yeast and human RRP41 (61). We also examined the impact on S-PTGS of impairing RRP44A, the *Arabidopsis* homolog of the RRP44 (DIS3) 3′–5′ RNAse responsible for nearly all of the catalytic activity of the yeast exosome (62,63). Finally, we examined the impact on S-PTGS of a mutation in RRP6L1 [also known as RRP6A; Supplementary Figure S1C (32,64)], one of two *Arabidopsis* homologs of the yeast and human RRP6 exoribonuclease. Although the nuclease function of *Arabidopsis* RRP6L1 has not been confirmed, expression of *Arabidopsis RRP6L1* complements the growth defects of a yeast *rrp6* mutant strain (64). Because *rrp4* and *rrp41* mutants are seedling lethal, we analysed PTGS in the previously characterized *rrp4* and *rrp41* iRNAi lines, which silence *RRP4* and *RRP41* after estradiol treatment owing to the induced expression of an IR transgene targeting *RRP4* and *RRP41*, respectively [Supplementary Figure S1D and (32)]. Furthermore, because *35S*-free loss-of-function mutants in *rrp44A* were not available, we generated *Arabidopsis* lines expressing an artificial miRNA (65) (*amiR-RRP44A*) under the ubiquitin promoter, and analysed PTGS in line *Hc1*.

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

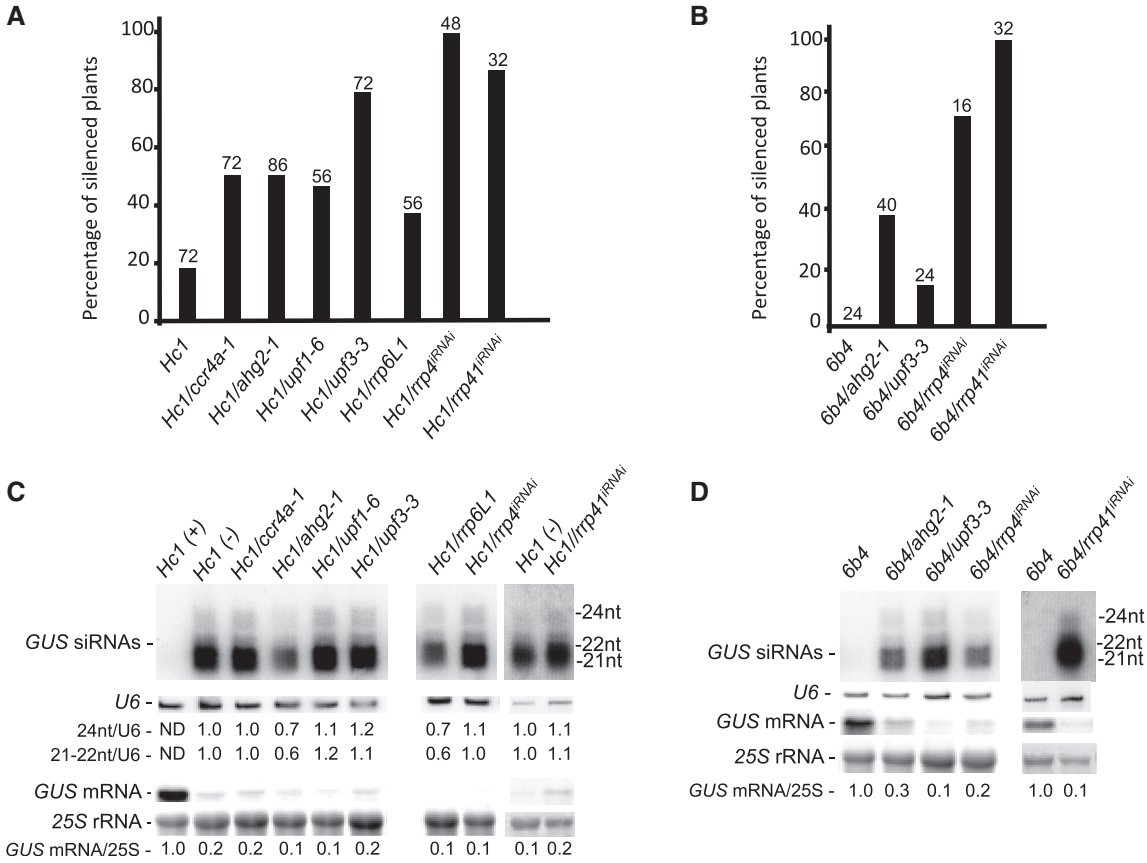


Figure 1. NMD, deadenylation and exosome mutants enhance transgene S-PTGS. (**A** and **B**) The percentage of silenced plants determined by GUS quantitative protein assays in the indicated mutant and control lines. The number of plants analysed is indicated above each bar. (**C** and **D**) RNA gel blot analyses of the indicated mutant and control lines. High molecular weight RNA and siRNA gel blots were hybridized with a *GUS* DNA probe. *25S* ribosomal RNA (rRNA) and *U6* small nucleolar RNA (snRNA) served as loading controls, respectively. *Hc1* plants that were expressing (+) and silenced (−) for *GUS* were analysed. The position of *GUS* 24, 22 and 21 nt siRNAs is noted. Normalized values of *GUS* mRNA to *25S* rRNA (with either *Hc1* (+) or *6b4* levels set at 1.0) and *GUS* 24 nt and *GUS* 21–22 nt siRNA to *U6* snRNA [with *Hc1* (−) levels set at 1.0] are indicated. ND = non-detectable.

Quantitative GUS assays indicated that loss-of-function of *rrp6L1* and down-regulation of *rrp4*^{iRNAi} and *rrp41*^{iRNAi} enhanced PTGS in line *Hc1* from the 20% baseline to 30, 90 and 80%, respectively (Figure 1A). Furthermore, analysis of S-PTGS in *Hc1*/*amiR-RRP44A* plants revealed that line 46, which accumulated more *amiR-RRP44A* (Figure 2A) and less *RRP44A* mRNA (Figure 2B) than lines 43 and 53, triggered PTGS in 100% of *Hc1* plants analysed (Figure 2C) and accumulated *GUS* siRNAs (Figure 2B). Moreover, the *rrp4*^{iRNAi} and *rrp41*^{iRNAi} lines triggered PTGS in nearly 70 and 100% of *6b4* plants, respectively (Figure 1B), whereas *rrp6L1* mutants did not trigger silencing of *6b4* (*GUS* silencing was not observed in 64 *6b4*/*rrp6L1* plants). The effect of the expression of the artificial *amiR-RRP44A* on *6b4* PTGS was not tested. Collectively these data indicate that mutations in a variety of components involved in RQC and exonucleolytic RNA turnover have the capacity to enhance S-PTGS. As all these pathways act on single-stranded (ss)RNA, these results suggest that modulation of ssRNA abundance is a key element controlling entry into PTGS.

To broaden our S-PTGS analysis to an endogenous silencing system that, like S-PTGS, requires RDR6 and SGS3 for dsRNA production, we examined the effect of RQC mutants on the ta-siRNA pathway (66–69). We did not observe any changes in tasiRNA levels arising from the *TAS1* or *TAS2* locus in any of our RQC mutants (Supplementary Figure S2).

Mutations in NMD, deadenylation and exosome components do not impact IR-PTGS

Next, we examined the impact of mutations in these NMD, deadenylation and exosome components on a PTGS system that produces a stem-loop dsRNA directly through transcription and, thus, does not rely on the RDR6- and SGS3-dependent conversion of ssRNA to dsRNA to become a substrate of DCL proteins. The line *JAP3* expresses a *PHYTOENE DESATURASE (PDS)* inverted repeat under the control of the phloem-specific *Suc2* promoter (70) and initiates from the veins PDS silencing, which can be easily traced owing to the photobleaching phenotype.

Mutations in NMD, deadenylation and the core exosome complex did not appear to impact the initiation or spreading of IR-PTGS in the line *JAP3* (Figure 3). It was shown previously that UPF1 influenced RNAi

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

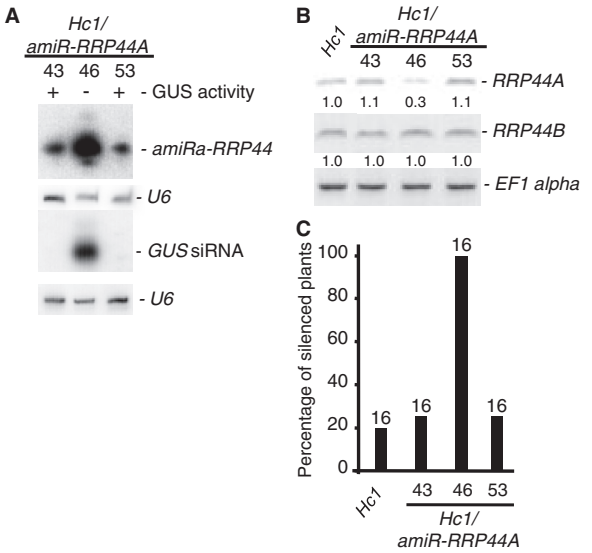


Figure 2. Expression of an artificial miRNA targeting *RRP44A* leads to enhanced S-PTGS. (**A**) RNA gel blot analyses of three different *Hc1* plant lines expressing the artificial *RRP44A* miRNA *amiR-RRP44A*. Small RNA gel blots were hybridized with a *GUS* DNA probe or an oligonucleotide antisense to the amiR. *U6* served as a loading control for small RNA. (**B**) Reverse transcriptase-PCR of *RRP44A* and *RRP44B* transcripts in the corresponding *Hc1*/*amiR*-RRP44A and control *Hc1* seedlings. *EF1alpha* was used as an amplification control. Normalized values of *RRP44A* and *RRP44B* mRNA to *EF1 alpha* mRNA (with *Hc1* levels set at 1.0) are indicated. (**C**) The percentage of silenced plants determined by GUS quantitative protein assays in the indicated mutant and control lines. The number of plants analysed is indicated above each bar.

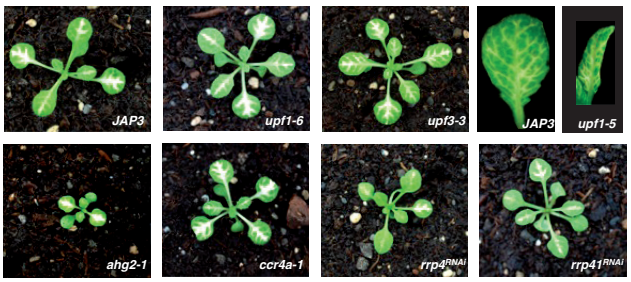


Figure 3. NMD, deadenylation and exosome mutants do not impact *JAP3* IR-PTGS. Eighteen-day-old control *JAP3* plants and *JAP3* plants containing the indicated mutations. The photo is representative of a minimum of 20 plants screened for each genotype.

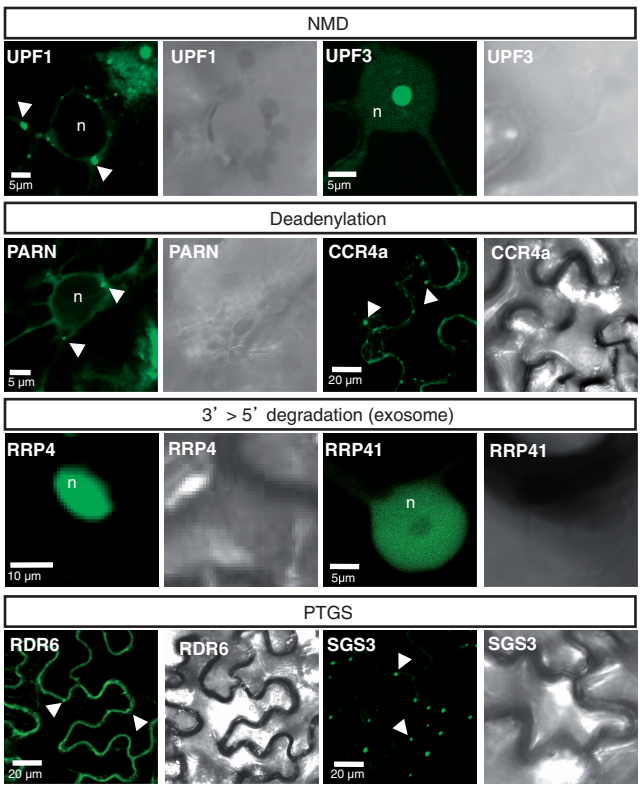


Figure 4. Subcellular localization of NMD, deadenylation, exosome and PTGS components. Confocal sections and their corresponding bright-field images of *N. benthamiana* leaves expressing the indicated proteins fused to GFP. The arrowheads indicate cytoplasmic foci, whereas 'n' labels the nucleus. Scale bars are shown on the images.

persistence in *Caenorhabditis elegans* and IR-PTGS in *Arabidopsis*, but UPF1 did not appear to affect RNAi in Drosophila (71–73). The analysis in *Arabidopsis* examined the effect of the *upf1-5* mutant, a SALK T-DNA insertion line containing a *35S* promoter, on an IR of the endogenous *APETALA 3* (*AP3*) gene that was expressed under the *35S* promoter (71). Given the report of *35S* interference on PTGS observed when combining two transgenes each containing the *35S* promoter (59), we re-examined the effect of the *upf1-5* mutant on IR-PTGS using the *JAP3* *35S*-free IR-PTGS system. Similar to what we observed for the *upf1-6* mutant, the *upf1-5* mutant did not appear to impact the initiation or spreading of *JAP3* IR-PTGS (Figure 3), indicating that UPF1 likely does not play a role in IR-PTGS in *Arabidopsis* and that the initial report likely was hampered by *35S* interference.

These results indicate that deadenylation, NMD and exosome components impinge on PTGS at a step unique to S-PTGS that is not required for IR-PTGS. It is interesting to speculate that this step is linked to aberrant ssRNA protection or dsRNA generation, processes accomplished by the SGS3 and RDR6 proteins, respectively (74–76).

Both nuclear and cytoplasmic RNA decay proteins regulate S-PTGS

To determine where within the cell the different exonucleolytic RNA decay and S-PTGS pathways could overlap, we first expressed a subset of the components for which mutations were shown to alter S-PTGS as translational fusions to fluorescent reporters in *N. benthamiana* leaves (Figure 4). The S-PTGS components RDR6 and SGS3 were confirmed to localize in cytoplasmic foci. We also confirmed the previously reported subcellular localization of UPF3 and UPF1 in the nucleus and cytoplasmic foci, respectively (77,78). RRP44A was previously reported to be predominantly nuclear (35), and we observed that the core subunits of the exosome, RRP4 and RRP41, also were detected primarily in the nucleus, with only a weak diffuse signal present in the cytoplasm (Figure 4). Finally, we showed that the deadenylation

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

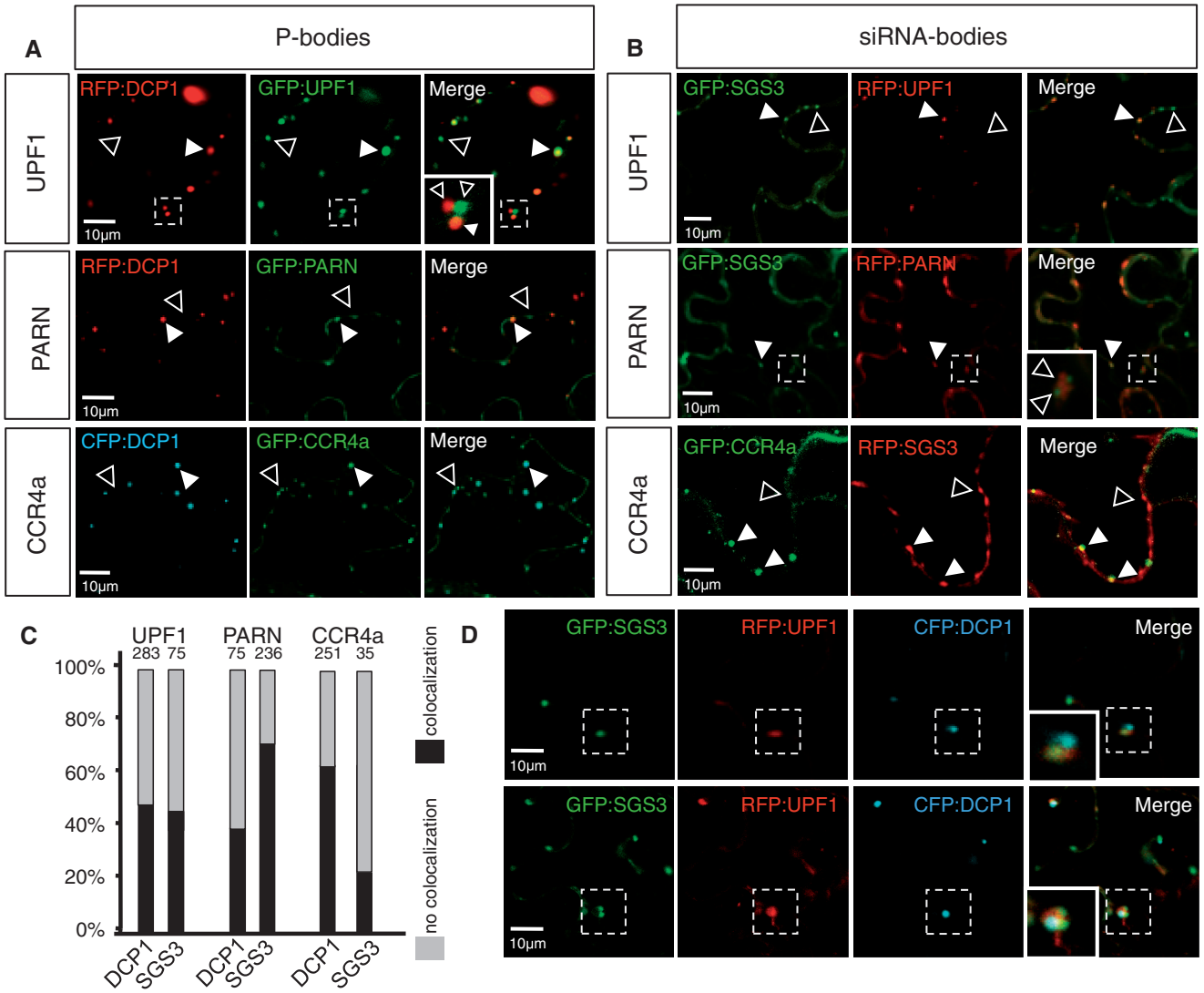


Figure 5. UPF1, PARN and CCR4a associate with both P- and siRNA-bodies. (**A** and **B**) Confocal sections of *N. benthamiana* leaves co-expressing the indicated fluorescent fusion proteins. Co-expression of UPF1, PARN and CCR4a with DCP1, a P-bodies marker (A), or with SGS3, a siRNA-bodies marker (B). White arrowheads indicate co-localization, and open arrowheads indicate foci positive for only one of the two fusion proteins. The area enclosed in the dashed box is shown in the close-up view. (**C**) Percentage of UPF1, PARN and CCR4a foci that co-localize with either P-bodies (as marked by DCP1) or siRNA-bodies (as marked by SGS3). Percentage of foci co-localizing (black) or not co-localizing (grey) with DCP1 and SGS3. The total number of foci counted is indicated above each bar. DCP1 and SGS3 foci were never observed to overlap. (**D**) Confocal sections of *N. benthamiana* leaves co-expressing SGS3, DCP1 and UPF1 fluorescent fusion proteins. Upper row: UPF1 is associated with a siRNA-body that is located adjacent to a P-body. Lower row: UPF1 is associated with a P-body that is located adjacent to two siRNA-bodies. The area enclosed in the dashed box is shown in the close-up view. Scale bars are shown on the images.

factors CCR4a and PARN both accumulated in cytoplasmic foci (Figure 4). The subcellular localizations of UPF3, RRP41, CCR4a, PARN, RDR6 and SGS3 observed in transient expression were confirmed in stable *Arabidopsis* lines expressing the fusion proteins at low levels (Supplementary Figure S3), indicating that the subcellular localizations observed in *N. benthamiana* leaves are not artifacts caused by over-expression.

Although RDR6 has been reported in both the nucleus and the cytoplasm, it only co-localizes with SGS3 in cytoplasmic foci called siRNA-bodies (75,79–81). Another class of cytoplasmic foci involved in mRNA degradation, distinct from siRNA-bodies, is the P-bodies where the decapping complex protein DCP1 accumulates (29,80). We therefore investigated whether the cytoplasmic foci observed for UPF1, PARN and CCR4a were siRNA-bodies or P-bodies or these proteins shuttle between them. To this end, we co-expressed these tagged proteins with either fluorescently tagged DCP1 or SGS3. We observed that tagged UPF1, PARN and CCR4a co-localized with both DCP1 and SGS3 (Figure 5A and B). Quantification of the proportion of UPF1, PARN and CCR4a bodies co-localizing with DCP1 (P-bodies) or SGS3 (siRNA-bodies) indicated that while nearly 70% of PARN foci co-localized with siRNA-bodies and >60% of CCR4a foci were associated with P-bodies (Figure 5A–C), UPF1 was found nearly equally associated with both siRNA- and P-bodies. The fraction of UPF1

Downloaded from http://nar.oxfordjournals.org/ at CNRS (1-20 DAW) on April 25, 2013

that co-localized with P- or siRNA-bodies nearly equaled the fraction of UPF1 that was non-co-localized to the other body (siRNA- or P-bodies, respectively, Figure 5C); thus, we more precisely examine these associations through a triple localization experiment among UPF1, DCP1 and SGS3. In the triple localization, we examined 32 adjacent P- and siRNA-body clusters and observed that for a given group of closely associated P- and siRNA-bodies, the UPF1 protein was either associated with the P-body or the siRNA-body but never with both bodies in the same cluster at the same time (Figure 5D and Supplementary Table S1).

DISCUSSION

Our results hint to a multi-layered regulatory network governing RQC and PTGS in different subcellular compartments. It was shown previously that mutations in the cytoplasmic exoribonuclease XRN4, the cytoplasmic decapping component DCP2 and the nuclear exoribonuclease XRN2 and XRN3 enhance PTGS (49,82). Here, we show that, in addition to mutations in several cytoplasmic deadenylation and NMD components, mutations in essentially nuclear RQC components (UPF3, RRP44A and RRP6L1) enhance PTGS. These results are in agreement with the existence of both a cytoplasmic and a nuclear arm to the PTGS pathway (79,83,84) and suggest that nuclear RNAs, in addition to cytoplasmic RNAs, are instrumental during S-PTGS. However, it remains unknown if these nuclear localized proteins are spatially associated with nuclear components of PTGS. Indeed, the DCL proteins responsible for siRNA generation are nuclear localized (85). Additional work is needed to examine these putative associations.

Although it is intriguing to imagine a nuclear interface among these pathways, we cannot exclude the possibility that RNA substrates that evade elimination by these nuclear RQC components are exported from the nucleus where they trigger S-PTGS in the cytoplasm. Moreover, it is also possible that at least a fraction of these primarily nuclear RQC proteins can be shuttled to the cytoplasm at certain times. Indeed, in yeast, UPF3 is shuttle protein operating in NMD, which involves both nuclear-localized steps and a cytoplasmic-localized translation termination coupled step (86).

Our observations that UPF1, CCR4a and PARN co-localize with both P- and siRNA-body markers suggest that exchange of ribonucleoparticle substrates between the two RNA degradation bodies can occur. We propose a model of ssRNA tug-of-war between RQC and S-PTGS that ensures the correct partitioning of aberrant RNA substrates among these RNA degradation mechanisms, potentially contributing to the discrimination of dysfunctional self-RNA and invading non–self-RNA from functional self-RNA (87). We assert that this discrimination allows a cell to efficiently clear undesirable RNAs without triggering PTGS, which, owing to the amplification of siRNA production, would lead to the unregulated trans-degradation of any RNA transcripts sharing homology with the dsRNA trigger. Indeed, it is interesting to speculate that the existence of the S-PTGS pathway may serve to reinforce the efficiency of RQC pathways to eliminate defective RNAs.

We recognize that this system of checks and balances between PTGS and RQC pathways was revealed in RQC mutant plants, and, thus, contend that these pathways may normally act independently, and that RNA substrate sharing may only occur when RQC pathways are rendered inefficient or compromised. Indeed, it is highly plausible that, in normal conditions, defective endogenous transcripts are eliminated efficiently by RQC pathways so as to prevent their 'auto-death' by PTGS.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2 and Supplementary Figures 1–3.

ACKNOWLEDGEMENTS

We thank Bruno Letarnac, Herve Ferry, Philippe Marechal and Fabrice Petitpas for plant care. This work has benefited from the facilities and expertise of the Imagif Cell Biology Unit of the Gif campus (www.imagif.cnrs.fr), which is supported by the Conseil Général de l'Essonne. A.B.M., F.B. and A.M. carried out the cloning and the imaging experiments; A.E.Md.A., H.V. and A.C.M. carried out the genetic and molecular PTGS analyses; M.D.C., A.C.M., A.M. and H.V. designed and coordinated the study. All authors contributed to the writing and editing of the manuscript and read and approved the final manuscript.

FUNDING

Agence Nationale de la Recherche [ANR-08-BLAN-0082 to A.C.M. and A.M. and ANR-10-LABX-40 to M.C. and H.V.]; Land Baden-Württemberg, Chica und Heinz Schaller Stiftung and the CellNetworks cluster of excellence (to A.M.). Funding for open access charge: Agence Nationale de la Recherche [ANR-08-BLAN-0082] the Chica und Heinz Schaller Stiftung.

Conflict of interest statement. None declared.

REFERENCES

1. Chen,C.Y. and Shyu,A.B. (2011) Mechanisms of deadenylation-dependent decay. *Wiley Interdiscip. Rev. RNA*, **2**, 167–183.
2. Isken,O. and Maquat,L.E. (2008) The multiple lives of NMD factors: balancing roles in gene and genome regulation. *Nat. Rev. Genet.*, **9**, 699–712.
3. Shoemaker,C.J. and Green,R. (2012) Translation drives mRNA quality control. *Nat. Struct. Mol. Biol.*, **19**, 594–601.
4. Staiger,D., Korneli,C., Lummer,M. and Navarro,L. (2012) Emerging role for RNA-based regulation in plant immunity. *New Phytol.*, **197**, 394–404.
5. Chen,X. (2012) Small RNAs in development – insights from plants. *Curr. Opin. Genet. Dev.*, **22**, 361–367.
6. Parent,J.S., Martínez de Alba,A.E. and Vaucheret,H. (2012) The origin and effect of small RNA signaling in plants. *Front. Plant Sci.*, **3**, 179.

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

7. Rubio-Somoza,I., Cuperus,J.T., Weigel,D. and Carrington,J.C. (2009) Regulation and functional specialization of small RNA-target nodes during plant development. *Curr. Opin. Plant Biol.*, **12**, 622–627.
8. Voinnet,O. (2008) Post-transcriptional RNA silencing in plant-microbe interactions: a touch of robustness and versatility. *Curr. Opin. Plant Biol.*, **11**, 464–470.
9. Hansen,K.D., Lareau,L.F., Blanchette,M., Green,R.E., Meng,Q., Rehwinkel,J., Gallusser,F.L., Izaurralde,E., Rio,D.C., Dudoit,S. *et al.* (2009) Genome-wide identification of alternative splice forms down-regulated by nonsense-mediated mRNA decay in Drosophila. *PLoS Genet.*, **5**, e1000525.
10. Kurihara,Y., Matsui,A., Hanada,K., Kawashima,M., Ishida,J., Morosawa,T., Tanaka,M., Kaminuma,E., Mochizuki,Y., Matsushima,A. *et al.* (2009) Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in *Arabidopsis*. *Proc. Natl Acad. Sci. USA*, **106**, 2453–2458.
11. Weischenfeldt,J., Waage,J., Tian,G., Zhao,J., Damgaard,I., Jakobsen,J.S., Kristiansen,K., Krogh,A., Wang,J. and Porse,B.T. (2012) Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. *Genome Biol.*, **13**, R35.
12. Amrani,N., Ganesan,R., Kervestin,S., Mangus,D.A., Ghosh,S. and Jacobson,A. (2004) A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature*, **432**, 112–118.
13. Brogna,S. and Wen,J. (2009) Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat. Struct. Mol. Biol.*, **16**, 107–113.
14. Kashima,I., Yamashita,A., Izumi,N., Kataoka,N., Morishita,R., Hoshino,S., Ohno,M., Dreyfuss,G. and Ohno,S. (2006) Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.*, **20**, 355–367.
15. Le Hir,H. and Seraphin,B. (2008) EJCs at the heart of translational control. *Cell*, **133**, 213–216.
16. van Hoof,A. and Green,P.J. (1996) Premature nonsense codons decrease the stability of phytohemagglutinin mRNA in a position-dependent manner. *Plant J.*, **10**, 415–424.
17. Buhler,M., Paillusson,A. and Muhlemann,O. (2004) Efficient downregulation of immunoglobulin mu mRNA with premature translation-termination codons requires the 5′-half of the VDJ exon. *Nucleic Acids Res.*, **32**, 3304–3315.
18. Chan,M.T. and Yu,S.M. (1998) The 3′ untranslated region of a rice alpha-amylase gene functions as a sugar-dependent mRNA stability determinant. *Proc. Natl Acad. Sci. USA*, **95**, 6543–6547.
19. Hong,X., Scofield,D.G. and Lynch,M. (2006) Intron size, abundance, and distribution within untranslated regions of genes. *Mol. Biol. Evol.*, **23**, 2392–2404.
20. LeBlanc,J.J. and Beemon,K.L. (2004) Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging. *J. Virol.*, **78**, 5139–5146.
21. Rajavel,K.S. and Neufeld,E.F. (2001) Nonsense-mediated decay of human HEXA mRNA. *Mol. Cell. Biol.*, **21**, 5512–5519.
22. Gatfield,D. and Izaurralde,E. (2004) Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. *Nature*, **429**, 575–578.
23. Stalder,L. and Muhlemann,O. (2008) The meaning of nonsense. *Trends Cell Biol.*, **18**, 315–321.
24. Chiba,Y., Johnson,M.A., Lidder,P., Vogel,J.T., van Erp,H. and Green,P.J. (2004) AtPARN is an essential poly(A) ribonuclease in *Arabidopsis*. *Gene*, **328**, 95–102.
25. Dupressoir,A., Morel,A.P., Barbot,W., Loireau,M.P., Corbo,L. and Heidmann,T. (2001) Identification of four families of yCCR4- and Mg2+-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. *BMC Genomics*, **2**, 9–25.
26. Reverdatto,S.V., Dutko,J.A., Chekanova,J.A., Hamilton,D.A. and Belostotsky,D.A. (2004) mRNA deadenylation by PARN is essential for embryogenesis in higher plants. *RNA*, **10**, 1200–1214.
27. Walley,J.W., Kelley,D.R., Nestorova,G., Hirschberg,D.L. and Dehesh,K. (2010) *Arabidopsis* deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. *Plant Physiol.*, **152**, 866–875.
28. Goeres,D.C., Van Norman,J.M., Zhang,W., Fauver,N.A., Spencer,M.L. and Sieburth,L.E. (2007) Components of the *Arabidopsis* mRNA decapping complex are required for early seedling development. *Plant Cell*, **19**, 1549–1564.
29. Xu,J. and Chua,N.H. (2009) *Arabidopsis* decapping 5 is required for mRNA decapping, P-body formation, and translational repression during postembryonic development. *Plant Cell*, **21**, 3270–3279.
30. Xu,J., Yang,J.Y., Niu,Q.W. and Chua,N.H. (2006) *Arabidopsis* DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. *Plant Cell*, **18**, 3386–3398.
31. Kastenmayer,J.P. and Green,P.J. (2000) Novel features of the XRN-family in *Arabidopsis*: evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. *Proc. Natl Acad. Sci. USA*, **97**, 13985–13990.
32. Chekanova,J.A., Gregory,B.D., Reverdatto,S.V., Chen,H., Kumar,R., Hooker,T., Yazaki,J., Li,P., Skiba,N., Peng,Q. *et al.* (2007) Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. *Cell*, **131**, 1340–1353.
33. Hooker,T.S., Lam,P., Zheng,H. and Kunst,L. (2007) A core subunit of the RNA-processing/degrading exosome specifically influences cuticular wax biosynthesis in *Arabidopsis*. *Plant Cell*, **19**, 904–913.
34. Lange,H. and Gagliardi,D. (2011) The exosome and 3′-5′ RNA degradation in plants. *Adv. Exp. Med. Biol.*, **702**, 50–62.
35. Zhang,W., Murphy,C. and Sieburth,L.E. (2010) Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses *Arabidopsis* decapping mutant phenotypes. *Proc. Natl Acad. Sci. USA*, **107**, 15981–15985.
36. Ding,S.W. and Voinnet,O. (2007) Antiviral immunity directed by small RNAs. *Cell*, **130**, 413–426.
37. Mallory,A.C., Elmayan,T. and Vaucheret,H. (2008) MicroRNA maturation and action–the expanding roles of ARGONAUTEs. *Curr. Opin. Plant Biol.*, **11**, 560–566.
38. van Mierlo,J.T., van Cleef,K.W. and van Rij,R.P. (2011) Defense and counterdefense in the RNAi-based antiviral immune system in insects. *Methods Mol. Biol.*, **721**, 3–22.
39. Wang,M.B., Masuta,C., Smith,N.A. and Shimura,H. (2012) RNA silencing and plant viral diseases. *Mol. Plant Microbe Interact.*, **25**, 1275–1285.
40. Waterhouse,P.M. (2006) Defense and counterdefense in the plant world. *Nat. Genet.*, **38**, 138–139.
41. Skopelitis,D.S., Husbands,A.Y. and Timmermans,M.C. (2012) Plant small RNAs as morphogens. *Curr. Opin. Cell Biol.*, **24**, 217–224.
42. Willmann,M.R. and Poethig,R.S. (2005) Time to grow up: the temporal role of smallRNAs in plants. *Curr. Opin. Plant Biol.*, **8**, 548–552.
43. Chapman,E.J. and Carrington,J.C. (2007) Specialization and evolution of endogenous small RNA pathways. *Nat. Rev. Genet.*, **8**, 884–896.
44. Chen,X. (2010) Small RNAs - secrets and surprises of the genome. *Plant J.*, **61**, 941–958.
45. Mallory,A. and Vaucheret,H. (2010) Form, function, and regulation of ARGONAUTE proteins. *Plant Cell*, **22**, 3879–3889.
46. Brosnan,C.A. and Voinnet,O. (2011) Cell-to-cell and long-distance siRNA movement in plants: mechanisms and biological implications. *Curr. Opin. Plant Biol.*, **14**, 580–587.
47. Melnyk,C.W., Molnar,A. and Baulcombe,D.C. (2011) Intercellular and systemic movement of RNA silencing signals. *EMBO J.*, **30**, 3553–3563.
48. Gazzani,S., Lawrenson,T., Woodward,C., Headon,D. and Sablowski,R. (2004) A link between mRNA turnover and RNA interference in *Arabidopsis*. *Science*, **306**, 1046–1048.
49. Gy,I., Gasciolli,V., Lauressergues,D., Morel,J.B., Gombert,J., Proux,F., Proux,C., Vaucheret,H. and Mallory,A.C. (2007) *Arabidopsis* FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *Plant Cell*, **19**, 3451–3461.
50. Herr,A.J., Molnar,A., Jones,A. and Baulcombe,D.C. (2006) Defective RNA processing enhances RNA silencing and influences

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

flowering of *Arabidopsis*. *Proc. Natl Acad. Sci. USA*, **103**, 14994–15001.
51. Luo,Z. and Chen,Z. (2007) Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in *Arabidopsis*. *Plant Cell*, **19**, 943–958.
52. Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C. *et al*. (2002) A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell*, **14**, 2985–2994.
53. Nishimura,N., Okamoto,M., Narusaka,M., Yasuda,M., Nakashita,H., Shinozaki,K., Narusaka,Y. and Hirayama,T. (2009) ABA hypersensitive germination2-1 causes the activation of both abscisic acid and salicylic acid responses in *Arabidopsis*. *Plant Cell Physiol.*, **50**, 2112–2122.
54. Clough,S.J. and Bent,A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis* thaliana. *Plant J.*, **16**, 735–743.
55. Grefen,C., Donald,N., Hashimoto,K., Kudla,J., Schumacher,K. and Blatt,M.R. (2010) A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. *Plant J.*, **64**, 355–365.
56. Karimi,M., Bleys,A., Vanderhaeghen,R. and Hilson,P. (2007) Building blocks for plant gene assembly. *Plant Physiol.*, **145**, 1183–1191.
57. Beclin,C., Boutet,S., Waterhouse,P. and Vaucheret,H. (2002) A branched pathway for transgene-induced RNA silencing in plants. *Curr. Biol.*, **12**, 684–688.
58. Elmayan,T., Balzergue,S., Beon,F., Bourdon,V., Daubremet,J., Guenet,Y., Mourrain,P., Palauqui,J.C., Vernhettes,S., Vialle,T. *et al*. (1998) *Arabidopsis* mutants impaired in cosuppression. *Plant Cell*, **10**, 1747–1758.
59. Daxinger,L., Hunter,B., Sheikh,M., Jauvion,V., Gasciolli,V., Vaucheret,H., Matzke,M. and Furner,I. (2008) Unexpected silencing effects from T-DNA tags in *Arabidopsis*. *Trends Plant Sci.*, **13**, 4–6.
60. Nishimura,N., Kitahata,N., Seki,M., Narusaka,Y., Narusaka,M., Kuromori,T., Asami,T., Shinozaki,K. and Hirayama,T. (2005) Analysis of ABA hypersensitive germination2 revealed the pivotal functions of PARN in stress response in *Arabidopsis*. *Plant J.*, **44**, 972–984.
61. Chekanova,J.A., Shaw,R.J., Wills,M.A. and Belostotsky,D.A. (2000) Poly(A) tail-dependent exonuclease AtRrp41p from *Arabidopsis* thaliana rescues 5.8 S rRNA processing and mRNA decay defects of the yeast ski6 mutant and is found in an exosome-sized complex in plant and yeast cells. *J. Biol. Chem.*, **275**, 33158–33166.
62. Dziembowski,A., Lorentzen,E., Conti,E. and Seraphin,B. (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.*, **14**, 15–22.
63. Liu,Q., Greimann,J.C. and Lima,C.D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell*, **127**, 1223–1237.
64. Lange,H., Holec,S., Cognat,V., Pieuchot,L., Le Ret,M., Canaday,J. and Gagliardi,D. (2008) Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in *Arabidopsis* thaliana. *Mol. Cell. Biol.*, **28**, 3038–3044.
65. Felippes,F.F., Wang,J.W. and Weigel,D. (2012) MIGS: miRNA-induced gene silencing. *Plant J.*, **70**, 541–547.
66. Allen,E., Xie,Z., Gustafson,A.M. and Carrington,J.C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, **121**, 207–221.
67. Vazquez,F., Vaucheret,H., Rajagopalan,R., Lepers,C., Gasciolli,V., Mallory,A.C., Hilbert,J.L., Bartel,D.P. and Crete,P. (2004) Endogenous trans-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs. *Mol. Cell.*, **16**, 69–79.
68. Yoshikawa,M., Peragine,A., Park,M.Y. and Poethig,R.S. (2005) A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. *Genes Dev.*, **19**, 2164–2175.
69. Peragine,A., Yoshikawa,M., Wu,G., Albrecht,H.L. and Poethig,R.S. (2004) SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in *Arabidopsis*. *Genes Dev.*, **18**, 2368–2379.
70. Smith,L.M., Pontes,O., Searle,I., Yelina,N., Yousafzai,F.K., Herr,A.J., Pikaard,C.S. and Baulcombe,D.C. (2007) An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. *Plant Cell*, **19**, 1507–1521.
71. Arciga-Reyes,L., Wootton,L., Kieffer,M. and Davies,B. (2006) UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in *Arabidopsis*. *Plant J.*, **47**, 480–489.
72. Domeier,M.E., Morse,D.P., Knight,S.W., Portereiko,M., Bass,B.L. and Mango,S.E. (2000) A link between RNA interference and nonsense-mediated decay in Caenorhabditis elegans. *Science*, **289**, 1928–1931.
73. Rehwinkel,J., Letunic,I., Raes,J., Bork,P. and Izaurralde,E. (2005) Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. *RNA*, **11**, 1530–1544.
74. Dalmay,T., Hamilton,A., Rudd,S., Angell,S. and Baulcombe,D.C. (2000) 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**, 543–553.
75. Elmayan,T., Adenot,X., Gissot,L., Lauressergues,D., Gy,I. and Vaucheret,H. (2009) A neomorphic sgs3 allele stabilizing miRNA cleavage products reveals that SGS3 acts as a homodimer. *FEBS J.*, **276**, 835–844.
76. Mourrain,P., Beclin,C., Elmayan,T., Feuerbach,F., Godon,C., Morel,J.B., Jouette,D., Lacombe,A.M., Nikic,S., Picault,N. *et al*. (2000) *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell*, **101**, 533–542.
77. Kim,S.H., Koroleva,O.A., Lewandowska,D., Pendle,A.F., Clark,G.P., Simpson,C.G., Shaw,P.J. and Brown,J.W. (2009) Aberrant mRNA transcripts and the nonsense-mediated decay proteins UPF2 and UPF3 are enriched in the *Arabidopsis* nucleolus. *Plant Cell*, **21**, 2045–2057.
78. Merai,Z., Benkovics,A.H., Nyiko,T., Debreczeny,M., Hiripi,L., Kerenyi,Z., Kondorosi,E. and Silhavy,D. (2013) The late steps of plant nonsense-mediated mRNA decay. *Plant J.*, **73**, 50–62.
79. Hoffer,P., Ivashuta,S., Pontes,O., Vitins,A., Pikaard,C., Mroczka,A., Wagner,N. and Voelker,T. (2011) Posttranscriptional gene silencing in nuclei. *Proc. Natl Acad. Sci. USA*, **108**, 409–414.
80. Jouannet,V., Moreno,A.B., Elmayan,T., Vaucheret,H., Crespi,M.D. and Maizel,A. (2012) Cytoplasmic *Arabidopsis* AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. *EMBO J.*, **31**, 1704–1713.
81. Kumakura,N., Takeda,A., Fujioka,Y., Motose,H., Takano,R. and Watanabe,Y. (2009) SGS3 and RDR6 interact and colocalize in cytoplasmic SGS3/RDR6-bodies. *FEBS Lett.*, **583**, 1261–1266.
82. Thran,M., Link,K. and Sonnewald,U. (2012) The *Arabidopsis* DCP2 gene is required for proper mRNA turnover and prevents transgene silencing in *Arabidopsis*. *Plant J.*, **72**, 368–377.
83. Le Masson,I., Jauvion,V., Bouteiller,N., Rivard,M., Elmayan,T. and Vaucheret,H. (2012) Mutations in the *Arabidopsis* H3K4me2/3 demethylase JMJ14 suppress posttranscriptional gene silencing by decreasing transgene transcription. *Plant Cell*, **24**, 3603–3612.
84. Morel,J.B., Mourrain,P., Beclin,C. and Vaucheret,H. (2000) DNA methylation and chromatin structure affect transcriptional and post-transcriptional transgene silencing in *Arabidopsis*. *Curr. Biol.*, **10**, 1591–1594.
85. Xie,Z., Johansen,L.K., Gustafson,A.M., Kasschau,K.D., Lellis,A.D., Zilberman,D., Jacobsen,S.E. and Carrington,J.C. (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.*, **2**, E104.
86. Shirley,R.L., Ford,A.S., Richards,M.R., Albertini,M. and Culbertson,M.R. (2002) Nuclear import of Upf3p is mediated by importin-alpha/-beta and export to the cytoplasm is required for a functional nonsense-mediated mRNA decay pathway in yeast. *Genetics*, **161**, 1465–1482.
87. Christie,M., Brosnan,C.A., Rothnagel,J.A. and Carroll,B.J. (2011) RNA decay and RNA silencing in plants: competition or collaboration? *Front Plant Sci.*, **2**, 99–106.

Downloaded from http://nar.oxfordjournals.org/ at CNRS (I-20 DAW) on April 25, 2013

Annexe 3:

Charon, C., Moreno, A.B., Bardou, F., and Crespi, M. (2010). Non-protein-coding RNAs and their interacting RNA-binding proteins in the plant cell nucleus. Molecular plant **3**: 729–39.

Molecular Plant Advance Access published July 5, 2010

Molecular Plant • Pages 1–11, 2010 REVIEW ARTICLE

Non-Protein-Coding RNAs and their Interacting RNA-Binding Proteins in the Plant Cell Nucleus

Celine Charon[a,b], Ana Beatriz Moreno[c], Florian Bardou[c] and Martin Crespi[c,1]

a University Paris-Sud, Institut de Biologie des Plantes, UMR8618, Orsay Cedex, F-91405, France
b CNRS, Orsay Cedex, F-91405, France
c Institut des Sciences du Végétal, CNRS, F-91198 Gif-sur-Yvette Cedex, France

ABSTRACT The complex responses of eukaryotic cells to external factors are governed by several transcriptional and post-transcriptional processes. Several of them occur in the nucleus and have been linked to the action of non-protein-coding RNAs (or npcRNAs), both long and small npcRNAs, that recently emerged as major regulators of gene expression. Regulatory npcRNAs acting in the nucleus include silencing-related RNAs, intergenic npcRNAs, natural antisense RNAs, and other aberrant RNAs resulting from the interplay between global transcription and RNA processing activities (such as Dicers and RNA-dependent polymerases). Generally, the resulting npcRNAs exert their regulatory effects through interactions with RNA-binding proteins (or RBPs) within ribonucleoprotein particles (or RNPs). A large group of RBPs are implicated in the silencing machinery through small interfering RNAs (siRNAs) and their localization suggests that several act in the nucleus to trigger epigenetic and chromatin changes at a whole-genome scale. Other nuclear RBPs interact with npcRNAs and change their localization. In the fission yeast, the RNA-binding Mei2p protein, playing pivotal roles in meiosis, interact with a meiotic npcRNA involved in its nuclear re-localization. Related processes have been identified in plants and the *ENOD40* npcRNA was shown to re-localize a nuclear-speckle RBP from the nucleus to the cytoplasm in *Medicago truncatula*. Plant RBPs have been also implicated in RNA-mediated chromatin silencing in the FLC locus through interaction with specific antisense transcripts. In this review, we discuss the interactions between RBPs and npcRNAs in the context of nuclear-related processes and their implication in plant development and stress responses. We propose that these interactions may add a regulatory layer that modulates the interactions between the nuclear genome and the environment and, consequently, control plant developmental plasticity.

Key words: Abiotic/environmental stress; gene silencing; *Arabidopsis*; RNA-binding proteins; non-protein-coding RNAs; nucleus.

INTRODUCTION

In recent years, RNA researchers have shown a growing interest in a hidden part of the transcriptome: the non-protein-coding RNAs (npcRNAs). This group of RNAs does not encode proteins but their function is associated with the RNA molecule itself. Although some npcRNAs can code for small functional peptides, they do not contain long ORFs and, consequently, they have eluded bioinformatic searches mainly based on coding capacity. Nonetheless, new bioinformatical and experimental strategies as high-throughput sequencing of RNAs and microarray analysis have revealed an outstanding number of novel npcRNA candidates in various model organisms from yeast or plants to *Homo sapiens* (He et al., 2006; Hüttenhofer and Vogel 2006; Mattick and Makunin, 2006; Mercer et al., 2009; Yasuda and Hayashizaki, 2008). Apart from the well known housekeeping non-protein-coding RNAs like ribosomal RNA (rRNA), transfer RNA (tRNA), small nucleolar RNA (snoRNA), and small nuclear RNA (snRNA), many new regulatory npcRNAs or riboregulators have been discovered and characterized (Mercer et al., 2009; Wilusz et al., 2009). NpcRNAs can be transcribed from intergenic regions, but they also include a surprising number of antisense RNA transcripts, pseudogenes, and truncated transcripts in eukaryotes. The transcriptome is surprisingly complex, with long npcRNAs often overlapping with or interspersed in between coding transcripts. This complexity has created a shift in our comprehension

[1] To whom correspondence should be addressed. E-mail Martin.Crespi@isv.cnrs-gif.fr, fax 331 69823695, tel. 33 1 69823703.

© The Author 2010. Published by the Molecular Plant Shanghai Editorial Office in association with Oxford University Press on behalf of CSPP and IPPE, SIBS, CAS.

doi: 10.1093/mp/ssq037
Received 29 April 2010; accepted 31 May 2010

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

of gene expression as a DNA sequence can be transcribed in multiple sense and antisense transcripts, intronic npcRNAs, intergenic, or promoter-associated RNAs (Mercer et al., 2009). In *Arabidopsis thaliana*, whole-genome mapping based on the use of tiling arrays revealed that >50% of observed transcription was intergenic and that numerous antisense RNA transcripts exist (Yamada et al., 2003).

Several well known nuclear RNAs (snRNA, snoRNAs) are mainly involved in ribosome biogenesis and splicing mechanisms, including alternative splicing, and will not be described here (reviewed in Simpson et al., 2010). Nevertheless, introns may themselves code for non-protein-coding RNAs such as the intronic microRNA (miRNA) or mirtrons (Ying et al., 2010), as also shown in plants for some miRNAs (Brown et al., 2008; Hirsch et al., 2006). Globally, npcRNAs have been far less studied in plants than in animals, in which diverse mechanisms have been discovered (for review, see Prasanth and Spector, 2007; Voinnet, 2009; Wilusz et al., 2009).

All aspects of RNA metabolism are accompanied by the activities of a myriad of RNA-binding proteins (RBPs) (Burd and Dreyfuss, 1994). Most RBPs contain one or more conserved domains, such as the RNA-recognition motif (RRM), the K-homology (KH) motif, RGG (Arg-Gly-Gly) boxes, and double-stranded RNA-binding domains (dsRBDs). A survey of the *A. thaliana* genome for RNA-binding proteins revealed 196 RRM- and 26 KH-containing proteins (Barta et al., 2008; Lorkovic, 2009). Although most of these proteins have not been characterized experimentally, forward and reverse genetic approaches are beginning to reveal a requirement for proteins that have roles in RNA metabolism in plant development (Lorkovic, 2009). Several of these RBPs localize in the nucleus and have been linked to diverse nuclear activities (e.g. nuclear splicing) or sub-nuclear domains (e.g. nuclear speckles, nucleolus or Cajal bodies). In this review, we will discuss the interplay between npcRNAs and RNA-binding proteins that may be involved in environmental stress responses or plant development through their action in the plant cell nucleus.

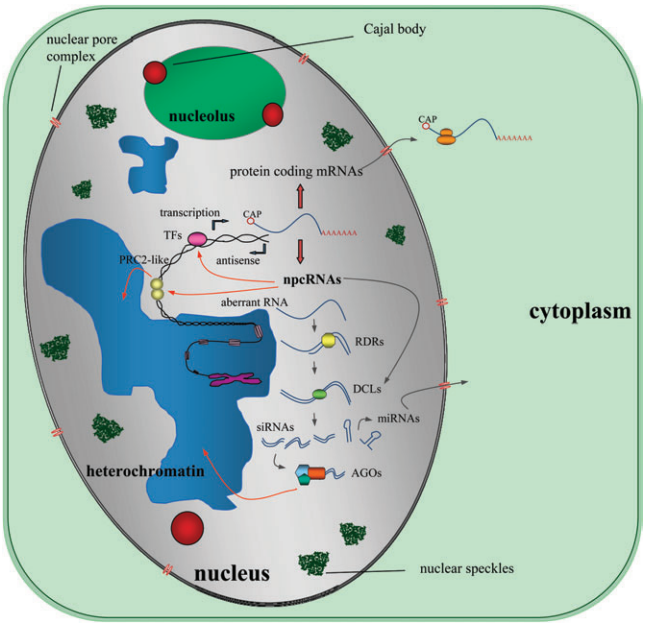


Figure 1. Schematic Representation of a Plant Nucleus Showing the Diversity of npcRNAs, their Biogenesis and Potential Mode of Action.

After transcription, a small percentage of the total RNA pool are protein coding mRNAs that are exported to the cytoplasm and subsequently translated to give rise to the cellular proteins. Another large portion of the transcriptome is composed of the so-called npcRNAs. One part of the npcRNAs will trigger regulatory mechanisms inside the nucleus as indicated by red arrows: examples of long npcRNAs have been shown to affect transcription, act as antisense RNAs, or be involved in recruitment of Polycomb complexes and heterochromatin formation. We also show a simplified scheme of the biogenesis of small npcRNAs, miRNAs, and siRNAs. The latter exert their function in *de novo* heterochomatin generation through the action of AGO-containing complexes or regulate mRNA stability or translation in the cytoplasm. TFs, transcription factors; PRC2-like, Polycomb repressive complex 2-like; npcRNAs, non-protein-coding RNAs; RDRs, RNA dependent RNA polymerases; DCLs, Dicer-like proteins; AGOs, argonaute proteins; siRNAs, small interfering RNAs; miRNAs, microRNAs. Different sub-nuclear particles (nuclear speckles, cajal bodies, nucleolus) are indicated and heterochromatin is indicated in dark blue.

REGULATORY NON-PROTEIN-CODING RNAs COMING TO LIGHT

Regulatory npcRNAs or riboregulators include npcRNAs expressed at a certain stage of development, during cell differentiation, or as a response to external stimuli, and can affect transcription or translation of other genes (Mattick and Makunin, 2006; Yasuda and Hayashizaki, 2008). According to their size, regulatory npcRNAs are classified as small regulatory npcRNAs (<40 bp) or long regulatory npcRNAs (>40 bp). NpcRNAs have been implicated in different regulatory mechanisms in plant development (Brown et al., 2008; Voinnet, 2009), in environmental biotic interactions, and abiotic stress responses (Ben Amor et al., 2009; Jay et al., 2010; Sunkar, 2010), and/or shown to have specific localization at tissular, cellular, and sub-cellular levels (Campalans et al., 2004; Zhan and Lukens, 2010). In Figure 1, we show a schematic representation of a plant nucleus indicating different sub-nuclear compartments and the biogenesis of both long and small npcRNAs.

The best characterized npcRNAs are the small si/miRNAs, key regulators of gene expression (Vaucheret, 2006; Voinnet, 2009). The small RNAs range from 20 to 40 nucleotides long and play a major role in gene silencing at transcriptional and post-transcriptional levels. There are many classes of small RNAs acting in the regulation of gene expression by different pathways, including small interfering RNAs (siRNAs), microRNAs (miRNAs), heterochromatic siRNAs (hc-siRNAs), Piwi-interacting RNAs (piRNAs), *trans*-acting siRNAs (ta-siRNAs), and naturally occurring antisense siRNAs (nat-siRNAs) (Jamalkandi and Masoudi-Nejad, 2009; MacLean et al., 2010). Small si/miRNAs induce messenger RNA (mRNA) cleavage and

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

translational inhibition through pairing with specific mRNA targets, mainly in the cytoplasm, or lead to transcriptional gene silencing (TGS), heterochromatin formation, and *de novo* DNA methylation in the nucleus (Jamalkandi and Masoudi-Nejad, 2009; Llave, 2004; Vaucheret, 2006; Verdel et al., 2009) (Figure 1). Although heterogeneous in size, sequence, genomic distribution, biogenesis and action, most of these molecules mediate repressive gene regulation through a mechanism often referred to as RNA silencing or RNA interference (RNAi). Their main role relies on the maintenance of genome integrity and developmental patterning, as well as on the generation of novel regulatory mechanisms to help plants to adapt and respond to adverse biotic and abiotic environmental conditions (Ruiz-Ferrer and Voinnet, 2009). Hence, the enormous diversity of small RNAs (more than 250 000 different sequences in *Arabidopsis*, the ASRP database, http://asrp.cgrb.oregonstate.edu/db/) may serve as a substrate to develop novel regulations for the protein-coding mRNAs as the enzymes involved in the processing of si/miRNAs are highly related, the RNAase III DICER and the DICER-like enzymes. The miRNA precursors are generally cleaved by DICER-like enzymes (four genes in *A. thaliana*, *DICER-like* (*DCL*)*1* to *DCL4*) and the resulting small RNA is incorporated in the RNA silencing complex (RISC) containing ARGONAUTE (AGO) proteins and determines the specificity of action of this complex (Vaucheret, 2006). The recognition of specific mRNA targets through the loaded miRNA will lead to the cleavage or the translational inhibition of the mRNA target (Xie et al., 2010). For siRNAs, the recognition of transgene mRNA targets leads to their degradation in a similar way (gene silencing) but also to the generation of secondary siRNAs (Vaucheret, 2006). In *A. thaliana*, DCL1 is the main enzyme responsible for miRNA production from imperfectly double-stranded RNA precursors arising from spontaneous folding of endogenous loci. Plant miRNAs are then exported to the cytoplasm and taken up in the RISC complex to exert their function on mRNAs, a well described post-transcriptional regulatory process (Voinnet, 2009). In fact, transcription factors and miRNAs are the major *trans*-acting regulators that determine the dynamic equilibrium of transcriptional networks at each developmental stage in eukaryotes (Hobert, 2008). As negative regulators of gene expression, they can act as developmental switches used to shut down gene expression programs. Alternatively, small RNAs can fine tune gene expression to quantitatively adapt the development to endogenous or environmental fluctuations and therefore act as canalization factors (Li et al., 2009).

Other endogenous pathways lead to gene silencing in plants. In contrast to miRNAs, siRNAs derive from dsRNA (double-stranded RNA) precursors, which originate from either convergent transcription of neighboring loci, inverted repeats, or from the action of RNA-dependent RNA polymerases (RDR) on precursor single-stranded RNAs, including transgenes or viruses, so-called aberrant RNAs (Chapman and Carrington, 2007; Vaucheret, 2006). The different DCL enzymes produce siRNA classes with different sizes. DCL4 produces 21-nt ta-siRNAs through recognition of the cleaved product of the long *TAS* npcRNAs by a specific miRNA (Chapman and Carrington, 2007). This cleaved product is converted into a dsRNA by the RDR6 polymerase and then cleaved by DCL4 to release ta-siRNAs loaded into AGO1 and guiding the cleavage of complementary mRNA in the cytoplasm. On the other hand, DCL3 produces 24-nt siRNAs from transposons and repeats RNA after they are converted into dsRNA by PolIV and the RDR2 polymerase (Vaucheret, 2006; Voinnet, 2009). These 24-nt siRNAs associate with AGO4 and other proteins, and induce transcriptional silencing through histone modification, DNA methylation, and chromatin remodeling. Finally, transcription of overlapping gene pairs in a convergent orientation therefore allowing the production of antisense transcripts can lead to the production of dsRNA and the so-called nat-siRNAs (natural antisense-siRNAs) through DCL enzymes (Borsani et al., 2005). In *A. thaliana*, the four members of the DICER family coupled to the 10 ARGONAUTE (AGO) members preclude an enormous complexity of the different pathways in which si/miRNAs can act (Chen, 2009; Höck and Meister, 2008; Margis et al., 2006). In all those pathways, there are some steps occurring in the nucleus, and all of them are tightly linked to the presence of specific RNA-binding proteins having different functions to ensure the functionality of the small RNAs in gene silencing, heterochromatin formation, or mRNA post-transcriptional regulation.

In contrast to small npcRNAs, much less is known about the diversity of biogenesis and action of long npcRNAs. These transcripts, as the miRNA precursors, are generally produced by RNA polymerase II, capped, and polyadenylated. Few data exist on long npcRNA activity, sub-cellular localization or molecular roles in plants; however, several of these npcRNAs have been identified (Ben Amor et al., 2009; Charon et al., 1999; Franco-Zorrilla et al., 2007; Hirsch et al., 2006; Ma et al., 2008). Although npcRNAs may play essential cytoplasmic regulatory roles such as inhibiting miRNA activity (Franco-Zorrilla et al., 2007) or affecting miRNA expression through other indirect means (Ben Amor et al., 2009), others seem to act as *cis*- or *trans*-antisense RNAs (Borsani et al., 2005; Katiyar-Agarwal et al., 2007)—a process that may be linked to the nucleus. In addition, as small siRNAs, long npcRNAs can also mediate epigenetic changes by recruiting chromatin remodeling complexes to specific genomic loci, as shown for the *HOTAIR* (*HOX* antisense RNA) that silences transcription across 40 kb of the *HOXD* locus in *Drosophila* (Rinn et al., 2007). This process is mediated by the Polycomb chromatin remodeling complex PRC2 and the *HOTAIR* RNA. Homologs of certain members of the PRC complex have been identified in plants and their mutations affect heterochromatin organization, cell proliferation and lead to spontaneous embryogenesis in plants (Chanvivattana et al., 2004), suggesting a link between heterochromatin regulation and plant development. Nevertheless, npcRNAs linked to the action of these plant *PRC*-like genes have not yet been identified. Other studies on npcRNA showed their ability to modulate nuclear activities of different

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

proteins. For example, in the presence of an npcRNA, the TLS (TRANSLOCATED IN LIPOSARCOMA) protein can change its conformation into an active form to inhibit the histone acetyl-transferases CBP and P300, and silence the cyclin D1 in human cells (Wang et al., 2008). More recently, the *GAS5* npcRNA (for *GROWTH ARREST SPECIFIC5*) has been shown to bind specifically to the GLUCOCORTICOID RECEPTOR (GR) protein, a transcription factor involved in cell growth, and to inhibit its activity in HeLa cells (Kino et al., 2010). Other mechanisms imply the transcription of an npcRNA across the promoter region of a downstream protein-coding gene to interfere with its expression pattern (Martens et al., 2004) or to induce histone modification leading to the repression of transcription initiation (Houseley et al., 2008) or, conversely, chromatin remodeling and opening to activate transcription (Hirota et al., 2008). The action of long npcRNAs that are not precursors of small RNAs, and act through direct interaction with the transcription machinery or via chromatin formation, are also depicted in Figure 1.

RNA-BINDING PROTEINS LINKED TO SILENCING MECHANISMS MAY ACT IN THE NUCLEUS

Identification of the RBPs with which each npcRNA is associated is at the core of understanding RNP interaction networks in the cell. Several RNA-binding proteins are involved in the biogenesis and the mechanisms of action of small RNAs. In plants, the RNAi pathways have been largely diversified and amplified, and several steps occur in the nucleus. In addition to the cleavage and the degradation of target homologous mRNAs in the cytoplasm (gene silencing), small siRNAs play a role in maintaining epigenetic marks in eukaryotic genomes (Valencia-Sanchez et al., 2006). Genome-wide analyses of DNA methylation at single-base resolution demonstrated that siRNAs direct approximately 30% of the cytosine methylation in *A. thaliana* (Cokus et al., 2008; Lister et al., 2008). Likely, long npcRNAs are substrates for DCL3 and are processed by the RNA silencing apparatus to generate 24-nt siRNAs that, in turn, will guide chromatin modifications to homologous regions of the genome (Bühler, 2009; Grewal and Elgin, 2007; Kloc and Martienssen, 2008; Matzke et al., 2009; Nagano et al., 2008; Zaratiegui et al., 2007). This last mechanism is known as RNA directed DNA methylation (RdDM) (Huettel et al., 2007), and was observed for the first time in tobacco plants infected with viroids, pathogen circular RNA molecules (Wassenegger et al., 1994). A proposed model for the RdDM pathway in *A. thaliana* begins with the recruitment of a form of the DNA-dependent RNA polymerase IV, the POL IVa, to a target genomic site (e.g. a transposon or DNA tandem repeats), through an unknown mechanism. The POLIV in *Arabidopsis* exists in two isoforms (POLIVa and POLIVb), with NRPD1a and NRPD1b as their respective largest subunits, and both isoforms are implicated in the production and activity of siRNAs and in RNA-directed DNA methylation (RdDM) (Mosher et al., 2008). The POLIVa isoform synthesizes a single-stranded RNA (ssRNA) that is recognized by RDR2 (RNA-DEPENDENT RNA POLYMERASE2) and/or POLIVa as an aberrant RNA, and subsequently converted by them into dsRNAs. The dsRNA is then digested by DCL3 to produce siRNAs that are loaded onto AGO4 proteins. An AGO4 protein bound to a siRNA may form a complex with POL IVb, and DRM2 (a *de novo* cytosine methyltransferase, DOMAIN REARRANGED METHYL-TRANSFERASE2) to initiate, in a sequence-specific manner, DNA methylation, histone methylation, and possibly ATP-dependent chromatin remodeling (Henderson and Jacobsen, 2007; Matzke et al., 2009). Other genetic approaches revealed that DRD1, a putative chromatin remodeling protein involved in RNA-directed DNA methylation, acts together with AGO4 to enable DRM2 to access the target DNA to carry out *de novo* DNA methylation (Huettel et al., 2006, 2007). Nevertheless, this is an active process, and the methylation status of a number of genes seems dynamically regulated by methylation and demethylation (Bei et al., 2007; Lister et al., 2008; Penterman et al., 2007a, 2007b; Zhu et al., 2007; Zhu, 2008). Active demethylation may naturally function to protect plant genes from genome defense pathways and/or reversibly modulate transcription in non-dividing cells (Huettel et al., 2006; Lister et al., 2008; Penterman et al., 2007a). In *A. thaliana*, this process is mediated by a subfamily of DNA glycosylases, including the ROS1 (REPRESSOR OF SILENCING1), that prevent DNA hyper-methylation and transcriptional gene silencing of the specific transgene, but it is still unknown how these proteins are targeted to specific sequences (Penterman et al., 2007a; Zhu et al., 2007). More recently, another element of this pathway was identified—the ROS3 protein that has an amino-terminal RNA-recognition motif (RRM) able to bind *in vivo* specific small RNAs (Zheng et al., 2008). ROS3 may direct demethylation of target sequences and further studies will clarify whether ROS3 can also bind larger npcRNAs and which specific sequence features it recognizes (Zheng et al., 2008). Interestingly, immuno-localization experiments revealed that ROS3 is localized in the nucleoplasm as well as in the nucleolus, as scattered speckle-like structures, and ROS1 was similarly found to be dispersed throughout the nucleoplasm and nucleolus. However, the ROS1 immuno-localization signals tended to appear more diffuse and smaller than the ROS3 nuclear foci (Zheng et al., 2008; Zhu et al., 2007). Taken together, the results indicate that ROS1 and ROS3 are interdependent for their nuclear, and especially nucleolar, co-localization, and that ROS3 functions in the same DNA demethylation pathway with ROS1 through the recognition of specific npcRNAs (small RNAs and/or larger RNAs) to guide sequence-specific DNA demethylation. In fact, several other components of the RdDM have been found in nuclear bodies or processing centers adjacent to or within the nucleolus in *A. thaliana* (Li et al., 2006; Pontes et al., 2006; Verdel et al., 2009) (Figure 2). Proteins that formed such discrete nuclear bodies include RDR2, DCL3, AGO4, and the subunit NRPD1b of POLIV.

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

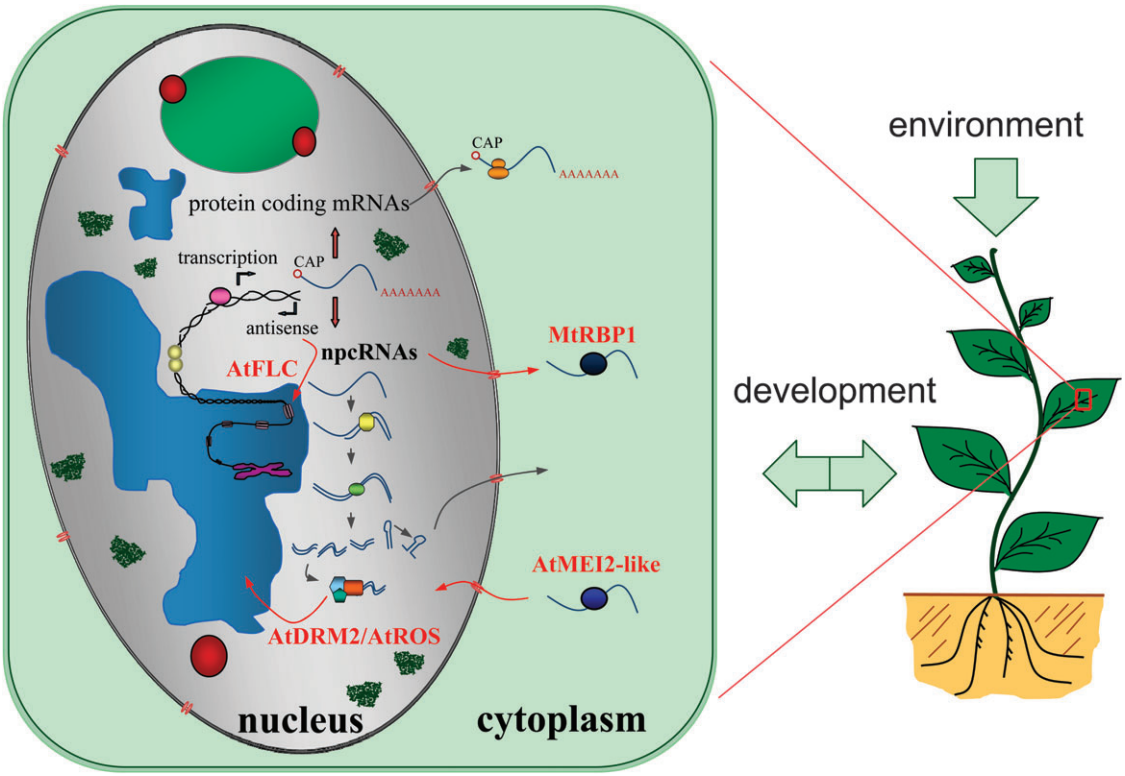


Figure 2. Regulatory Mechanisms Involving npcRNAs that May Affect Adaptive Developmental Responses to the Environment.

Diverse mechanisms linking certain npcRNAs and RBPs discussed in the text are schematized (arrows). MtRBP1 re-localizes from nuclear speckles to cytoplasmic particles in the presence the *ENOD40* npcRNA. Alternatively, AtMEI2-like proteins may be re-localized into the nucleus through interaction with npcRNAs, as shown in yeast cells. Another mechanism involves siRNAs loaded in AGO4-containing complexes in association with specific RBPs (ROS3/ROS1) and the demethylase DRM2 to trigger DNA methylation. Finally, 3′ processing of an antisense npcRNA results in epigenetic changes at the FLC locus. FLC, FLORAL LOCUS C; MtRBP1, *Medicago truncatula* RNA Binding Protein 1; AtMEI2-like, *Arabidopsis thaliana* MEIosis2-like; AtDRM2, *Arabidopsis thaliana* Domains Rearranged Methyl-transferase 2; AtROS, *Arabidopsis thaliana* Repressor Of Silencing.

Another mechanism involving RBPs in silencing pathways also linked to the nucleus is the miRNA biogenesis pathway. Some of its components, namely the processing factors DCL1 and HYL1, are present in nucleoplasm foci known as DICING bodies (D-bodies) (Fang and Spector, 2007; Song et al., 2007). However, contrarily to RDR2, DCL3, AGO4, and NRPD1b, they do not have an exclusive link with the nucleolus. Finally, a recent work demonstrated that AGO4 is present in two distinct types of bodies. The major part of AGO4 is found in Cajal bodies, whereas a smaller population is located in a second class of bodies called AB-bodies (Li et al., 2008). Those bodies contained some RdDM components as the DNA methyltransferase DRM2, and NRPD1b and NRPD2 (the smallest subunit of POLIV) were found to be adjacent to the 45S ribosomal DNA (rDNA) loci, suggesting that these bodies may be a site of active RNA-directed DNA methylation. Cajal bodies are nuclear compartments mainly adjacent to the nucleolus that can also be observed in the nucleoplasm. They serve as important centers for the maturation and processing of several ribonucleoparticles, which enter the Cajal body, are modified there, and then exit for further utilization in other processes elsewhere (Cioce and Lamond, 2005). Cajal bodies were shown to be dynamic structures in a constant state of assembly/disassembly implicated in diverse processes as in the assembly and trafficking of telomerase (Lukowiak et al., 2001), of U7 small nuclear ribonucleoprotein particule (snRNP) (Cioce and Lamond, 2005; Handwerger and Gall, 2006) and of RNA polymerase I, II, and III transcription factors (Pontes and Pikaard, 2008). Cajal bodies also play roles in assembling snoRNA ribonucleoprotein (snoRNP) complexes that are thought to accumulate in Cajal bodies before entering the nucleolus (Pontes and Pikaard, 2008). Their role in the regulation of processing or maturation of silencing complexes is still unknown.

RNA-BINDING PROTEIN AND LONG NON-PROTEIN-CODING RNA NUCLEAR INTERACTIONS MAY AFFECT PLANT DEVELOPMENT

Through interaction with specific RBPs, long npcRNAs may modulate the cellular RNP networks to determine new patterns of gene regulation similarly to small npcRNAs. One

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

sub-nuclear structure involved in multiple RNA-related processes is the nuclear speckle and one of its functions is the storage of the small nuclear ribonucleoprotein (snRNP) spliceosomal complexes (maturated in the Cajal bodies) until the appropriate molecular signals trigger their journey to transcribing genes in the spliceosomes, therefore acting in mRNA processing (Cioce and Lamond, 2005; Handwerger and Gall, 2006; Li et al., 2006). These nuclear speckles (or interchromatin granule clusters) are spotted shapeless structures containing elevated concentrations of splicing snRNPs and other splicing-related proteins that participate in the co-transcriptional splicing of mRNAs at the chromosomes, in addition to several molecules with structural roles in the cell (Shaw and Brown, 2004). In mammalian cells, changes in transcription and protein phosphorylation state perturb the composition and intra-nuclear localization of the speckles (Lamond and Spector, 2003). Besides, nuclear speckles may also supply a stopover and regulatory checkpoint for components traveling with mRNAs through the nuclear pore to the cytoplasm (Handwerger and Gall, 2006). Interestingly, RBPs present in nuclear speckles and interacting with npcRNAs may play a role in plant development. The npcRNA family *ENOD40* has been involved in the formation of symbiotic nitrogen-fixing nodules in legumes (Charon et al., 1999; Crespi and Gàlvez, 2000; Kouchi et al., 1999). Transgenic *Medicago truncatula* plants overexpressing or silenced for *ENOD40* exhibited accelerated nodulation or form only a few and modified nodule-like structures, respectively, suggesting that *MtENOD40* regulates nodule development (Charon et al., 1999; Wan et al., 2007). The *ENOD40* npcRNA is highly structured (Crespi et al., 1994; Girard et al., 2003), although a small peptide has been proposed to be translated from this transcript (Rohrig et al., 2002). Using the yeast three-hybrid system, a constitutively expressed RNA-binding protein, MtRBP1, localized in nuclear speckles, has been identified to interact with the *ENOD40* RNA (Campalans et al., 2004) (Figure 2). Immuno-localization experiments and transient assays demonstrated that the *MtENOD40* npcRNA seems to be required for the re-localization of MtRBP1, from nuclear speckles to cytoplasmic granules, during nodule organogenesis (Campalans et al., 2004).

Re-localization of RNP complexes has also been linked to the action of npcRNAs in the fission yeast. The *sme2/meiRNA* npcRNA was shown to bind the Mei2p protein, considered as a master regulator of meiosis (Watanabe and Yamamoto, 1994). The *mei2* gene encodes an RBP with three RNA-recognition motifs (RRMs), of which the C-terminal RRM3 is critical for its function (Watanabe et al., 1997). During mitosis, *mei2* transcripts are accumulated but Mei2p remains inactive within the cytoplasm. Under meiosis-inducing conditions (mainly nutrient starvation), Mei2p shuttles from the cytoplasm to the nucleus (Sato et al., 2001; Yamashita et al., 1998), where it binds to the *meiRNA* npcRNA at the *sme2* locus and forms a Mei2p dot structure (Shimada et al., 2003). Formation of this dot coincides with the onset of meiosis I and Mei2p may antagonize selective elimination of meiotic messenger RNAs by sequestering another RBP, Mmi1p, in this nuclear dot structure (Harigaya et al., 2006; Harigaya and Yamamoto, 2007). Whereas one unique *mei2* gene is present in *S. pombe* and does not exist in *S. cerevisiae*, the *mei2*-like family has undergone a great expansion in the vascular plant lineage (Anderson et al., 2004; Charon et al., 2010; Jeffares et al., 2004). All these plant *mei2*-like genes encode RBPs with three RRMs, particularly the third distinctive C-terminal RRM that is essential for Mei2p function and only present in the Mei2p-like proteins. Nevertheless, Mei2p-like TEL (TERMINAL EAR1-like) proteins are only found in land plants and can be distinguished from the other Mei2p-like proteins by the presence of a 14–18-amino acid specific insertion within this third RRM (Anderson et al., 2004; Jeffares et al., 2004). In monocots, *TEL* genes were shown to regulate leaf initiation and development, as well as flowering transition and inflorescence development (Veit et al., 1998; Kawakatsu et al., 2006; Xiong et al., 2006), whereas the other *mei2*-like genes, like the *AMLs* (*Arabidopsis mei2*-like) in *A. thaliana*, mainly seem to play a role in meiosis like *mei2* in fission yeast (Hirayama et al., 1997; Kaur et al., 2006). Expression analysis confirmed a conserved association of *TEL* expression with tissue and organ initiation (Paquet et al., 2005; Charon et al., 2010), suggesting a putative role for TEL RBPs in the cellular mitosis-differentiation transition. *Sme2*-like npcRNAs appear not to exist in plants (C. Charon, unpublished data) and plant RNA Mei2p-like RBP partners still remain unknown. However, npcRNA sequences can diverge rapidly even between closely related species (Mercer et al., 2009). These re-localization mechanisms are depicted in Figure 2. In the case of the *A. thaliana* Mei2p-like proteins, their nucleo-cytoplasmic translocations were hypothesized from the re-localization of Mei2p in *S. pombe*, since the sub-cellular localization of the plant Mei2p-related proteins remains unknown.

ANTISENSE RNAs MAY MODIFY EPIGENETIC PATTERNS OF EXPRESSION AND ARE ACTIVELY DEGRADED IN PLANT CELLS

Long npcRNAs produced in the nuclei such as processed 'aberrant' RNAs (RNAs without cap or polyadenylated tail) or antisense RNAs can generate double-stranded RNAs and trigger silencing mechanisms through the action of DCLs, leading to the generation of heterochromatin-siRNAs or nat-siRNAs as mentioned above. However, other RNA quality-control mechanisms occur in the cell. For example, the non-sense-mediated mRNA decay (or NMD) is a quality-control mechanism related to cytoplasmic foci known as P-bodies that recognizes premature nonsense or stop codons (PTC) within an mRNA by the action of the exon-junction complex (EJC) that marks correctly fused exons (Pontes and Pikaard, 2008). After recognition of an incorrectly positioned stop codon, the NMD system signals the elimination of the

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

mRNAs through de-capping, de-adenylation, and exonucleolytic degradation (Amrani et al., 2006; Conti and Izaurralde, 2005; Lejeune and Maquat, 2005). The UP-frameshift proteins (UPFs) are essential for NMD and three *UPF* genes exist in *A. thaliana*. A genome-wide analysis of these mutants revealed that in addition to the expected NMD substrates, most npcRNAs, including large numbers of antisense RNAs, are degraded by this pathway, suggesting that one of the most important roles of NMD is the genome-wide suppression of aberrant or antisense RNAs (Kurihara et al., 2009). Interestingly, a proteomic approach using *Arabidopsis* nucleoli revealed the presence of several unexpected proteins, notably some components of the post-splicing EJC involved in mRNA export and NMD, proposing an additional function of the nucleolus in mRNA surveillance (Pendle et al., 2005). More recently, two components of the NMD were shown to localize in the nucleolus: the UPF2 and UPF3 proteins (Kim et al., 2009), further reinforcing a role for the nucleolus in these processes. Similarly, a genome-wide atlas of exosome substrates in *A. thaliana* revealed, in addition to mRNA and miRNA processing intermediates, hundreds of npcRNAs not previously described (Chekanova et al., 2007). The exosome is a macromolecular complex that mediates RNA processing and degradation, in both the nucleus and the cytoplasm, and is generally essential for viability in eukaryotes. These npcRNAs only detected in exosome mutants include large numbers of antisense RNAs as they are rapidly and actively degraded in wild-type plants; however, certain of them may play post-transcriptional regulatory roles. Recently, it has been elegantly shown that the targeted 3′ processing of antisense transcripts triggers *Arabidopsis* chromatin silencing at the locus encoding the major flowering repressor FLC (FLOWERING LOCUS C; Liu et al., 2010). In fact, FLC represses several major floral regulators, and vernalization, the regulation of flowering competence through cold exposure of plants, leads to the deposition of epigenetic marks in this locus to activate early flowering (Simpson et al., 2003). This epigenetic control results in *FLC* transcriptional silencing through the activities of two RBPs (FCA and FPA), a member of a 3′ RNA processing complex, and a histone demethylase (Liu et al., 2007; Simpson et al., 2003). A suppressor mutagenesis screen and a detailed analysis of FLC locus transcription revealed the 3′ processing of *FLC* antisense (but not sense) transcripts. A specific RBP directs the 3′ processing activities to a proximal antisense polyadenylation site, a targeted processing triggering locally histone demethylation and leading to *FLC* sense silencing (Liu et al., 2010, Figure 2). As the RBPs involved in this mechanism also silence transposons and transgenes in *Arabidopsis* (Bäurle et al., 2007), the 3′ processing of antisense transcripts may be a general mechanism triggering chromatin silencing in eukaryotes (Liu et al., 2010). These nuclear roles of npcRNAs can have consequences on the generation of novel patterns of gene expression and, as shown for the FLC locus, can trigger environmentally driven epigenetic changes.

CONCLUDING REMARKS

Even though many nuclear RBPs have been identified as having critical roles during the development and in epigenetic remodeling of chromatin, it is largely unclear how their action may control development, primarily due to the difficulty in identifying their RNA partners (Lorkovic, 2009). Most RBPs are likely to have multiple RNA partners such as mRNAs and npcRNAs (e.g. antisense RNAs, various 'aberrant' RNAs or mRNA-like npcRNAs) that may compete in the different RNPs to generate RNA networks in which npcRNAs can act as competitors or activators and determine their localization or action.

The large diversity of npcRNAs identified in eukaryotes supports the notion that npcRNAs are important to explain, at least in part, the complexity of multi-cellular organisms, since the total number of protein coding genes in diverse organisms (from a sea anemone to humans) varies much less than the number of different transcripts along evolutionary scales (Mattick and Makunin, 2006; Yasuda and Hayashizaki, 2008). It has then been suggested that such npcRNAs may modulate the resulting proteome from a specific transcriptome and increases genome plasticity (Mercer et al., 2009) and partly explain the large variations observed between these genomic approaches (ENCODE Project Consortium, 2007).

Plants exhibit a remarkable flexibility in their architecture and developmental patterns in response to external conditions, in contrast to animals, due to the continuously active growth of shoot and root meristems and their capability to generate new organs after embryogenesis (Wolters and Jürgens 2009). This developmental plasticity is a major characteristic of higher plants allowing individuals having the same genotype to give rise to different phenotypes, depending on environmental conditions. We suggest that npcRNAs by modulating RNP networks could be involved in this plasticity and significantly impact the outcome of the transcriptome, notably in response to abiotic stresses, in order to adapt growth and development to the environment. Riboregulation mediated by npcRNA–RBPs interactions is emerging as an important determinant of differentiation in eukaryotes. Several of these mechanisms may be linked to the active transcription taking place in the nuclei and may use the large amount of npcRNAs accumulating in this compartment to introduce flexibility to the system, notably adding an epigenetic dimension to the genome (Figure 2). Future challenges lie in understanding the interplay of these molecular mechanisms with the RNP networks to determine growth and developmental outcomes under different environmental conditions.

FUNDING

Our recent studies cited in this review were supported by grants from ANR (RNAPATHS) and from the Institut Fédératif de Recherches IFR87, Gif sur Yvette. F.B. was supported by a grant from the Ministère Français de l'Enseignement Supérieur et de la Recherche.

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

ACKNOWLEDGMENTS

We apologize to our colleagues whose work we were unable to cite due to page limitations. No conflict of interest declared.

REFERENCES

Amrani, N., Sachs, M.S., and Jacobson, A. (2006). Early nonsense: mRNA decay solves a translational problem. Nat. Rev. Mol. Cell Biol. **7,** 415–425.

Anderson, G.H., Alvarez, N.D.G., Gilman, C., Jeffares, D.C., Trainor, V.C.W., Hanson, M.R., and Veit, B. (2004). Diversification of genes encoding *mei2*-like RNA binding proteins in plants. Plant Mol. Biol. **54,** 653–670.

Barta, A., Kalyna, M., and Lorkovic, Z.J. (2008). Plant SR proteins and their functions. Curr. Top Microbiol. Immunol. **326,** 83–102.

Bäurle, I., Smith, L., Baulcombe, D.C., and Dean, C. (2007). Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. Science. **318,** 109–112.

Bei, Y., Pressman, S., and Carthew, R. (2007). SnapShot: small RNA-mediated epigenetic modifications. Cell. **130,** 756.

Ben Amor, B., et al. (2009). Novel long non-protein-coding RNAs involved in *Arabidopsis* differentiation and stress responses. Genome Res. **19,** 57–69.

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs derived from a pair of natural *cis*-antisense transcripts regulate salt tolerance in *Arabidopsis*. Cell. **29,** 1279–1291.

Brown, J.W., Marschall, D.F., and Echeverria, M. (2008). Intronic non-coding RNAs and splicing. Trends Plant Sci. **13,** 335–342.

Bühler, M. (2009). RNA turnover and chromatin-dependent gene silencing. Chromosoma. **118,** 141–151.

Burd, C.G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science. **265,** 615–621.

Campalans, A., Kondorosi, A., and Crespi, M. (2004). *Enod40*, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*. Plant Cell. **16,** 1047–1059.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.H., Sung, Z.R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. Development. **131,** 5263–5276.

Chapman, E.J., and Carrington, J.C. (2007). Specialization and evolution of endogenous small RNA pathways. Nat. Rev. Genet. **8,** 884–896.

Charon, C., Sousa, C., Crespi, M., and Kondorosi, A. (1999). Alteration of *enod40* expression modifies *Medicago truncatula* root nodule development induced by *Sinorhizobium meliloti*. Plant Cell. **11,** 1953–1966.

Charon, C., Vivancos, J., Mazubert, C., Paquet, N., Pilate, G., and Dron, M. (2010). Structure and vascular tissue expression of duplicated *TERMINAL EAR1*-like paralogues in poplar. Planta. **231,** 525–535.

Chekanova, J.A., et al. (2007). Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. Cell. **131,** 1340–1353.

Chen, X. (2009). Small RNAs and their roles in plant development. Annu. Rev. Cell Dev. Biol. **35,** 21–44.

Cioce, M., and Lamond, A.I. (2005). Cajal bodies: a long history of discovery. Annu. Rev. Cell Dev. Biol. **21,** 105–131.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pelligini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature. **452,** 215–219.

Conti, E., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Curr. Opin. Cell Biol. **17,** 316–325.

Crespi, M., and Gàlvez, S. (2000). Molecular mechanisms in root nodule development. J. Plant Growth Regul. **19,** 155–166.

Crespi, M., Jurkevitch, E., Poiret, M., d'Aubenton-Carafa, Y., Petrovics, G., Kondorosi, E., and Kondorosi, A. (1994). *enod40*, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. EMBO J. **13,** 5099–5112.

ENCODE Project Consortium (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. **447,** 782–783.

Fang, Y., and Spector, D.L. (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living *Arabidopsis* plants. Curr. Biol. **17,** 818–823.

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., García, J.A., and Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet. **39,** 1033–1037.

Girard, G., Roussis, A., Gultyaev, A.P., Pleij, C.W., and Spaink, H.P. (2003). Structural motifs in the RNA encoded by the early nodulation gene *enod40* of soybean. Nucleic Acids Res. **31,** 5003–5015.

Grewal, S.I., and Elgin, S.C. (2007). Transcription and RNA interference in the formation of heterochromatin. Nature. **447,** 399–406.

Handwerger, K.E., and Gall, J.G. (2006). Subnuclear organelles: new insights into form and function. Trends Cell Biol. **16,** 19–26.

Harigaya, Y., and Yamamoto, M. (2007). Molecular mechanisms underlying the mitosis-meiosis decision. Chromosome Research. **15,** 523–537.

Harigaya, Y., Tanaka, H., Yamanaka, S., Tanaka, K., Watanabe, Y., Tsutsumi, C., Chikashige, Y., Hiraoka, Y., Yamashita, A., and Yamamoto, M. (2006). Selective elimination of messenger RNA prevents an incidence of untimely meiosis. Nature. **442,** 45–55.

He, H., et al. (2006). Profiling *Caenorhabditis elegans* non-coding RNA expression with a combined microarray. Nucleic Acids Res. **34,** 2976–2983.

Henderson, I.R., and Jacobsen, S.E. (2007). Epigenetic inheritance in plants. Nature. **447,** 418–424.

Hirayama, T., Ishida, C., Kuromori, T., Obata, S., Shimoda, C., Yamamoto, M., Shinozaki, K., and Ohto, C. (1997). Functional cloning of a cDNA encoding *Mei2*-like protein from *Arabidopsis thaliana* using a fission yeast pheromone receptor deficient mutant. FEBS Lett. **413,** 16–20.

Hirota, K., Miyoshi, T., Kugou, K., Hoffman, C.S., Shibata, T., and Ohta, K. (2008). Stepwise chromatin remodeling by a cascade of transcription initiation of non-coding RNAs. Nature. **456,** 130–134.

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

Hirsch, J., Lefort, V., Vankersschaver, M., Boualem, A., Lucas, A., Thermes, C., d'Aubenton-Carafa, Y., and Crespi, M. (2006). Characterization of 43 non-protein-coding mRNA genes in *Arabidopsis*, including the MIR162a-derived transcripts. Plant Physiol. **140,** 1192–1204.

Hobert, O. (2008). Gene regulation by transcription factors and microRNAs. Science. **319,** 1785–1786.

Höck, J., and Meister, G. (2008). The Argonaute protein family. Genome Biology. **9,** 210.

Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008). A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. Mol. Cell. **32,** 685–695.

Huettel, B., Kanno, T., Daxinger, L., Aufsatz, W., Matzke, A.J., and Matzke, M. (2006). Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. EMBO J. **25,** 2828–2836.

Huettel, B., Kanno, T., Daxinger, L., Bucher, E., van der Winden, J., Matzke, A.J., and Matzke, M. (2007). RNA-directed DNA methylation mediated by DRD1 and Pol IVb: a versatile pathway for transcriptional gene silencing in plants. Biochim. Biophys. Acta. **1769,** 358–374.

Hüttenhofer, A., and Vogel, J. (2006). Experimental approaches to identify non-coding RNAs. Nucleic Acids Res. **34,** 635–646.

Jamalkandi, S.A., and Masoudi-Nejad, A. (2009). Reconstruction of *Arabidopsis thaliana* fully integrated small RNA pathway. Funct. Integr. Genomics. **9,** 419–432.

Jay, F., Renou, J.P., Voinnet, O., and Navarro, L. (2010). Biotic stress-associated microRNAs: identification, detection, regulation, and functional analysis. Methods Mol. Biol. **592,** 183–202.

Jeffares, D.C., Phillips, M.J., Moore, S., and Veit, B. (2004). A description of the Mei2-like protein family; structure, phylogenetic distribution and biological context. Dev. Genes Evol. **14,** 149–158.

Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A., and Jin, H. (2007). A novel class of bacteria-induced small RNAs in *Arabidopsis*. Genes Dev. **21,** 3123–3134.

Kaur, J., Sebastian, J., and Siddiqi, I. (2006). The *Arabidopsis- mei2*-like genes play a role in meiosis and vegetative growth in *Arabidopsis*. Plant Cell. **18,** 545–559.

Kawakatsu, T., Itoh, J.-I., Miyoshi, K., Kurata, N., Alvarez, N., Veit, B., and Nagato, Y. (2006). *PLASTOCHRON2* regulates leaf initiation and maturation in rice. Plant Cell. **18,** 612–625.

Kim, S.H., Koroleva, O.A., Lewandowska, D., Pendle, A.F., Clark, G.P., Simpson, C.G., Shaw, P.J., and Brown, J.W.S. (2009). Aberrant mRNA transcripts and the nonsense-mediated decay proteins UPF2 and UPF3 are enriched in the *Arabidopsis* nucleolus. Plant Cell. **21,** 2045–2057.

Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci Signal. 3, doi: 10.1126/scisignal.2000568.

Kloc, A., and Martienssen, R. (2008). RNAi, heterochromatin and the cell cycle. Trends Genet. **24,** 511–517.

Kouchi, H., Takane, K., So, R.B., Ladha, J.K., and Reddy, P.M. (1999). Rice *ENOD40*: isolation and expression analysis in rice and transgenic soybean root nodules. Plant J. **18,** 121–129.

Kurihara, Y., et al. (2009). Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in *Arabidopsis*. Proc. Natl Acad. Sci. U S A. **106,** 2453–2458.

Lamond, A.L., and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. Nat. Rev. Mol. Cell Biol. **4,** 605–612.

Lejeune, F., and Maquat, L.E. (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Curr. Opin. Cell Biol. **17,** 309–315.

Li, C.F., Henderson, I.R., Song, L., Fedoroff, N., Lagrange, T., and Jacobsen, S.E. (2008). Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in. Arabidopsis thaliana. PLoS Genet. 4, e27.

Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y.V., Simon, W.-L.C., Lagrange, T., Pikaard, C.S., and Jacobsen, S.E. (2006). An ARGONAUTE4-containing nuclear processing centre colocalized with Cajal bodies in *Arabidopsis thaliana*. Cell. **126,** 93–106.

Li, X., Cassidy, J.J., Reinke, C.A., Fischboeck, S., and Carthew, R.W. (2009). A microRNA imparts robustness against environmental fluctuation during development. Cell. **137,** 273–282.

Liu, F., Marquardt, S., Lister, C., Swiezewski, S., and Dean, C. (2010). Targeted 3′ processing of antisense transcripts triggers *Arabidopsis* FLC chromatin silencing. Science. **327,** 94–97.

Liu, F., Quesada, V., Crevillén, P., Bäurle, I., Swiezewski, S., and Dean, C. (2007). The *Arabidopsis* RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate FLC. Mol. Cell. **28,** 398–407.

Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., and Ecker, J.R. (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. Cell. **133,** 523–536.

Llave, C. (2004). MicroRNAs: more than a role in plant development? Mol. Plant Pathol. **5,** 361–366.

Lorkovic, Z.J. (2009). Role of plant RNA-binding proteins in development, stress response and genome organization. Trends Plant Sci. **14,** 229–236.

Lukowiak, A.A., Narayanan, A., Li, Z.H., Terns, R.M., and Terns, M.P. (2001). The snoRNA domain of vertebrate telomerase RNA functions to localize the RNA within the nucleus. RNA. **7,** 1833–1844.

Ma, J., Yan, B., Qu, Y., Qin, F., Yang, Y., Hao, X., Yu, J., Zhao, Q., Zhu, D., and Ao, G. (2008). Zm401, a short-open reading-frame mRNA or non-coding RNA, is essential for tapetum and microspore development and can regulate the floret formation in maize. J. Cell Biochem. **105,** 136–146.

MacLean, D., Elina, N., Havecker, E.R., Heimstaedt, S.B., Studholme, D.J., and Baulcombe, D.C. (2010). Evidence for large complex networks of plant short silencing RNAs. PLoS One. **5,** e9901.

Margis, R., Fusaro, A.F., Smith, N.A., Curtin, S.J., Watson, J.M., Finnegan, E.J., and Waterhouse, P.M. (2006). The evolution and diversification of Dicers in plants. FEBS Lett. **580,** 2442–2450.

Martens, J.A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the *Saccharomyces cerevisiae SER3* gene. Nature. **429,** 571–574.

Mattick, J.S., and Makunin, I.V. (2006). Non-coding RNA. Hum. Mol. Genet. **15,** 17–29.

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

Matzke, M., Kanno, T., Daxinger, L., Huettel, B., and Matzke, A.J.M. (2009). RNA-mediated chromatin-based silencing in plants. Curr. Opin. Cell Biol. **21,** 367–376.

Mercer, T.R., Dinger, M.E., and Mattick, J.S. (2009). Long non-coding RNAs: insights into functions. Nat. Rev. Genet. **10,** 155–159.

Mosher, R.A., Schwach, F., Studholme, D., and Baulcombe, D.C. (2008). PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. Proc. Natl Acad. Sci. U S A. **105,** 3145–3150.

Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science. **322,** 1717–1720.

Paquet, N., Bernadet, M., Morin, H., Traas, J., Dron, M., and Charon, C. (2005). Expression patterns of *TEL* genes in Poaceae suggest a conserved association with cell differentiation. J. Exp. Bot. **56,** 1605–1614.

Pendle, A.F., Clark, G.P., Boon, R., Lewandowska, D., Lam, Y.W., Andersen, J., Mann, M., Lamond, A.I., Brown, J.W.S., and Shaw, P.J. (2005). Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. Mol. Biol. Cell. **16,** 260–269.

Penterman, J., Uzawa, R., and Fischer, R.L. (2007b). Genetic interactions between DNA demethylation and methylation in *Arabidopsis*. Plant Physiol. **145,** 1549–1557.

Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L. (2007a). DNA demethylation in the *Arabidopsis* genome. Proc. Natl Acad. Sci. U S A. **104,** 6752–6757.

Pontes, O., and Pikaard, C.S. (2008). siRNA and miRNA processing: new functions for Cajal bodies. Curr. Opin. Genet. Dev. **18,** 197–203.

Pontes, O., Li, C.F., Nunes, P.C., Haag, J., Ream, T., Vitins, A., Jacobsen, S.E., and Pikaard, C.S. (2006). The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell. **126,** 79–92.

Prasanth, K., and Spector, D. (2007). Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. Genes Dev. **21,** 11–42.

Rinn, J.L., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. **129,** 1311–1323.

Rohrig, H., Schmidt, J., Miklashevichs, E., Schell, J., and John, M. (2002). Soybean ENOD40 encodes two peptides that bind to sucrose synthase. Proc. Natl Acad. Sci. U S A. **99,** 1915–1920.

Ruiz-Ferrer, V., and Voinnet, O. (2009). Roles of plant small RNAs in biotic stress responses. Annu. Rev. Plant Biol. **60,** 485–510.

Sato, M., Shinozaki-Yabana, S., Yamashita, A., Watanabe, Y., and Yamamoto, M. (2001). The fission yeast meiotic regulator Mei2p undergoes nucleocytoplasmic shuttling. FEBS Lett. **499,** 251–255.

Shaw, P.J., and Brown, J.W. (2004). Plant nuclear bodies. Curr. Opin. Plant Biol. **7,** 614–620.

Shimada, T., Yamashita, A., and Yamamoto, M. (2003). The fission yeast meiotic regulator Mei2p forms a dot structure in the horsetail nucleus in association with the sme2 locus on chromosome II. Mol. Biol. Cell. **14,** 2461–2469.

Simpson, C.G., et al. (2010). Regulation of plant gene expression by alternative splicing. Biochem. Soc. Trans. **38,** 667–671.

Simpson, G.G., Dijkwel, P.P., Quesada, V., Henderson, I., and Dean, C. (2003). FY is an RNA 3' end-processing factor that interacts with FCA to control the *Arabidopsis* floral transition. Cell. **113,** 777–787.

Song, L., Han, M.-H., Lesicka, J., and Fedoroff, N. (2007). *Arabidopsis* primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. Proc. Natl Acad. Sci. U S A. **104,** 5437–5442.

Sunkar, R. (2010). MicroRNAs with macro effects on plant stress responses. Semin. Cell Dev. Biol. **14 April**.

Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. **20,** 515–524.

Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes Dev. **20,** 759–771.

Veit, B., Briggs, S.P., Schmidt, R.J., Yanofsky, M.F., and Hake, S. (1998). Regulation of leaf initiation by the *terminal ear1* gene of maize. Nature. **393,** 166–168.

Verdel, A., Vavasseur, A., Gorrec, M., and Touat-Todeschini, L. (2009). Common themes in siRNA-mediated epigenetic silencing pathways. Int. J. Dev. Biol. **53,** 245–257.

Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. Cell. **136,** 669–687.

Wan, X., Hontelez, J., Lillo, A., Guarnerio, C., van de Peut, D., Fedorova, E., Bisseling, T., and Franssen, H. (2007). *Medicago truncatula ENOD40-1* and *ENOD40-2* are both involved in nodule initiation and bacteroid development. J. Exp. Bot. **58,** 2033–2041.

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in *cis* to inhibit transcription. Nature. **454,** 126–130.

Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H.L. (1994). RNA directed *de novo* methylation of genomic sequences in plants. Cell. **76,** 567–576.

Watanabe, Y., and Yamamoto, M. (1994). *S. pombe* mei2p encodes an RNA-binding protein essential for premeiotic DNA synthesis and meiosis I, which cooperates with a novel RNA species meiRNA. Cell. **78,** 487–498.

Watanabe, Y., Shinozaki-Yabana, S., Chikashige, Y., Hiraoka, Y., and Yamamoto, M. (1997). Phosphorylation of RNA-binding protein controls cell cycle switch from mitotic to meiotic in fission yeast. Nature. **386,** 187–190.

Wilusz, J.E., Sunwoo, H., and Spector, D.L. (2009). Long non-coding RNAs: functional surprises from the RNA world. Genes Dev. **23,** 1494–1504.

Wolters, H., and Jürgens, G. (2009). Survival of the flexible: hormonal growth control and adaptation in plant development. Nat. Rev. Genet. **10,** 305–317.

Xie, Z., Khanna, K., and Ruan, S. (2010). Expression of MicroRNAs and its regulation in plants. Semin. Cell Dev. Biol. **16 April**.

Xiong, G.S., Hu, X.M., Jiao, Y.Q., Yu, Y.C., Chu, C.C., Li, J.Y., Qian, Q., and Wang, Y.H. (2006). *LEAFY HEAD2*, which encodes a putative RNA-binding protein, regulates shoot development of rice. Cell Research. **16,** 267–276.

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

Yamada, K., et al. (2003). Empirical analysis of transcriptional activity in the /*Arabidopsis*/ genome. Science. **302,** 842–846.

Yamashita, A., Watanabe, Y., Nukina, N., and Yamamoto, M. (1998). RNA-assisted nuclear transport of the meiotic regulator Mei2p in fission yeast. Cell. **95,** 115–123.

Yasuda, J., and Hayashizaki, Y. (2008). The RNA continent. Adv. Cancer Res. **99,** 77–112.

Ying, S.Y., Chang, C.P., and Lin, S.L. (2010). Intron-mediated RNA interference, intronic microRNAs, and applications. Methods Mol. Biol. **629,** 205–237.

Zaratiegui, M., Irvine, D.V., and Martienssen, R.A. (2007). Noncoding RNAs and gene silencing. Cell. **128,** 763–776.

Zhan, S., and Lukens, L. (2010). Identification of novel miRNAs and miRNA dependent developmental shifts of gene expression in *Arabidopsis thaliana*. PloS One. **5,** e10157.

Zheng, X., Pontes, O., Zhu, J., Miki, D., Zhang, F., Li, W.-X., Iida, K., Kapoor, A., Pikaard, C.S., and Zhu, J.K. (2008). ROS3 is an RNA-binding protein required for DNA demethylation in *Arabidopsis*. Nature. **455,** 1259–1263.

Zhu, J., Kapoor, A., Sridhar, V.V., Agius, F., and Zhu, J.K. (2007). The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in *Arabidopsis*. Curr. Biol. **17,** 54–59.

Zhu, J.K. (2008). Epigenome sequencing comes of age. Cell. **133,** 395–397.

Downloaded from http://mplant.oxfordjournals.org at CNRS (I-20 DAW) on July 13, 2010

Annexe 4:

Bardou, F., Merchan, F., Ariel, F., and Crespi, M. (2011). Dual RNAs in plants. Biochimie.

Biochimie 93 (2011) 1950–1954

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Review

Dual RNAs in plants

Florian Bardou [a], Francisco Merchan [b], Federico Ariel [a], Martin Crespi [a,*]

[a] Institut des Sciences du Végétal, Centre National de la Recherche Scientifique, 91198 Gif sur Yvette, France
[b] Dto Microbiología y Parasitología, Facultad de Farmacia, Universidad de Sevilla, C/Profesor García González 2, 41012 Sevilla, Spain

ARTICLE INFO

Article history:
Received 5 April 2011
Accepted 25 July 2011
Available online 31 July 2011

Keywords:
Plant development
Dual RNA
Non-protein-coding RNA
Peptides
Antisense RNA

ABSTRACT

Plants have remarkable developmental plasticity, and the same genotype can result in different phenotypes depending on environmental variation. Indeed, abiotic stresses or biotic interactions affect organogenesis and post-embryonic growth and significantly affect gene regulation. The large diversity of non-protein-coding RNAs (npcRNAs) and genes containing only short open reading frames that are expressed during plant growth and development, contribute to the regulation of gene expression. Certain npcRNAs code for oligopeptides and may possess additional biological activity linked to the RNA moiety. The *ENOD40* gene is a dual RNA that is activated during a symbiotic interaction leading to root nodule organogenesis. Both the oligopeptides encoded by *ENOD40* and the structured regions of the *ENOD40* RNA have been shown to interact with different proteins in the cell to control enzymatic activities or induce the relocalisation of ribonucleoproteins, respectively. Other npcRNAs encode for small signalling peptides or are the precursors of small RNAs involved in post-transcriptional or transcriptional gene silencing. They may have RNA-related activities or encode peptides (or even larger proteins), and therefore act as dual RNAs. In addition, long natural antisense RNAs with a coding function and a regulatory RNA-mediated action that are expressed in response to abiotic stress in plants have been identified. In certain cases, these RNAs lead to the synthesis of nat-siRNAs, that are small RNAs derived from the overlapping double-stranded RNA region of natural antisense RNAs, which facilitates the silencing of complementary mRNAs. Finally, the advent of deep sequencing technologies has identified a large number of non-protein-coding RNAs in plants, which could be a large reservoir for dual RNAs.

© 2011 Elsevier Masson SAS. All rights reserved.

1. Introduction

Over the last years, several non-protein-coding RNAs have been identified, and recently, several of these long RNAs have been shown to encode short peptides. Furthermore, protein-coding mRNAs can also exert other RNA-intrinsic functions, such as acting as antisense RNAs. In fact, during evolution, some genes have lost their protein-coding capacity following mutation or the insertion of transposable elements and, conversely, non-protein-coding RNAs (npcRNAs) or pseudogenes have gained the capacity to encode new proteins [1]. Several genes seem to have both an intrinsic RNA activity and protein-coding capacity; therefore, they may have dual activities and can be considered bi-functional RNAs. Generally, dual functional RNAs are RNAs that have two distinct mechanisms of action, e.g., a human small nucleolar RNA also has a microRNA-like function [2].

Many different npcRNAs give rise to a variety of small RNAs, such as microRNAs (miRNAs) and small interfering RNAs (siRNAs), which are RNA molecules less than 30 nucleotides long that repress gene expression at the transcriptional or post-transcriptional level. In some cases, these small RNA precursors may also encode proteins and, therefore, act as dual RNAs. Other fundamental mechanisms also depend on new classes of long npcRNAs whose sequence, secondary structure and interactions may allow them to interact with many ribonucleoproteins (RNPs) in the cell and affect their function. Furthermore, a significant number of antisense RNAs or NAT-RNAs (for Natural Antisense RNAs), which are transcripts complementary to *cis* or *trans* mRNAs that repress the expression of the transcripts to which they bind, have been identified. Several NAT-RNAs also have coding capacity, and these transcripts may also act as dual RNAs, depending on the timing and conditions in the cell. Hence, several of these long npcRNAs encode oligopeptides that may be translated under specific conditions and participate in signalling or endogenous regulatory mechanisms. The recent evolution of sequencing technologies has led to the discovery of large numbers of npcRNAs that can act either an inhibitory RNA (small RNAs, antisense RNAs or long npcRNAs) and/or the potentially encoded peptides. However, in the plant kingdom, there are few studies examining npcRNAs and/or dual RNAs.

* Corresponding author. Tel.: +33169823703; fax: +33169823695.
E-mail address: martin.crespi@isv.cnrs-gif.fr (M. Crespi).

0300-9084/$ – see front matter © 2011 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.biochi.2011.07.028

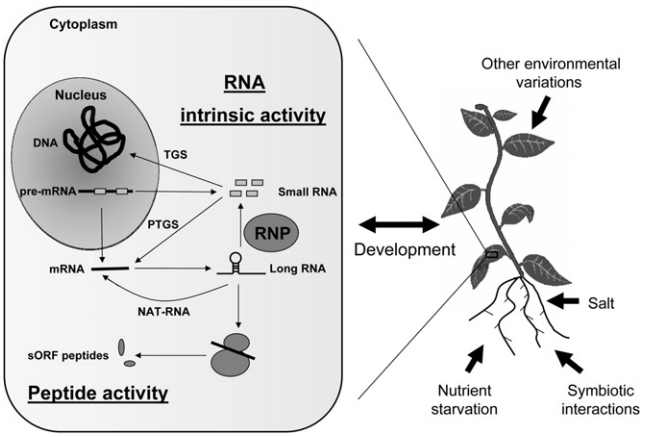


Fig. 1. Several dual RNAs participate in the modulation of plant development, which is highly sensitive to environmental conditions. Plant growth and development is highly influenced by environmental variations that have a strong impact on gene regulation. The response to these environmental variations, such as the interaction of roots with symbionts (symbiotic interactions), salt stress or nutrient starvation, has implicated dual RNAs. At the cellular level, pre-RNAs undergo different maturation steps, such as splicing, leading to the generation of mRNA. Dual RNAs can produce different small RNAs, mRNAs with a complex secondary structure or natural antisense RNAs (NAT-RNAs), which can regulate other mRNAs via integration into different ribonucleoproteins (RNP). Long non-protein-coding RNAs can additionally encode small peptides that can be translated in the ribosomes. Small RNAs and microRNAs associated with RNPs are able to induce gene silencing either transcriptionally in the nucleus (transcriptional gene silencing; TGS) or post-transcriptionally (PTGS). Hence, both the peptide and the intrinsic RNA activity may be associated with different classes of npcRNAs that modulate the expression of critical genes during plant development.

Plants are complex eukaryotic organisms that are highly sensitive to environmental variation, such as low or high temperatures, drought or nutrient starvation. Both biotic and abiotic stresses affect plant growth and development and contribute to their adaptation to specific environmental conditions. This so-called developmental plasticity allows plants with the same genotype to display different phenotypes depending on post-embryonic development in response to stress conditions. Plants are sessile organisms that constantly modulate their gene expression patterns in response to their environment, and several RNA-regulated mechanisms may be involved in this process. Indeed, transcriptional and post-transcriptional gene silencing (TGS and PTGS respectively) has been well documented in plants, and siRNAs were first discovered in plants. Here, we report some examples of bifunctional RNAs that have already been characterised in plants, and we also discuss several examples of antisense RNAs, long npcRNAs and short open reading frame (sORF)–encoded peptides, which may act in the future as dual RNAs Fig. 1.

2. *ENOD40*, a non-protein-coding RNA, may act as a dual RNA

The *ENOD40* genes code for highly structured plant mRNAs that contain a series of sORFs without any long ORF [3,4] and are involved in root nodule organogenesis, a process occurring essentially in legumes. Root nodules are nitrogen-fixing symbiotic plant organs that result from the interaction of soil bacteria of the genus *Rhizobium* with the roots of host legumes [5,6]. The development of root nodules depends on the coordinated expression of plant and bacterial genes, beginning with the induction of nodule formation by specific signals (Nod factors; [7,8]). This process initiates with cell-specific division in the roots, where *ENOD40* is strongly expressed [9,10]. The *ENOD40* gene is present in all legumes examined to date and are characterised by specific conserved nucleotide sequences that can be also found in some non-leguminous plants [4]. Furthermore, transgenic lines with increased or decreased levels of *ENOD40* exhibit accelerated or reduced nodulation, respectively [3]. It was first proposed that *ENOD40* was an npcRNA due to its highly stable RNA secondary structure, a characteristic of known non-coding RNAs [10,11] and due to the fact that only small peptides were encoded by these conserved transcripts. Furthermore, small peptides could also be produced using *in vitro* translation [12]. Translational analysis identified two sORFs (sORF I and II; 13 and 27 amino acids long, respectively) that could be translated from this transcript in *Medicago truncatula* [13]. sORF I contains a conserved nucleotide region that is also conserved in many legumes species, but not other plants, in contrast to the highly conserved stem-loops of the *ENOD40* RNA [14]. Hence, sORFs may play a role in symbiotic nodule formation. A cell-specific assay for the action of ENOD40 in *Medicago sativa* was developed using a biolistic process. Microtargeting of *ENOD40* specifically induced the division of root inner cortex cells 2 days after expression [13], which is consistent with the *in situ* localisation of *ENOD40* expression in this cell layer during nodule organogenesis [9]. The inner cortex cell layer corresponds to the first observable division during early nodule development. Mutations in the start codons of the each sORF strongly reduced *ENOD40* activity in this layer. Interestingly, mutations in the predicted structured RNA region also strongly inhibited induced cell division [13]. These results confirm the importance of both the sORF peptides and the RNA secondary structure of *ENOD40* in its activity and indicate that *ENOD40* encodes a bi-functional or dual RNA.

Data on the biological effects of *ENOD40* de-regulation suggest that the functions of this gene may not be restricted to the regulation of symbiosis. However, the molecular mechanisms underlying its activity are unclear. To gain further insight into the action of *ENOD40*, molecules that interact with the peptides or RNA were identified. A novel RNA-binding protein MtRBP1 (for *M. truncatula* RNA-Binding Protein1), which interacts with the *ENOD40* RNA, was identified using a yeast three-hybrid screen. Immunolocalisation studies and the use of an MtRBP1-DsRed2 fusion construct showed that MtRBP1 localised to nuclear "speckles," which are nuclear ribonucleoprotein complexes known to house the splicing machinery in plant cells. Interestingly, MtRBP1 was located in the cytoplasm of *ENOD40*-expressing cells in *M. truncatula* nodules. The direct involvement of the enod40 RNA in MtRBP1 relocalisation into cytoplasmic granules was confirmed using a transient expression assay and an MS2 bacteriophage system to tag the *ENOD40* RNA [15]. This *in vivo* approach to monitor RNA–protein interactions demonstrated that the cytoplasmic relocalisation of MtRBP1 was mediated by *ENOD40* and suggested that the relocalisation of nuclear RNA-binding proteins could be a new function mediated by npcRNAs [15]. However, the ENOD40 peptides expressed in soybeans were shown to bind to sucrose synthase (SUC1). This interaction suggests that the sORF peptides could play a role in the regulation of sucrose utilisation in the nodules [12]. *In vitro* experiments revealed that SUC1 possesses two phosphorylation sites that are targeted by a calcium-dependent protein kinase. Interestingly, the phosphorylation of SUC1 regulates its targeting to the proteasome for degradation, and the non-phosphorylated SUC1 is mainly soluble. The binding of both ENOD40 peptides to SUC1 inhibits its phosphorylation at one site, thereby inhibiting its degradation by the proteasome [16]. These data indicate that the ENOD40 peptides may regulate SUC1 turnover and proteolysis through an atypical regulatory mechanism [12,17]. These results further highlight that *ENOD40* genes may act as bi-functional RNAs in plants.

Many transcripts in plants may be dual RNAs as many genes contain potentially active sORF-encoded peptides. In *Arabidopsis,* more than 3000 sORFs are transcribed and/or under purifying selection, suggesting that huge numbers of sORF-encoded peptides are still hidden in genomic regions that have not yet been annotated [18]. For example, the POLARIS (PLS) gene in *Arabidopsis* was

identified experimentally using a promoter-trap approach and shown to have a root-specific expression pattern [19]. The PLS gene is first expressed in the basal region of embryos at the heart stage and is later expressed strongly in the seedling root and weakly in the vascular tissues of the lamina and petiole. *pls* mutants have short roots with radially expanded cells and reduced leaf vascularisation [20]. The PLS gene is transcribed as a relatively short, 500-nucleoide mRNA, which contains three short ORFs that encode putative peptides of 9, 8 and 36 amino acids. Over-expression of the ORF encoding the 36-amino acid peptide partially rescues the short-root phenotype, whereas over-expression of the 9- and 8-amino acid sORF products has not been analysed. Although the function of PLS has not been fully elucidated, it is involved in hormonal homeostasis, including ethylene signalling and auxin transport, and in the regulation of microtubule cytoskeletal dynamics [21]. Another example of a sORF-encoded active peptide is *ROTUNDIFOLIA4* (*ROT4*), which encodes a small, 53-amino acid peptide [22] with a conserved 29-amino acid region (the RTF domain) that is sufficient for ROT4 function. The rot4 mutant shows a peculiar leaf development phenotype that is similar to the over-expression phenotype [23]. However, to characterise the potential dual function of these genes, it will be necessary to examine the role of the mRNAs in the regulation of sORF translation.

In other cases, the sORF-encoded peptides present in the 5′ region of mRNAs have been shown to act in *trans* to regulate other processes independently of the encoded long ORF and can, therefore, be considered bi-functional RNAs. The *MtHAP2-1* gene is a CCAAT-binding transcription factor from *M. truncatula* that potentially plays a key role during nodule development by controlling nodule meristem function [24]. *MtHAP2-1* is regulated by miRNA169 during nodule differentiation. However, a novel regulatory mechanism controlling *MtHAP2-1* expression has been identified [25]. Alternative splicing of an intron in the *MtHAP2-1* 5′ leader sequence generates the predominant mRNA isoform present during nodule development. This isoform facilitates the production of a small peptide, uORF1p, that can bind the *MtHAP2-1* 5′ leader sequence and reduce the accumulation of the *MtHAP2-1* transcript. This regulatory mechanism contributes to the spatial restriction of *MtHAP2-1* expression within the nodule meristem. Hence, uORF1p is able to act in *trans* to down-regulate the mRNA accumulation of a specific mRNA isoform, suggesting a dual function for the *MtHAP2-1* gene.

3. Natural antisense transcripts in gene regulation

Natural antisense transcripts (NATs) are another candidate dual RNA, as they can both encode proteins and regulate complementary RNA transcripts. They have been identified in multiple eukaryotes, including humans, mice, yeasts and plants. This class of RNAs can be grouped into two classes: *cis*-NATs, which are generated by antisense transcription at the same genomic locus, and *trans*-NATs, which are generated from different loci. Large-scale genome projects have revealed the common occurrence of overlapping gene pairs in most species analysed [26–28]. In the model plant *Arabidopsis thaliana*, there are discrepancies in the number of NATs that likely result from the use of alternative methods to map NATs onto the genome and from the changes in gene annotation introduced in different genome releases. Wang et al. [27] identified 1340 potential *cis*-NAT pairs in *Arabidopsis* and confirmed the expression of sense and antisense transcripts of 957 *cis*-NAT pairs. Another screening of the *Arabidopsis* genome for protein-coding genes with overlapping orientations identified 1083 *cis*-NAT pairs [29]. However, a genome-wide screening for *trans*-NATs *in silico* identified 1320 putative *trans*-NAT pairs [30]. Interestingly, a large number of transcripts were predicted to have both *trans*- and *cis*-NATs, suggesting that antisense transcripts can form a complex regulatory network [28].

Overlapping transcripts can comprise two protein-coding transcripts, one protein-coding transcript and one non-coding transcript, or two non-coding transcripts. In the first case, each transcript from the two protein-coding and overlapping genes has a dual function, where one gene acts as a sense mRNA and an antisense transcript to the overlapping partner gene, potentially interfering with transcription or translation. In *A. thaliana*, a *cis*-NAT pair, which encodes SRO5 and P5CDH, was shown to not only negatively regulate translation, but to also antagonise *Arabidopsis* salt tolerance [31]. P5CDH is constitutively expressed and encodes Δ1-pyrroline-5-carboxylate dehydrogenase, which prevents proline accumulation, whereas SRO5 is induced by salt stress and encodes an unknown protein. In *Arabidopsis* plants growing in high NaCl conditions, both genes form a natural double-stranded pair of transcripts that are cleaved by DCL2 and DCL1 to generate 24- and 21-nucleotide nat-siRNAs that cleave the constitutively expressed P5CDH mRNA, resulting in salt tolerance. SRO5 could also play a role in counteracting the increased production of reactive oxygen species (ROS) caused by decreased P5CDH activity, either by blocking ROS production or increasing ROS detoxification [31].

The *Sho* locus in *Petunia hybrida* is another example of a NAT-RNA with dual functions. *Sho* encodes an enzyme responsible for the synthesis of plant cytokinins [32], and its 3′ region contains a promoter in the opposite orientation of a NAT that is induced by high cytokinin concentrations. This NAT encodes a 639-bp ORF that does not resemble any known protein in the database, although we cannot exclude a possible function of this protein in the cell. This NAT seems to participate in a tissue-specific manner to adjust local cytokinin synthesis via degradation of *Sho* through the formation of a double-stranded RNA (dsRNA). Similarly, in barley (*Hordeum vulgare*), the levels of siRNAs derived from the 3′-protein-coding region of the HvCesA6 gene increased markedly, and this increase was concomitant with a decrease in the expression of *CesAs*, *Csls*, and a *GT8* glycosyl transferase [33]. An *HvCesA6* antisense transcript acts in *cis* and *trans* on other related *Csl* transcripts through nat-siRNAs generated after the cleavage of the dsRNA formed by *HvCesA6* and its natural antisense transcript. This down-regulation could be a key requirement for the transition to new wall biogenesis programs during differentiation.

Apparently, antisense transcript functions do not always lead to the formation of nat-siRNAs, as the majority of *Arabidopsis cis*-NATs do not generate the small RNAs reported in databases [28]. Furthermore, by characterising the organisation and expression profiles of 956 convergent overlapping protein-encoding *A. thaliana* gene pairs, Jen et al. demonstrated that both transcripts are present in the same tissue and that the overlapping protein-coding gene pool is not significantly depleted under a variety of conditions [29]. However, as occurs in mammalian cells, at least for a subgroup of plant *cis*-NATs (with coding capacity), antisense expression may induce alternative splicing or polyadenylation [34]. This observation could indicate a third role for antisense transcripts in alternative splicing and polyadenylation apart from RNA degradation through dsRNA formation.

In many cases, the functions or coding capacities of NAT-RNAs remain largely unknown. In rice, three PHO1 genes generate *cis*-NATs and are all non-protein-coding. It is striking that while the *cis*-NAT associated with the OsPHO1;3 gene does not appear to be regulated either developmentally or in response to Pi deficiency, the *cis*-NATs of OsPHO1;1 and OsPHO1;2 are strongly up-regulated by Pi deficiency, while the expression of the complementary sense transcript remains relatively stable [35]. However, because the analysis was performed at the whole organ level, the expression of the sense and antisense OsPHO1 transcripts may not occur in the

same cells or tissues in all cases. In *A. thaliana*, a search for npcRNAs [36] allowed us to identify 13 antisense npcRNAs complementary to protein-coding transcripts. One of these (npc536) forms a *cis*-NAT with AT1G67930, and its over-expression allowed plants to grow under salt stress without modifying AT1G67930 mRNA accumulation. Furthermore, *npc536* mutants do not show misregulation of the antisense transcript. As this NAT contains a sORF conserved in rice, npc536 may act through this encoded peptide. Alternatively, npc536 may regulate translation of the AT1G67930 mRNA or act as a *trans*-NAT, with an unidentified complementary target that plays a role in the salt stress response.

In monocots, another class of *cis*-NAT pairs can generate a class of miRNAs called nat-miRNAs (for natural antisense miRNAs). To date, all nat-miRNAs have large introns in their precursors [37], and this intron seems to play an essential role to limit the potential base pairing of the pre-nat-miRNA with the sense transcript to better favour hairpin formation. The mature nat-miRNAs subsequently direct the cleavage of the target sense transcripts. Although the transcripts at the RNA level only share a small overlapping region, these *cis*-NAT pairs may have an miRNA-like or an antisense RNA-like activity.

4. Long non-protein-coding RNAs in plants: a potential reservoir of dual RNAs

Non-protein-coding RNAs are a class of RNAs with poor protein-coding potential, suggesting that there is a function associated with the RNA molecule itself. They form a heterogeneous group of RNAs that can be divided into three different classes based on their length and function: 18–25 nucleotides for microRNAs and small interfering RNAs linked to post-transcriptional and transcriptional gene silencing, 20–300 nucleotides for small RNAs commonly found as transcriptional and translational regulators (e.g., small nuclear RNAs), or several hundreds to over 10,000 nucleotides for medium and large npcRNAs that are mainly linked to epigenetic mechanisms. In recent years, bioinformatics and experimental strategies have revealed a remarkable number of novel npcRNA candidates in various organisms from yeasts to *Homo sapiens*, including plants [38–40]. In fact, transcriptomes have been found to be surprisingly complex, with long npcRNAs often overlapping with or interspersed between coding transcripts. Considering that a single DNA sequence can be transcribed as multiple sense and antisense transcripts, intronic npcRNAs, and intergenic or promoter-associated RNAs [40], this dynamic molecular picture significantly changes our understanding of gene expression in eukaryotes. In *A. thaliana*, whole-genome mapping based on the use of tiling arrays revealed that over 30% of the observed transcription was intergenic, including many antisense RNA transcripts [41].

The npcRNA-associated transcriptome complexity has been hypothesised to play a regulatory role that is required for the development and function of higher organisms [42]. Certain npcRNAs have been implicated in different regulatory mechanisms in plant development [43,44], in environmental biotic interactions and in the response to abiotic stresses [36,45,46]. These transcripts are generally produced by RNA polymerase II and are predominantly capped and polyadenylated, suggesting that they may lead to the production of sORF-encoded peptides under specific conditions. Several long npcRNAs are processed into small RNAs due to their folding as dsRNA loops derived from endogenous loci (the miRNAs) or due to the action of RNA-dependent RNA polymerases that generate long dsRNAs that are processed into siRNAs [47]. It is well known that small si/miRNAs induce mRNA cleavage and translational inhibition by pairing with specific mRNA targets, mainly in the cytoplasm, or lead to transcriptional gene silencing, heterochromatin formation and de novo DNA methylation in the nucleus [47–49]. However, there are specific npcRNAs that may have dual functions. npc78 is a structured RNA generated via alternative splicing, and one of these spliced introns contains MIR162a, a MIR162 isoform. This miRNA specifically targets the DICER1 RNA (DCL1) responsible for the synthesis of all miRNAs, thereby creating a negative feedback loop [50]. A bioinformatic analysis of npcRNA78 revealed the presence of five sORF that generate 115-, 76-, 59-, 45- and 16-residue peptides, and we cannot discount that this npcRNA, or any of its different spliced isoforms, may be translated under specific conditions. Hence, it has the potential to encode a protein/peptide and a miRNA and, therefore, to act as a dual RNA.

Among the best-studied non-coding RNAs in animals, several have been detected in the nucleus and act either as *cis*- or *trans*-acting epigenetic regulators of chromatin structure [51–54] by recruiting chromatin-remodelling complexes to specific genomic loci [55]. In *Arabidopsis*, cold temperatures result in the accumulation of tri-methylated (Lys 27) histone H3 at the floral repressor, FLOWERING LOCUS C (FLC), resulting in epigenetically stable repression of FLC. A long intronic npcRNA (designated COLDAIR) physically associates with a component of Polycomb Repressive Complex 2 (PRC2) and targets it to FLC, allowing the vernalisation-mediated epigenetic repression of the locus [56]. In addition, a group of related antisense npcRNAs (termed COOLAIR) from FLC has been proposed to be involved in vernalisation-mediated FLC repression [57]. Again these transcripts are polyadenylated and capped and can be translated in specific conditions. Although several of these long npcRNAs have been experimentally identified in plants [3,36,58–60], their activity, subcellular localisation and molecular roles remain largely unknown.

In recent years, it has become clear that RNAs can not only encode proteins, but can also exert a wide range of molecular functions, including the modulation of mRNA expression and RNA processing or localisation; the regulation of protein activity and structure; and as precursors to small RNAs or sORF-encoded peptides. The remarkable size of the non-protein-coding transcriptome indicates that many of the roles played by these RNAs may remain largely unknown. The few examples of RNAs having dual roles suggest that future work will likely identify many more transcripts that fit into this category. This will likely occur in plants whose developmental plasticity in response to environmental changes depends entirely on molecular flexibility in gene expression.

References

[1] D. Ulveling, C. Francastel, F. Hubé, When one is better than two: RNA with dual functions, Biochimie 93 (2011) 633–644.
[2] C. Ender, A. Krek, M.R. Friedlander, M. Beitzinger, L. Weinmann, W. Chen, S. Pfeffer, N. Rajewsky, G. Meister, A human snoRNA with microRNA-like functions, Mol. Cell 32 (2008) 519–528.
[3] C. Charon, C. Sousa, M. Crespi, A. Kondorosi, Alteration of enod40 expression modifies *Medicago truncatula* root nodule development induced by *Sinorhizobium meliloti*, Plant Cell 11 (1999) 1953–1966.
[4] A.P. Gultyaev, A. Roussis, Identification of conserved secondary structures and expansion segments in *ENOD40* RNAs reveals new *ENOD40* homologues in plants, Nucleic Acids Res. 35 (2007) 3144–3152.
[5] G.E. Oldroyd, J.A. Downie, Coordinating nodule morphogenesis with rhizobial infection in legumes, Annu. Rev. Plant Biol. 59 (2008) 519–546.
[6] M. Crespi, F. Frugier, De novo organ formation from differentiated cells: root nodule organogenesis, Sci. Signal. 1 (2008) re11.
[7] J. Denarié, J. Cullimore, Lipo-oligosaccharide nodulation factors: a new class of signaling molecules mediating recognition and morphogenesis, Cell 74 (1993) 951–954.
[8] M. Schultze, A. Kondorosi, Regulation of symbiotic root nodule development, Annu. Rev. Genet. 32 (1998) 33–57.
[9] W.C. Yang, P. Katinakis, P. Hendriks, A. Smolders, F. de Vries, J. Spee, A. Van Kammen, T. Bisseling, H. Franssen, Characterization of *GmENOD40*, a gene showing novel patterns of cell-specific expression during soybean nodule development, Plant J. 3 (1993) 573–585.
[10] M. Crespi, E. Jurkevitch, M. Poiret, Y. d'Aubenton-Carafa, G. Petrovics, E. Kondorosi, A. Kondorosi, *ENOD40*, a gene expressed during nodule

organogenesis, codes for a non-translatable RNA involved in plant growth, EMBO J. 13 (1994) 5099–5112.
[11] I.L. Hofacker, M. Fekete, P.F. Stadler, Secondary structure prediction for aligned RNA sequences, J. Mol. Biol. 319 (2002) 1059–1066.
[12] H. Rohrig, J. Schmidt, E. Miklashevichs, J. Schell, M. John, Soybean ENOD40 encodes two peptides that bind to sucrose synthase, Proc. Natl. Acad. Sci. U S A 99 (2002) 1915–1920.
[13] C. Sousa, C. Johansson, C. Charon, H. Manyani, C. Sautter, A. Kondorosi, M. Crespi, 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 (2001) 354–366.
[14] G. Girard, A. Roussis, A.P. Gultyaev, C.W.A. Pleij, H.P. Spaink, Structural motifs in the RNA encoded by the early nodulation gene *ENOD40* of soybean, Nucleic Acids Res. 31 (2003) 5003–5015.
[15] A. Campalans, A. Kondorosi, M. Crespi, *ENOD40*, a short open reading frame–containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*, Plant Cell 16 (2004) 1047–1059.
[16] S.C. Hardin, G.Q. Tang, A. Scholz, D. Holtgraewe, H. Winter, S.C. Huber, Phosphorylation of sucrose synthase at serine 170: occurrence and possible role as a signal for proteolysis, Plant J. 35 (2003) 588–603.
[17] H. Röhrig, M. John, J. Schmidt, Modification of soybean sucrose synthase by S-thiolation with ENOD40 peptide A, Biochem. Biophys. Res. Commun. 325 (2004) 864–870.
[18] K. Hanada, X. Zhang, J.O. Borevitz, W.H. Li, S.H. Shiu, A large number of novel genes encoding small open reading frames in the intergenic regions of the *Arabidopsis thaliana* genome are transcribed and/or under purifying selection, Genome Res. 17 (2007) 632–640.
[19] J.F. Topping, K. Lindsey, Promoter trap markers differentiate structural and positional components of polar development in *Arabidopsis*, Plant Cell 9 (1997) 1713–1725.
[20] S.A. Casson, P.M. Chilley, J.F. Topping, I.M. Evans, M.A. Souter, K. Lindsey, The POLARIS gene of *Arabidopsis* encodes a predicted peptide required for correct root growth and leaf vascular patterning, Plant Cell 14 (2002) 1705–1721.
[21] P.M. Chilley, S.A. Casson, P. Tarkowski, N. Hawkins, K.L.-C. Wang, P.J. Hussey, M. Beale, J.R. Ecker, G.K. Sandberg, K. Lindseya, The POLARIS peptide of *Arabidopsis* regulates auxin transport and root growth via effects on ethylene signaling, Plant Cell 18 (2006) 3058–3072.
[22] N.N. Narita, S. Moore, G. Horiguchi, M. Kubo, T. Demura, H. Fukuda, J. Goodrich, H. Tsukaya, Over-expression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in *Arabidopsis thaliana*, Plant J. 38 (2004) 699–713.
[23] L. De Veylder, T. Beeckman, G.T. Beemster, J. de Almeida Engler, S. Ormenese, S. Maes, M. Naudts, E. Van Der Schueren, A. Jacqmard, G. Engler, D. Inzé, Control of proliferation, endoreduplication and differentiation by the *Arabidopsis* E2Fa-DPa transcription factor, EMBO J. 21 (2002) 1360–1368.
[24] J.P. Combier, F. de Billy, P. Gamas, A. Niebel, S. Rivas, *Trans*-regulation of the expression of the transcription factor MtHAP2-1 by a uORF controls root nodule development, Genes Dev. 22 (2008) 1549–1559.
[25] J.P. Combier, H. Küster, E.P. Journet, N. Hohnjec, P. Gamas, A. Niebel, Evidence for the involvement in nodulation of the two small putative regulatory peptide-encoding genes *MtRALFL1* and *MtDVL1*, Mol. Plant. Microbe. Interact. 21 (2008) 1118–1127.
[26] R. Yelin, D. Dahary, R. Sorek, E.Y. Levanon, O. Goldstein, A. Shoshan, A. Diber, S. Biton, Y. Tamir, R. Khosravi, S. Nemzer, E. Pinner, S. Walach, J. Bernstein, K. Savitsky, G. Rotman, Widespread occurrence of antisense transcription in the human genome, Nat. Biotechnol. 21 (2003) 379–386.
[27] X.J. Wang, T. Gaasterland, N.H. Chua, Genome-wide prediction and identification of *cis*-natural antisense transcripts in *Arabidopsis thaliana*, Genome Biol. 6 (2005) R30.
[28] S.R. Henz, J.S. Cumbie, K.D. Kasschau, J.U. Lohmann, J.C. Carrington, D. Weigel, M. Schmid, Distinct expression patterns of natural antisense transcripts in *Arabidopsis*, Plant Physiol. 144 (2007) 1247–1255.
[29] C.H. Jen, I. Michalopoulos, D.R. Westhead, P. Meyer, Natural antisense transcripts with coding capacity in *Arabidopsis* may have a regulatory role that is not linked to double-stranded RNA degradation, Genome Biol. 6 (2005) R51.
[30] H. Wang, N.H. Chua, X.J. Wang, Prediction of *trans*-antisense transcripts in *Arabidopsis thaliana*, Genome Biol. 7 (2006) R92.
[31] O. Borsani, J. Zhu, P.E. Verslues, R. Sunkar, J.K. Zhu, Endogenous siRNAs derived from a pair of natural *cis*-antisense transcripts regulate salt tolerance in *Arabidopsis*, Cell 123 (2005) 1279–1291.
[32] E. Zubko, P. Meyer, A natural antisense transcript of the *Petunia hybrida Sho* gene suggests a role for an antisense mechanism in cytokinin regulation, Plant J. 52 (2007) 1131–1139.
[33] M.A. Held, B. Penning, A.S. Brandt, S.A. Kessans, W. Yong, S.R. Scofield, N.C. Carpita, Small-interfering RNAs from natural antisense transcripts derived from a cellulose synthase gene modulate cell wall biosynthesis in barley, Proc. Natl. Acad. Sci. U S A 105 (2008) 20534–20539.
[34] E. Zubko, A. Kunova, P. Meyer, Sense and antisense transcripts of convergent gene pairs in *Arabidopsis thaliana* can share a common polyadenylation region, PLoS ONE 6 (2011) e16769.
[35] D. Secco, A. Baumann, Y. Poirier, Characterization of the rice *PHO1* gene family reveals a key role for OsPHO1;2 in phosphate homeostasis and the evolution of a distinct clade in dicotyledons, Plant Physiol. 152 (2010) 1693–1704.
[36] B. Ben Amor, S. Wirth, F. Merchan, P. Laporte, Y. d'Aubenton-Carafa, J. Hirsch, A. Maizel, A. Mallory, A. Lucas, J.M. Deragon, H. Vaucheret, C. Thermes, M. Crespi, Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses, Genome Res. 19 (2009) 57–69.
[37] C. Lu, D.H. Jeong, K. Kulkarni, M. Pillay, K. Nobuta, R. German, S.R. Thatcher, C. Maher, L. Zhang, D. Ware, B. Liu, X. Cao, B.C. Meyers, P.J. Green, Genome-wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs), Proc. Natl. Acad. Sci. U S A 105 (2008) 4951–4956.
[38] J.S. Mattick, I.V. Makunin, Non-coding RNA, Hum. Mol. Genet. 15 (2006) 17–29.
[39] J. Yasuda, Y. Hayashizaki, The RNA continent, Adv. Cancer Res. 99 (2008) 77–112.
[40] T.R. Mercer, M.E. Dinger, J.S. Mattick, Long non-coding RNAs: insights into functions, Nat. Rev. Genet. 10 (2009) 155–159.
[41] K. Yamada, J. Lim, J.M. Dale, H. Chen, P. Shinn, Empirical analysis of transcriptional activity in the *Arabidopsis* genome, Science 302 (2003) 842–846.
[42] P. Carninci, RNA dust: where are the genes? DNA Res. 17 (2010) 51–59.
[43] J.W. Brown, D.F. Marshall, M. Echeverria, Intronic noncoding RNAs and splicing, Trends Plant Sci. 13 (2008) 335–342.
[44] O. Voinnet, Origin, biogenesis, and activity of plant microRNAs, Cell 136 (2009) 669–687.
[45] F. Jay, J.P. Renou, O. Voinnet, L. Navarro, Biotic stress-associated microRNAs: identification, detection, regulation, and functional analysis, Methods Mol. Biol. 592 (2010) 183–202.
[46] R. Sunkar, MicroRNAs with macro-effects on plant stress responses, Semin. Cell. Dev. Biol. 21 (2010) 805–811.
[47] H. Vaucheret, Post-transcriptional small RNA pathways in plants: mechanisms and regulations, Genes Dev. 20 (2006) 759–771.
[48] S.A. Jamalkandi, A. Masoudi-Nejad, Reconstruction of *Arabidopsis thaliana* fully integrated small RNA pathway, Funct. Integr. Genomics 9 (2009) 419–432.
[49] A. Verdel, A. Vavasseur, M. Le Gorrec, L. Touat-Todeschini, Common themes in siRNA-mediated epigenetic silencing pathways, Int. J. Dev. Biol. 53 (2009) 245–257.
[50] Z. Xie, K.D. Kasschau, J.C. Carrington, Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation, Curr. Biol. 13 (2003) 784–789.
[51] T. Nagano, J.A. Mitchell, L.A. Sanz, F.M. Pauler, A.C. Ferguson-Smith, R. Feil, P. Fraser, The air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin, Science 322 (2008) 1717–1720.
[52] R.R. Pandey, T. Mondal, F. Mohammad, S. Enroth, L. Redrup, J. Komorowski, T. Nagano, D. Mancini-Dinardo, C. Kanduri, Kcnq1ot1 antisense non-coding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation, Mol. Cell 32 (2008) 232–246.
[53] R. Terranova, S. Yokobayashi, M.B. Stadler, A.P. Otte, M. van Lohuizen, S.H. Orkin, A.H. Peters, Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos, Dev. Cell 15 (2008) 668–679.
[54] L. Redrup, M.R. Branco, E.R. Perdeaux, C. Krueger, A. Lewis, F. Santos, T. Nagano, B.S. Cobb, P. Fraser, W. Reik, The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing, Development 136 (2009) 525–530.
[55] J.L. Rinn, M. Kertesz, J.K. Wang, S.L. Squazzo, X. Xu, S.A. Brugmann, L.H. Goodnough, J.A. Helms, P.J. Farnham, E. Segal, H.Y. Chang, Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs, Cell 129 (2007) 1311–1323.
[56] J.B. Heo, S. Sung, Vernalization-mediated epigenetic silencing by a long intronic non-coding RNA, Science 331 (2010) 76–79.
[57] S. Swiezewski, F. Liu, A. Magusin, C. Dean, Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target, Nature 462 (2009) 799–802.
[58] J. Hirsch, V. Lefort, M. Vankersschaver, A. Boualem, A. Lucas, C. Thermes, Y. d'Aubenton-Carafa, M. Crespi, Characterization of 43 non-protein-coding mRNA genes in *Arabidopsis*, including the MIR162a-derived transcripts, Plant Physiol. 140 (2006) 1192–1204.
[59] L. Ma, B. Xie, Z. Hong, D.P.S. Verma, Z. Zhang, A novel RNA-binding protein associated with cell plate formation, Plant Physiol. 148 (2008) 223–234.
[60] J.M. Franco-Zorrilla, F.J. Del Toro, M. Godoy, J. Perez-Perez, I. Lopez-Vidriero, J.C. Oliveros, G. Garcia-Casado, C. Llave, R. Solano, Genome-wide identification of small RNA targets based on target enrichment and microarray hybridizations, Plant J. 59 (2009) 840–850.

Abstract:
In eukaryotes, several RNA binding proteins (RBPs) act on mRNA at various levels from splicing to translation. Recently a large number of non-protein coding RNAs (npcRNAs) have been identified in eukaryotes and shown to integrate into a variety of ribonucleoproteins (RNP) to control posttranscriptional gene expression. Our laboratory has identified a plant Nuclear-Speckle RBP (or NSR) that interacts with an npcRNA, *ENOD40* that accumulates during lateral root and nodule formation in legumes. NSR is relocalised into a cytoplasmic RNP in the *ENOD40*-expressing cells. During this PhD, we have analysed the role of NSRs in Arabidopsis thaliana and its link with npcRNAs.
Two AtNSR homologs from Arabidopsis thaliana, named AtNSRa and AtNSRb, code for proteins also localised in nuclear speckles together with certain splicing-related proteins. Interestingly, AtNSR-GFP fusions are relocalised into cytoplasmic granules in certain differentiated root cells and by ectopic expression of the *ENOD40* RNA. The AtNSRb gene is regulated by auxin whereas AtNSRa is constitutive. Root growth and lateral root formation of double nsra/nsrb mutants is partially insensitive to auxin. The localisation of these proteins prompted us to explore roles in splicing. No defects in general splicing were observed however analysis of 288 alternatively spliced genes in WT and nsra/nsrb roots in response to auxin revealed 77 changes in splicing profiles in response to auxin from which 51 required AtNSRs. In order to validate the interaction of NSRs with alternatively spliced mRNAs and npcRNAs, we have co-immunoprecipitated NSRs and identified at least 5 interacting alternatively spliced mRNAs and 2 npcRNAs. Expression of the *ENOD40* RNA or one interacting ncRNA modulate alternatively splicing in Arabidopsis. In a second chapter, we explored the role of NSRs in the modulation of PTGS triggered by intron-containing transgenes allowing us to link alternatively splicing and silencing. We propose that NSRs may link alternative splicing and the action of non-coding RNA, notably during root growth and development.

Résumé:
Chez les eucaryotes, plusieurs protéines liant l'ARN ou RBPs agissent sur l'ARNm à différents niveaux, de l'épissage à la traduction. Récemment, un grand nombre d'ARN non-codant des protéines (npcRNAs) ont été identifiés chez les eucaryotes et ont été montré comme interagissant avec une variété de ribonucléoprotéines (RNP) pour contrôler l'expression des gènes au niveau post-transcriptionnel. Nous avons identifié une Nuclear-Speckle RBP (ou NSR) qui interagit avec le npcRNA, *ENOD40*, un lncARN qui s'accumule au cours de la formation des racines latérales et des nodules chez les légumineuses. Durant cette thèse nous avons analysé le rôle des NSR d'*Arabidopsis thaliana* ainsi que leur lien avec les npcARN.
Deux gènes AtNSRs homologues existent chez Arabidopsis nommés *NSRa* et *NSRb*, ces gènes codent des protéines localisées dans des speckles nucléaires avec certaines protéines apparentées à l'épissage. Fait intéressant, les fusions AtNSR-GFP sont relocalisées dans des granules cytoplasmiques dans certaines cellules des racines différenciées ainsi que lors d'une co-expression éctopique de *ENOD40*. Le gène AtNSRb est régulé par l'auxine alors AtNSRa est constitutif. Les simples mutants Atnsr ne montrent pas de phénotype, mais la croissance des racines des doubles mutants est partiellement insensible à l'auxine, ce qui suggère une fonction redondante de ces protéines dans les racines. La localisation observée pour ces protéines nous a mené à explorer un rôle des NSRs dans l'épissage, nous avons donc analysé le profil d'épissage de 288 gènes en réponse à l'auxine chez Arabidopsis et comparé ces profils entre le WT et les mutants nsra/nsrb. Tout d'abord nous avons remarqué que l'épissage général ne variait pas, en revanche, l'analyse de 288 gènes alternativement épissés montre que le profil d'épissage de 77 gènes semble être modifié durant la réponse à l'auxine et 51 gènes nécessitent les protéines AtNSR pour ce changement. Afin de vérifier l'interaction des NSRs avec les cibles d'AS et avec les npcARN nous avons co-immunoprécipité les NSRs et nous avons identifié au moins 5 cible d'AS et 2 npcARN. L'expression de l'ARN *ENOD40* ainsi que du partenaire npcARN module L'AS chez Arabidopsis. Dans un deuxième chapitre, nous avons exploré le rôle des NSRs dans le PTGS déclenché par un transgène contenant un intron ce qui nous a permis de lier l'épissage alternatif et le silencing. Nous proposons donc que les NSRs pourraient lier l'épissage alternatif et l'action des ARN non codants, notamment lors de la croissance de la racine.