



HAL
open science

Diversité des voies de régulation des gènes chez les végétaux.

Natacha Bies-Etheve

► **To cite this version:**

Natacha Bies-Etheve. Diversité des voies de régulation des gènes chez les végétaux.. Botanique. Université de Perpignan Via Domitia (UPVD), 2016. tel-02194163

HAL Id: tel-02194163

<https://univ-perp.hal.science/tel-02194163>

Submitted on 26 Jul 2019

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.

UNIVERSITÉ
PERPIGNAN
VIA
DOMITIA



MÉMOIRE DE RECHERCHE

HABILITATION à DIRIGER les RECHERCHES

Natacha Bies-Ethève

Laboratoire Génome et Développement des Plantes

UPVD/UMR5096

24 Juin 2016

Membres du Jury:

Martin Crespi : Rapporteur
François Parcy : Rapporteur
Damien Garcia : Rapporteur
Christoph Grunau : Examineur
Jean-Marc Deragon : Examineur

REMERCIEMENTS

Je tiens tout d'abord à remercier Michel Delseny, de m'avoir accueillie dans son laboratoire et encadrée durant mon travail de Doctorat. J'ai ainsi pu m'initier au monde de la recherche et faire de très intéressantes rencontres scientifiques.

Au cours de mon stage post-doctoral, j'ai aussi bénéficié d'une collaboration très enrichissante avec Ralph Quatrano. Sa confiance et son professionnalisme m'ont beaucoup apporté. Je tiens également à remercier les enseignants de la Washington University qui m'ont permis de vivre une belle expérience en me recrutant pour assurer une partie des enseignements de biologie.

A mon retour en France, c'est avec beaucoup de plaisir que j'ai intégré le Laboratoire Génome et Développement des Plantes de l'Université de Perpignan. J'ai pu y retrouver un certain nombre de collègues mais aussi de nombreux nouveaux arrivants. Je tiens à remercier, parmi ces nouveaux arrivants, Thierry Lagrange de m'avoir accueillie dans sa jeune équipe et ainsi permis de réaliser le travail de recherche qui m'a occupé ces dernières années. Merci bien sûr à toutes les personnes de l'équipe "Régulation génique et RNA silencing chez les plantes", mais aussi à tous ceux qui travaillent dans le laboratoire.

Je tiens aussi à remercier tous les étudiants que j'ai pu encadrer. C'est toujours un réel plaisir de pouvoir partager, transmettre et guider de nouvelles vocations.

Enfin, je remercie chaleureusement Martin Crespi, François Parcy, Damien Garcia, Christoph Grunau et Jean-Marc Deragon de me faire l'honneur de participer à mon jury de soutenance.

SOMMAIRE

RESUME	p3
CURRICULUM VITAE	p6
1/ ACTIVITES D'ENCADREMENT ET D'ANIMATION DE LA RECHERCHE	p8
2/ ACTIVITES D'ENSEIGNEMENT ET TACHES ADMINISTRATIVES	p9
3/ PUBLICATION ET COMMUNICATIONS	p9
4/ ACTIVITE DE RECHERCHE DOCTORALE ET POST DOCTORALE: CONCERNANT DES GENES SPECIFIQUEMENT EXPRIMES AU COURS DE LA MATURATION DE LA GRAINE	p13
4-1/ Contexte scientifique et objectifs	p13
4-2/ Caractérisation des gènes <i>ATEM</i>	p13
4-3/ Caractérisation de la famille des gènes <i>LEA</i>	p23
5/ ACTIVITE DE RECHERCHE ACTUELLE AXEE SUR L'ETUDE DE MECANISMES DE "SILENCING" CHEZ LES VEGETAUX	p35
5-1/ Contexte scientifique: la voie du RdDM et les plateformes à protéines AGO	p35
5-2/ Objectifs, stratégie et motivations scientifiques.	p39
5-3/ Résultats.	p43
5-3-1/ Caractérisation du gène <i>AT5G04290 (SPT5like)</i>	p43
5-3-2/ Caractérisation du gène <i>AT1G65440 (SPT6like)</i>	p59
6/ CONCLUSIONS ET PROJETS	p71
7/ BIBLIOGRAPHIE	p82

RESUME

Ce manuscrit relate les principaux travaux de recherche effectués depuis l'obtention de mon diplôme de DEA, délivré par l'Université de Montpellier.

J'ai réalisé mon travail de recherche doctorale dans le laboratoire de Physiologie Végétale dirigé par Michel Delseny à l'Université de Perpignan. Ce travail a porté sur la caractérisation des conditions de l'accumulation des protéines EM, l'investigation de la fonction de ces protéines et la recherche de nouveaux mutants affectés dans l'expression des gènes *ATEM* (*AtEm6/AtLEA1-1* et *AtEm1/AtLEA1-2*) chez la plante modèle *Arabidopsis thaliana*. Les protéines EM ont initialement été décrites comme étant fortement accumulées au cours des dernières étapes de la maturation des graines. Les résultats obtenus par mon travail sont multiples et peuvent être résumés en quatre volets. Le premier concerne la réalisation d'un crible génétique en partie réalisé au Pays Bas dans le cadre d'un séjour de recherche et d'une collaboration avec Marteen Koornneef. Ce crible avait pour objectif d'identifier de nouvelles protéines impliquées dans la régulation des gènes *ATEM*. Cette approche a conduit à l'identification d'un nouvel allèle du mutant *abi3*, altéré dans la réponse à l'hormone végétale acide abscissique (Bies-Etheve et al., 1999) et également par celle, concomitante aux travaux d'un autre laboratoire, de la protéine ABI5 impliquée, elle aussi, dans la réponse à l'acide abscissique et la régulation des gènes *ATEM* (Carles et al., 2002). Le second volet a consisté à analyser l'expression des gènes *ATEM* dans des mutants altérés dans la maturation des graines. Cette approche a permis de montrer que l'importance des facteurs de transcription LEC1 et FUS3 (Vicent et al., 2000). Le troisième volet concerne la caractérisation détaillée des conditions d'expression des gènes *ATEM* d'une part et de l'accumulation de leurs produits d'autre part. Ce travail a mis en évidence une régulation post-traductionnelle qui empêche l'accumulation de protéines *ATEM* dans les tissus végétatifs (Hull et al., 1996 ; Bies et al., 1998). Un quatrième volet a consisté à produire des plantes transgéniques spécifiquement affectées dans la production des protéines *ATEM* de façon à obtenir des indices sur la fonction de ces protéines. Le fait qu'aucun phénotype révélateur n'ait pu être observé indique que les deux protéines AtLEA1-1 et AtLEA1-2 pourraient avoir un rôle redondant. Au cours de mon travail doctoral, j'ai eu par ailleurs la chance de bénéficier de la formation de moniteur. Cela m'a permis d'assister à différentes journées de formation à Montpellier mais aussi à me confronter à l'univers de l'enseignement à l'Université de Perpignan.

A la suite de l'obtention du diplôme de docteur, je suis partie en post-doc aux Etats Unis dans l'équipe dirigée par Ralph Quatrano tout d'abord à l'Université de Chapell Hill en Caroline du Nord puis à l'Université Washington à Saint Louis dans le Missouri. Mon activité de recherche a principalement porté sur la mise au point d'une technique permettant de visualiser *in vivo* les microfilaments dans les cellules via la production de plantes transgéniques produisant une protéine fusion ACTIN2::GFP. Les plantes générées ont montré un important problème de développement dont la caractérisation complète n'a pas pu être terminée avant mon retour en France. Cette partie n'est pas développée dans ce mémoire et sera brièvement abordée lors de la présentation orale. Au cours de mon séjour, mon expérience de recherche s'est enrichie avec un poste temporaire d'enseignante à l'université privée de médecine de Saint Louis.

En mai 2000, j'ai été recrutée comme maître de conférence à l'Université de Perpignan. J'ai alors rejoint l'équipe dirigée par Michel Delseny et j'ai focalisé mon travail sur l'identification de tous les gènes *LEA* (*Late Embryogenesis Abundant*) présents dans le génome de *Arabidopsis thaliana*. Cette analyse a permis la caractérisation de 50 gènes. L'étude des positions des différents gènes et celle des domaines répétés, ou non, des protéines correspondantes a permis de proposer un certain nombre d'hypothèses sur leur évolution. Une dernière analyse a consisté à étudier la régulation des gènes identifiés, en présence ou non d'acide abscisique et dans différents mutants affectés dans le développement de la graine. L'ensemble de ce travail a conduit à la publication d'un article (Bies-Etheve et al., 2008) et d'une revue (Delseny et al., 2001). Il faut noter que rapidement après mon recrutement il m'a été proposé d'être présidente des premières années de licence de BioEco de l'Université de Perpignan, tâche assez conséquente que j'ai assurée pendant 6 années.

En 2005, Michel Delseny a pris sa retraite et j'ai alors décidé d'intégrer une équipe de recherche qui venait d'être rattachée au laboratoire par le biais d'une ATIP. Les thématiques de recherche de cette équipe dirigée par Thierry Lagrange m'intéressaient particulièrement, même si elles étaient relativement différentes de ce que j'avais pu étudier auparavant. Il s'agissait de s'intéresser à des mécanismes nouveaux dits de silencing ou de répression des gènes dans lesquels les protéines ARGONAUTES (AGO) jouent des rôles majeurs. Le travail dans lequel je me suis rapidement intégrée consistait à étudier et à caractériser de nouvelles protéines identifiées comme étant potentiellement capables d'interagir avec des protéines AGO. La première protéine que j'ai caractérisée, la protéine SPT5like, présente des similarités de séquences avec SPT5 un facteur d'élongation de la transcription associé à l'ARN polymérase II. Chez *Arabidopsis*, il y a deux gènes qui codent des protéines SPT5 classiques et celui codant la protéine SPT5like qui se différencie par la présence d'une plateforme à AGO hook. Une première analyse a permis de confirmer la fonctionnalité de la plateforme en démontrant une interaction avec les protéines AGO4. Cette information a conduit à tester, dans des mutants d'insertion, les voies du RdDM (RNA directed DNA méthylation) qui impliquent AGO4 et à démontrer que SPT5like (KTF1) affecte la production de petits ARN (siARN) impliqués dans la méthylation ciblée de séquences répétées du génome (Bies-Etheve et al., 2009). La caractérisation approfondie de la plateforme AGO de SPT5like a de plus permis de montrer que cette plateforme et celle de la protéine NRPE1 (grande sous-unité de l'ARN polymérase V), impliquée elle aussi dans la même voie, sont redondantes. L'une de ces plateformes est suffisante pour recruter assez de protéines AGO4 pour assurer le recrutement des enzymes nécessaires à la méthylation de l'ADN. L'ensemble de ce travail, qui permet de proposer un nouveau modèle d'action du complexe POLV, est en cours de rédaction pour publication et une revue concernant les deux polymérases POLIV et POLV a été publiée (Lahmy et al., 2010). J'ai également en parallèle co-encadré le travail d'une étudiante en thèse Emilie Cavel. Le sujet de la thèse était "Caractérisation fonctionnelle de deux gènes nucléaires codant des variants du facteur général de transcription TFIIB chez *Arabidopsis thaliana*". Ce travail a fait l'objet d'une publication portant sur la protéine pBRP2 (Cavel et al., 2011).

Le dernier volet abordé concerne la caractérisation d'une seconde protéine, SPT6like codée par le gène *At1g65440*, qui a elle aussi été identifiée comme possédant potentiellement une plateforme à AGO hook. Cette protéine est comparable aux protéines SPT6 classiques qui ont été décrites comme se liant à la grosse sous-unité du complexe de polymérase II et étant impliquées, elles aussi dans l'élongation de la transcription. Les protéines SPT6 ont également été décrites comme capables de se lier à des histones. Mon analyse montre que la protéine SPT6like d'*Arabidopsis* assure probablement les mêmes fonctions que les protéines SPT6 classiques, car son

absence est létale. Un second gène codant une protéine similaire, mais sans plateforme est présent et actif uniquement dans les fleurs. L'inactivation de ce second gène semble n'avoir que peu d'effets sur les plantes. Des purifications de complexes par GST pull down, ainsi que des immunoprécipitations, m'ont permis de montrer que la protéine SPT6like est capable de lier les protéines AGO1. J'ai également pu observer que SPT6like est la première protéine, possédant une plateforme à AGO hook, capable de se lier au complexe de la polymérase II. Des mutant *Spt6like* complémentés avec une version SPT6like avec une plateforme intacte ou mutée (plateforme incapable de lier des protéines AGO) ne semblent montrer aucun phénotype particulier. Une analyse par séquençage d'ARN a été réalisée et indique un possible effet de la plateforme sur l'épissage alternatif de certains gènes. Une conclusion et différents projets d'analyses à court terme sur la poursuite de l'étude fonctionnelle de cette protéine sont présentés à la fin de ce rapport.

Enfin deux projets sont présenté en dernières pages du manuscrit. Ils concernent la caractérisation de nouvelles protéines à AGO hook. Le premier porte sur l'étude de protéines possédant en plus de la plateforme à AGO un domaine RRM de liaison aux ARN. Le second projet concerne des protéines qui appartiennent à la famille des AGD ou ArfGAP (Arf GAP Domain, GAP; GTPases activating protein).

CURRICULUM VITAE

Natacha Bies-Ethève

Née le 22-04-1968 à Mortagne Au Perche, Orne (61) ; **Nationalité** : Française

Situation : Divorcée, deux enfants, 13 et 16 ans à charge

Position actuelle : Maître de Conférence, équipe Régulation génique et RNA silencing chez les plantes.

Adresse professionnelle.

Laboratoire Génome et Développement des Plantes, UMR5096, CNRS

Université de Perpignan Via Domitia, 58 Av. Paul Alduy, 66860 Perpignan

tel : 04.68.66.22.26 ; 06.51.02.48.08

email : natacha.bies-etheve@univ-perp.fr

Formation:

- **1997 Doctorat** Sciences de la Vie et Santé, spécialité Bases de la production végétale option Biotechnologies et amélioration des Plantes, Université de Montpellier. Titre de la thèse "Caractérisation des conditions de l'accumulation des protéines Em, investigation sur la fonction de ces protéines et recherche de nouveaux mutants affectés dans l'expression des gènes *Em* chez *Arabidopsis thaliana*" diplôme obtenu avec la mention très honorable et les **félicitations du jury**.

- **1993 Diplôme d'Etudes Approfondies**: Bases de la production végétale option Biotechnologie et Amélioration des plantes. Université de Montpellier. **Mention B**.

- **1992 Maîtrise de sciences et Techniques**: Physiologie Végétale appliquée, Université de Montpellier. **Mention B**.

- **1991 Licence de Sciences Appliquées**: Physiologie végétale appliquée. Université de Montpellier. **Mention AB**

- **1990 Diplôme d'Etudes Universitaires Générales (DEUG B2)**: Option Biochimie, Université de Montpellier. **Mention AB**.

Activités de Recherche:

- **Depuis 2000. Maître de conférence rattachée au Laboratoire Génome et Développement des Plantes UMR 5096 à l'Université de Perpignan.**

- **de 2007 à ce jour. Rattachée à l'équipe de recherche dirigée par Thierry Lagrange. Régulation génique et RNA silencing chez les plantes.** Mon travail a principalement porté sur la caractérisations de deux protéines (SPT5like et SPT6like) d'*Arabidopsis thaliana* initialement identifiées comme étant potentiellement capables d'interagir avec des protéines Argonautes et donc potentiellement impliquées dans des mécanismes dits de répression de gènes ou "gene silencing".

- **de 2000 à 2007. Rattachée à l'équipe de recherche dirigée par Michel Delseny, directeur du laboratoire.** La thématique de l'équipe était la maturation de la graine et mon travail a principalement été axé sur l'inventaire des gènes *LEA* chez *Arabidopsis thaliana* et l'étude de leur régulation chez différents mutants affectés dans la maturation des graines ou la réponse à l'hormone végétale ABA.

- **de 1997 à 2000. Stage Post doctoral dans le laboratoire dirigé par le Pr. Ralph Quatrano à l'Université de North Carolina,Chapell Hill, Caroline du Nord USA de 1997 à 1999 puis à Washington University, Saint Louis, Missouri USA de 1999 à 2000.** Mon travail avait pour objectif de mettre au point une technique permettant de visualiser *in vivo* les microfilaments d'actine chez *Arabidopsis thaliana*, .

- **1993/1997. Stage de Doctorat dirigé par Michel Delseny dans le Laboratoire Génome et Développement des plantes UMR 5096 à l'Université de Perpignan.** Sujet: Caractérisation des conditions de l'accumulation des protéines Em, investigation sur la fonction de ces protéines et recherche de nouveaux mutants affectés dans l'expression des gènes *Em* chez *Arabidopsis thaliana*.
- **1993 Stage de DEA dirigé par Michel Delseny dans le Laboratoire Génome et développement des plantes UMR CNRS à l'Université de Perpignan.** Sujet: Production par voie recombinante et purification des protéines Em d'*Arabidopsis thaliana*.
- **1992 Rapport bibliographique dans le cadre de l'obtention du diplôme de maîtrise.** Sujet: Apports de la biologie moléculaire pour l'étude des effets des auxines sur l'expression du génome.
- **1991 Stage de Licence dirigé par Olivier Turc dans le laboratoire d'Agrophysiologie des légumineuses à l'INRA (LESCA) ENSA.M de Montpellier.** Sujet: Etude de la conséquence de stress hydriques sur la qualité des semences

Séjours de recherche effectués à l'étranger

- **2004 Mission à Taiwan.** Visite de nombreux laboratoires et mise en place de collaborations Perpignan/Taiwan.
- **2001 Séjour d'un mois à Taiwan** dans le cadre d'une collaboration avec Yue-Ie Hsing à Academia Sinica à Taipei.
- **1999 Séjour d'un mois à Cambridge en Angleterre** dans le cadre d'une collaboration avec Jim Hasselof.
- **1995 Séjour d'un mois à Wageningen aux Pays Bas** dans le laboratoire de génétique dirigé par Marteen Koornneef.

1- Activités d'encadrement et d'animation de la recherche

Activités d'encadrement et de formation par la recherche:

- **Mars 2015 Tutrice du stage de Licence 3 de Isabelle Sauquet.** Sujet: Régulation génique et gene silencing chez les plantes.
- **Avril 2014 Tutrice de deux Stages de Licence 3 STE effectués dans des laboratoires autres que le laboratoire génome et développement des plantes.**
- **Jane Jardine**, Département des neurosciences moléculaire, University college of London. Co-direction Dr. H Plun-Favreau. Sujet: Dysfonctionnement des mitochondries dans la dégénérescence neuronale, Implication de la voie PINK1-Parkin lors de la mitophagie.
- **Michel Kluba**, Laboratoire de Toxicologie, Université Montpellier 1. Co-direction Pr S Galas. Sujet: Étude des mécanismes induits par l'Acide Chicorique sur la durée de vie de *C. elegans*
- **Janvier-Juin 2013 Encadrement du stage de Master 2 BIMoPoDD Génomique Environnementale de Kelly Brener.** Sujet: Study of importance of SPT5-like and POLV WG-platform in the RDDM and its redundancy
- **Janvier-juin 2012 Encadrement du stage de Master 2 Dinev-GDP de Gwenole Prats.** Sujet: Etude d'une protéine à domaine WG: Spt6-like chez *Arabidopsis thaliana*
- **Avril 2012 Encadrement du stage de licence 3 de Alexandre Nore,** Sujet : Caractérisation de la protéine SPT6 d'*Arabidopsis thaliana*
- **Janvier-Juin 2011 Encadrement du stage de Master 2 Dinev-GDP de Hadjera Touchene.** Sujet: Characterization of a new WG-containing protein : SPT6-like in *Arabidopsis thaliana*
- **2010-2011 Tutrice d'un travail d'étude bibliographique de Elsa Habig et Bruno Miléna.** Sujet: Impact de l'environnement sur les marques épigénétiques.
- **Février-Avril 2010 Encadrement du stage de Master 1 Dinev-GDP de Hadjera Touchene.** Sujet: Production de facteurs de transcription recombinants dans le cadre de l'étude du variant pBrp2
- **2009-2010 Tutrice du travail d'étude bibliographique de Doberva Margot et Prats Gwenole.** Sujet: Rôle et fonction des ARN polymérases IV et V des végétaux.
- **2009 co-encadrement de la thèse de Emilie Cavel.** Caractérisation fonctionnelle de deux gènes nucléaires codant des variants du facteur général de transcription TFIIB chez *Arabidopsis thaliana*.
- **1999 Encadrement d'une étudiante undergraduate à l'université de Washington à Saint Louis Missouri, USA**

Activités d'évaluation de la recherche:

- **2010 à 2013, Membre du jury des stage de master 1 et 2, Dinev et BIMoPoDD.**
- **2012, Membre d'un comité de recrutement d'un maître de conférence**
- **2002, Rapportrice d'un comité de thèse à Poitiers.**

Financement des travaux de recherche.

- **Travail sur les LEA en partie financé par un Projet LEA (laboratoire associé européen)**
- **Travail sur les plateforme à AGO hook en partie finéncé par uneANR AGO hook**

Collaborations scientifiques:

Klaus Grasser, Steve Jacobsen, Rebecca Schawb, Jérôme Giraudat, Marteen Koornneef, Ralph Quatrano, David Ho, Yue-Ie Hsing.

2- Activités d'enseignement et tâches administratives

Activités d'enseignement:

- 2000 à ce jour. Enseignement en tant que Maître de conférence à l'Université de Perpignan. Responsable de différents modules de la licence au master.
- 2001 Enseignement temporaire à Academia Sinica à Taipei (Taiwan).
- 1999-2000 Chargée d'enseignement à Medical Washington University (Saint-Louis, U.S.A).
- 1996-1997 Poste ATER à l'Université de Perpignan. Biochimie et biologie moléculaire.
- 1993-1996 Monitorat à l'Université de Perpignan. Biochimie et biologie moléculaire.

Tâches administratives:

- 2016 Elue au conseil académique de l'université ainsi qu'au comité restreint de ce conseil.
- 2000 à ce jour. Participation aux jury restreints de formations de la licence de biologie de première année au master Dinev et BIMoPoDD de seconde année. Présidence de jury de BAC à plusieurs reprises.
- 2014 à ce jour, élue au comité technique de l'Université de Perpignan.
- 2014 à ce jour, élue au comité d'hygiène, de sécurité et des conditions de travail (CHSCT) de l'Université de Perpignan.
- 2009 à ce jour, Représentante des enseignants chercheurs au conseil de laboratoire.
- 2008 à ce jour, Présidente du CE de l'université de Perpignan après avoir été membre du bureau depuis 2002.
- 2005-2011, Présidente des premières années de licence bioeco et STE. Création des emplois du temps, répartition des étudiants en groupes en fonction de leur choix d'options et organisation des examens et jury d'examens.

3- Publications et communications

Publications scientifiques:

- Köllen, K., Dietz, L., **Bies-Etheve, N.**, Lagrange, T., Grasser, M., and Grasser, K.D. (2015). The zinc-finger protein SPT4 interacts with SPT5L/KTF1 and modulates transcriptional silencing in Arabidopsis. **FEBS Lett.** 589, 3254–3257.
- Cavel, E., Pillot, M., Pontier, D., Lahmy, S., **Bies-Etheve, N.**, Vega, D., Grimanelli, D., and Lagrange, T. (2011). A plant-specific transcription factor IIB-related protein, pBRP2, is involved in endosperm growth control. **PLoS One** 6, e17216.
- Lahmy, S., **Bies-Etheve, N.**, and Lagrange, T. (2010). Plant-specific multisubunit RNA polymerase in gene silencing. **Epigenetics** 5, 4–8.
- **Bies-Etheve, N.**, Pontier, D., Lahmy, S., Picart, C., Vega, D., Cooke, R., and Lagrange, T. (2009). RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. **EMBO Rep.** 10, 649–654. **Article cité deux fois par le Faculty of 1000 Biology et ayant fait l'objet d'un littérature report dans EMBO reports avril 2009.**
- **Bies-Ethève N.**, Gaubier-Comella P., Debures A., Lasserre E., Jobet E., Raynal M., Cooke R. and Delseny M. (2008). Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. **Plant Molecular Biology**.
- Carles, C., **Bies-Etheve, N.**, Aspart, L., Leon-Kloosterziel, K.M., Koornneef, M., Echeverria, M.,

and Delseny, M. (2002). Regulation of *Arabidopsis thaliana* Em genes: role of ABI5. *Plant J. Cell Mol. Biol.* *30*, 373–383.

- Delseny, M., **Bies-Etheve, N.**, Carles, C., Hull, G., Vicient, C., Raynal, M., Grellet, F., and Aspart, L. (2001). Late Embryogenesis Abundant (LEA) protein gene regulation during *Arabidopsis* seed maturation. *J. Plant Physiol.* *158*, 419–427.

- **Bies-Etheve, N.**, da Silva Conceicao, A., Giraudat, J., Koornneef, M., Leon-Kloosterziel, K., Valon, C., and Delseny, M. (1999). Importance of the B2 domain of the *Arabidopsis* ABI3 protein for Em and 2S albumin gene regulation. *Plant Mol. Biol.* *40*, 1045–1054.

- Vicient, C.M., **Bies-Etheve, N.**, and Delseny, M. (2000). Changes in gene expression in the leafy cotyledon1 (*lec1*) and *fusca3* (*fus3*) mutants of *Arabidopsis thaliana* L. *J. Exp. Bot.* *51*, 995–1003.

- **Bies, N.**, Aspart, L., Carles, C., Gallois, P., and Delseny, M. (1998). Accumulation and degradation of Em proteins in *Arabidopsis thaliana*: evidence for post-transcriptional controls. *J. Exp. Bot.* *49*, 1925–1933.

- Hull, G.A., **Bies, N.**, Twell, D., and Delseny, M. (1996). Analysis of the promoter of an abscisic acid responsive late embryogenesis abundant gene of *Arabidopsis thaliana*. *Plant Sci.* *114*, 181–192.

Participation à des congrès:

- Présentation de posters.

- **Bies-Ethève N.**, Hadjera TOUCHENE H., Michèle Laudie M., Picard C. and Lagrange T. Small RNAs and Polymerase II in plants. Initiation au traitement des données de génomique obtenues par séquençage à haut débit. Du 5 au 10 octobre 2014 à Roscoff.

- **Bies-Etheve N.**, Pontier D., Lahmy S., Picart C., Vega D., Cooke R. and Lagrange T. 2012 Saint Louis. Characterization of SPT5LIKE, an AGO4 interacting transcription elongation factor involved in RNA directed DNA methylation. IX International Plant Molecular Biology Congress (IPMB). 25-30 Octobre 2009. Saint Louis, USA.

- **Bies-Ethève N.**, Raynal M., Cooke R., Aspart L., Gaubier P. and Delseny M. Late embryogenesis abundant (*LEA*) protein genes in *Arabidopsis thaliana*. 7th International Congress of Plant Molecular Biology, 23-28 juin 2003, Barcelone, Espagne.

- **Bies-Ethève N.**, Comella P., Anne Debures and Michel Delseny. A post genomic study on *Arabidopsis* Late Embryogenesis Abundantly (*LEA*) expressed genes: inventory, classification and profiling. Gatersleben, Allemagne.

- Communications orales

- **Bies-Ethève N.**, Comella P. et Delseny M. **La transduction du signal ABA dans la graine, Etude du facteur de transcription ABI5.** 2003, Montpellier.

- **Bies-Ethève N.**, Comella P. et Delseny M. **Etudes de la fonction et de la régulation des LEA** (Late embryogenesis abundant protein). 2002 Canet,

- **Bies-Etheve N.** **Effects of cytoskeleton alterations.** 2000, Rencontres Franco Espagnoles, Sitges.

- **Bies-Ethève N.** **Effects of actin alterations.** 1999, Nashville, USA .

Participation à des rencontres ou formations.

- Ecole de bioinformatique (AVIESAN), **Initiation au traitement des données de génomique obtenues par séquençage à haut débit.** Du 5 au 10 octobre 2014 à Roscoff.

- **Meeting Chromatin Meets South**, du 23 au 28 juin 2014, Montpellier.

- **International meeting IFR83. Epigenetics: From Bases to Pathology.** 12-14 oct, 2011, Paris.

ATEM1 (152 acides aminés)

MASKQLSREEEKAKQGETVVPGGTGGHSLEAQEHLAEGRSKGGQTRKEQLGHEGYOEIGH
KGGEARKEQLGHEGYOEMGHKGGEARKEQLGHEGYOEMGHKGGEARKEQLGHEGYKEMGR
KGGLSTMEKSGGERAEIEIDESKFTNK

ATEM6 (92 acides aminés)

MASQQEKKQLDERAKKGETVVPGGTGGKSFEAQQHLAEGRSRGGQTRKEQLGTEGYOQMG
RKGGLSTGDKPGGEHAEIEIDESKFRTKT

Figure 1. Séquences des protéines ATEM1 et ATEM6. Les domaines conservés potentiellement répétés (ATEM1) sont sous-lignés.

4/ Recherche Doctorale et Post-Doctorale concernant des gènes spécifiquement exprimés au cours de la maturation de la graine

La première partie de ce chapitre porte sur la caractérisation des conditions de l'accumulation des protéines EM, l'étude de leur fonction et l'identification de nouveaux mutants affectés dans leur expression chez *Arabidopsis thaliana*. L'ensemble de ce travail a été finalisé par l'obtention du diplôme de doctorat délivré par l'Université de Montpellier et 4 publications.

La seconde partie présente une étude exhaustive des gènes *LEA* (*Late Embryogenesis Abundant*), qui incluent les gènes *EM*, et constitue une partie de mon travail post-doctoral (2 publications).

4-1/ Contexte scientifique.

Les protéines Em (Early methionine) doivent leur nom au fait de leur première caractérisation chez le blé par marquage à la méthionine radioactive au cours des dernières étapes de la maturation de la graine (Cumming and Lane, 1979). Ces gènes appartiennent à la famille des gènes *LEA* (*Late Embryogenesis Abundant*) dont les membres, regroupés en fonction de leur similarité de séquence, sont en général fortement exprimés au cours des dernières étapes de la maturation de la graine et dans de nombreux cas sous le contrôle de l'acide abscissique (ABA). Les gènes *Em* constituent le premier groupe de gènes *LEA* et sont chez *Arabidopsis thaliana* représentés par les gènes initialement nommés *GEA1* et *GEA6* et désormais appelés *ATEM1* ou *AtLEA1-1* (*At3g51810*) et *ATEM6* ou *AtLEA1-2* (*At2g40170*). Les produits de ces deux gènes se différencient, notamment, par le nombre de répétitions d'un domaine de 20 acides aminés: une copie unique dans la protéine *ATEM6* et quatre copies en tandem dans la protéine *ATEM1* (figure 1). Les transcrits *ATEM6* et *ATEM1* ne sont détectés que dans les graines matures (Gaubier et al., 1993). Un apport en ABA permet cependant de les induire précocement au cours du développement de la graine. Des analyses par hybridation *in situ* ont montré une expression localisée dans les tissus provasculaires des cotylédons ainsi qu'au niveau de l'axe et l'épiderme de l'embryon (Gaubier et al., 1993). C'est principalement dans le but de compléter ces informations que les deux axes de mon travail doctoral ont été définis. Le premier axe concerne l'étude des mécanismes de régulation des gènes *EM*. Une première stratégie a consisté à réaliser un crible génétique en utilisant une lignée transgénique *ATEM1::GUS* et une seconde à analyser l'expression des gènes *ATEM* dans les graines de mutants affectés dans le développement de la graine. Le second axe, destiné à identifier la fonction des protéines *ATEM*, a consisté à mieux caractériser les conditions d'accumulation de ces protéines et à analyser les effets de leur absence par utilisation de constructions antisens. Ce travail était soutenu par le fait que certaines des protéines *LEA* et donc peut être les protéines *EM*, très hydrophiles, pourraient être impliquées dans la capacité des graines matures à résister à une très faible teneur en eau, de l'ordre de 6%.

4-2/ Recherche Doctorale : Caractérisation des gènes *ATEM*

La majeure partie de ce travail a été réalisée dans le Laboratoire Génome et Développement des Plantes UMR CNRS/UPVD 5096 à l'Université de Perpignan et un séjour d'un mois a été effectué, dans le cadre d'une collaboration, dans le laboratoire de génétique dirigé par Marteen Koornneef à Wageningen (Pays Bas).

L'ensemble du travail présenté a été effectué sous la direction de Michel Delseny et en collaboration avec Marteen Koornneef et Jérôme Giraudat.

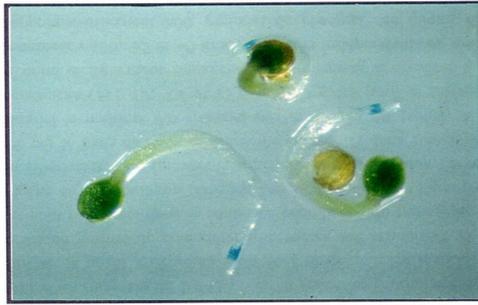


Figure 2. Détection de la β glucuronidase. Photo de plantules transgéniques de 3 jours incubées en présence de 2mM de X-Gluc pendant 12h à une température de 37°C.

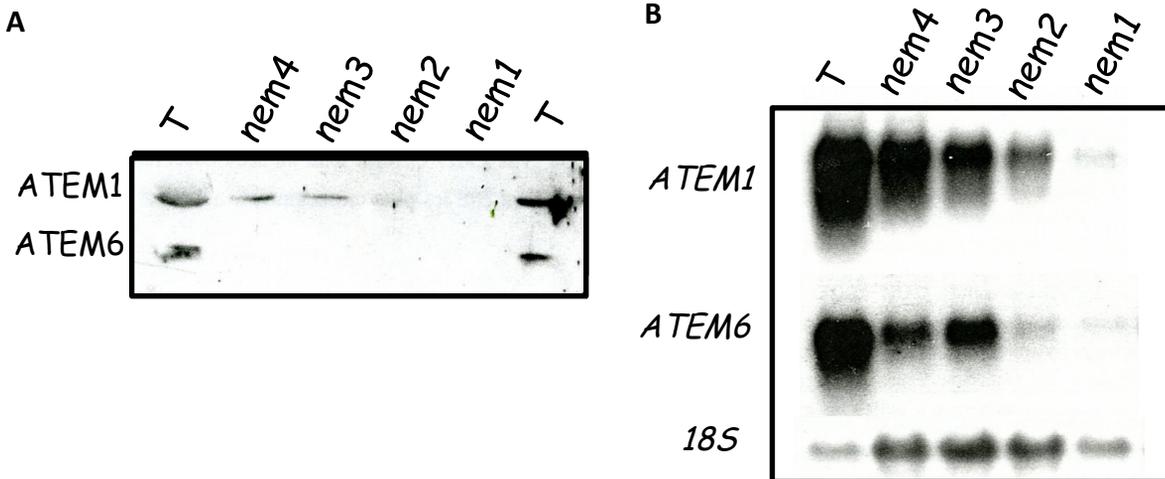


Figure 3. Accumulation des protéines ATEM et des messagers correspondants dans les graines de mutants *nem*. **A, Détection des protéines ATEM dans les extraits de graines matures des mutants *nem1*, *nem2*, *nem3* et *nem4*.** 15 μ g de protéines ont été utilisés. Un extrait protéique de graines transgéniques (T) non mutagénisées représente un contrôle. La détection a été faite avec le sérum 6. **B, Analyse de l'accumulation des transcrits ATEM.** Northern blot réalisés avec des ARNm de graines matures hybridés en présence de sondes ADN *ATEM1*, *ATEM6* et *18S*.

4-2-1/ Recherche de mutants affectés dans l'expression du gène *ATEM1*

4-2-1-1/ Mise en place d'un crible génétique

Une des stratégies choisies pour identifier de nouveaux régulateurs de l'expression des gènes *EM* a consisté à réaliser un crible génétique basé sur l'utilisation de plantes transgéniques possédant un fragment d'ADN de 1143bp de la séquence promotrice du gène *ATEM1* fusionné à la séquence codante du gène rapporteur codant pour la β -glucuronidase (GUS). Deux lots de graines ont été utilisés, un contenant uniquement le transgène et l'autre contenant le transgène et une mutation dans le gène *ABI3* (background *abi3-1*) qui est un facteur de transcription impliqué dans la réponse des graines à l'ABA et dans l'induction des gènes *ATEM*. La mutation *abi3-1*, considérée comme hypomorphe n'affecte pas l'expression des gènes *ATEM*. Les deux lots de graines ont été soumis à une mutagénèse par rayonnements gamma (30 ou 35 krad). De manière à optimiser le crible une technique de détection non létale de l'expression du gène rapporteur codant la β -glucuronidase a été mise au point. Après différents essais, il s'est avéré que l'incubation de plantules de 3 jours en présence de 2 mM de X-Gluc, substrat de l'enzyme, à 37 degrés pendant 12h, entraîne une coloration bleue au niveau de l'extrémité de la racine et permet de repiquer les plantules en terre, sans altérer leur viabilité (figure 2). C'est pour réaliser ce crible que j'ai effectué un séjour de recherche dans le laboratoire de génétique de Marteen Koornneef (Pays Bas).

4-2-1-2/ Analyses préliminaires

Les premières étapes du crible, réalisé sur des pools de graines de seconde génération, ont permis l'identification, parmi plus de 52000 graines, de 321 plantules affectées dans l'expression du gène rapporteur et donc, potentiellement dans celle du gène *ATEM1*. L'analyse de la transmission du phénotype sélectionné à la descendance a permis de conserver 67 lignées. Des dosages fluorométriques de l'activité du gène rapporteur en comparaison avec l'activité mesurée à partir de plantes contrôles, non irradiées et possédant le transgène *pATEM1::GUS* (lignée *T*) ont permis de classer les lignées sélectionnées en 4 familles :

- 4 lignées ne présentent aucune activité GUS (lignées *nem* pour no Em)
- 29 lignées avec une activité GUS inférieure à 25% par rapport à celle des graines contrôles (lignées nommées *lem* pour low Em)
- 33 avec une activité GUS comprise entre 25 et 80%
- 2 lignées avec une activité GUS supérieure à 150%.

Dans le but d'identifier les lignées affectées dans l'expression du gène *ATEM1* l'intégrité du transgène a tout d'abord été contrôlée. L'étape suivante a consisté à tester si la mutation affectant l'expression du gène rapporteur était ou non à l'état homozygote. Pour cela, seules les descendances des lignées ayant une activité GUS inférieure à 25% et celles ayant une activité supérieure à 150% du contrôle, ont été analysées et l'état homozygote de la cause du phénotype a été confirmé pour chacune d'entre elles. En raison du nombre de lignées sélectionnées, la suite du travail s'est limitée à l'analyse des lignées *nem* (activité GUS de 0%). Pour ces lignées, l'accumulation des protéines *ATEM* a été analysée grâce à des anticorps spécifiques: le premier, sérum 6, permet la détection des deux protéines *ATEM* d'*Arabidopsis* et le second, sérum 1, permet uniquement la détection de la protéine *ATEM1*. Il s'est avéré que les protéines *ATEM* s'accumulent significativement moins dans les graines des quatre lignées *nem* (figure 3A). En adéquation avec cette observation l'expression des gènes *ATEM1* et *ATEM6* est aussi réduite dans leur graines matures (figure 3B). L'effet le plus important est observé pour la lignée *nem1*.

Il était, à ce stade, important de vérifier et confirmer par test d'allélisme que les lignées identifiées ne présentaient pas de mutations aux mêmes loci.

Une dernière étape a consisté à tester la capacité des lignées sélectionnées à produire de

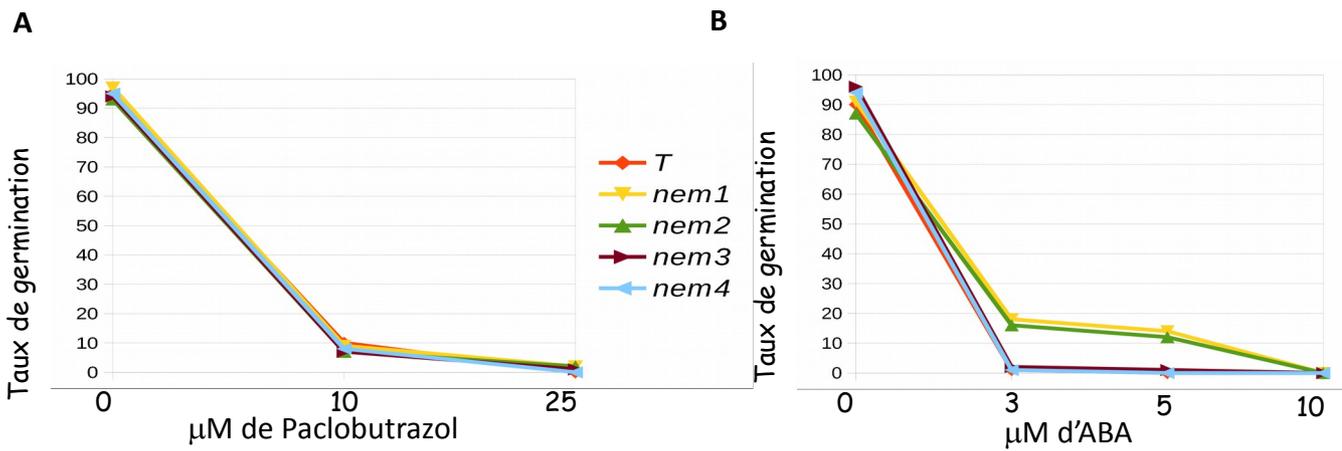


Figure 4. Sensibilité des graines de mutants *nem* à la présence de paclobutrazol et d'ABA. Les graines germées ont été comptabilisées après 6 jours en chambre de culture en présence de différentes concentration en paclobutrazol (A) ou d'ABA (B). Des graines T (transgéniques non mutées) ont été utilisées comme contrôle.

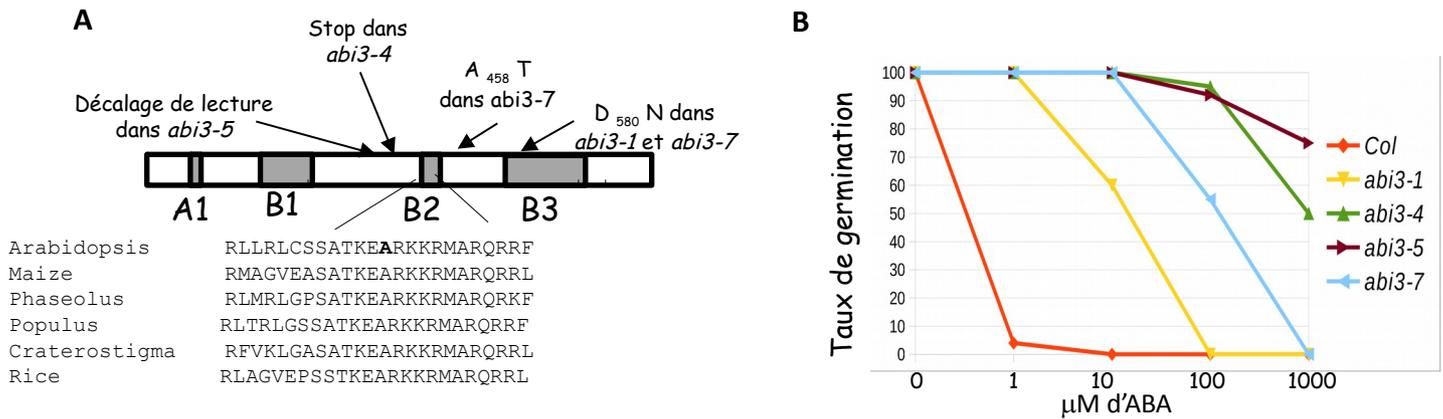


Figure 5. Caractéristiques du mutant *nem2*. A, Structure de la protéine ABI3. Les domaines conservés, les effets des mutations des différents allèles et les séquences du domaine B2 de différentes espèces sont indiqués B, Comparaison de la sensibilité à l'ABA de différents allèles *abi3*. Les taux de germination de graines placées en présence de différentes concentration en ABA ont été comptabilisés après 6 jours en chambre de culture. Des graines Col ont été utilisées comme contrôle.

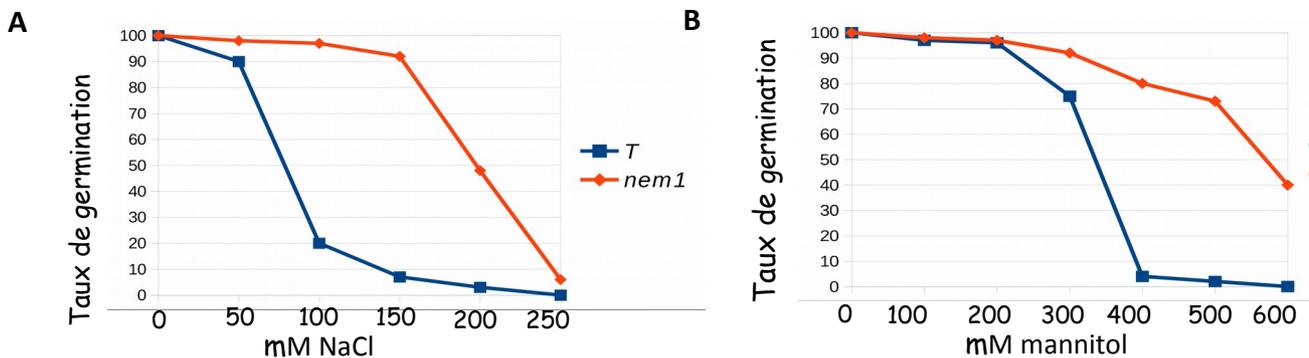


Figure 6. Résistance des graines *nem1* à la présence de sels (NaCl) ou de mannitol. Les taux de germination ont été comptabilisés après 6 jours en chambre de culture en présence de différentes concentrations en NaCl (A) ou en mannitol (B). Des graines T ont été utilisées comme contrôle.

l'ABA et à répondre à cette hormone. Information importante du fait que l'expression des gènes *ATEM* est en partie sous le contrôle de l'ABA. Des graines ont donc été placées en conditions de germination *in vitro* en présence de paclobutrazol qui est un inhibiteur de la synthèse des gibbérellines. Dans ces conditions, seules des graines de mutants déficients dans la production de l'ABA (mutants *aba*) sont capables de germer (Koornneef et al., 1982). Cette analyse a révélé que chacune des lignées testées semble ne pas être déficiente en ABA (figure 4A). Pour tester la sensibilité des graines à l'ABA, des graines de chaque lignée ont été placées en condition de germination *in vitro*, en présence de 3, 5 et 10 mM d'ABA. Deux des lignées ont montré un taux de germination en présence d'ABA supérieur à celui observé avec des graines sauvages (figure 4B).

Ce sont ces deux mutants qui ont fait l'objet d'une caractérisation complète. En raison du fait que trois mutants insensibles à l'ABA (*abi1*, *abi2* et *abi3*) avaient été isolés par d'autres laboratoires des tests d'allélismes entre ces mutants et les lignées *nem* ont été réalisés.

4-2-1-3/ Caractérisation de la lignée *nem2*.

Les tests d'allélismes avec les différents mutants *abi* ont montré que le mutant *nem2* était un nouvel allèle du mutant *abi3*. Le gène *ABI3* code un facteur de transcription orthologue au facteur de transcription VP1 (VIVIPAROUS1) de maïs. La protéine ABI3 est impliquée dans la transduction de l'ABA et dans l'expression de nombreux gènes dits de stress, en se fixant principalement sur des motifs de type RY (CATGCA(TG)) (Ezcurra et al., 2000a). Le séquençage du locus *ABI3* chez *nem2* (*abi3-7*) indique, en plus de la mutation *abi3-1*, une mutation ponctuelle qui entraîne le remplacement de l'Alanine 458 en Thréonine. Cette Alanine est conservée chez tous les orthologues et située dans le domaine B2 de ces protéines (figure 5A). Le phénotype du mutant *abi3-7* est, au regard de sa réponse à l'ABA et à l'expression de gènes modèles (i.e. *ATEM1*, *ATEM6*, *AT2S1*, *At2S2*), intermédiaire à celui du mutant faible *abi3-1* et des mutants sévères *abi3-4* et *abi3-5* (figure 5B). La caractérisation complète de ce mutant a permis de démontrer l'importance du domaine B2 dans la fonction des protéines de type ABI3. Ce travail a fait l'objet d'une publication dans Plant Mol Biol. 1999, Importance of the B2 domain of the Arabidopsis ABI3 protein for *EM* and *2S* albumin gene regulation. (Bies-Etheve et al., 1999)

4-2-1-4/ Caractérisation de la lignée *nem1*.

Le second mutant, fortement affecté dans l'expression des gènes *ATEM1* et *ATEM6* et présentant une certaine insensibilité de ses graines à la présence d'ABA, s'est avéré ne pas être allélique aux mutants *abi1*, *abi2* ou *abi3*. En plus de cette insensibilité à l'ABA les graines *nem1* ont montré une meilleure capacité à germer en présence de sels (NaCl) et de mannitol que des graines sauvages (figure 6). Cette observation suggérait que la mutation responsable du phénotype observé était probablement impliquée dans la transduction de l'ABA, mais aussi dans la réponse au sel et au mannitol, transductions qui peuvent se faire par des voies communes ou distinctes.

Une approche par CAPS (Co-dominant Amplified Polymorphic Sequences) a permis de déterminer que la mutation *nem1* se trouve sur le chromosome 2 entre les marqueurs m323 et ve017. De manière intéressante, deux gènes *ABI4* et *ABI5*, impliqués dans la transduction du signal ABA venaient d'être identifiés à proximité de cette région (Finkelstein and Lynch, 2000). Des tests d'allélismes avec ces deux mutants ont permis de conclure que le mutant isolé par notre crible était un nouvel allèle du mutant *abi5*. Le mutant *nem2*, par la suite nommé *abi5-5*, se caractérise par une délétion de deux bases à la position 949 de la séquence codante (figure 7A). Cette délétion entraîne un décalage de phase de lecture qui conduit à l'apparition d'un codon stop, 30 bases en aval de la délétion et se traduit par la production d'une protéine tronquée de 327 acides aminés au lieu de 442. Cette protéine tronquée ne contient plus son domaine basique ni son domaine riche en

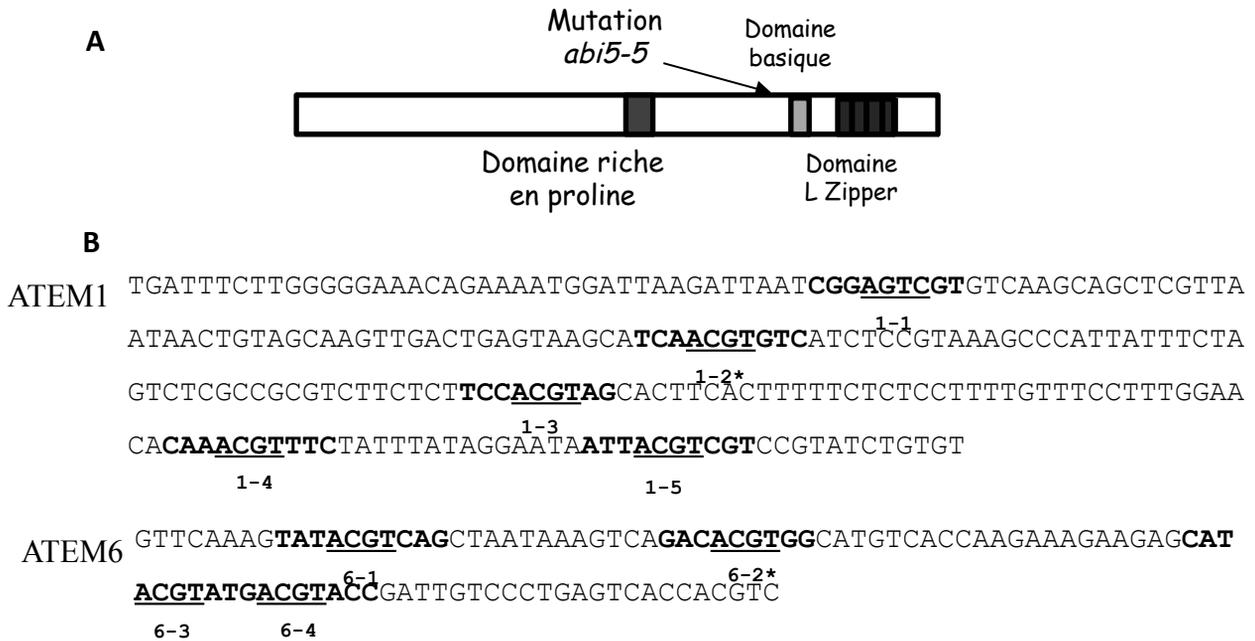


Figure 7 Caractéristiques du mutant *nem1*. **A**, Représentation schématique de la protéine ABI5. La présence de domaines conservés ainsi que la localisation de la délétion de 2 nucléotides dans le mutant *nem1* (*abi5-5*) sont indiquées. **B**, Séquences promotrices des gènes *ATEM* sur lesquelles la protéine ABI5 se fixe *in vitro*. Les différentes régions en gras ont été utilisées pour la réalisation de gels retard. Celles notées avec une astérisque sont celles pour lesquelles l'interaction est la plus forte.

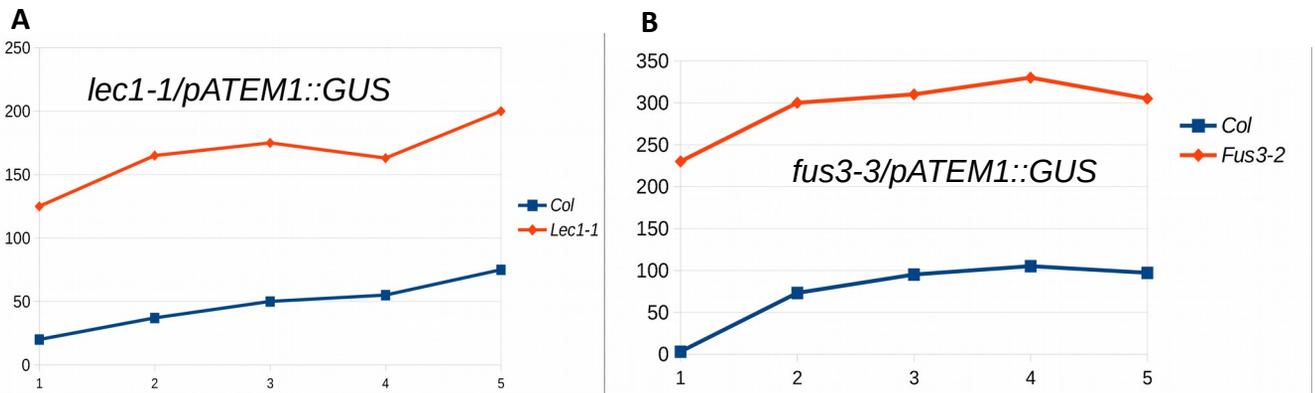


Figure 8. Activités du promoteur *ATEM1* dans les mutants *lec1-1* et *fus3-2*. Les activités GUS ont été mesurées à partir de graines, de différents stades au cours de la maturation, de mutants *lec1-1* (A) et *fus3-3* (B) croisés avec des plantes transgéniques contenant la séquence promotrice du gène *ATEM1* fusionnée à la séquence codante du gène GUS

leucine (L zipper). Par réalisation de gels retards en présence de fragment de promoteur des gènes *ATEM1* et *ATEM6*, Cristel Carles, qui a terminé ce travail, a montré que la protéine ABI5 produite *in vitro* est capable de se fixer sur les séquences de types G-BOX (CACGTG) des promoteurs *ATEM*. La fixation est particulièrement importante avec deux séquences (1-2 et 6-2), l'une présente dans la région promotrice du gène *ATEM1* et l'autre, dans celle du gène *ATEM6* (figure 7B). En utilisant la protéine ABI5 tronquée aucune interaction n'est observée indiquant que le domaine basique et le domaine leucine-zipper situés à l'extrémité C terminale de la protéine ABI5, sont nécessaires pour l'interaction avec les séquences G-BOX présentes au niveau des promoteurs des gènes *ATEM*. Ce travail a fait l'objet d'une publication dans Plant J. 2002, Regulation of *Arabidopsis thaliana* EM genes: role of ABI5. (Carles et al., 2002)

4-2-1-5/ Analyse de l'expression des gènes *ATEM* dans les graines de mutants affectés dans la maturation des graines.

Parallèlement au crible génétique et afin d'identifier de nouvelles protéines impliquées dans la régulation des gènes *ATEM*, leur expression a été analysée chez différents mutants affectés dans le développement de la graine. Les mutants choisis étaient les mutants *lec1* (*leafy cotyledons*) et *fus3* qui présentent un défaut de la mise en place du programme de maturation des graines (Meinke, 1992 ; Keith et al., 1994 ; Misera et al., 1994). La protéine FUS3 code un facteur de transcription qui possède un domaine B3 similaire à celui de la protéine ABI3 (Luerssen et al., 1998). Nos analyses ont montré que les effets des deux mutations sur les profils d'expression des deux gènes *ATEM* ne sont pas identiques. Chez le mutant *lec1-1* (considéré comme un allèle nul), *ATEM1* est exprimé plus tôt et plus fortement au cours du développement que dans des graines sauvages alors que *ATEM6* est réprimé. Chez le mutant *fus3*, on observe, là encore, une expression plus importante du gène *ATEM1*. Les observations concernant *ATEM1* ont été confirmées par croisements des mutants avec des plantes transgéniques contenant la construction promoteur *ATEM1::GUS* (figure 8A et 8B). Malgré les fortes expressions observées les deux protéines ATEM ne sont pratiquement pas détectées dans les graines du mutant *lec1-1*. Une explication à cette observation sera donnée dans un paragraphe suivant démontrant la régulation post traductionnelle des protéines ATEM. Il est intéressant, de noter que chez le mutant *lec1-1*, l'expression du gène *ATEM1* peut être augmentée par un apport d'ABA. Par croisements avec des plantes transgéniques contenant différentes longueurs du promoteur *ATEM1* fusionnées à la séquence codante du gène rapporteur GUS, deux régions du promoteur impliquées dans la régulation *via* LEC1 ont pu être identifiées. L'ensemble de ce travail a été publié dans J Exp Bot. 2000, Changes in gene expression in the leafy cotyledon1 (*lec1*) and *fusca3* (*fus3*) mutants of *Arabidopsis thaliana* (Vicient et al., 2000).

4-2-2/ Caractérisation des conditions d'accumulation des protéines ATEM et analyse des effets d'une diminution de la quantité de ces protéines.

4-2-2-1/ Profil d'accumulation des protéines ATEM et étude des facteurs impactant leur accumulation.

Au cours de la maturation de la graine, l'accumulation des protéines ATEM corrèle celle des transcrits (figure 9). La protéine ATEM1 est détectée dès 17 JAP (jours après pollinisation) et ATEM6 dès 20 JAP (figure 9A). Un maximum d'accumulation de ces deux protéines est observé

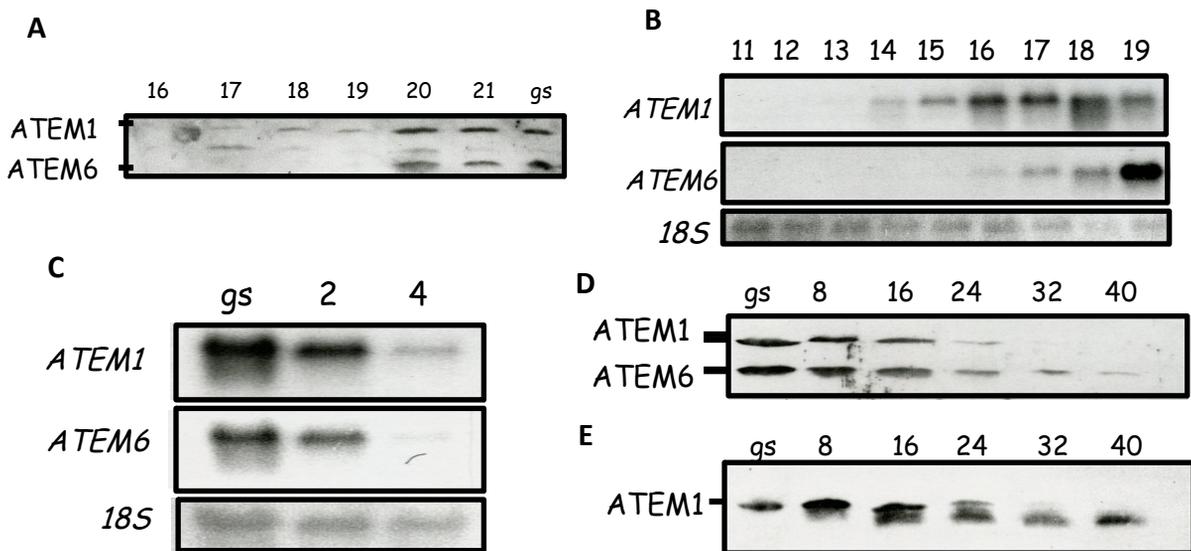


Figure 9. Accumulation des protéines et transcrits ATEM au cours du développement de graines. A, D et C Accumulation des protéines ATEM. 15 μ g de protéines de graines récupérées de 16 à 21 jours après la pollinisation (A) ou de 8 à 40 heures après le début de l'imbibition (D et E) ont été analysés par western blot avec le sérum 6 (A et D) ou le sérum 1 (E) B et C, Accumulation des transcrits ATEM. Northern blot réalisés avec des ARNm de graines sèches (gs) ou récupérées de 11 à 19 jours après la pollinisation (B) ou quelques heures après l'imbibition (C) et hybridés avec des sondes ADN *ATEM1*, *ATEM6* et *18S*.

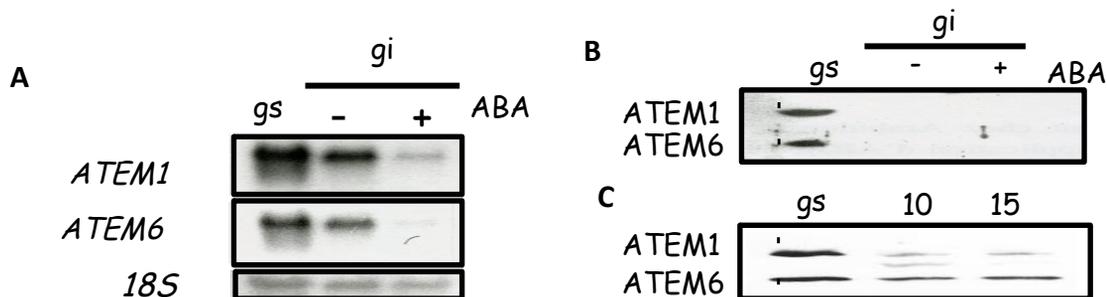


Figure 10 Stabilité des protéines ATEM dans des graines immatures ou en présence de protéines de feuilles. A et B, Accumulation des transcrits et protéines ATEM en réponse à un apport en ABA. Des graines sèches (gs) et des graines immatures (gi) récupérées 7 jours après la pollinisation et placées pendant 5 jours en présence de 50 mM d'ABA ont été utilisées pour analyser l'accumulation des transcrits ATEM (A) et celle des protéines correspondantes (B). C, Stabilité des protéines ATEM en présence de protéines de feuilles. 10mg de protéines de graines sèches (gs) ont été incubés en présence de 10 (10) ou 15 (15) mg de protéines de feuilles. La détection des protéines ATEM est réalisée avec le sérum 6.

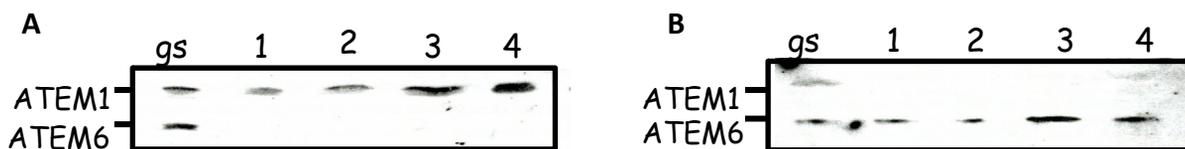


Figure 11. Accumulation des protéines ATEM dans les graines de lignées antisens *ATEM6* (A) ou antisens *ATEM1* (B). 15 μ g de protéines ont été analysés par western par utilisation du sérum 6 permettant la détection des deux protéines ATEM.

dans la graine sèche. Les premiers transcrits *ATEM1* sont détectés dès 14 jours après la pollinisation (JAP) et les transcrits *ATEM6*, dès 17 JAP (figure 9B). Quatre heures après le début de l'imbibition de la graine, les transcrits *ATEM6* ne sont pratiquement plus détectés et les transcrits *ATEM1* ne sont que très faiblement mis en évidence (figure 9C). Les protéines *ATEM1* sont décelables jusqu'à 24 HAI (heures après début de l'imbibition) et les protéines *ATEM6* jusqu'à 40 HAI (figure 9D). De manière intéressante, des signaux de tailles intermédiaires à celles des deux protéines *ATEM* sont détectés avec l'anticorps spécifique de la protéine *ATEM1* (figure 9E). Il est possible que ce signal, qui apparaît transitoirement au moment de la diminution du signal caractéristique de la protéine *ATEM1*, corresponde à des produits de clivages spécifiques. Parce que des analyses *in vitro* indiquent que la protéine EM de blé subit un clivage spécifique par action d'une endoprotéinase de type cystéine protéinase (Taylor and Cuming, 1993) et que le site de clivage décrit est présent dans la séquence de la protéine *ATEM1*, il est possible que la protéine *ATEM1* subisse elle aussi un clivage spécifique avant sa dégradation.

Les effets d'une induction précoce des gènes *ATEM* en réponse à un apport d'ABA dans des graines immatures ont été testé au niveau protéique. Des graines récupérées 7 jours après la pollinisation ont été incubées pendant 5 jours en présence ou non d'ABA. Dans ces conditions on observe bien une accumulation de transcrits *ATEM* mais aucun signal protéique (figure 10A et 10B). Cette observation ainsi que celles obtenues avec les mutants *lec1* et *fus3* suggèrent que les protéines *ATEM* font potentiellement l'objet d'une importante régulation post-transcriptionnelle. Dans le but de tester cette hypothèse, des extraits protéiques de graines sèches contenant les protéines *ATEM* ont été incubés en présence d'extraits protéiques de feuilles. Cette analyse a été réalisée en faisant varier la quantité d'extraits de protéines de feuilles. On observe alors que la présence de protéines de feuilles entraîne la dégradation des protéines *ATEM* (figure 10C). Ce travail a fait l'objet d'une publication dans *Journal of Experimental Botany*, 1998, Accumulation and degradation of Em proteins in *Arabidopsis thaliana*: evidence for post-transcriptional controls, (Bies et al., 1998)

4-2-2-2/ Analyse des effets d'une diminution de la quantité des protéines ATEM dans les graines.

Afin d'appréhender le rôle physiologique des protéines *ATEM*, des plantes transgéniques contenant des constructions antisens correspondant à chacun des gènes *ATEM* ont été générées. Deux constructions (*pa1* et *pa6*) ont été réalisées avec des fragments de promoteurs des gènes *ATEM* préalablement caractérisés et entraînant une expression similaire à celle du gène endogène. Deux autres constructions (*a1* et *a6*) possèdent le promoteurs constitutifs 35S. Les séquences correspondant aux ADNc entiers ont été placées en orientation inversée sous le contrôle des deux types de promoteurs. L'accumulation des protéines *ATEM* dans les graines de lignées transgéniques générées indique que une forte diminution de la teneur ou une absence de protéine *ATEM6* dans les graines des plantes qui contiennent des constructions antisens *AtEm6* (figure 11A). La majorité des plantes contenant la construction antisens *ATEM1* ne semble plus du tout accumuler de protéine *ATEM1* dans leurs graines (figure 11B). Les lignées pour lesquelles l'accumulation de l'une des protéines *ATEM* est la plus affectée, ont été utilisées pour tenter de mettre en évidence un phénotype relié à cet état. Des essais de germination à différents stades du développement de la graine, en présence de sels de sodium et de mannitol, n'ont pas permis d'identifier un quelconque phénotype révélateur de la fonction des protéines *ATEM*. Il est par conséquent possible, que la fonction de ces deux protéines soient redondantes et que l'analyse de plantes contenant les deux types de construction soit nécessaire. L'ensemble de ces différents travaux a permis la rédaction d'une revue dans *Journal of Plant Physiology*, 2001, Late Embryogenesis Abundant (LEA) protein gene regulation during *Arabidopsis* seed maturation, (Delseny et al., 2001)

Origine	Nom du gène ou de l'EST	Groupe LEA	Références ou locus
Arabidopsis	AtEm6	1	Gaubier et al. 1993
	Cor47	2	Gilmour et al. 1992
	Lti30	2	Welin et al. 1994
	Ecp63	3	Yang et al. 1997
	EST PAP310	2	At2g21490
	EST PAP240	3	At1g52690
	EST PAP260 et PAP051	4	At1g32560 and At5g06760
	EST PAP140 et PAP140b	6	At3g22490 and At3g22500
	M10	New	Raynal et al. 1999
M17	New	Raynal et al. 1999	
Glycine max	PM27	3	Y. I. Hsing (personal communication)
Coton	D29	5	Galau et al. 1986
	D73 (Lea5)	7	Galau et al. 1986
	D95 (Lea14)	8	Galau et al. 1986
Haricot	PvLea-18	LEA-18	Colmenero-Flores et al. 1999

Tableau 1. Liste des séquences de références utilisées pour l'identification de tous les gènes *LEA* d'*Arabidopsis*.

4-3/ Recherche Post Doctorale : Caractérisation de la famille des gènes *LEA*

Lors de mon recrutement en tant que Maître de conférence à l'Université de Perpignan, j'ai intégré l'équipe de recherches dirigée par Michel Delseny, alors directeur du laboratoire Génome et Développement des Plantes. Nous avons décidé de poursuivre les travaux concernant la maturation de la graine et plus particulièrement orienté notre approche sur une étude exhaustive des gènes *LEA* (Late Embryogenesis Abundant). Nos objectifs étaient d'identifier tous les gènes *LEA* chez *Arabidopsis*, de les classer en fonction de la présence de domaines conservés, d'analyser leur position dans le génome et enfin, d'identifier ceux qui étaient régulés par les mêmes voies.

4-3-1/ Identification de tous les gènes *LEA* chez *Arabidopsis thaliana*.

Au moment de ce travail, le séquençage du génome d'*Arabidopsis* était complet et nous avons focalisé notre étude sur la recherche de tous les gènes présentant des similarités avec les gènes *LEA* décrits dans la littérature, ou possédant un ou plusieurs domaines caractéristiques de gènes *LEA*. Pour cela, nous avons utilisé comme référence différents sets de gènes représentant un total de 17 séquences (tableau 1). Quatre gènes *LEA* connus chez *Arabidopsis* et deux gènes supplémentaires *MI0* et *MI7* caractérisés et correspondant à une nouvelle classe de *LEA* (Raynal et al., 1999). Nous avons également utilisé quatre gènes qui correspondent à des *EST* d'*Arabidopsis* ayant un profil d'expression caractéristique des *LEA* et des similarités de séquences avec des gènes *LEA* d'autres espèces, et enfin, cinq séquences de gènes *LEA* d'autres espèces. Parmi ces dernières se trouve celle du gène *PvLEA-18* qui avait été identifié chez *Phaseolus vulgaris* et décrit comme constituant un nouveau groupe de gènes *LEA* (Colmenero-Flores et al., 1999). A partir de ces gènes de référence, un algorithme a été défini afin de les aligner avec le génome d'*Arabidopsis* (TIGR version 4) en utilisant le programme TBLASTN (Altschul et al., 1997). Une seconde étape réciproque a été réalisée par le biais de l'outil BLASTX et enfin, une comparaison par TBLASTN avec des bases de données d'*EST* et d'*ARNm* a complété notre analyse. Les séquences identifiées ont été alignées avec la base de données génomiques par utilisation du logiciel sim4 (Florea et al., 1998). La position des introns et exons a été déterminée et confirmée avec le programme Artemis (Rutherford et al., 2000). Ce travail a permis l'identification de 50 gènes codant potentiellement des protéines *LEA* chez *Arabidopsis*. Certains de ces gènes, comme les gènes *Kin*, avaient déjà fait l'objet de publications et avaient été identifiés comme des gènes dont l'expression peut être induite en réponse à de faibles températures (Kurkela and Franck, 1990; Gilmour et al., 1992). Il est intéressant de noter que plus de 80% des protéines codées par les gènes identifiés sont de petite taille, souvent inférieure à 30 kDa et que la majorité d'entre elles sont neutres ou légèrement acides.

4-3-2/ Classification des gènes *LEA* d'*Arabidopsis thaliana*

Après leur identification, les différents gènes ont été classés en fonction des similarités de séquences des protéines qu'ils codent et une étude des domaines conservés a été réalisée. L'ensemble de ce travail a nécessité l'utilisation des interfaces ClustalW (http://www.infobiogen.fr/services/analyseq/cgi-bin:clustalw_in.pl), multalin (<http://prodes.toulouse.inra.fr/multalin.multalin.html>) et MEME (<http://meme.sdsc.edu/meme>). L'objectif était de donner une classification la plus rigoureuse possible et une dénomination explicite. Nous avons ainsi renommé les gènes *ATEM1* et *ATEM6*, *AtLEA1-1* et *AtLEA1-2*. Cette dénomination signifie qu'ils appartiennent au premier groupe des gènes *LEA*. Les spécificité de toutes les

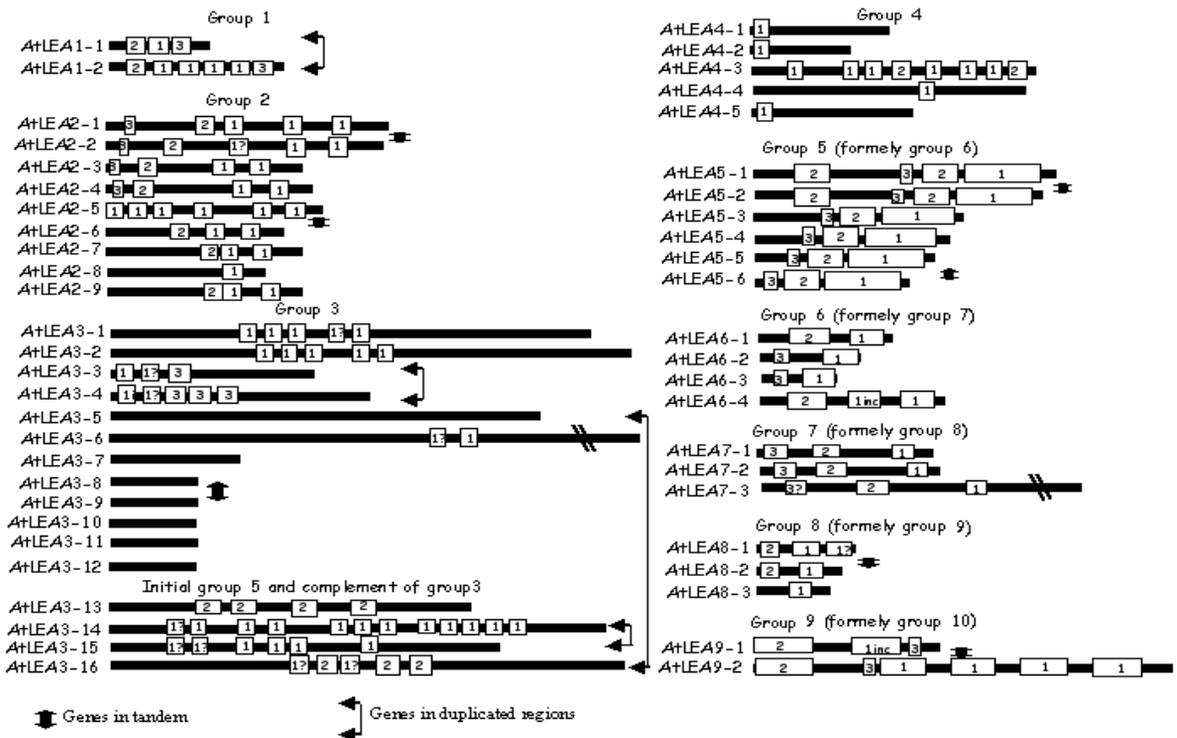


Figure 12. Représentation schématique des protéines LEA chez *Arabidopsis thaliana*. Les domaines conservés sont indiqués. Les points d'interrogations indiquent des domaines dégénérés ou des domaines incomplets. Les protéines correspondant à des gènes en tandem ou résultant de duplication du génome sont également indiquées.

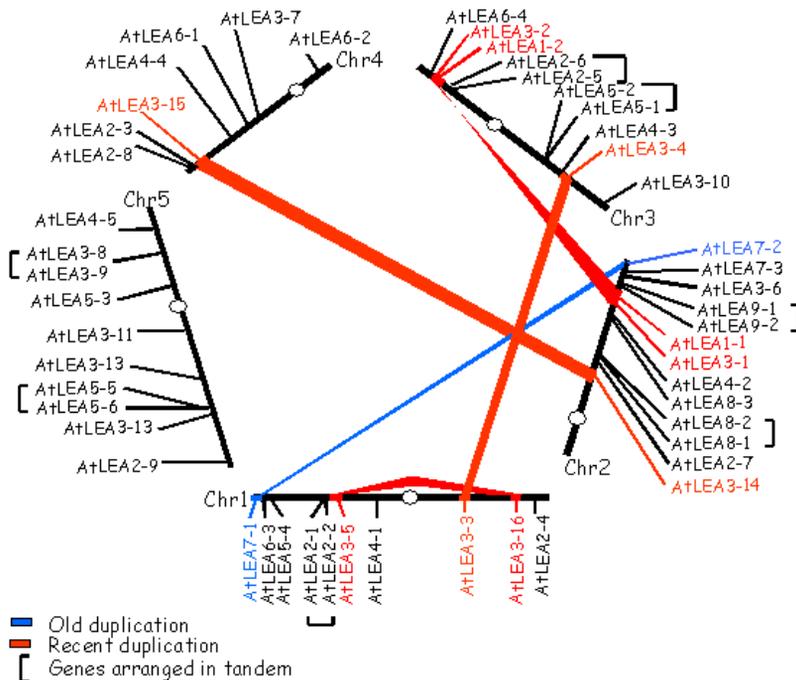


Figure 13. Position des différents gènes *LEA* dans le génome d'*Arabidopsis thaliana*. Les gènes positionnés en tandem résultant de duplication sont indiqués.

protéines identifiées sont indiquées dans la figure 12. Le choix des numéros attribués à chaque domaine identifié est relatif au nombre de répétitions, à la conservation du domaine dans le groupe et à leur taille. Nous avons ainsi pu déterminer que le groupe LEA2, qui est relatif aux déhydrines, contient 9 représentants chez *Arabidopsis*, notés de LEA2-1 à LEA2-9, trois sont similaires au gène *COR47*, un au gène *LT10* et deux à l'EST *PAP310*. Le groupe 3 est constitué de 16 représentants. De manière intéressante, quatre gènes, de *AtLEA3-13* à *AtLEA3-16*, précédemment décrits comme formant le groupe 5, du fait de leur similarité avec le gène *D29*, possèdent au moins un domaine similaire à ceux observés dans les protéines du groupe 3. Cette observation nous a conduit à incorporer ces gènes au groupe *LEA3*. Aucun domaine conservé n'est présent dans toutes les protéines codées par les gènes du groupe 3, mais toutes possèdent au moins un domaine en commun avec des protéines de ce groupe. Le groupe des *LEA4* contient cinq gènes qui codent des protéines qui peuvent contenir jusqu'à six copies du domaine conservé SAKEKIS (Figure 12). Du fait de la suppression du groupe 5 par son inclusion dans le groupe 3, le groupe 6 devient le groupe 5. Ce groupe contient six gènes codant des protéines qui possèdent toutes trois domaines conservés et présents dans chacune d'entre elles. Le groupe *LEA6* (anciennement nommé groupe 7) est composé de quatre gènes similaires au gène *D73* de coton. Le groupe *LEA7* contient trois gènes similaires au gène *D95* de coton. Le groupe 8 est composé de trois gènes qui présentent de fortes similarités avec le gène *PvLEA-18* de *Phaseolus vulgaris*. Enfin, le dernier groupe de gènes, le groupe LEA9, est celui composé des deux gènes initialement nommés *M10 (LEA9-1)* et *M17(LEA9-2)*.

Notre analyse met en évidence un certain nombre de domaines présents dans la base de données de Prosite. Cependant, pour aucun de ces domaines une fonction n'a été décrite. L'ensemble de notre travail a permis de montrer que l'utilisation de la similarité des domaines conservés, en plus de la similarité de séquence globale, est relativement bien adaptée à l'étude des gènes *LEA* et permet d'avoir une vision des domaines potentiellement importants chez ces protéines. Le fait que certains de ces domaines sont aussi conservés chez d'autres espèces souligne leur importance.

4-3-3/ Analyse de la localisation des gènes *LEA* dans le génome d'*Arabidopsis thaliana*: rôle des duplications

Il est intéressant de noter que 14 des gènes identifiés sont positionnés en tandem (figure 13). Chaque tandem étant composé de gènes appartenant à un même groupe, leur présence pourrait résulter d'une duplication locale. En utilisant les bases de données décrivant les régions du génome dupliquées (<http://wolfe.gen.tcd.ie/athal/dup/>), nous avons observé que douze des gènes identifiés sont positionnés, par paires, dans des régions dupliquées du génome (*AtLEA1-1* et *AtLEA1-2*, *AtLEA3-1* et *AtLEA3-2*, *AtLEA3-3* et *AtLEA3-4*, *AtLEA3-5* et *AtLEA3-16*, *AtLEA3-14* et *AtLEA3-15*, *AtLEA7-1* et *AtLEA7-2*)(figure 13). Cinq paires résultent de duplications dites récentes (25 à 40 milliards d'années) et la dernière, d'une duplication plus ancienne (85 à 100 milliards d'années). Douze autres gènes *LEA* (*AtLEA3-13*, *AtLEA4-1*, *AtLEA4-2*, *AtLEA4-3*, *AtLEA4-4*, *AtLEA4-5*, *AtLEA5-3*, *AtLEA5-4*, *AtLEA5-5*, *AtLEA7-3*, *AtLEA9-1* et *AtLEA9-2*) sont localisés à des loci issus de duplication du génome, mais un seul représentant est présent. Il est possible dans ces derniers cas que le second gène ait été perdu au moment, ou à la suite de la duplication. Une autre possibilité est que le gène *LEA* présent ait été intégré après la duplication. La localisation par paires au niveau de régions dupliquées, ou en tandem, s'accompagne de taux de similarités plus importants des gènes concernés. Une seule exception à cela, les gènes *AtLEA7-1* et *AtLEA7-2* qui eux résultent, d'après les informations disponibles, d'une duplication plus ancienne.

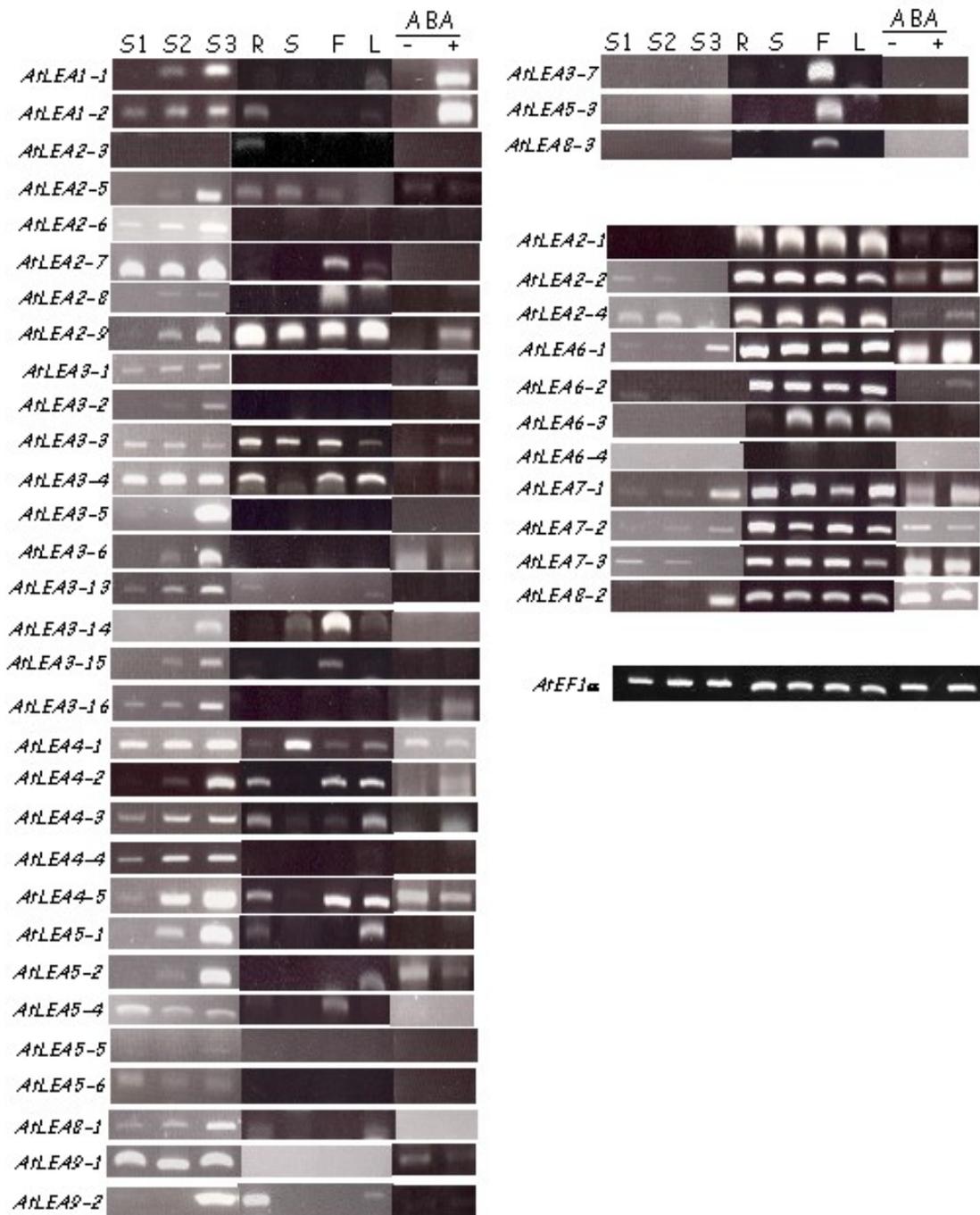


Figure 14. Profil d'expression des 50 gènes LEA d'*Arabidopsis* au cours du développement de la graine, dans différents tissus de la plante et en réponse à un apport d'ABA. Les stades S1, S2 et S3 correspondent respectivement à des graines récupérées de 10 à 13 jours après la pollinisation, 14 à 17 jours et de 18 à 21. R, ARNm de racines, S, ARNm de graines sèches, F, ARNm de fleurs et L, ARNm de feuilles. Le traitement en présence d'ABA a été réalisé sur des graines immatures prélevées 5 jours après la pollinisation et placées pendant 24h en présence ou non de 50 mM d'ABA. Les quantités de transcrits ont été analysés par RT-PCR

4-3-4/ Comparaison des répétitions de domaines conservés des protéines LEA.

Parce que les protéines LEA possèdent souvent des domaines répétés, il était intéressant de comparer ces domaines entre eux. Cette analyse a permis de proposer différentes hypothèses sur l'historique des domaines répétés de ces gènes qui pourraient, d'après leur similarités résulter de duplication locales de séquences intragéniques.

4-3-5/ Analyse des profils d'expression des gènes *LEA* chez *Arabidopsis thaliana*.

Nous avons cherché à classer les gènes *LEA* selon leurs voies de régulation. Pour cela la quantité de transcrits de chacun de ces gènes au cours des différentes étapes de la maturation de la graine (10-13, 14-17 et 18-21 JAP) ainsi que dans différents tissus de la plante a été analysée (figure 14). Nous avons ainsi pu définir trois groupes de gènes. Le premier groupe, composé de 31 gènes (*AtLEA1-1* à *AtLEA9-2*), est le plus important en nombre de gènes, et se caractérise par des profils d'accumulation de transcrits spécifiques des gènes *LEA*, à savoir une forte accumulation au cours du développement de la graine (figure 14). Le second groupe, composé de trois gènes (*AtLEA3-7*; *AtLEA5-3* et *AtLEA8-3*), ne semble être exprimé que dans les fleurs. Le dernier groupe est composé de onze gènes (*AtLEA2-1* à *AtLEA8*) qui présentent une accumulation de transcrits plus importante dans les tissus végétatifs par rapport à celle observée dans les graines (figure 14).

4-3-6/ Etude de facteurs impactant l'expression des gènes *LEA* chez *Arabidopsis thaliana*.

Afin d'améliorer nos connaissances sur les voies de régulation de l'expression des gènes *LEA* chez *Arabidopsis*, leur expression en réponse à un apport en ABA dans des plantules de cinq jours a été analysé et montre que seulement dix d'entre eux ont une expression augmentée par l'hormone (figure 14). Leur niveau d'expression chez deux mutants affectés dans la transduction du signal de l'ABA (le mutant *abi3-6* et le mutant *abi5-5 (nem2)*), ainsi que dans les graines des mutants *lec2-1* et *fus3-3* permet d'observer que plus de la moitié des gènes *LEA* chez *Arabidopsis* sont affectée par au moins une des mutations testées. Certains nécessitent la protéine ABI3, d'autres ABI3 et ABI5. Certains pouvant être induit par l'ABA ne sont pas affectés par les mutations *abi3* et *abi5*. Seule l'expression de 14 gènes *LEA* est affectée par la mutation *fus3-3*. Très peu de gènes ont une expression anormale dans les graines du mutant *lec2-1*. En effet, seul le gène *AtLEA3-5* semble être fortement affecté par cette mutation. Il aurait été intéressant de tester le mutant *lec1*. Enfin neuf gènes ne sont pas ou peu exprimés dans les graines matures sauvages et sont plus fortement exprimés dans les graines du mutant *abi3-6*. Cette observation est assez surprenante, car elle suggère que pour ces gènes, la protéine ABI3 aurait un effet inhibiteur sur leur expression. Enfin pour dix gènes, aucun effet des mutations testées n'a pu être observé dans la graine mature. L'ensemble de ces observations permet d'avoir une meilleure vision de la régulation de l'expression des différents gènes *LEA*. Elle indique aussi que les voies de régulation de chacun de ces gènes sont parfois distinctes, probablement complexes et impliquent probablement d'autres protéines que ABI3, ABI5, FUS3 et LEC2 .

Conclusion

En ce qui concerne l'étude des gènes *ATEM* (*ATLEA1-1* et *AtLEA1-2*) les analyses réalisées ont permis de mieux caractériser leurs conditions d'expression ainsi que celles de l'accumulation des protéines correspondantes. Une régulation post traductionnelle potentielle a pu être mise en évidence. Un autre point important a été la réalisation d'un crible génétique basé sur l'utilisation d'une construction transgénique. Cette approche assez novatrice, à l'époque, a permis de mieux caractériser les domaines importants, tel que le domaine B2, de la protéine ABI3 et d'identifier une nouvelle protéine ABI5 impliquée dans la transduction du signal ABA. Des interactions de la protéine ABI5 avec des régions cis-régulatrices des gènes *ATEM* ont également pu être mises en évidence. L'implication des protéines FUS3 et LEC1 a également été démontré et des séquences cis-régulatrices du promoteur *ATEM1* impliquées dans la régulation par LEC1 ont pu être identifiées. Il est intéressant de signaler que toutes les lignées sélectionnées lors du crible n'ont pas fait l'objet d'une étude complète et que certaines pourraient permettre encore d'augmenter nos connaissances concernant la régulation des gènes *ATEM* et plus généralement peut être celles des gènes exprimés au cours des dernières étapes de la maturation de la graine.

L'identification exhaustive et la classification de tous les gènes *LEA* d'*Arabidopsis thaliana* selon leur similarité, l'analyse de leur localisation dans le génome, et leur profil d'expression ont permis de réunir un certain nombre de données générales. Notre analyse a entre autre permis de classer tous ces gènes en fonction de leur similarités et de proposer des hypothèses sur la diversité des protéines LEA et sur l'origine possible des domaines conservés et répétés en tandem. Le travail réalisé montre également la complexité de la régulation de ces gènes. La majorité des voies de régulation impliquent la protéine ABI3, certaines également la protéine ABI5 et une fraction moins importante les protéines FUS3-3 et LEC1. La protéine LEC2 ne semble, elle, être impliquée que dans la régulation du gène *AtLEA3-5*. Il est important de noter que nos analyses ont été réalisées par RT-PCR et par conséquent seules des variations assez importantes d'expression ont pu être mises en évidence. Pour la première fois cependant un effet inhibiteur de la transcription réalisé par la protéine ABI3 a été mis en évidence. Des analyses plus précises par RT-PCR quantitative et avec un plus grand nombre de mutants pourraient permettre de mieux comprendre les voies de transduction impliquées dans la maturation de la graine.

L'ensemble de ce travail a permis la publication de 8 articles scientifiques. Depuis de nombreux autres articles ont été publiés et plusieurs ont utilisé comme référence notre analyse exhaustive des gènes *LEA* chez *Arabidopsis*, ainsi que celle réalisée par Hundertmark et Hinchin pour identifier tous les gènes *LEA* chez d'autres espèces comme le peuplier, le citronnier, le maïs, la tomate, la pomme de terre, le mil, la cacahuète et l'algue verte *Klesormidium* (Hundertmark and Hinchin, 2008 ; Su et al., 2011; Reddy et al., 2012 ; Holzinger et al., 2014 ; Pedrosa et al., 2015 ; Li et al., 2015 ; Cao and Li, 2015 ; Charfeddine et al., 2015). Chez le citronnier par exemple tout comme ce que nous avons observé chez *Arabidopsis*, 33% des gènes *LEA* sont positionnés en tandem dans le génome (Pedrosa et al., 2015). Toutes les analyses réalisées confirme la conservation de ces gènes et de certains de leur domaines entre les espèces végétales. Il est important de signaler à ce sujet que des gènes similaires ont depuis été identifiés chez des procaryotes et de petits arthropodes (Wise and Tunnacliffe, 2004; Bahrndorff et al., 2009).

De nombreuses analyses ont porté sur la fonction des protéines LEA. Il a été montré par exemple qu'une protéine Em de blé confère lorsqu'elle est produite dans des levures une meilleure croissance en conditions de stress osmotique (Swire-Clark and Marcotte, 1999). Une autre

Figure 15. Phénotype de mutant d'insertion *ATLEAI-1*. A, Siliques de plantes sauvages Col. B, Siliques du mutant d'insertion, C, Siliques de mutants complémentés. Issu de Manfre et al, 2006.

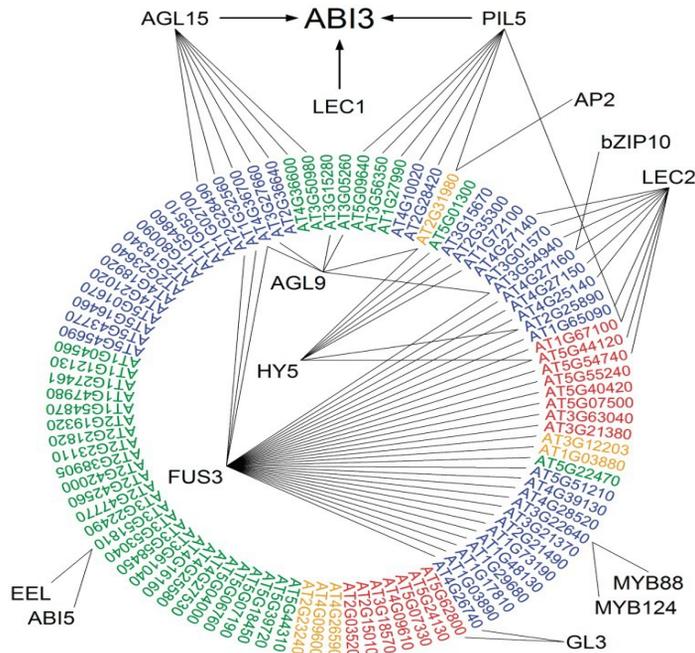
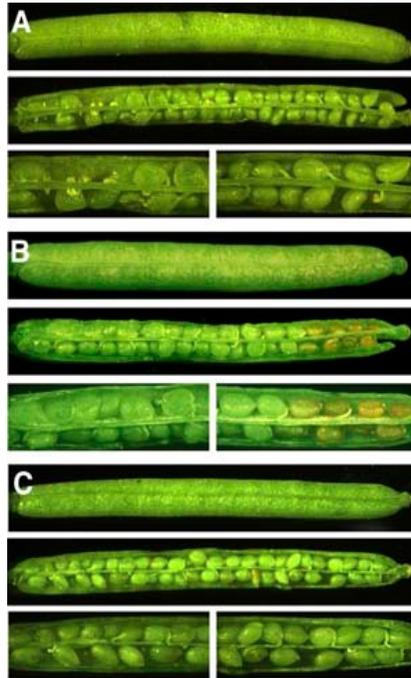


Figure 16, Représentation des interactions des différents facteurs impliqués dans la maturation de la graine. Les gènes régulés par ABI3 sont représentés en cercle et classés en fonction des autres facteurs impliqués dans leur régulation. Les couleurs indiquent la présence de séquences cis-régulatrices particulières. Bleu, RY+G-box; rouge, Rybos; vert, G-box et jaune sans motif particulier. Issu de Mönke G, *et al.* 2012.

publication indique qu'un mutant d'insertion dans le gène *AtLEA1-1* se caractérise par une maturation et une déshydratation prématurées des graines (figure 15) (Manfre et al., 2006).

Les protéines LEA pour lesquelles les fonctions ont été les plus étudiées sont celles du groupe 2 ou déhydrines. L'ensemble des travaux réalisés *in vitro* indique que ces protéines sont impliquées dans la protection des enzymes, une possible protection des membranes et de l'ADN ainsi qu'une interaction avec des formes toxiques d'oxygène (Cuevas-Velazquez et al., 2014; Graether and Boddington, 2014; Clarke et al., 2015). Des analyses *in vivo* elles montrent une corrélation entre la production de protéines LEA du groupe 2 et la résistance de plantes transgéniques à des stress thermiques (froid) ou à une déshydratation (Hara et al., 2003; Puhakainen et al., 2004; Shekhawat et al., 2011; Zhang et al., 2012; Shi et al., 2015).

Finalement la majorité des travaux portant sur l'étude de la fonction des protéines LEA montrent très souvent un effet sur la résistance aux stress relatifs à l'eau, la concentration en sels ou la température Wang et al, 2008; Li et al, 2015 (Wang et al., 2014; Jia et al., 2014 ; Liu et al., 2015; Yokota et al., 2015; Gao and Lan, 2016). Malgré l'ensemble de ces informations il reste cependant difficile de définir une fonction particulière à chacun des groupes de protéines.

De nombreux travaux ont, par ailleurs, porté sur le rôle des protéines ABI3, ABI4, LEC1, LEC2 et FUS3 dans la régulation des gènes *LEA* et la maturation des graines.

Différentes analyses montrent une fonction similaire des protéines ABI3 de différentes espèces végétales. Il a de plus été établis que les domaines B2 et B3 de ABI3 interagissent avec différentes séquences cis-régulatrices de promoteurs (Ezcurra et al., 2000b). Une analyse par RNA seq des cibles de ABI3 montre la diversité des gènes régulés et confirme la complexité du contrôle de la maturation des graines (figure 16) (Monke et al., 2012). Différents facteurs de transcription tels que WRKY41 et WRKY2 se fixent sur le promoteur du gène *ABI3*, WRKY2 régule aussi *ABI5* (Jiang and Yu, 2009; Ding et al., 2014). De manière intéressante des protéines similaires telles que WRKY18, WRKY40 et WRKY60 sont, elles, capables de réguler l'expression des gènes *ABI4* et *ABI5* (Liu et al., 2012). Des modifications épigénétiques semblent aussi affecter l'expression du gène *ABI3* il serait intéressant de savoir si certaines de ces modifications agissent comme régulateur spécifiques (Perruc et al., 2007; Tanaka et al., 2008; Ryu et al., 2014)

De nombreux travaux ont aussi porté sur le gène *ABI5*. Il est établi que la protéine HY5 est, sous le contrôle de l'ABA, impliquée dans l'induction de *ABI5* (Brocard et al., 2002; Piskurewicz et al., 2008; Chen and Xiong, 2008; Xu et al., 2014). Nous savons désormais aussi que la protéine ABI5 peut être activée par phosphorylation. Différentes kinases impliquées ont été identifiées (Hu and Yu, 2014; Zhou et al., 2015). En plus des Sérines, l'Alanine en position 214 est importante pour l'activité de la protéine ABI5 (Tezuka et al., 2013). Il a, par ailleurs été montré que l'association de ABI5 avec MED25 inhibe l'action activatrice de ABI5 sur l'expression des gènes (Chen et al., 2012). Une protéine AFP est elle responsable par interaction avec ABI5 d'entraîner sa dégradation (Lopez-Molina et al., 2003). Enfin la protéine ABI5 est conservée dans le règne végétal et son orthologue chez *Gladiolus hybridus*, par exemple, présente les mêmes conditions d'induction et probablement les mêmes fonctions (Wu et al., 2015). De manière assez surprenante la protéine ABI5 régule l'expression du gène *FLC* impliqué dans la floraison (Wang et al., 2013b).

Enfin de nombreuses analyses confirment les inter-connections des protéines ABI3, ABI5, FUS3, LEC1, LEC2 dans le contrôle de la maturation des graines. L'analyse par "differential

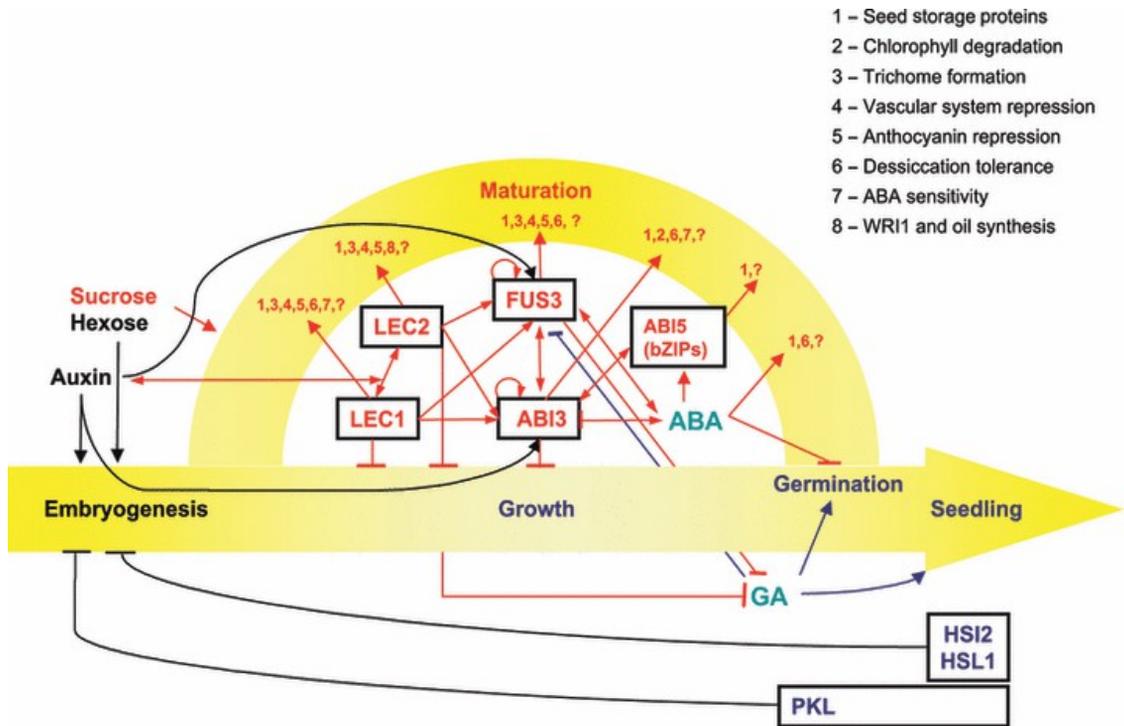


Figure 17, Modèle d'inter-connection des différents facteurs impliqués dans le contrôle de la maturation des graines. Issu de Santos-Mendoza et al, 2008.

display" des gènes affectés par les mutations *abi3* et *fus3* a par exemple permis de confirmer une de nos observations à savoir l'effet inhibiteur, dans certains cas, de ces facteurs de transcription (Nambara et al., 2000). Tous les travaux réalisés indiquent que les régulations mises en jeu sont complexes comme cela est représenté dans le modèle de la figure 17 (Santos-Mendoza et al., 2008). De nombreux nouveaux facteurs ont de plus été identifiés au cours de ces dernières années. On peut citer par exemple NFYA1 et RAV1 (related to ABI3/VP1) qui est une cible de kinases SnRK2s et qui a été identifié comme capable de se lier sur les promoteurs des gènes *ABI3*, *ABI4* et *ABI5* et de réprimer leur expression (Li et al., 2013b ; Feng et al., 2014; Kim et al., 2016).

Malgré tout ce qui a été effectué de nombreux travaux restent à réaliser et les techniques d'études actuelles, plus globales, devraient permettre la proposition de modèles ou de cascades d'induction plus complètes et précises.

Pol IV-RDR2-dependent RdDM

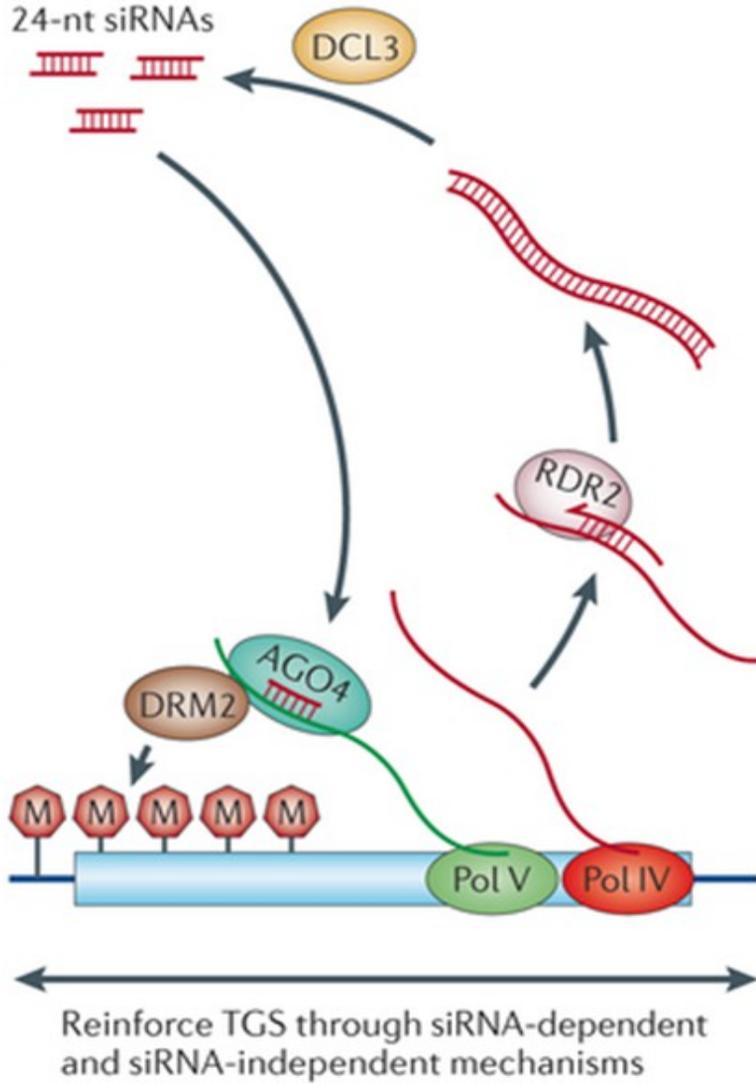


Figure 18. Représentation schématique du modèle RdDM. Réalisé à partir de Matzke et Mosher, 2014

5/ Activité de recherche actuelle axée sur l'étude des mécanismes dits de répression ou de silencing chez les végétaux

En 2005, au moment du départ à la retraite de Michel Delseny, deux possibilités se présentaient à moi. La première consistait à poursuivre sur la même thématique, c'est à dire la régulation des gènes par l'ABA et les gènes *LEA*, et la seconde à intégrer une autre équipe du laboratoire. En raison de mon intérêt pour l'étude de la régulation de l'expression des gènes j'ai décidé de me pencher sur un nouvel aspect de cette régulation, celle abordée par l'équipe dirigée par Thierry Lagrange. Ce choix fut le fruit d'une longue réflexion car il imposait une réorientation de mon travail sur une toute nouvelle thématique, à savoir l'étude des mécanismes de répression des gènes ou "gene silencing" et de la mise en place de marques épigénétiques chez les végétaux.

5-1/ Contexte scientifique

5-1-1/ Identification des ARN Polymérase IV et V et de la voie du RdDM (RNA directed DNA Methylation) chez *Arabidopsis thaliana*.

Au moment de ma décision, l'équipe de Thierry travaillait sur la caractérisation de deux grandes sous-unités de nouvelles ARN polymérase IV et V chez *A. thaliana* (NRPD1 et NRPE1). L'identification de ces deux protéines résultait d'une recherche, dans les séquences du génome d'*Arabidopsis*, de gènes codant des protéines similaires aux grandes sous-unités des trois ARN polymérase connues, à savoir NRPA1 (ARN polymérase I), NRPB1 (ARN polymérase II), et NRPC1 (ARN polymérase III). Au même moment, deux autres équipes, celle de Craig Pikaard et de David Baulcombe, travaillaient elles aussi sur l'étude de grandes sous-unités d'ARN polymérase. Ces deux derniers groupes ont principalement montré que la protéine sur laquelle ils travaillaient, NRPD1, est impliquée dans la mise en place de l'hétérochromatine et permet la formation d'un nouveau complexe d'ARN polymérase spécifique des végétaux nommé ARN polymérase IV. Leur travaux ont permis de démontrer que NRPD1 (POLIV) est nécessaire à la production de petits ARN, les siARN (short interfering RNA) ou hcsiRNA (heterochromatine siRNA), molécules clé guidant la méthylation de l'ADN de façon séquence spécifique par une voie baptisée RdDM (RNA directed DNA Méthylation) (Onodera et al., 2005; Herr et al., 2005). L'équipe de Thierry a plus particulièrement montré que la sous-unité, NRPE1, permet la formation d'une autre polymérase l'ARN polymérase V (POLV), et est également impliquée dans la voie du RdDM (Pontier et al., 2005; Pontes et al., 2006). Cette étude, ainsi que celle menée de façon concomittante par l'équipe de Matzke par la voie génétique, a permis de proposer que PolV, bien que non nécessaire à la formation des siRNAs, agit en aval de l'ARN polymérase IV, en permettant aux siRNA de guider la méthylation de l'ADN (Pontier et al., 2005; Matzke et al., 2006;).

C'est à partir de L'ensemble de ces travaux qu'un modèle de fonctionnement du RdDM a pu être proposé (figure 18). Ce modèle suggère que l'ARN polymérase IV synthétise de longs transcrits non codant qui sont pris en charge par RDR2, une ARN polymérase ARN dépendante, conduisant ainsi à la synthèse de molécules d'ARN "double brin". Ces derniers vont être coupés par DCL3, une ribonucléase à motif RNA III de la famille des Dicer, en petits ARN de 20 à 25 nucléotides qui interagissent alors avec une protéine ARGONAUTE, AGO4, cette association s'accompagne de la dégradation de l'un des brins du duplex. C'est à ce stade que l'on parle de

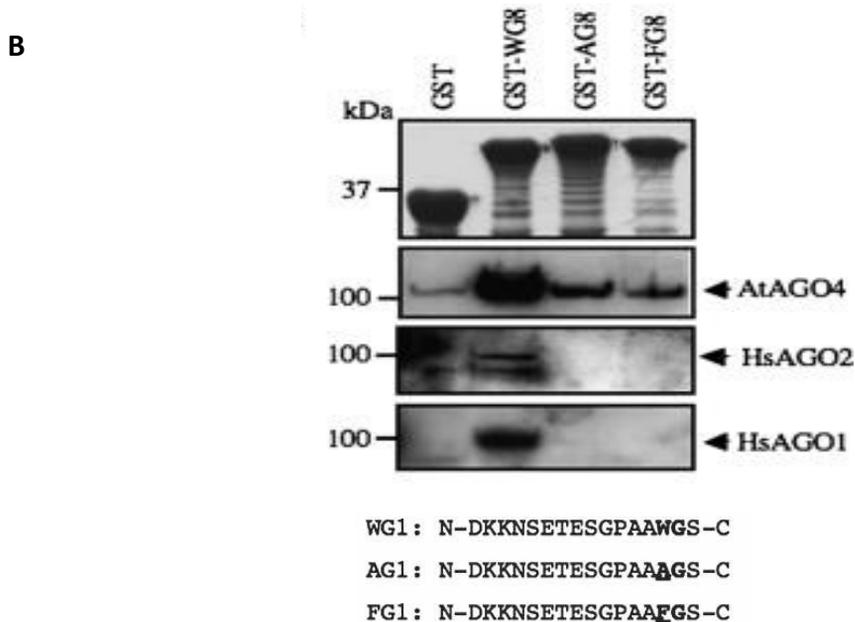
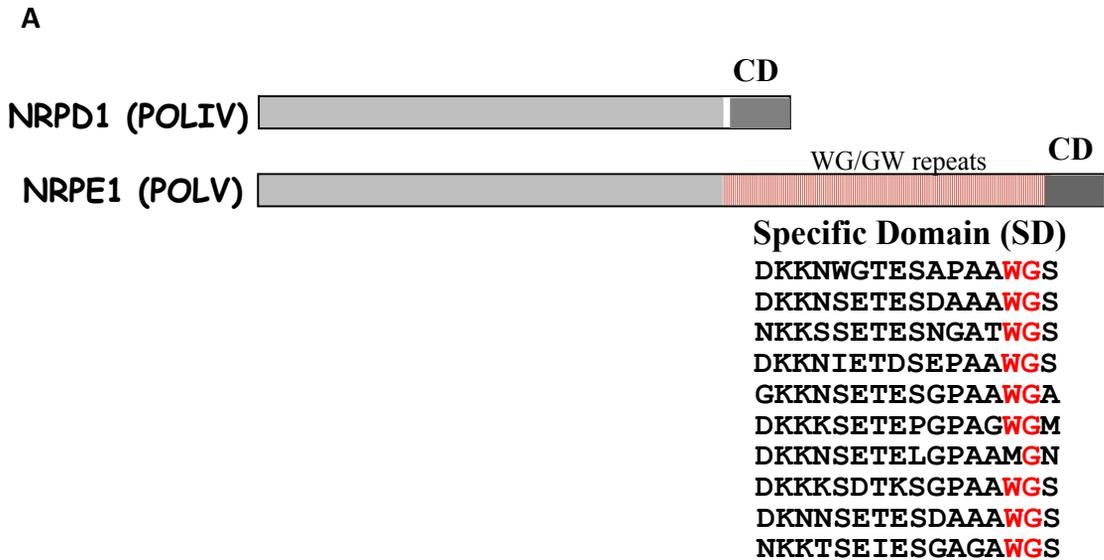


Figure 19. Caractérisation de la protéine NRPE1. A, Représentation schématique des protéines NRPD1 et NRPE1. Les séquences en acides aminés des régions répétées du domaine SD sont indiqués. B, Interaction du domaine AGO hook de NRPE1 avec les protéines Argonautes (AGO). GST pull down réalisé avec la GST ou des fusions entre la GST et huit copies du domaine AGO hook muté ou non. Les incubations ont été réalisées en présence d'extraits protéiques de plantes Myc::AGO4, HA::HsAGO1 ou HA::HsAGO2. Une coloration coomassie est utilisée comme contrôle de quantité. Résultats présentés par El-Shami et al., 2007.

siARN, *hcsiARN* (heterochromatin siARN) ou encore *easiARN* (epigenetically active siARN) associés à AGO4. La protéine AGO4 et les *siARN* se lient ensuite avec les transcrits de l'ARN polymérase V et permettent le recrutement d'ADN méthylases conduisant à la méthylation spécifique de cytosines de l'ADN. Cette méthylation entraîne alors une inhibition de la transcription.

Ce qui était particulièrement intéressant dans la caractérisation de ces deux nouveaux complexes de polymérase est qu'ils sont spécifiques aux végétaux et agissent dans la même voie de régulation de la chromatine et de la répression de l'expression de séquences en général répétées, souvent des transposons, du génome.

5-1-2/ Identification et caractérisation de la plateforme AGO hook de NRPE1 (POLV).

La comparaison des séquences des deux sous-unités NRPE1 et NRPE1 met en évidence que leurs régions C-terminales sont différentes notamment en raison de la présence d'un domaine particulier (SD Specific Domain) uniquement présent chez la protéine NRPE1 (figure 19A). L'équipe de Thierry a montré que le domaine SD est indispensable au RdDM (El-Shami et al., 2007). En effet contrairement à une version complète, une version tronquée, délétée de ce domaine (NRPE1 Δ SD) ne permet pas la restauration de la méthylation sur les cibles du RdDM dans un mutant *nrpe1*. Une analyse plus approfondie du domaine SD a permis de mettre en évidence la présence de séquences répétées (figure 19A) ainsi qu'un biais important dans la proportion de certains acides aminés, notamment un taux particulièrement élevé en Glycine (G), Tryptophane (W) et Serine (S) ainsi qu'une fréquence d'acide Glutamique, d'acide Asparagique et d'Asparagine légèrement supérieure à la normale (El-Shami et al., 2007). De manière intéressante, un domaine semblable de par ses caractéristiques biochimiques est présent chez d'autres grandes sous-unité d'ARN polymérase végétales qui pourraient par conséquent être des orthologues de NRPE1. Un domaine caractérisé par les mêmes particularité est aussi présent dans la protéine humaine GW182 (Homo sapiens Glycine/Tryptophane-rich protein). Cette protéine est capable d'interagir avec les protéines AGO1 et AGO2 et est aussi impliquée dans des mécanismes de RNAi (Jakymiw et al., 2005; Liu et al., 2005). De manière intéressante des approches de co-immunoprécipitation et de GST pull down ont permis de démontrer que c'est le domaine C terminal de la protéine NRPE1 qui est responsable de l'interaction avec AGO4 (Li et al., 2006). Cette observation et le fait que les domaines C terminaux de NRPE1 et GW182 se caractérisent par la présence conservée d'association de résidus Tryptophane/Glycine ou Glycine/Tryptophane espacés régulièrement, sans toutefois définir des répétitions totalement identiques, ont conduit à penser que ces résidus pouvaient être responsables de l'interaction avec des protéines AGO. Un consensus du domaine SD de NRPE1 a été multimérisé afin de produire différentes fusions protéiques GST composées de 8 répétitions de cette séquence. Pour une de ces versions, les résidus Tryptophane ont été remplacés par des Alanine ou Phénylalanine. Les résultats montrent une forte interaction de la version non mutée exclusivement avec les protéines AGO4 d'Arabidopsis et AGO1 et AGO2 humaines (figure 19B). Cela démontre le rôle clé des résidus Tryptophane dans l'interaction avec les protéines AGO (El-Shami et al., 2007). Ces analyses ainsi que d'autres observations ont permis de proposer que des domaines présentant des caractéristiques biochimiques identiques à celles du domaines SD de NRPE1, de la protéine humaine GW182 et de la protéine TAS3 de levure sont responsables de l'interaction avec des protéines AGO (Till et al., 2007). De tels domaines composés de résidus Tryptophane associés à des résidus Glycine et présents en plusieurs copies ont alors été appelés plateforme à protéines AGO ou AGO hook.

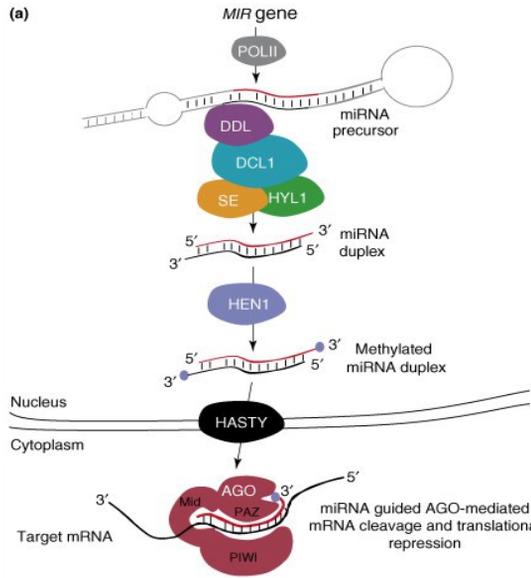


Figure 20. Représentation schématique de la voie des *miARN* impliqués dans le PTGS. Les différentes protéines impliquées dans cette voie sont indiquées. Issu de Mallory et al, 2008

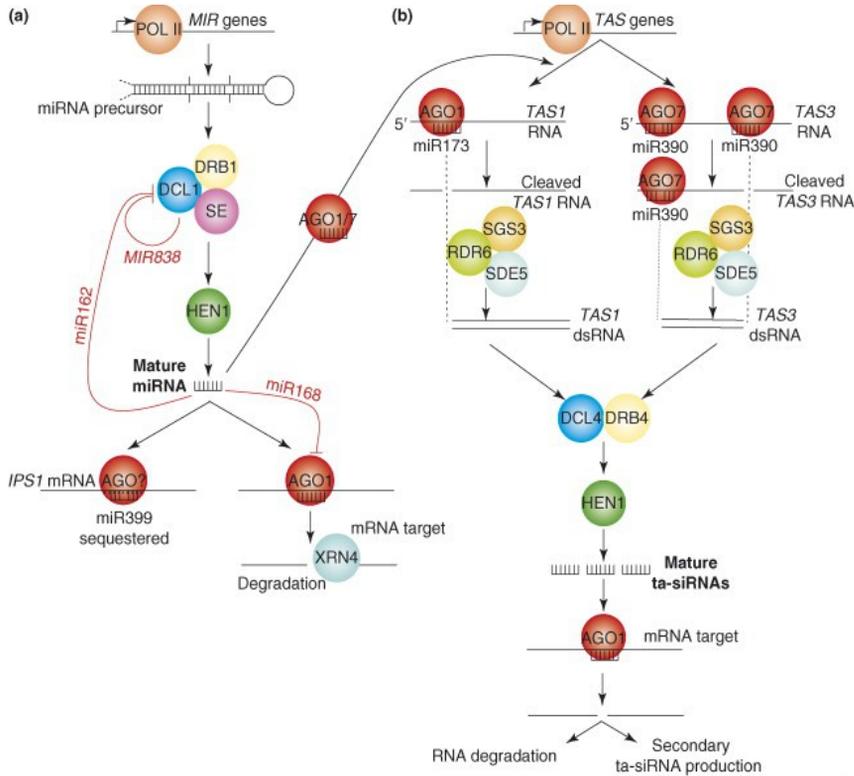


Figure 21 Représentation schématique de la voie des *tasiARN*. Issu de Mallory et Bouché, 2008

5-2/ Objectifs, stratégie et motivations scientifiques.

5-2-1/ Objectifs et motivations scientifiques.

Au regard des précédentes informations et parce que toutes les voies de répression liées à de petits ARN impliquent des protéines AGO il semblait particulièrement intéressant d'envisager d'identifier et de caractériser de nouvelles protéines d'*Arabidopsis* possédant des plateformes à protéines AGO. Il est important de noter que l'importante diversité des protéines AGO chez les végétaux, 10 chez *Arabidopsis*, et celle des petits ARN permettent de penser que les voies impliquant ces molécules ne sont peut être pas encore toutes identifiées ou totalement caractérisées. Les objectifs de notre approche consistent par conséquent à compléter nos connaissances et potentiellement identifier de nouvelles voies impliquant de petits ARN et des protéines AGO.

Point sur les différentes voies régulatrices ARN dépendantes déjà identifiées.

La voie des *hcsiARN* et des *راسيARN*.

Les premiers *siARN* identifiés et désormais nommés *hcsiARN* ou *easiARN* sont impliqués dans la voie du RdDM déjà explicitée et qui correspond à une répression transcriptionnelle, aussi nommée TGS (Transcriptional Gene Silencing). Des transcrits synthétisés par la polymérase IV conduisent par le biais des actions des protéines RDR2 et DCL3 à la production de petits *siARN* ou *hcsiARN* qui après association avec AGO4 sont recrutés au niveau du complexe POLV et permettent de diriger la méthylation de l'ADN. Parmi les *hcARN* se trouve une catégorie particulière les *راسيARN* (repeat associated small RNA) qui sont plus particulièrement impliqués dans la méthylation de l'ADN de répétitions présentes dans le génome. La voie du RdDM est particulièrement importante au niveau des transposons et des séquences répétées de génome et est considérée comme permettant une certaine protection du génome.

La voie des *miARN* ou *microARN*.

La voie des *miARN* correspond à une répression post transcriptionnelle aussi nommée extinction post transcriptionnelle des gènes ou PTGS (Post Transcriptional Gene Silencing). Cette voie est initiée à partir d'ARN particuliers qui forment des structures "tiges boucles" qui, par l'action de DCL1 permettent la libération de petits ARN simples brins pris en charge par des protéines AGO, très souvent AGO1. Ces complexes ciblent alors des ARNm et entraînent leur dégradation ou inhibent leur traduction (figure 20) (Mallory et al., 2008; Brodersen et al., 2008; Garcia, 2008)

La voie des *tasiARN*.

La voie des *tasiARN*, représentée dans la figure 21 débute de manière similaire à celle des *microARN* mais les produits de clivage des ARNm cibles (messagers TAS) vont par action des protéines RDR6, SDE5, SGS3 et SDE3 générer la formation d'ARN double brins qui par interaction avec DCL4 et DRB4 permettent la production de nouveaux petits ARN nommés *tasiRNA* (*trans cis acting RNA*). Ces derniers, se lient alors à des protéines AGO et ciblent des ARNm, correspondant souvent à des facteurs de transcription, et entraînent leur dégradation (Mallory and Bouche, 2008; Ramachandran and Chen, 2008; Mallory and Vaucheret, 2010).

La voie des *natsiARN*.

La production de *natsiRNA* résulte de la transcription à un même locus ou à des loci différents de transcrits qui sont complémentaires et donc capable de former des duplex. Ces duplex vont par action de protéines DCL (DCL1, DCL2, DCL3 et DCL4), de protéines RDR (RDR1,

RDR2 et RDR6) et des protéines SDE4 et SGS3 permettre la production de *natsiARN* de 21 nucléotides (Borsani et al., 2005; Yuan et al., 2015). Il semblerait que les *natsiARN* régulent leurs cibles à plusieurs niveaux, celui de leur transcription, de leur maturation, de leur épissage, de leur stabilité, de leur transport et ou de leur traduction. Une étude récente consistant à rechercher dans le génome des séquences transcrites capables de s'apparier a permis de proposer la présence de plus de 5000 *natsiARN* potentiels chez *Arabidopsis* (Yuan et al., 2015). Les auteurs de ces travaux ont de plus démontré que la production de nombreux transcrits permettant d'initier la voie des *natsiARN* est dépendante de stress environnementaux (Yuan et al., 2015).

Observations qui indiquent que les mécanismes identifiés sont probablement plus complexes et que tous les mécanismes régulés par les protéines AGO et les petits ARN n'ont probablement pas encore tous été caractérisés.

Il est important de noter tout d'abord que chacun des gènes codant des protéines AGO chez *Arabidopsis thaliana* se caractérise par un profil d'expression spécifique et que certains, notamment les gènes codant les protéines AGO1, AGO4 et AGO5, peuvent être différentiellement régulés en condition de stress (Mallory and Vaucheret, 2010) (Li et al., 2012). Il est à ce sujet intéressant de noter que les protéines AGO1, AGO2 et AGO5 sont, elles mêmes, impliquées dans des mécanismes de résistance des plantes lors d'infection par des pathogènes principalement viraux (Harvey et al., 2011) (Pumplin and Voinnet, 2013) (Garcia-Ruiz et al., 2015). En fonction du virus, le mécanisme d'interférence, de résistance de la plante hôte, n'est pas toujours le même mais dans la plupart des cas le virus réprime les voies de silencing de la plante afin de permettre sa multiplication et contourner ainsi un mécanisme de défense (Pumplin and Voinnet, 2013). Une autre information troublante est que certaines protéines AGO peuvent avoir des rôles similaires et antagonistes. C'est le cas par exemple des protéines AGO1 et AGO10 qui ont un effet similaire sur la production du *microARN 172* et un rôle antagoniste sur la régulation des *microARN 165* et *166* (Ji et al., 2011). Il est important aussi de prendre en compte la diversité de taille des petits ARN (de 21 à 25 nucléotides) caractérisés à ce jour. Il est, en général, décrit que les ARN de petite taille (21 nt) sont principalement associés aux protéines AGO1 et AGO2 et ceux de plus grande taille (24 nt) majoritairement associés aux protéines AGO4 et AGO5 (Kim, 2008) (Mi et al., 2008) (Wang et al., 2011). On observe cependant que cette règle n'est pas exclusive et que, par exemple, de petits ARN (21 ou 22 nucléotides) peuvent être associés à AGO5 (Mi et al., 2008). Une autre observation est une potentielle préférence des protéines AGO selon le nucléotide en 5' du petit ARN (Mi et al., 2008). Majoritairement les petits ARN possédant une uridine en 5' s'associent avec la protéine AGO1, ceux possédant une adénine s'associent avec les protéines AGO2 ou AGO4 et enfin ceux possédant une cytosine s'associent plutôt avec AGO5 (Kim, 2008). De nombreux travaux montrent cependant que là encore cette règle n'est pas toujours respectée, et que AGO1, par exemple, peut interagir avec des ARN possédant n'importe quel nucléotide en 5' (Mi et al., 2008).

L'ensemble de ces différentes observations souligne que les mécanismes qui impliquent des protéines AGO et des petits ARN sont complexes et qu'il est nécessaire de multiplier les travaux de recherches et les approches afin de parvenir à tous les identifier et mieux les comprendre.

5-2-2/ Mise en place de la stratégie.

Une des premières difficultés de l'approche choisie, qui consiste à identifier de nouvelles protéines possédant des plateformes à AGO, était que nous ne recherchions pas des protéines possédant un domaine conservé mais plutôt des domaines présentant un biais particulier dans leur composition ainsi que des résidus Tryptophane associés à des résidus Glycine à des intervalles plus

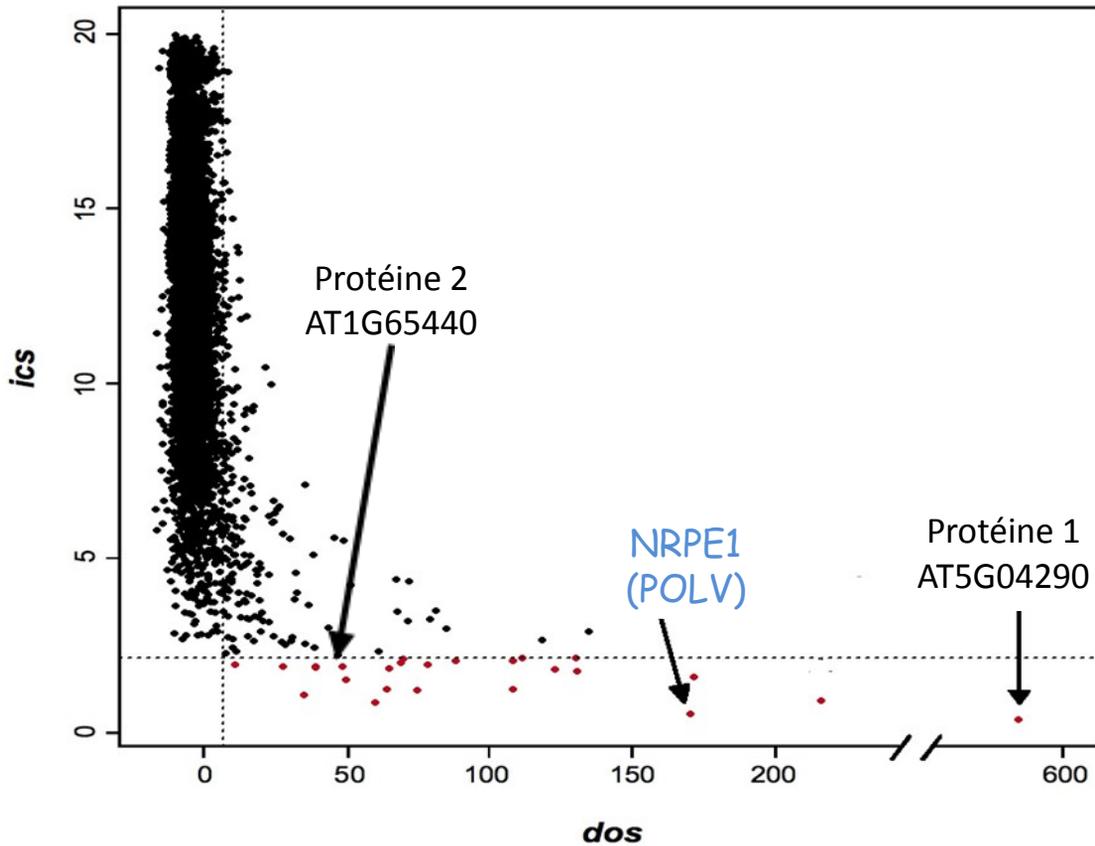


Figure 22 Visualisation, en fonction des caractéristiques recherchées, des protéines d'*Arabidopsis* qui possèdent une potentielle plateforme AGO hook. La position de chaque spot est relative aux caractéristiques de chaque protéine en fonction des critères recherchés (présence de Tryptophane associés à des Glycine, présence de répétitions et biais dans la présence de certains Acides aminés).

ou moins réguliers. La seule manière d'aborder cette problématique était de réaliser une recherche de type Psi-Blast permettant de rechercher dans des bases de données des séquences présentant des caractéristiques similaires à des séquences de référence. La stratégie choisie, en partie réalisée par Richard Cooke, a permis d'identifier plusieurs candidats qui ont été classés en fonction des caractéristiques recherchées et des domaines qu'ils possédaient et qui pouvaient laisser supposer une fonction particulière (figure 22) (Karlowski et al., 2010). Parmi toutes les protéines candidates, mon travail a principalement porté sur la caractérisation de deux d'entre elles, indiquées sur la figure 22, qui en plus des potentielles plateformes à AGO présentent des similarités avec des protéines impliquées dans la transcription.

5-3/ Résultats

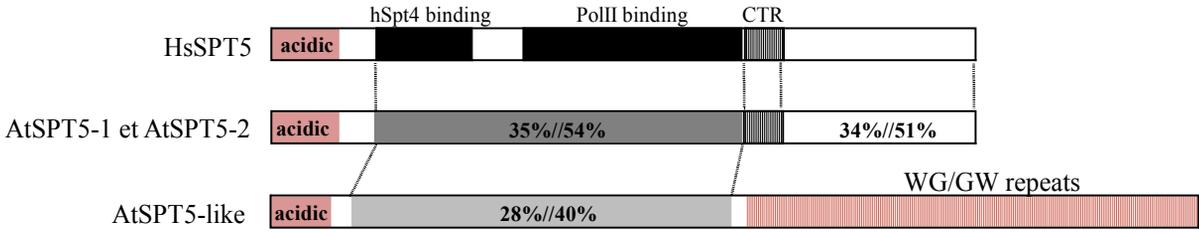
5-3-1/ Etude du gène *AT5G04290*

5-3-1-1 Analyse de la séquence.

- Similarité de la séquence du gène *AT5G04290* avec celles codant des protéines SPT5.

Le gène *At5g04290* code une protéine présentant de très fortes similarités avec la protéine SPT5/NusG ou SuPresseur d'insertions Ty qui a initialement été identifiée à partir d'un cribble génétique, réalisé chez *S. cerevisiae*, qui avait pour objectif d'identifier des mutations qui suppriment le défaut de transcription causé par des insertions Ty dans le promoteur du gène *HIS4* (Winston et al., 1984 ; Ponting, 2002). Les protéines SPT5 sont très conservées et sont composées d'un domaine N terminal acide ou domaine NusG N (NGN), de plusieurs domaines KOW (Kyprides, Ouzounis, Woese), potentiellement impliqués dans les interactions protéines/protéines ou protéines/nucléotides, et enfin de petites répétitions présentes à proximité de l'extrémité C terminale (CTR) (Martinez-Rucobo et al., 2011). Les protéines SPT5 forment, via leur domaine NGN, un complexe avec des protéines SPT4 et ce dernier s'associe avec la polymérase II, et potentiellement avec la polymérase I, au moment du passage d'un état d'initiation de la transcription à un état d'élongation et reste associé au complexe de transcription tout au long de son déplacement le long des séquences codantes (Hartzog et al., 1998; Schneider et al., 2006; Guo et al., 2008; Glover-Cutter et al., 2008; (Wenzel et al., 2010)Martinez-Rucobo et al., 2011; Hartzog and Fu, 2013) Les domaines KOW sont responsables de l'interaction avec la polymérase II (Hartzog and Fu, 2013). De manière intéressante une liaison directe de SPT5 à l'ARN peut aussi être observée (Glover-Cutter et al., 2008).

De nombreux travaux, souvent basés sur la caractérisation de mutants de levure, indiquent que les protéines SPT5 sont impliquées dans l'élongation des transcrits et sont importantes pour coupler les modifications de la chromatine avec la vitesse de transcription et la maturation des ARN (Wen and Shatkin, 1999; Zorio and Bentley, 2001; Winston, 2001; Hartzog, 2003; Svetlov and Nudler, 2011; Werner, 2012; Hartzog and Fu, 2013). Certains mutants *spt5* se caractérisent par exemple par des défauts d'épissage (Hartzog and Fu, 2013) (Xiao et al., 2005). Il a de plus été décrit que SPT5 est impliquée dans la réparation de l'ADN couplée à la transcription (Hartzog and Fu, 2013). Un rôle dans le développement a également été proposé (Guo et al., 2000).

A**B**

MDRKGKGVAGSDSYSGGQKRKNSVEFRDEGLRIKKRKNPEVLQFFFEESAIEVGYGGSSDEDDDLGFLNDMEDEPEVEESSKAGK
 GEKGSFVFPKEEDLNEEEFDRIMEERYKPGSGFLRYADDDIKDAIEMDALAPT SKDPP IWKVKCAIGRERRSVFCLMHKFVLELRK
 IGTKLEIISVFSVDHVKGFI FIEADKEHDVLEACKSLVGIYATRMVLLPKAETPNLLTVQKKTKKVSEGTWARVKNKGYKGDLAQIV
 AVSDTRNKALIKLIPRIDIQALTQYGGGVTVQKQTPAPRLISSSELEEFRLIQVRRDRDTGITFEHLDSLMLKDG YLYKKVSLD
 SISSWGVIPTKDELKFTPVDRKETGDVWEISEIYGEERKKKILPTCREGGKGEKGGGGKGGSGGGKGGESRGGKGGSSDFKSES
 SYELYNLVCFSRKDFGLIVGVDDKGDGKYVLKEGIDGPVVVTVGKKEMQNGPFD SKFTALDLNKKQISVNDVVKISKGPSEKQGVV
 RQVYRGIIFLYDESEEENGGYFCCSKSQSCEKVKLFTEESNEKTGGFDGTA FEFVSSPKSPLSPEKEWQPRERYNSSNQGDIGSTYS
 IGQKLRIRVGPLKGYLCRVIALRYSVTVKLDLSQHKIFTVKSEHLAEVRDRNTVLS TSGDAGTGSFQPF GMLGTESSTGDWAIGAGT
 SSEGGNWNIGGPSTSDSNPWGSGKPTSDVSVSTLQPTVADDNTSAWANA AENKPASASDQPGGWNPWGKTPASEAGTVSGWGDTSAS
 NVEASSWEKQGASTSNVADLGSWGTHGSSGGNKQDEDSVWGKLCEASESSQKKEESSWGKKGSDGESSWGNKDGNS SASKKDGVS
 WGQQDKGSDSKGSAWSNQCGDFSGGKKKDGSSGWNKSAEDSNANSKGVDPWGWQPN DGSSWGKKGDGAASWGKKDDGSSWGKKDDG
 NKDDGSSWGKKDDGQKDDGSSWEKFFDGGSSWGKKDDGSSWGKKDDGSSLWGKKDDGSSWGKEDDGGSLWGKKDDGSSWGKK
 DDGESSWGKKDDGSSWGKKDEGGYSEQTFDRGGRGFGRRRGGRRGRDQFGRGSSFGNSEDPAPWSKPSGSSWGKQDGDGGSS
 WGKENDAGGGSSWGKQDNGVSSWGKQNDGSGGGSSWGKQNDAGGGSSWGKQDSGGDSSWGKQDGGGDSGSAWGKQNTSGSSWG
 KQSDAGGGSSWGKQDGGGGSSWGKQDGGGGSSAWGKQNETSNGSSWGKQNDSSGGSSWGKQDGGGGSSWGKONDGGGGSSWGKQ
 GDGGSKPWNEHSGGGRGFGERRGGGFRGGRNQSGRGRSFDGGRSSWKTDNQENTWKSQDQSGGSDWKKGWGEDSNNSKPSGSSAG
 GCAGNWP SWDTNSKKTNDKPGDDSKSAWGTSTNDQVNTDNNNDSWNKPNNDVGTSGEADNAWGGKTNVAVAPSPSGSAAWGTGDKKT
 GW

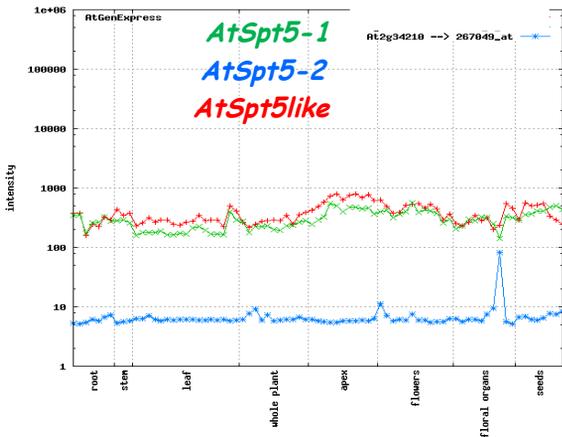
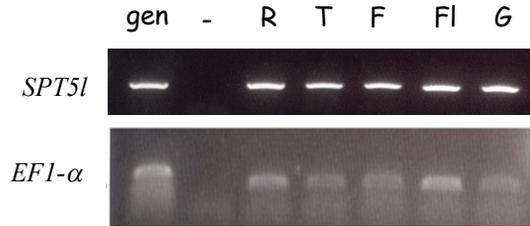
C**D**

Figure 23. Caractéristiques du gène *At5g04290*. **A**, Comparaison de la structure de la protéine SPT5 humaine avec les protéines SPT5 présentes chez *Arabidopsis*. **B**, Sequence de la protéine AT5G04290 (SPT5like). **C**, Profils d'expression des gènes *AT4G08340* (*AtSpt5-1*), *AT2G34210* (*AtSPT5-2*) et *At5g04290* (*SPT5like*). Données disponibles sur le site Atgene express. **D**, Profil d'expression du gène *At5g04290* (*SPT5like*). Des RT-PCR ont été réalisées à partir d'ADNc de racines (R), tiges (T), feuilles (F), fleurs (Fl) et graines (G). L'échantillon gen correspond à de l'ADN génomique. Les amorces correspondant au gène *Efla* ont été utilisées comme contrôle d'équilibre et d'absence de contamination, l'ADNc amplifié est de taille inférieure à celle du fragment génomique.

Le mode d'action de la protéine SPT5 a principalement été décrit à partir de l'étude des fractions protéiques, DSIF (DRB Sensitivity Inducing Factor), qui contient SPT5, et NELF, qui avaient été toutes les deux identifiées comme responsables de l'inhibition de l'élongation de la transcription par le 5,6-dichloro-1- β -d-ribofuranosylbenzimidazole (DRB) (Chodosh et al., 1989 ; Wada et al., 1998; Yamaguchi et al., 1999; Kaplan et al., 2000; Palangat et al., 2005; Wenzel et al., 2008; Mayer et al., 2010, Hartzog and Fu, 2013). Le DRB agit en inhibant la kinase pTEF qui est responsable de la phosphorylation de SPT5 et de la serine 2 de la grande sous-unité de la polymérase II qui permet la transition de l'état d'initiation à celui d'élongation (Ni et al., 2008; Zhou et al., 2009; Hartzog and Fu, 2013). La phosphorylation de SPT5, pour laquelle des kinases ont été identifiées chez de nombreux organismes, n'est en fait pas nécessaire pour son interaction avec la grande sous-unité de la polymérase II mais est nécessaire pour son action positive sur l'élongation (Zorio and Bentley, 2001; Larochelle et al., 2006 ; Zhou et al., 2009 ; Hartzog and Fu, 2013). La présence de SPT5 permet, au moins chez les levures, le recrutement du complexe PafI qui entraîne la méthylation de la lysine 4 de l'histone H4 (Jaehning, 2010; Hartzog and Fu, 2013). Une interaction de SPT5 avec PRMT5, une arginine méthyltransférase a par ailleurs été mise en évidence (Amente et al., 2005).

Certaines analyses suggèrent aussi que les protéines SPT5 pourraient avoir un rôle ou des effets "gène spécifique" (Komori et al., 2009). Les données actuellement disponibles à ce sujet ne permettent cependant pas d'établir de généralité et il est possible que la taille des gènes, leur niveau d'expression ou leur densité en nucléosome soient liés aux effets observés.

Ces différentes observations ainsi que le fait que les protéines SPT5 sont présentes chez les eucaryotes, les bactéries et les archées démontrent leur rôle capital et il était par conséquent intéressant de découvrir pourquoi chez *Arabidopsis thaliana* une telle protéine possède une plateforme à AGO.

- Diversité des gènes *SPT5* chez *Arabidopsis* et les autres organismes.

De manière intéressante, chez *Arabidopsis thaliana*, trois gènes codent des protéines de type SPT5, celui identifié par l'approche psi-blast (*AT5G04290*) et les gènes *AT4G08340* (*ATSPT5-1*) et *AT2G34210* (*ATSPT5-2*). Ces deux derniers codent des protéines très similaires entre elles avec une structure classique de protéine SPT5 alors que celle codée par le gène *AT5G04290* se caractérise par la présence d'une extension C terminale de type AGO hook (figure 23A). C'est la présence de cette extension, qui est codée comme pour la protéine NRPE1 (grande sous unité de la polymérase V) par deux longs exons, qui nous a conduit à nommer le gène identifié (*AT5G04290*) *SPT5like*. L'extension de la protéine SPT5like se caractérise par un nombre de résidus Glycine et Tryptophane supérieur à la normale et même supérieur à celui observé au niveau de la plateforme AGO hook de NRPE1 (44 résidus GW ou WG contre 17 chez NRPE1)(figure 23B).

D'après les bases de données de microarrays, les gènes *SPT5-1* et *SPT5like* ont une expression quasi constitutive alors que le gène *SPT5-2* ne semble être exprimé qu'au cours des dernières étapes du développement de la fleur (figure 23C). Le profil d'expression du gène *SPT5like* a été confirmé par RT-PCR (figure 23D). Il est important de signaler que des séquences similaires à la séquence *Spt5like* sont présentes dans les bases de données génomiques de tous les végétaux dont les génomes ont été totalement séquencés comme le riz, le colza, le citron, le cacao, le ricin, la vigne, l'eucalyptus, le peuplier et le prunier. En plus de ces versions codant des protéines SPT5like avec une potentielle plateforme AGO hook, un gène codant une protéine SPT5 classique sans extension est présent dans les différents génomes analysés. Par contre, aucun gène codant des protéines de type STP5like, avec une potentielle plateforme, n'a été identifié chez d'autres organismes que les végétaux.

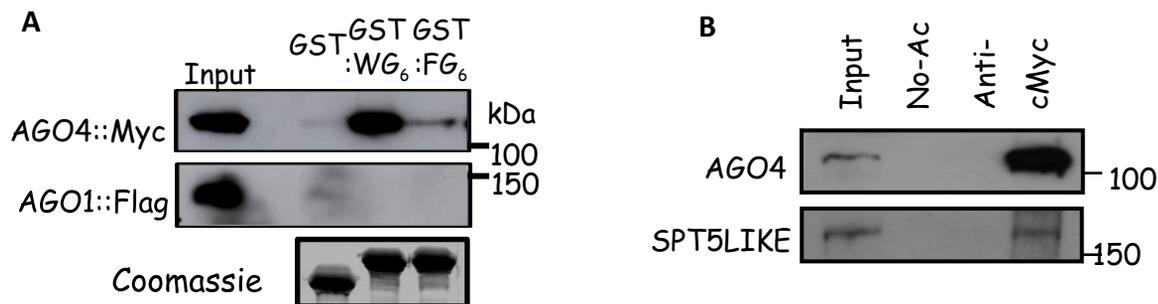


Figure 24. Test d'interaction de l'extrémité C-terminale de SPT5like avec des protéines AGO. **A**, GST pull down réalisé avec des protéines recombinantes (GST::WG₆) ou (GST::FG₆) incubées avec des extraits de plantes transgéniques (AGO4::CMyc ou AGO1::Flag). Les révélations ont été réalisées avec des anticorps anti-Myc et anti-Flag. **B**, Immuno-précipitation à partir de plantes produisant la protéine CMyc::AGO4. Une piste sans anticorps est utilisée comme contrôle négatif.

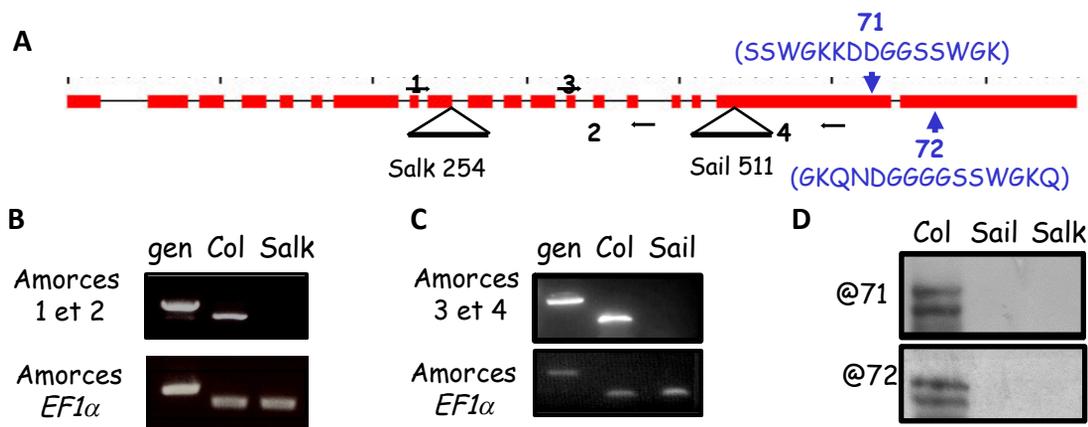


Figure 25. Caractérisation des mutants *spt5like*. **A**, Représentation schématique des sites d'insertion et sites d'hybridation des amorces utilisées pour caractériser les mutants d'insertion. Les positions des domaines et séquences utilisés pour la production d'anticorps spécifiques sont également indiquées. **B** et **C**, RT-PCR avec des amorces spécifiques encadrant les sites d'insertion des lignées Salk254 (**B**) et Sail (**C**). Les amorces correspondant au gène *Ef1a* ont été utilisées comme contrôle d'équilibrage et d'absence de contamination. **D**, Détection de la protéine SPT5like. Western blot réalisé avec les anticorps 71 et 72 et des extraits protéiques de fleurs sauvages (Col) ou des mutants *Sail254* et *Salk511*.

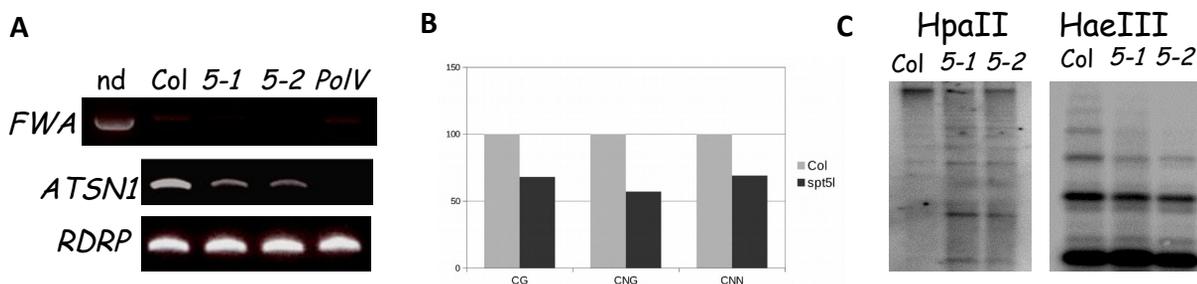


Figure 26. Analyse de la méthylation de cibles du RdDM chez les mutants *spt5like*. **A**, Analyse de la méthylation de la cible *ATSN1* révélée par Chop PCR. **A**, Chop PCR après digestion avec HaeIII. l'ADN de plantes sauvages, des mutants *spt5l-1*, *spt5l-2* et du mutant *nripe1 (polV)* ont été utilisés. Les amorces *FWA* sont utilisés comme contrôle de digestion et les amorces *RDRP* comme contrôle d'équilibrage. De l'ADN de plantes sauvages non digéré (nd) a également été utilisé comme contrôle. **B**, Analyse du locus *AtSn1* après traitement au bisulfite. **C**, Méthylation des cibles 5S révélées par Southern blot. Les ADN de plantes sauvages et de mutants *spt5l-1*, *spt5l-2* ont été digérés par les enzymes de restriction HpaII et HaeIII. La révélation est réalisée avec une sonde 5S.

5-3-1-2 Caractérisation de la protéine SPT5like.

- La protéine SPT5like est-elle capable d'interagir avec des protéines AGO ?

Nous avons tout d'abord analysé la capacité de l'extrémité C-terminale de la protéine SPT5like à interagir avec des protéines Argonautes. Pour cela des analyses de liaison par purification de complexes ont été effectuées avec deux types de fusions GST. La première correspond à une fusion de la séquence GST à un fragment d'ADN correspondant à 6 répétitions d'un domaine consensus KKDDGSSWGG présent plusieurs fois dans la protéine SPT5like (GST::WG6). La seconde fusion correspond à une forme mutée consistant en un remplacement de tous les résidus Tryptophane par des Phénylalanine (GST::FG6). Les protéines recombinantes ont été incubées en présence d'extraits protéiques de plantes transgéniques produisant soit la protéine fusion cMyc::AGO4 soit la protéine Flag::AGO1. Les analyses montrent que seule la version non mutée (WG6) interagit avec la protéine AGO4 (figure 24A). Aucune version n'interagit avec la protéine AGO1. Afin de confirmer l'interaction avec la protéine AGO4, *in vivo*, des co-immunoprécipitations ont été réalisées en utilisant des extraits de plantes transgéniques produisant la protéine fusion cMyc::AGO4. Après immunoprécipitation d'AGO4 par l'anticorps cMyc, la présence de SPT5like est révélée par détection avec un anticorps dirigé contre un peptide de SPT5like (figure 24B). Cette analyse confirme observations *in vitro* et permet de conclure que la protéine SPT5like possède un domaine d'interaction avec la protéine AGO4. Cette interaction suggère une possible implication de SPT5like dans la voie du RdDM, voie de silencing dans laquelle la protéine AGO4 est majoritairement impliquée.

- La protéine SPT5like est elle impliquée dans la voie du RdDM ?

Dans le but de progresser dans l'analyse il était alors important de caractériser des mutants d'insertion dans le gène *SPT5like*. Deux lignées SALK_001254C nommée *spt5l-1* et SAIL_362_G07 nommée *spt5l-2* étaient disponibles. Un schéma du lieu d'intégration de l'ADN-T dans ces lignées est fourni dans la figure 25A. Aucune de ces deux lignées ne produit d'ARNm *SPT5like* entier ou des protéines SPT5like (figure 25B, C et D). Il est important de signaler qu'aucune de ces deux lignées ne présente un phénotype visuel anormal.

Afin de déterminer si la voie du RdDM est affectée dans ces lignées, le niveau de méthylation de sites dépendant de cette voie a été analysé par CHOP PCR. La première cible qui a été testée est le site correspondant au rétrotransposon *AtSNI* (figure 26). On observe pour les deux mutants *spt5like*, après digestion de l'ADN avec une enzyme sensible à la méthylation et amplification par PCR, un signal d'intensité intermédiaire à celui observé à partir d'ADN de plantes sauvages et de mutants *nrpe1-1* (*POLV*) (figure 26). Cette observation indique que le degré de méthylation du locus *AtSNI* chez les mutants *spt5like* est sensiblement inférieur à celui observé chez des plantes sauvages mais supérieur à celui observé pour le mutant *nrpe1-1*. L'analyse du même locus par séquençage après traitement bisulfite (permettant de détecter la méthylation des cytosines) confirme pour le mutant *spt5like* un niveau de méthylation inférieur à celui observé avec l'ADN de plantes sauvages (figure 26B). Cette dernière analyse permet en plus d'observer que les trois contextes de cytosines CG, CNG et CNN présentent une méthylation diminuée de 20 à 40% chez les mutants *spt5like* comparée à celle observée dans les plantes sauvages. L'analyse du locus codant les ARN 5S, par Southern Blot réalisé après digestion de l'ADN avec les enzymes HpaII ou HaeIII, montre également un défaut de méthylation chez les deux mutants *spt5like* (figure 26C). L'analyse d'une autre cible du RdDM, le locus *soloLTR*, permet d'observer dans ce cas précis un taux de méthylation similaire chez les mutants *spt5like* et les plantes sauvages (figure 27A). La

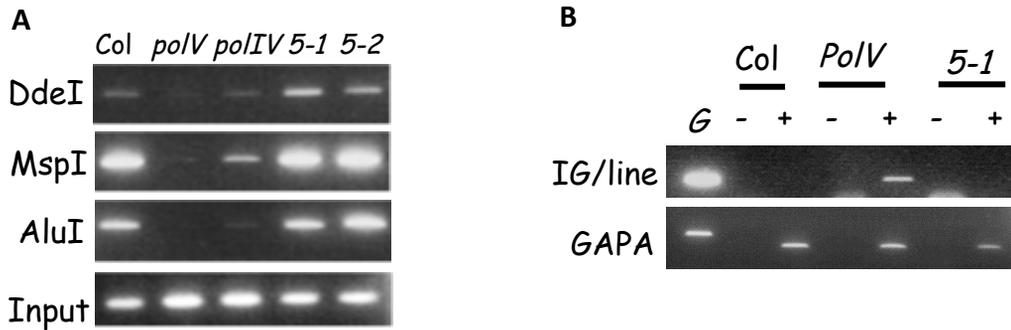


Figure 27. Analyse de la méthylation et de l'expression du locus soloLTR chez les mutants *spt5like*. **A**, Méthylation de *SOLOLTR* révélée par Chop PCR après digestion DdeI, MspI ou AluI. d'ADN génomique de plantes sauvages, des mutants *polV*, *spt5l-1* et *spt5l-2*. Une PCR sur ADN non digéré (input) constitue un contrôle d'équilibrage. **B**, Expression du locus *IG/Line*. RT-PCR réalisée à partir d'ARN de fleurs de plantes sauvages (Col), de mutants *polV* et *spt5l*. De l'ADN génomique de plantes sauvages (G) a été utilisé comme contrôle positif. Les ARN traités par la transcriptase reverse (+) ou non (-). Une amplification avec des amorces GAPDH permet de contrôler l'équilibrage.

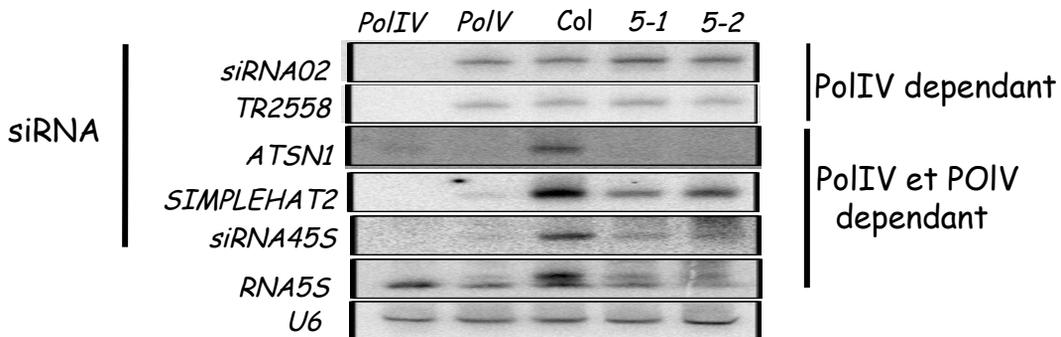


Figure 28. Analyse de l'accumulation des *siRNA* dans les mutants *spt5l*. Les petits ARN de fleurs de plantes sauvages (Col) et de mutants *polIV*, *polV*, *spt5l-1* et *spt5l-2* ont été hybridés par la technique du Northern blot avec des sondes correspondantes à différents *siRNA*. Une hybridation avec une sonde *U6* a été réalisée comme contrôle d'équilibrage.

même analyse chez le mutant *polIV* et *polV* indique, par contre, un important défaut de méthylation à ce locus. L'analyse de la dé-répression de ce locus indique que seul le mutant *polV* (utilisé comme contrôle) qui est déméthylé au locus *soloLTR* permet sa transcription ou dé-répression (figure 27B). L'ensemble de ces observations indique que certaines des cibles du RdDM (*AtSN1*, *5S*) sont affectées, peu méthylées, chez les mutants *spt5like* mais que d'autres comme le locus *soloLTR* ne sont pas altérées. Il semble donc que la protéine SPT5like n'est nécessaire que pour la méthylation complète de certaines cibles du RdDM. La raison de cette observation reste relativement obscure, nous n'avons pas trouvé de critères qui permettent de différencier les deux types de cibles analysés.

- La protéine SPT5like est-elle nécessaire pour la production des siRNA ?

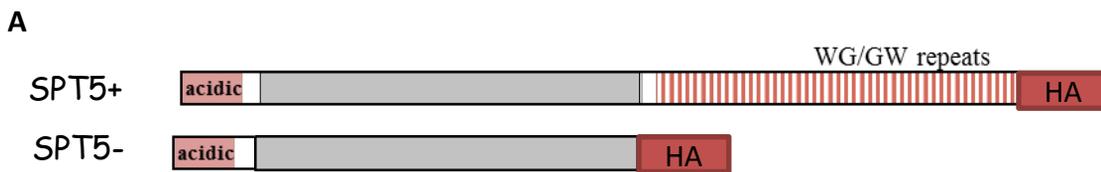
Dans le but de progresser dans nos travaux, nous avons tenté de déterminer à quel niveau de la voie du RdDM la protéine SPT5like est impliquée. Dans cet objectif nous avons analysé l'accumulation de *siARN* spécifiques de la voie du RdDM chez les mutants *spt5like*. Ce travail a permis d'observer que tous les *siARN* POLIV-dépendants décrits (*siRNA02* et *TR2558*) sont correctement produits chez les mutants *spt5like* (figure 28). Par contre les *siARN* POLIV et POLV dépendants (*AtSN1*, *simpleHAT2*, *45S* et *5S*) sont absents ou moins abondants chez les mutants *spt5like* que chez des plantes sauvages (figure 28). Les profil d'accumulation des *siARN* observés permettent de proposer que la protéine SPT5like agit à une étape proche de celle de l'implication de POLV, mais que son absence conduit à moins d'effets que l'absence de la grande sous-unité de POLV.

L'ensemble de cette première partie a été publiée dans *EMBO Reports*, RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family, (Bies-Etheve et al., 2009) et a fait l'objet d'un *literature reports* dans le même journal (Wang and Dennis, 2009). Par ailleurs ce travail a également été reconnu par le biais de Faculty of 1000. Une minirevue basée sur l'implication de POLV et SPT5like a également pu être publiée dans *Epigenetics*, Plant-specific multisubunit RNA polymerase in gene silencing (Lahmy et al., 2010).

- La protéine SPT5like est-elle capable d'interagir avec NRPE1 (POLV) ?

Du fait des premières observations et parce que les protéines SPT5 sont des facteurs d'élongation de la transcription, il nous semblait intéressant de tester l'interaction de SPT5like avec l'ARN polymérase V. Il avait été montré par des expériences d'immuno-localisations que ces deux protéines sont dans le noyau (He et al., 2009). Une purification du complexe POLV n'avait cependant pas permis de détecter la protéine SPT5like (Ream et al., 2009). De notre côté, nous n'avons jamais pu mettre en évidence, par immuno-précipitation, l'interaction de SPT5like avec NRPE1. Il est possible que cette liaison ne soit pas stable dans les conditions utilisées ou que la liaison soit très faible. Il a cependant été montré, par chromatine immunoprécipitation, que SPT5like se lie aux cibles du RdDM de façon dépendante de la présence de POLV (Rowley et al., 2011).

Au regard de l'ensemble de nos résultats et observations se posait la question du rôle de la plateforme à AGO de la protéine SPT5like. Ce rôle est-il lié à celui de la plateforme de NRPE1 ? Dans ce cas, pourquoi avoir dans un même complexe deux protéines capables de recruter des protéines AGO ?



B

MDRKGGKQVAGSDSYSGGQKRKNSVEFRDEGLRIKKRKNPEVLQFFEESEAEGVYGGSSDEDDDGFLNDMEDEPEVEESS
 KAGKGEKGS SFVFPKEEDLN EEEFDRIMEERYKPGSGFLRYADDDIKDAIEMDALAPT SKDPPIWKVKCAIGRERRSVFCLM
 HKFVELRKIGTKLEIIISVFSVDHVKGFI FIEADKEHDVLEACKSLVGIYATRMVLLPKAETPNLLTVQKTKKVVSEGTWARVK
 NGKYKGDLAQIVAVSDTRNKALIKLIPRIDIQALTQKYGGGVTVQKQTPAPRLISSSELEEFRLIQVRRDRDTGITFEHLD
 SLMLKDG YLYKKVSLDSISSWGVIPTKDELLKFTPVDRKETGDVEWIEIYGEERKKKILPTCREGGKGE GSGGGKGE GSGGG
 KEGSRGGKGE GSSDFKSESSYELYNLVCFSRKDFGLIVGVDDKGDGYKVLKEGIDGPVVVTVGKKEMQNGPFD SKFTALDLN
 KKQISVNDVVKISKGPSEGKQGVVRQVYRGIIFLYDESEENGGYFCCKSQSCEKVKLFTEESNEKTGGFDGTA FEFVSSPK
 SPLSPEKEWQPRERYNSSNQGDIGSTYSIGQKLRI RVGPLKGYLCRVIALRYS DVTVKLDSQH KIFTVKSEHLAEVRDRNTVL
 STSGDAGTGSFQPFMGLGTESS TGDWAIGAGTSSEGGNWNIGGPSTDSNP*

Figure 29. Informations concernant les construits utilisés pour compléter les mutants *spt5l*. A, Représentation schématique et dénomination des construits utilisés. B, Séquences en acides aminés de la forme SPT5-.

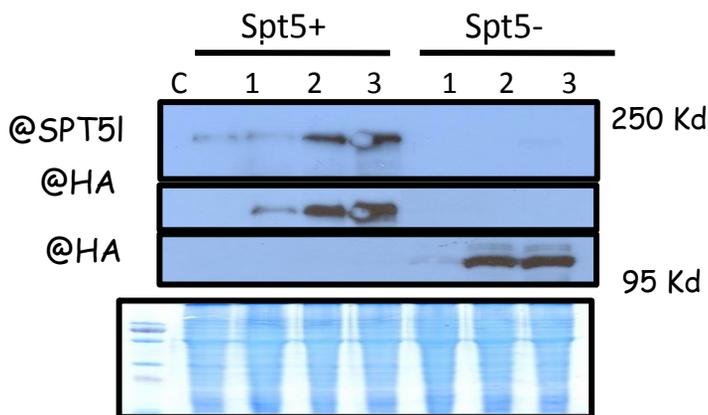
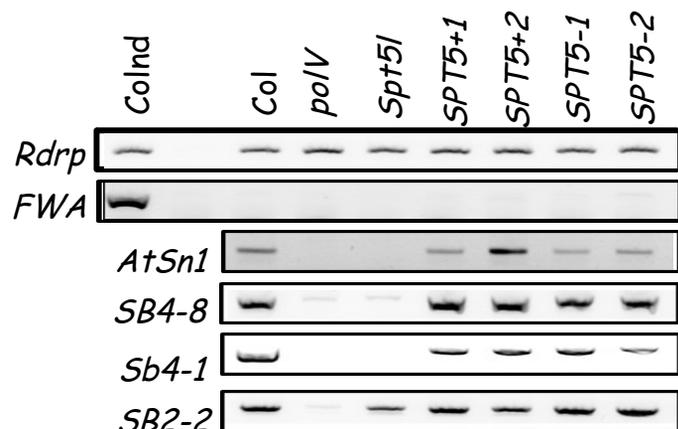


Figure 30. Accumulation des protéines SPT5+ ou SPT5- dans les mutants *spt5like* complémentés. Des extraits protéiques de fleurs de plantes sauvage (C) et de différentes lignées complémentées ont été analysés par western blot. Les détections ont été réalisées avec l'anticorps anti-SPT51 (72) et un anticorps anti-HA. Une migration sur gel d'acrylamide coloré au bleu de coomassie a été réalisée comme contrôle d'équilibre.

Figure 31. Méthylation des cibles du RdDM chez les mutants *spt5like* complémentés. L'ADN génomique de fleurs a été extrait et digéré par l'enzyme *HaeIII* puis utilisé pour amplification avec des amorces spécifiques de loci cible de RdDM tel que *AtSn1*, *SB4-8*, *SB4-1* et *SB2-2*. Des amplifications ont également été réalisées avec les amorces *Rdrp*, permettant un contrôle d'équilibre et les amorces *FWA* permettant de contrôler la digestion de l'ADN.



5-3-1-3 Quelle est la spécificité de la plateforme AGO hook de la protéine SPT5like

- La plateforme AGO hook de la protéine SPT5like est-elle nécessaire pour sa fonction dans la voie du RdDM ?

Pour déterminer si la plateforme AGO de la protéine SPT5like, impliquée dans le recrutement de la protéine AGO4, est indispensable à sa fonction dans la voie du RdDM des mutants *spt5like* ont été complémentés avec des fragments d'ADN permettant la production de protéines SPT5like entière et étiquetée HA (SPT5l+) ou celle d'une protéine tronquée, sans l'extension AGO Hook (SPT5l-) (figure 29). Les deux constructions ont été placées sous le contrôle de la séquence promotrice du gène *SPT5like*. Des plantes transgéniques ont été sélectionnées sur la base de leur niveau d'accumulation des protéines étiquetées, une lignée avec un niveau d'accumulation de la protéine SPT5like équivalent à celui observé chez des plantes sauvages et une avec un niveau supérieur (figure 30). Pour les lignées produisant la forme tronquée de la protéine SPT5like, nous n'avons pu nous baser que sur les signaux obtenus avec l'anticorps HA car l'anticorps anti-SPT5l reconnaît un épitope absent de la forme tronquée. Il est important de noter que les deux lignées de plantes transgéniques ne présentent, comme les mutants de la voie du RdDM, aucun phénotype anormal.

Différentes cibles du RdDM ont été analysées dans ces mutants complémentés (*spt5l+* et *spt5l-*). Nous avons en priorité testé les deux cibles *AtSN1* et *SoloLTR* qui avaient déjà été analysées chez les mutants *spt5like*. Nous avons observé que seule la première de ces cibles était affectée par la mutation *spt5like*. Notre analyse démontre que quelle que soit la plante complémentée, avec ou sans l'extension C-terminale de la protéine SPT5like, et quel que soit le niveau d'accumulation de la protéine étiquetée, les deux cibles du RdDM sont correctement méthylées chez les mutants complémentés (figure 31). Afin de confirmer ces premières observations, d'autres cibles du RdDM telles que *IGN23* et *IGN25* ont également été analysées et montrent qu'aucune d'entre elles n'est affectée dans les plantes complémentées. Ces résultats suggèrent que l'extension C-terminale de la protéine SPT5like, capable de recruter la protéine AGO4, n'est pas indispensable au RdDM. Ce résultat était relativement inattendu et assez déroutant. Quelle pouvait être la fonction de cette plateforme si ce n'est pas de recruter les protéines AGO4 associées aux siARN impliqués dans la méthylation ciblée de l'ADN ? Comment pouvait-elle alors affecter la voie du RdDM ?

Il est important de préciser que dans l'équipe, Sylvie Lahmy travaillait sur une caractérisation plus approfondie du domaine SD ou AGO hook de la protéine NRPE1, la grande sous unité de l'ARN polymérase V. Son approche a consisté à complémenter le mutant *nrpe1* avec une première version NRPE1 qui possède à la place de l'extension C-terminale 8 copies d'un consensus contenant les résidus Tryptophane (POLV+) qui sont indispensables pour l'interaction avec la protéine AGO4. Une seconde version était identique mais les résidus Tryptophane étaient remplacés par des résidus Arginine dans les 8 copies (POLV-). Les mutants *nrpe1* complémentés avec ces deux versions POLV+ ou POLV- ne présentaient, tout comme le mutant *nrpe1*, pas de phénotype apparent anormal. Une analyse de différentes cibles du RdDM dans ces plantes, montraient d'une manière assez surprenante que l'extension C-terminale de la grande sous-unité de la polymérase V n'est pas indispensable pour le RdDM. En effet quelle que soit la lignée complémentée utilisée, toutes les cibles du RdDM présentent le même niveau de méthylation que les plantes sauvages. L'accumulation des siARN ne montrait également pas de variations. Ces observations surprenantes permettaient de conclure que seule la partie N-terminale de la protéine NRPE1 est nécessaire au RdDM.

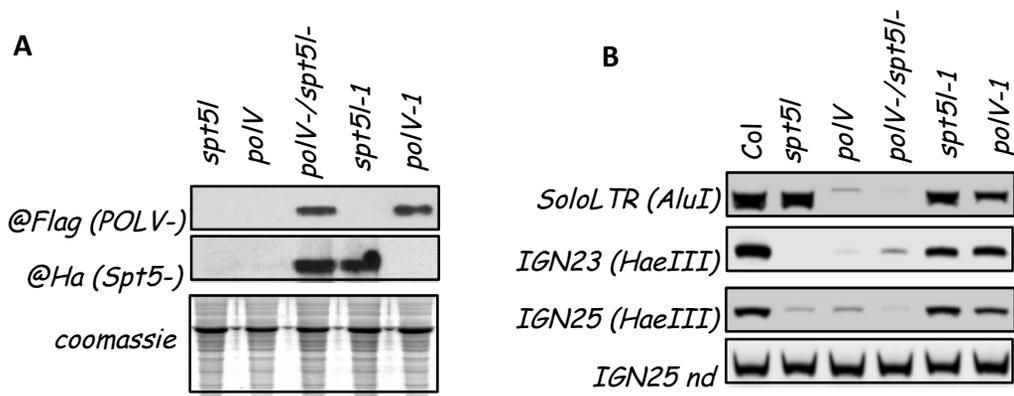


Figure 32. Méthylation de cibles de RdDM dans les plantes POLV-/SPT5I-. **A, Accumulation des protéines étiquetées dans la lignée issue du croisement des plantes complémentées POL- et SPT5I-.** Les détections sont réalisées à partir d'extraits protéiques de fleurs avec l'anticorps anti-Flag (détection POLV-) et avec l'anticorps anti-Ha (détection SPT5I-). Une migration sur gel d'acrylamide colorée au bleu de coomassie a été réalisée comme contrôle. **B, Méthylation de cibles du RdDM.** L'ADN génomique de fleurs de plantes sauvages (Col), de mutant *spt5l*, de mutant *nrpe1* (*polV*), de mutants complémentés *spt5l-1* et *PolV-1* et de plantes *PolV-/SPT5I-* ont été extraits et digérés par les enzymes AluI ou HaeIII puis utilisés pour amplification avec des amorces spécifiques des loci *SoloLTR*, *IGN23* et *IGN25*. Une amplification a également été réalisée avec de l'ADN non digéré (nd) permettant un contrôle d'équilibre

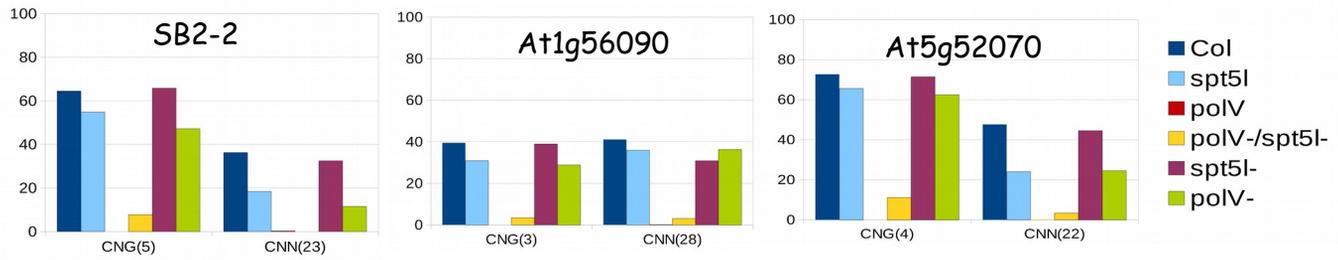
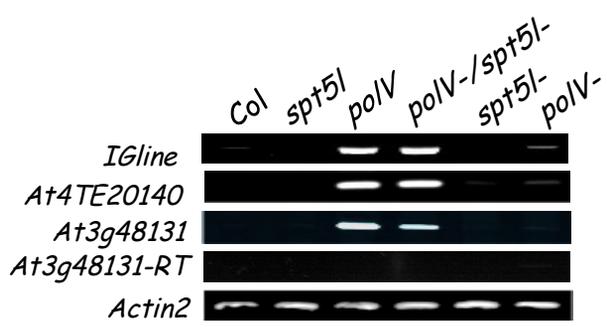


Figure 33. Effets de l'absence des plateformes AGO hook des protéines NRPE1 et SPT5like sur la méthylation de cibles du RdDM. Différentes cibles ont été analysées par séquençage après traitement au bisulfite. Le nombre d'occurrence de chaque contexte (CNG) ou (CNN) est indiqué pour chacune d'entre elle. Les valeurs indiquent la nombre moyen de cytosines méthylées. Résultats obtenus par Sylvie Lahmy.

Figure 34. Effets de l'absence des plateformes des protéines NRPE1 et SPT5like sur la repression de certains loci. Les ARN totaux de fleurs de plantes sauvages (Col), de mutants *spt5l* et *polV*, de lignées complémentées *spt5l-1* et *polV-1* et de plantes *PolV-/SPT5I-* ont été analysés par RT-PCR avec des amorces spécifiques des loci *IGline*, *At4TE20140* et *At3g48131*. Une PCR sans reverse transcriptase et été réalisée comme contrôle d'absence d'ADN génomique et une RT-PCR avec les amorces du gène constitutif *Actin2* confirme l'équilibre des échantillons.



Nous savions par conséquent que la grande sous-unité de la polymérase V (NRPE1) est capable de recruter par sa plateforme C-terminale des protéines AGO4 qui sont, a priori, indispensable pour la méthylation de nombreuses cibles du RdDM. Pourtant l'absence de la plateforme de NRPE1 n'affecte pas le RdDM. Nous savions, par ailleurs, que la protéine SPT5like possède elle aussi une plateforme capable d'interagir avec la protéine AGO4, et que l'absence de cette plateforme n'affecte pas les cibles du RdDM. Enfin, nous savions également que l'action des protéines NRPE1 et SPT5like se situait au même niveau dans la voie du RdDM. Il était par conséquent possible d'envisager que les protéines NRPE1 et SPT5like qui appartiennent à un même complexe soient redondantes dans leur capacité à recruter des protéines AGO. En effet, on pouvait penser que si une de ces deux protéines possède son extension AGO hook, cela suffise à recruter suffisamment de protéines AGO et de petits ARN pour permettre la méthylation ciblée de l'ADN par la voie du RdDM. Afin de tester cette hypothèse les mutants *nrpe1* et *spt5l* complémentés ont été croisés et les cibles du RdDM analysées dans les descendance.

- Les plateformes de SPT5like et NRPE1 sont elles redondantes ?

Après croisement des différentes plantes complémentées, des lignées homozygotes pour les quatre ADN-T, celui responsable de la mutation *nrpe1*, celui responsable de la mutation *spt5like*, celui correspondant à la version POLV- et enfin celui correspondant à la version SPT5like tronquée (SPT5l-) ont été sélectionnées. Les plantes homozygotes pour ces quatre ADN-T seront par la suite nommées POLV-/SPT5l-. La présence des différentes protéines étiquetées a été contrôlée (figure 32A) et différentes cibles du RdDM ont été testées. Les premières analyses montrent que les cibles *SoloLTR*, *IGN23* et *IGN25* sont hypométhylées dans les plantes POLV-/SPT5like- à un niveau équivalent à celui du mutant *nrpe1(polV)* (figure 32B). Le mutant *spt5like* lui n'est affecté que pour les cibles *IGN23* et *IGN25* (figure 32B). Par conséquent seules les plantes dépourvues des plateformes AGO des protéines POLV- ou SPT5l- simultanément se caractérisent par une méthylation défectueuse des cibles du RdDM. La même conclusion peut être donnée aux résultats obtenus après traitement bisulfite sur d'autres cibles du RdDM (figure 33). On observe alors que les deux contextes CNG et CNN sont hypométhylés chez le mutant *nrpe1* ou les plantes POLV-/SPT5l-. Une analyse bisulfite à l'échelle du génome entier, réalisée en collaboration avec Steve Jacobsen (UCLA, USA) a confirmé nos observations. En adéquation avec ces observations on observe la dé-répression transcriptionnelle des cibles *IGline*, *At4TE20140* et *At3g48131* dans les plantes POLV-/SPT5l- (figure 34).

Ces résultats valident notre hypothèse et permettent de conclure qu'une seule plateforme AGO hook, celle de NRPE1 ou de SPT5like, est nécessaire et suffisante pour permettre de cibler la méthylation de l'ADN par la voie RdDM. Les plateformes AGO hook de POLV et SPT5like sont donc redondantes. Ces deux protéines présentent toutefois des fonctions très distinctes, NRPE1 permettant la synthèse d'un brin d'ARN et SPT5like étant connue pour son action de facteur d'élongation de la transcription. C'est probablement leurs fonctions spécifiques qui font que les mutants *spt5like* sont moins affectés que les mutants *nrpe1*, en particulier pour la cible *SoloLTR*.

Afin de confirmer notre hypothèse nous avons testé si la présence de la version NRPE1WG8 serait suffisante pour rétablir le RdDM dans des plantes SPT5l-. Les croisements nécessaires à l'obtention de ce matériel ont été réalisés et des plantes homozygotes pour les deux mutations *nrpe1* et *spt5like*, possédant la version NRPE1WG8 (POL+) et la protéine SPT5like tronquée (SPT5l-) ont été sélectionnées. Ces plantes seront par la suite nommées POL+/SPT5l-. Une première analyse a permis de confirmer la présence des protéines étiquetées dans les plantes issues du croisement

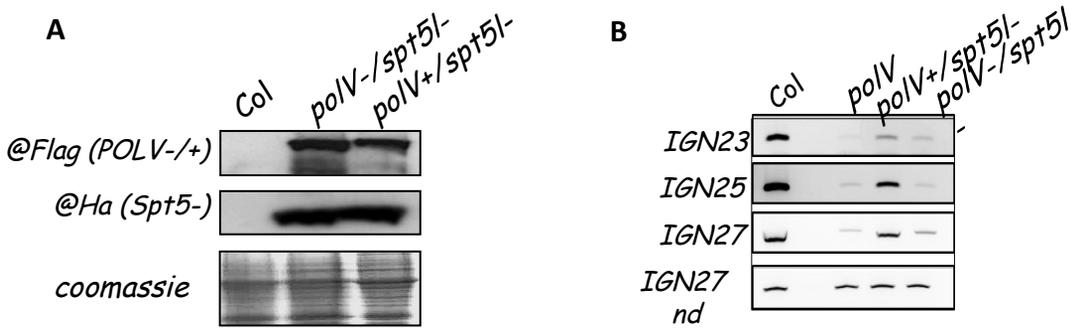


Figure 35. Effet de la présence de 8 répétitions de la protéine NRPE1(POL+) sur la méthylation de cibles du RdDM. **A, Accumulation des protéines étiquetées dans les plantes POIV-/SPT51- ou POLV+/SPT51-.** Des extraits protéiques de fleurs ont été analysés par western blot avec l'anticorps anti-Flag (détection des protéines POLV- et POLV+) et l'anticorps anti- Ha (détection SPT51-). Une migration sur gel d'acrylamide colorée au bleu de coomassie a été réalisée comme contrôle. **B, Analyse de la méthylation de cibles du RdDM.** L'ADN génomique de fleurs de plantes sauvages (Col), de mutant *nrpe1* (*polV*), et de plantes POLV-/SPT51- et POLV+/SPT51- ont digérés par l'enzyme HaeIII puis utilisés pour amplification avec des amorces spécifiques des loci *IGN23*, *IGN25* et *IGN27*. Des PCR ont été réalisées avec de l'ADN non digéré (nd) pour contrôler l'équilibrage des échantillons.

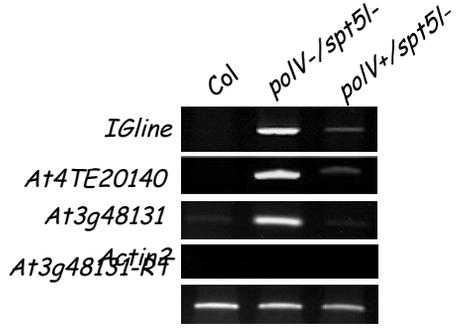


Figure 36. Effet de la présence de 8 répétitions de la protéine NRPE1 sur la répression transcriptionnelle de cibles du RdDM. Les ARN totaux de fleurs de plantes sauvages (Col) et de plantes POL-/SPT51- et POL+/SPT51- ont été analysés par RT-PCR avec des amorces spécifiques de loci cible de répression transcriptionnelle par RdDM (*IGline*, *At4TE20140* et *At3g48131*). Une PCR sans reverse transcriptase et été réalisée comme contrôle d'absence d'ADN génomique et une RT-PCR avec les amorces du gène constitutif *actin2* indique un bon équilibrage des échantillons.

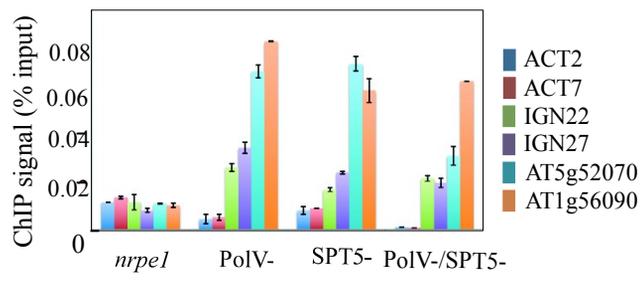


Figure 37. Liaison de la Polymérase V sur l'ADN. Des immuno-précipitations de chromatine réalisées à partir de mutant *nrpe1*, de plantes POLV-, SPT51- et POLV-/SPT51- avec l'anticorps anti NRPE1. La présence de fragments correspondants à différents fragments d'ADN a été testé par qPCR. Résultat obtenu par Dominique Pontier.

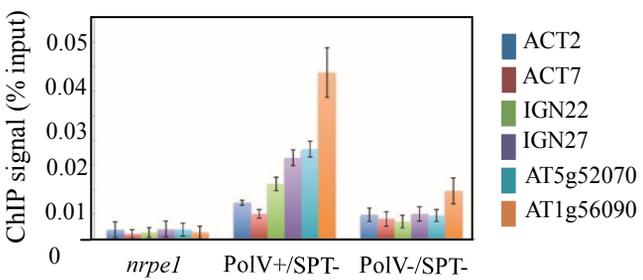
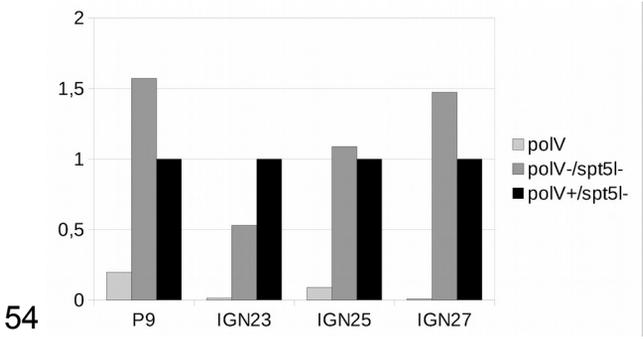


Figure 38. Analyse de la liaison des protéines AGO4 sur l'ADN. Des immunoprécipitations de chromatine ont été réalisées à partir de mutant *nrpe1*, de plantes POLV+/SPT51- et POLV+/SPT51- avec l'anticorps anti-AGO4. La présence de fragments correspondants à différents fragments d'ADN a été testé par qPCR. Résultat obtenu par Dominique Pontier.

Figure 39. Accumulation des lncARN. Quantification par qRT-PCR des lncARN produits dans les plantes mutant *polV*, et les plantes POLV-/SPT51- et POLV+/SPT51-.



(figure 35A). L'analyse de différentes cibles du RdDM a permis de mettre en évidence la méthylation de différentes cibles comparable à celle observée chez les plantes sauvages (figure 35B) ainsi qu'une répression des cibles similaire à celle observée chez des plantes sauvages (figure 36). Ces observations indiquent par conséquent que la plateforme synthétique suffit à restaurer l'activité RdDM. Ce résultat a été confirmé et affiné, par une analyse après traitement au bisulfite, et les résultats indiquent que, pour les plantes POL+/SPT51-, la méthylation des cibles testées est restaurée avec un niveau de méthylation légèrement inférieur à celui observé chez les plantes sauvages, mais supérieur à celui observé dans les plantes POL-/SPT51-. Cette analyse met également en évidence que les Cytosine dans les plantes POL+/SPT51- dans les contextes CNG ou CNN sont méthylées au même niveau que dans les plantes sauvages. Même si la méthylation n'est pas totale elle est largement supérieure à celle observée en absence des plateformes AGO hook des deux protéines NRPE1 et SPT5like.

- L'absence de plateforme a-t-elle un effet sur l'interaction de l'ARN POLV avec l'ADN ?

L'absence de méthylation dans les plantes POLV-/SPT51- pourrait s'expliquer par une incapacité de la POLV- de se fixer sur l'ADN. Des expériences par Chromatin immuno précipitation (CHIP) réalisées par Dominique Pontier à partir de plantes complémentées ont montré que dans les contextes POLV-, SPT51- et POL-/SPT51- la protéine NRPE1 modifiée ou non se fixe sur les cibles testées (IGN22, IGN27, At5g52070 et At1g56090) du RdDM (figure 37). On peut par conséquent conclure que l'absence de plateforme AGO hook n'a pas d'impact sur la fixation de la protéine NRPE1 sur l'ADN.

- L'absence de plateforme a-t-elle un effet sur le recrutement de protéines AGO4 ?

Après s'être assurée que l'ensemble des mutants complémentés présentait une quantité équivalente de protéine AGO4, Dominique Pontier a testé la présence de protéines AGO4 par CHIP, sur les mêmes cibles que celles précédemment analysées. Les résultats observés démontrent une détection de protéine AGO4 beaucoup moins importante pour les plantes POLV-/SPT51- ou *nrpe1* que pour les plantes POLV+/SPT51- (figure 39). Ces observations ont permis de conclure que le défaut de méthylation observé dans les plantes POLV-/SPT51- est corrélé à un défaut de recrutement de protéines AGO4.

- L'absence de plateforme a-t-elle un effet sur la production de transcrits POLV ?

Un autre point important à analyser était celui des transcrits non codants POLV (lncRNA). Les modèles suggèrent que ce sont ces transcrits qui permettent le recrutement des associations AGO4/siARN et donc la méthylation ciblée de l'ADN. De manière intéressante nos analyses montrent que la quantité de lncRNA polV est comparable chez les plantes POL+/SPT51-, pour lesquelles la méthylation des cibles du RdDM est normale, et les plantes POLV-/SPT51- qui présentent un défaut de méthylation des mêmes loci (figure 39). Par contre chez le mutant *nrpe1* aucun lncARN ou IGN n'est détecté. Cette observation assez inattendue indique que la seule production des lncARN ou IGN n'est pas suffisante pour recruter les protéines AGO4. Il peut par conséquent être conclu que c'est la présence de plateformes AGO qui semble seule être nécessaire au recrutement des complexes AGO4/siARN. Pour confirmer cette observation Dominique Pontier a montré que la présence ou non d'ARN (traitement RNase avant CHIP) n'a aucun effet sur le recrutement des protéines AGO4 sur les cibles du RdDM. Il peut par conséquent être conclu que les lncARN POLV ne sont pas responsables du recrutement des protéines AGO4 et donc de la méthylation de l'ADN.

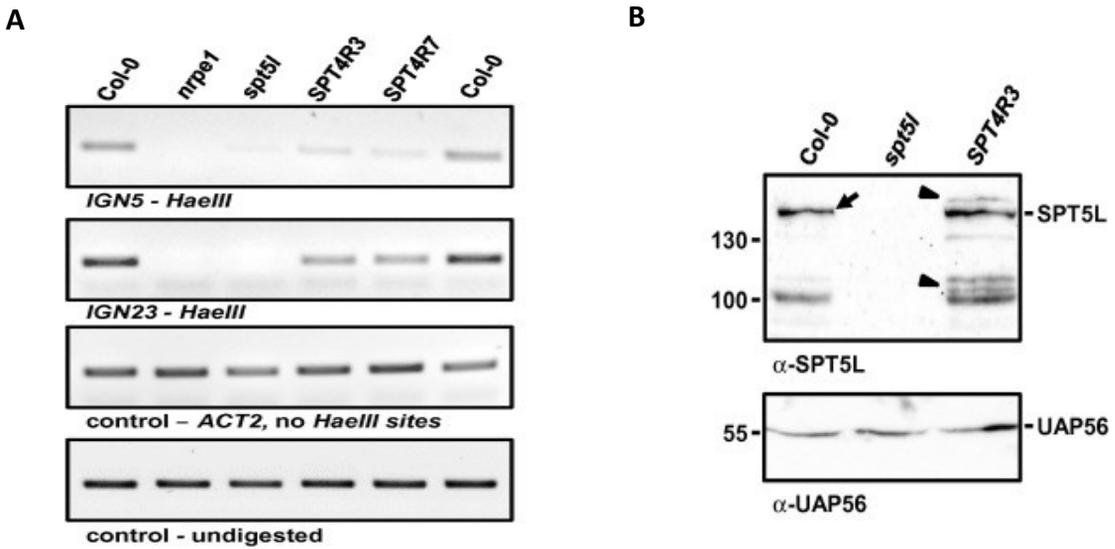


Figure 40. Caractérisation de lignées RNAi SPT4. **A, Méthylation de différentes cibles du RdDM.** Les ADN de plantes sauvages (Col-O), de mutant *nrpe1* (*polV*), de mutant *spt5like* et de lignées RNAi SPT4 (SPT4R3 et SPT4R7) ont été extraits et digérés par l'enzyme *HaeIII* puis utilisés pour amplification avec des amorces spécifiques de loci cible de RdDM telles que *IGN5* et *IGN23*. Des amplification avec des amorces *ACTINE2* (*ACT2*) et de l'ADN non digéré sont utilisées comme contrôle. **B, Analyse de l'accumulation de protéines SPT5like.** La teneur en protéines SPT5like est estimée par western blot (anticorps 72) à partir d'extrait protéiques de plantes sauvages (Col-O), de mutant *spt5like* et de plantes RNAiSPT4.

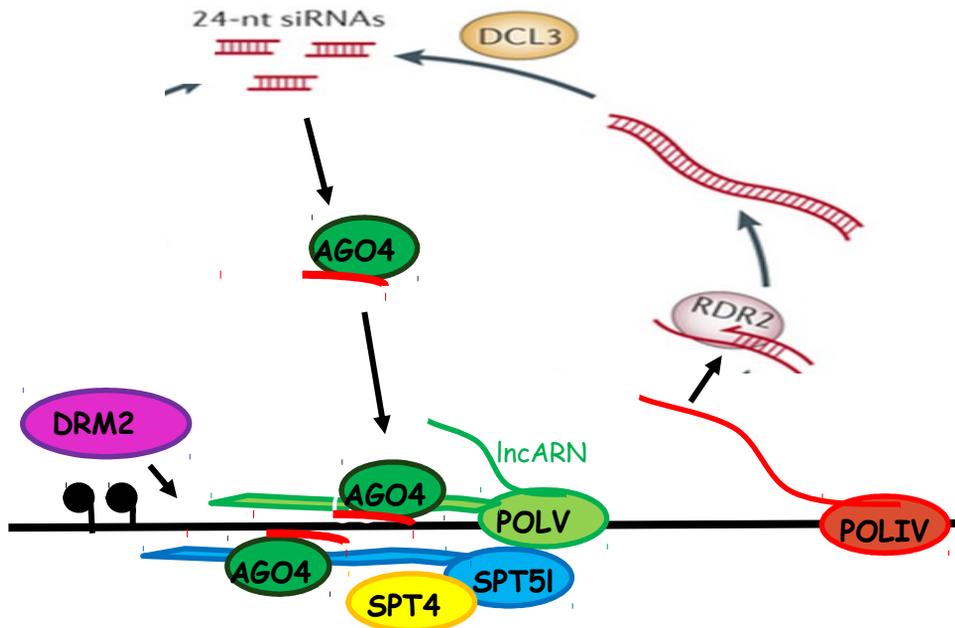


Figure 41. Modèle revisité du RdDM. Ce sont les plateformes à AGO hook de NRPE1 (*POLV*) et SPT5like qui recrutent les protéines AGO4 liées aux siARN qui vont permettre le recrutement de DRM2 qui va pouvoir méthyler de manière séquence spécifique l'ADN.

5-3-1-4/ La protéine SPT4 est associée à la protéine SPT5like et agit sur le RdDM.

Chez de nombreux organismes il a été établi que les protéines SPT5 interagissent avec les protéines SPT4 (Glover-Cutter et al., 2008; Wenzel et al., 2008; Martinez-Rucobo et al., 2011; Hartzog and Fu, 2013). L'équipe de Klaus Grasser (Allemagne) a montré qu'il en est de même chez *Arabidopsis* (Durr et al., 2014). SPT4 interagit avec les protéines SPT5 classiques mais aussi avec Spt5like. De par la fonction de SPT5like il était alors tentant de tester si la protéine SPT4 avait un effet sur la voie de répression RdDM. Cette analyse a été réalisée en collaboration avec Klaus Grasser (Allemagne) et il s'avère que des lignées RNAi *spt4* qui sont affectées dans la division cellulaire, la croissance et le développement se caractérisent aussi par une hypométhylation des cibles *IGN5* et *IGN23* du RdDM (figure 40A). Parce que il avait été montré que chez de nombreux organismes le taux d'accumulation des protéines SPT4 agit sur celui des protéines SPT5 la teneur en protéine SPT5like a été contrôlée dans les lignées RNAi *spt4* (figure 40B). Les résultats observés indiquent une quantité de protéine SPT5like similaire dans les plantes RNAi et dans des plantes sauvages. A partir de ces observations nous avons proposé que l'effet observé sur les cibles du RdDM dans les lignées RNAi devait résulter du fait que la protéine SPT4 est associée à SPT5like et agit sur le RdDM. Ce travail a fait l'objet d'une publication dans FEBS letter, The zinc-finger protein SPT4 interacts with SPT5L/KTF1 and modulates transcriptional silencing in *Arabidopsis* (Kollen et al., 2015).

Conclusion

L'ensemble du travail réalisé a permis de compléter et d'affiner de manière importante le modèle de fonctionnement de la voie de répression des transposons ou séquences répétées du génome par le RdDM. Il a permis tout d'abord d'établir que la protéine SPT5like est associée au complexe de la polymérase V et a un effet partiel sur le RdDM en participant au recrutement des protéines AGO4 associées à des siARN. Nos analyses démontrent par ailleurs que les plateformes de NRPE1 et SPT5like sont redondantes pour cette fonction de recrutement qui est nécessaire pour la méthylation ciblée de l'ADN. Ces observations remettent en question le modèle suggérant que ce sont les lncARN POLV qui recrutent les protéines AGO4. En effet même en présence de ces transcrits on observe un défaut de méthylation de l'ADN. C'est uniquement en absence des plateformes AGO hook que les protéines AGO4 ne sont pas recrutées et qu'un défaut de méthylation est observé. L'ensemble des analyses réalisées permet par conséquent de proposer un modèle plus précis et plus détaillé des dernières étapes du RdDM en prenant en compte également la participation de la protéine SPT4 (figure 41). Ce travail est actuellement poursuivi dans le but de tenter de déterminer si après avoir été recruté les siARN et AGO4 se lient aux transcrits ou à l'ADN. Nos analyses sur NRPE1 et SPT5like ont également permis de publier un article portant sur l'évolution des variants de sous unités de complexes d'ARN polymérases (Lahmy et al., 2010)

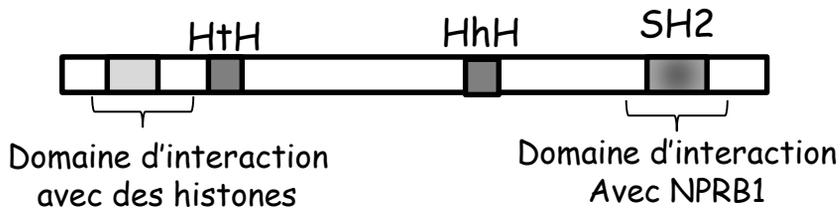


Figure 42. Représentation schématique des protéines SPT6. Les domaines conservés et ceux impliqués dans des interactions avec les histones ou la polymérase II sont indiqués. HtH, helix turn helix domain; HhH, helix hairpin helix domain. SH2 domaine impliqué dans l'interaction avec NPRB1 (polymérase II). Représentation inspirée de Dronamraju et Strahl, 2014.

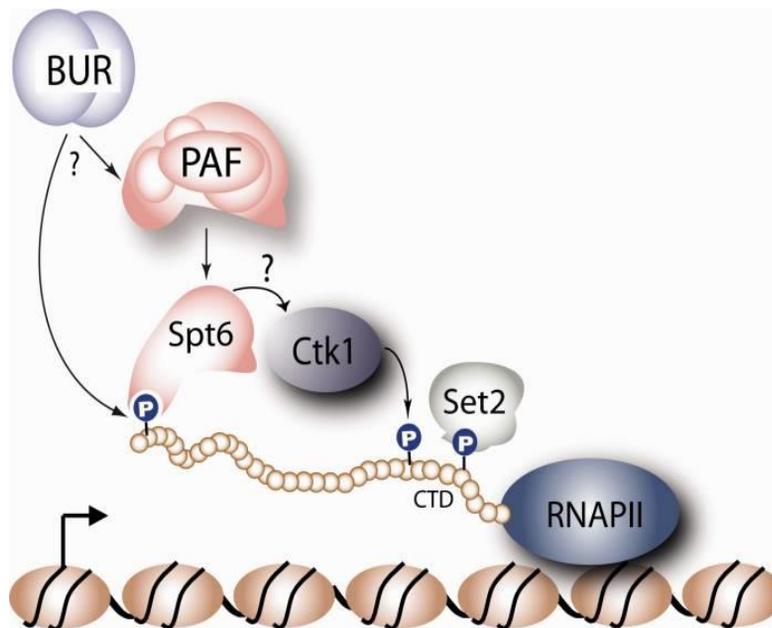


Figure 43. Représentation graphique du mode d'action des protéines SPT6. Issu de Dronamraju et Strahl, 2014.

5-3-2/ Caractérisation du gène *At1g65440*

5-3-2-1/ Analyse de la séquence.

- Similarité de la séquence du gène *At1g65440* avec celles codant des protéines SPT6.

Le gène *At1g65440* code une protéine similaire aux protéines SPT6. De manière intéressante, la première protéine SPT6 a été identifiée à partir du crible de levure ayant permis l'identification du gène codant la protéine SPT5. Tout comme les protéines SPT5, les protéines SPT6 sont associées au complexe de polymérase II et sont elles aussi décrites comme facteur d'élongation de la transcription. Ces protéines auraient également une activité chaperone d'histones (Mayer et al., 2010; Kwak and Lis, 2013).

Les protéines SPT6 sont composées de différents domaines conservés représentés dans la figure 42. Le domaine SH2 est impliqué dans l'interaction avec la protéine NRPB1 (grande sous-unité de l'ARN polymérase II) (Dengl et al., 2009; Dronamraju and Strahl, 2014). Ce domaine se lie, au moins chez la levure, aux répétitions YSPTSPS de la CTD de NRPB1, notamment lorsque celles-ci sont phosphorylées au niveau de la sérine 2 (S2) par action de la kinase BUR1 (Dengl et al., 2009; Burugula et al., 2014). Il a été montré, par ailleurs, que différents mutants *spt6* sont affectés dans le recrutement du complexe PAF, contiennent moins de protéine SET2 (histone méthyltransferase) et de kinase CTK1 qui phosphoryle aussi NRPB1 (DeGennaro et al., 2013; Dronamraju and Strahl, 2014). Ces observations pourraient s'expliquer par le fait que SPT6 interagit physiquement avec CTK1 et le complexe PAF (Dronamraju and Strahl, 2014). C'est principalement à partir de ces indications qu'il a été proposé que la kinase BUR, phosphoryle à la fois la CTD de NRPB1 et le complexe PAF, que ces phosphorylations modulent le recrutement de SPT6 qui pourrait ensuite agir sur CTK1 et le recrutement de SET2 (figure 43) (Dronamraju and Strahl, 2014). Il a, par ailleurs, été démontré que le domaine N terminal acide des protéines SPT6, celui qui serait impliqué dans la fonction chaperone d'histones, interagit avec des protéines IWS qui ont été décrites comme capables de recruter SET2, responsable de la tri-méthylation H3K36 (Yoh et al., 2008).

Différents travaux suggèrent que les protéines SPT6 ont un rôle dans le recrutement de protéines impliquées dans la maturation des messagers. Un exemple concret est celui résultant de l'interaction de protéines SPT6 avec des protéines ALY qui sont impliquées dans le transport et l'export des ARN messagers hors du noyau (Yoh et al., 2008). Un autre exemple est celui qui indique que des mutants *spt6* sont altérés dans les événements d'épissages (DeGennaro et al., 2013).

La fonction de chaperone d'histones a principalement été mise en évidence à partir de mutants *spt6* chez lesquels le positionnement des nucléosomes est altéré ce qui entraîne des transcriptions cryptiques intergéniques, des sur-expressions de transcrits antisens et pourrait en partie expliquer la fonction des protéines SPT6 dans le contrôle de la vitesse de la transcription des gènes (DeGennaro et al., 2013). Les mutants *spt6* se caractérisent aussi par un recrutement moins important, au niveau des gènes, à la fois du complexe COMPASS qui est impliqué dans la méthylation H3K4, et de SET2, qui est impliqué dans la méthylation H3K36 (DeGennaro et al., 2013). En conséquence les quantités d'histones H3K4me3 et H3K36me3 sont diminuées dans ces mutants (DeGennaro et al., 2013). D'autres études montrent que SPT6 a aussi un effet sur la méthylation H3K9, sur la déméthylation H3K27 et est également capable d'interagir avec les histones H2A et H2B (Kato et al., 2013; Wang et al., 2013a; Mc Cullough et al., 2015). L'ensemble

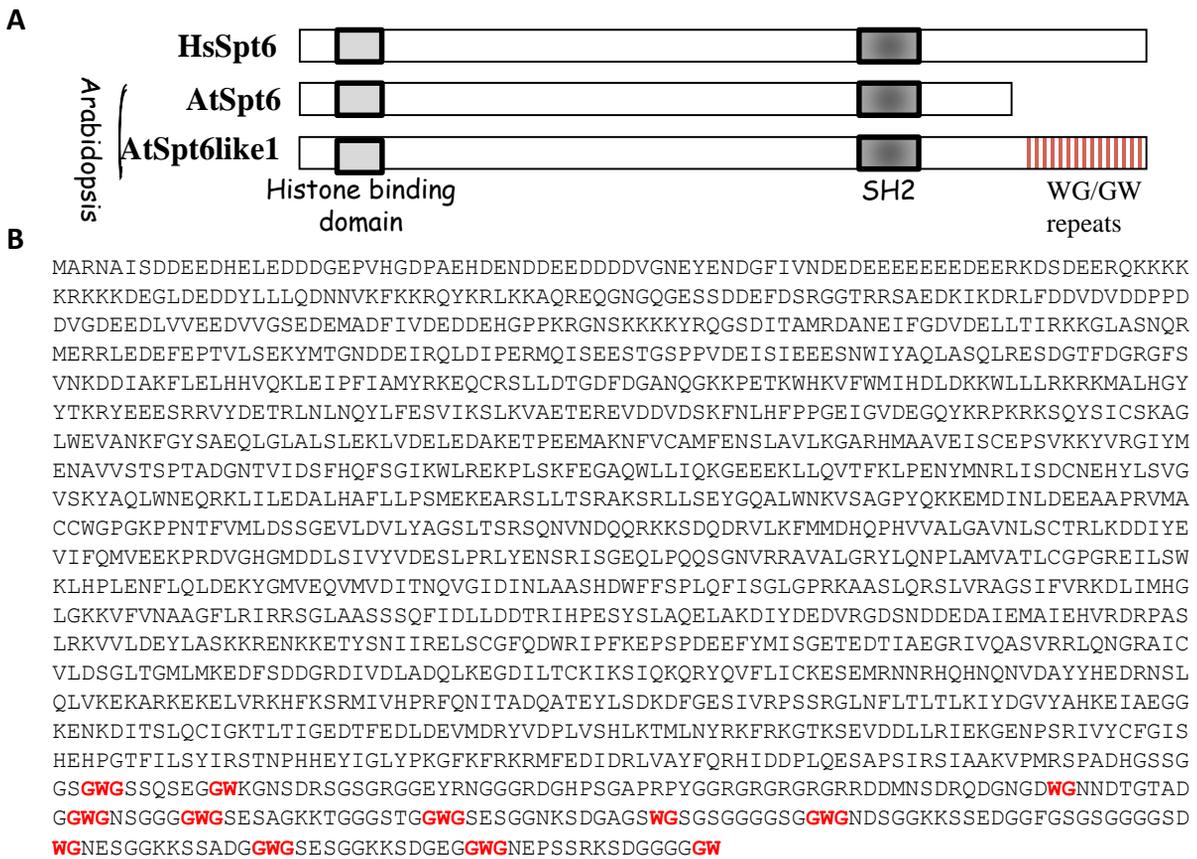


Figure 44. Caractéristiques des protéines SPT6 d'Arabidopsis thaliana. A, Représentation schématique des protéines SPT6. Le gène *At1g63210* code une protéine SPT6 classique et le gène *At1g65440* code une protéine qui possède dans sa partie C terminale une plateforme AGO hook. Les extrémité N terminales de ces protéines sont très conservées. B, Séquence en Acides aminés de la protéine SPT6like (*At1g65440*). Les répétitions de résidus Tryptophane et Glycine sont indiqués.

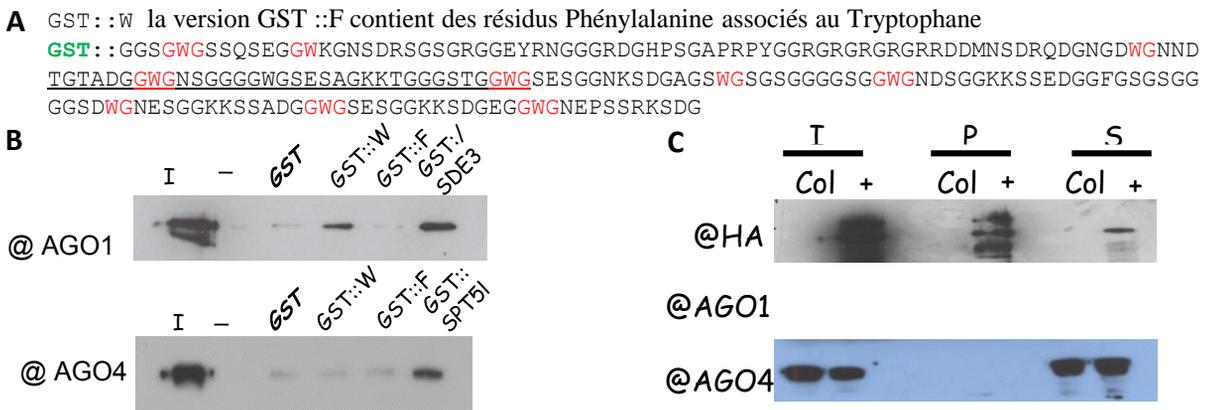


Figure 45. Interaction de la plateforme SPT6like avec des protéines AGO. A, Séquence des plateformes utilisées. B, Test d'interaction des plateformes W et F avec des protéines AGO. GST pull down réalisé avec GST::W ou GST::F incubées avec des extraits de plantes transgéniques (AGO4::Myc ou AGO1::Flag). Les révélations ont été réalisées avec des anticorps anti-Myc (détection AGO4) et anti-Flag (détection AGO1). C, Immunoprécipitation de protéines SPT6+ (+). Des extraits protéiques de mutants complémentés avec une forme étiquetée de la protéine SPT6like ont été utilisés pour tester après immuno-précipitation la présence de protéines AGO. Des anticorps anti AGO1 et AGO4 ont indépendamment été utilisés.

de ces observations indique que les protéines SPT6 sont probablement impliquées dans le maintien ou la mise en place de certaines marques épigénétiques relatives aux histones.

- Diversité des gènes *SPT6* chez *Arabidopsis* et les autres organismes.

Dans le génome d'*Arabidopsis* deux gènes *At1g63210* et *At1g65440*, que nous avons identifié par psi-blast, codent des protéines similaires aux protéines SPT6 (figure 44). Le premier de ces gènes n'est actif que dans les fleurs, alors que le second présente une expression à la fois plus forte et constitutive. Ces données d'expression ont été confirmées par RT-PCR. Il est important de noter que seul le second de ces gènes code une protéine qui possède, dans son extrémité C terminale, un domaine pouvant constituer une plateforme à protéines AGO (figure 44B). La présence de ce domaine, qui contient douze motifs Tryptophane/Glycine, nous a conduits à nommer ce gène *SPT6like*. La recherche de séquences similaires dans d'autres génomes montre que, comme cela avait été le cas pour *SPT5like*, seuls les végétaux (riz, peuplier, vigne, tabac et sélaginelle) possèdent des gènes similaires au gène *At1g65440*. Chacun de ces gènes code des protéines possédant de 7 à 9 associations de Tryptophane/Glycine. La conservation des protéines SPT6like avec une potentielle plateforme AGO hook chez les végétaux indique un rôle particulier probablement important.

5-3-2-2 Caractérisation de la protéine SPT6like.

- La protéine SPT6like est-elle capable d'interagir avec des protéines AGO ?

Dans un premier temps il était nécessaire de tester si la protéine SPT6like est capable de se lier à des protéines AGO. Pour cela, nous avons réalisé des fusions GST avec un fragment correspondant à une partie de l'extension C-terminale de la protéine SPT6like, version GST::W dont les caractéristiques sont indiquées dans la figure 45A. Une version comparable, GST::F, dans laquelle les Tryptophane ont été remplacés par des Phénylalanine, a également été utilisée (figure 45A). Les protéines recombinantes GST, GST::W et GST::F ont permis de mettre en évidence une interaction de la version GST::W avec la protéine AGO1 (figure 45B). Pour les échantillons contenant la GST ou GST::F, aucune interaction avec AGO1 n'a été observée. Une fusion GST avec la plateforme de la protéine SDE3, fixant AGO1, a été utilisée comme contrôle positif (figure 45B) (Garcia et al., 2012). La protéine AGO4, elle, ne semble pas se fixer aux fusions GST utilisées (figure 45B). Dans les mêmes conditions un échantillon contenant une fusion GST avec l'extension C-terminale de la protéine SPT5like, démontre une rétention de protéines AGO4 (figure 45B). De manière à confirmer l'interaction observée *in vivo* une construction permettant la production de protéine SPT6like fusionnée à une étiquette FLAG/HA a été introduite dans une lignée d'insertion d'ADN-T dans le gène *SPT6like*. La construction utilisée a été placée sous le contrôle d'un fragment de 1456 nucléotides correspondant à la région présente en amont du site d'initiation de la transcription du gène *SPT6like*. Cette version sera nommée SPT6+. L'analyse réalisée a permis de confirmer par immuno-précipitation l'interaction de la protéine SPT6like avec la protéine AGO1 (figure 45C). Ces premières observations permettent de conclure que la protéine SPT6like interagit avec la protéine AGO1 et pourrait par conséquent être impliquée dans une voie AGO1 dépendante.

Les voies connues impliquant AGO1 sont principalement celles relatives aux microARNs, impliquées dans le PTGS, ainsi que la voie particulière des tasiARNs. Il est important de noter que des mutants affectés dans ces voies présentent en général des phénotypes morphologiques assez

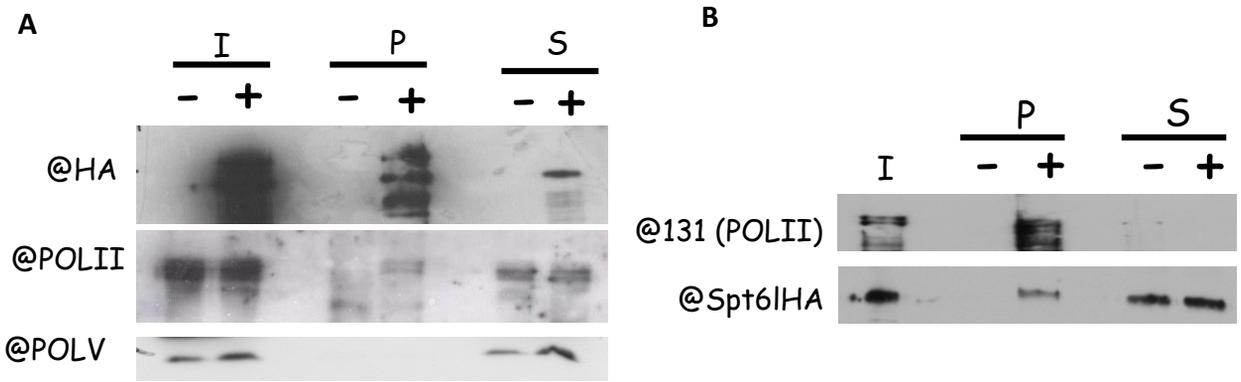


Figure 46. Identification des ARN polymérase liant SPT6like. **A, Immuno-précipitation de la protéine SPT6l+.** Des extraits protéiques de plantes SPT6+ (+) ou sauvage (-) ont été utilisés pour tester après immuno-précipitation la présence des grosses sous unités des complexes de ARN polymérase II et V. I, input, P, précipité et S surnageant. L'anticorps 8WG16 (POLII) et D5 (POLV) ont respectivement été utilisés **B, Immuno-précipitation de la grosse sous unité du complexe Polymérase II.** Un anticorps permettant d'immuno-précipiter NRPB1 a été utilisé à partir d'extraits protéiques de plantes SPT6l+ et les différents extraits ont été analysés avec un anticorps anti-HA afin de tester la présence de protéine SPT6+.

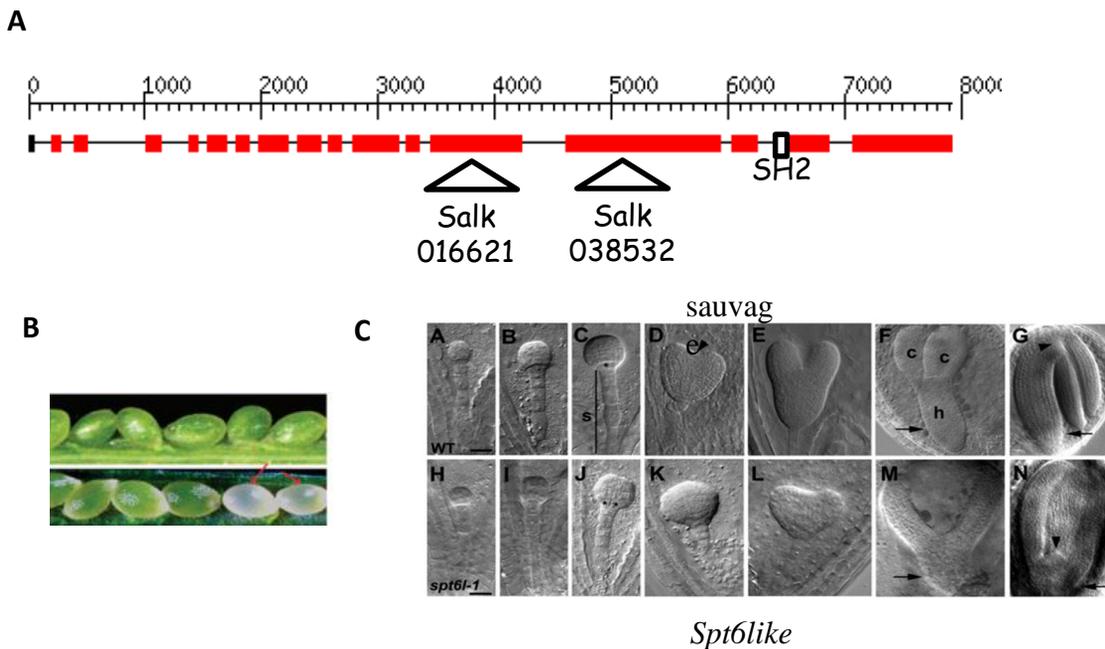


Figure 47. Caractérisation de mutants d'insertion dans le gène SPT6like. **A, Représentation schématique des insertions d'ADN dans les différents mutants spt6like.** **B, Comparaison de graines sauvages et de mutants spt6like.** Les graines blanches correspondent au mutant et vont rapidement stopper leur développement. **C, Comparaison du développement d'embryons sauvages et homozygotes pour la mutation spt6like.** Issu de Gu et al, 2011).

caractéristiques (Mallory et al., 2008; Vaucheret, 2008; Matsui et al., 2014) La protéine AGO1 est aussi impliquée dans la réponse à certains stress environnementaux. Les graines du mutant *ago1-27*, par exemple, sont plus sensibles que des graines sauvages à la présence d'ABA ou à celle de sels (NaCl) (Li et al., 2012). Une réponse similaire est observée dans le cas d'une sur-expression du *microARN 168* qui cible les transcrits AGO1 (Li et al., 2012). La protéine AGO1 constitue également une cible pour certains pathogènes comme le virus turnip crinkle, le polérovirus et le potatovirus. Dans chacun des cas étudiés, des protéines virales, suppresseurs du RNA silencing, interagissent avec AGO1 modifiant ainsi le pool de protéine AGO1 libre ou entraînant sa dégradation ce qui a pour effet de limiter la réponse de l'hôte (Pumplin and Voinnet, 2013).

Ces informations démontrent que différentes voies impliquent AGO1 et que la caractérisation fonctionnelle de la protéine SPT6like devrait, par conséquent, permettre de compléter nos connaissances sur certains de ces mécanismes ou potentiellement en identifier de nouveaux.

- Avec quelles ARN Polymérases la protéine SPT6like est-elle capable de se lier ?

Il est établi que chez de nombreux organismes, les protéines SPT6 s'associent principalement avec le complexe ARN polymérase II. Il était important de déterminer s'il en est de même pour la protéine SPT6like. Les plantes SPT6+ ont été utilisées pour réaliser des co-immunoprécipitations et les résultats démontrent une interaction de la protéine SPT6like avec la polymérase II (figure 46A). Ce résultat a été confirmé par co-immunoprécipitation inverse (figure 46B). En revanche, aucune interaction n'a pu être mise en évidence avec la polymérase V. Ces résultats indiquent que la protéine SPT6like est présente dans le complexe de la polymérase II et absente de celui de la polymérase V.

- Quels sont les effets résultant d'une absence de protéine SPT6like ?

Deux lignées d'insertion d'ADN-T dans le gène *SPT6like* (Salk016621 et Salk038532), dont les sites d'insertion sont indiqués dans la figure 47A, étaient disponibles mais aucune ne permettait d'obtenir des descendances homozygotes pour l'insertion indiquant un problème de létalité. La caractérisation complète de l'un de ces mutants, réalisée par Cui et son équipe, montre que les premiers défauts apparaissent au stade 16 cellules de l'embryon avec une orientation anormale des divisions des cellules entraînant une organisation cellulaire atypique (figure 47B et C) (Gu et al., 2012). En effet on observe chez les mutants *spt6like* la présence de deux cellules au niveau du centre quiescent à la base du futur embryon, alors qu'une seule est visible chez un embryon sauvage. Il y a également une sur-prolifération des cellules de l'hypophyse conduisant à des embryons de formes anormales (figure 48). L'hypocotyle est en général plus court que celui observé chez des embryons sauvages, la taille des cotylédons est souvent disproportionnée et le méristème est élargi et sur-développé (figure 48).

Dans le but de comprendre ce qui ne s'effectue pas correctement chez les mutants *spt6like*, différentes analyses ont été réalisées principalement par hybridation *in situ* (Gu et al., 2012). De manière intéressante, le profil d'expression des gènes *STM*, *CUC1*, *WUS*, *FIL* et *WOX5* est moins restreint chez le mutant que chez des embryons sauvages. Au contraire, l'expression du gène *PLT1*, qui code un facteur de transcription impliqué dans la formation de la base de l'embryon, n'est pas détectée chez le mutant *spt6like*, alors qu'elle l'est chez les embryons sauvages à un stade de développement comparable. Au cours du développement embryonnaire, la protéine *PLT1* agit de manière antagoniste à celles de facteurs de transcription HD-ZIPIII impliqués dans la régulation de

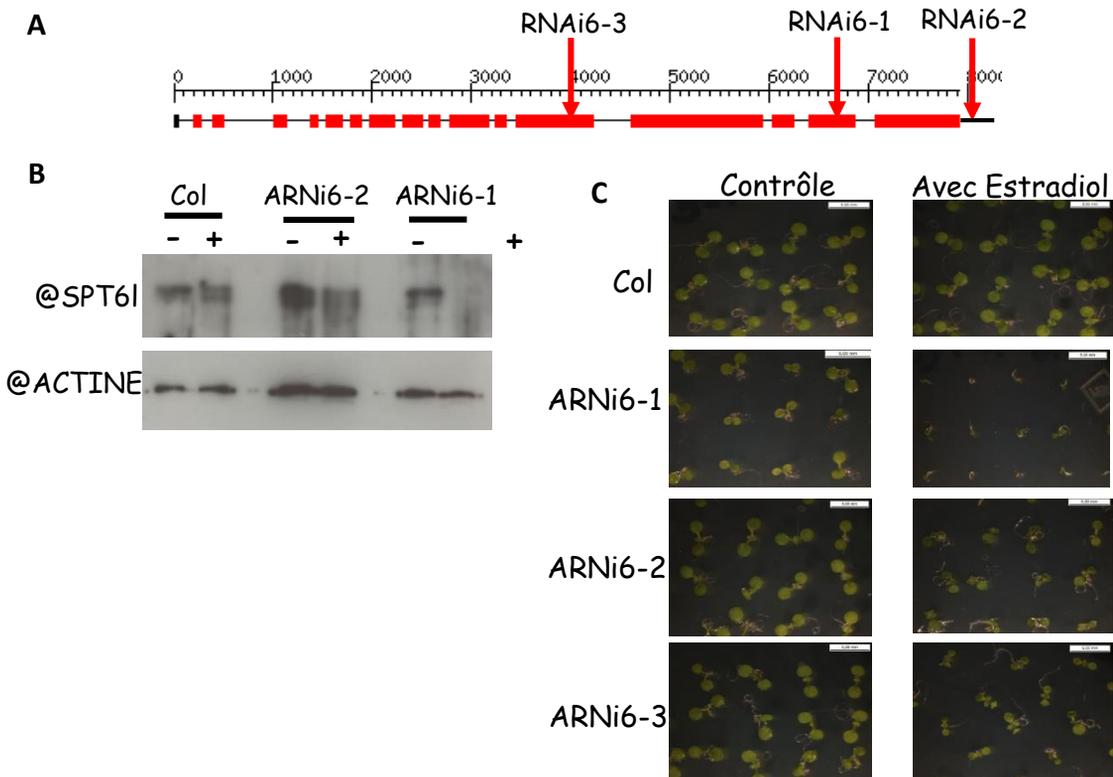


Figure 48. Caractérisation de lignées RNAi *SPT6like*. **A**, Information sur les régions ciblées par les différents ARNi synthétiques *spt6like*. **B**, Teneur en protéine SPT6like dans différentes lignées. Des extraits protéiques de plantules de 7 jours placées (+) ou non (-) en présence de 20mM d'estradiol ont été analysés par western blot en utilisant des anticorps antiSPT6like dirigés contre un peptide de la protéine SPT6like. **C**, Phénotype de plantules de différentes lignées RNAi traitées ou non pendant 7 jours avec 20mM d'estradiol.

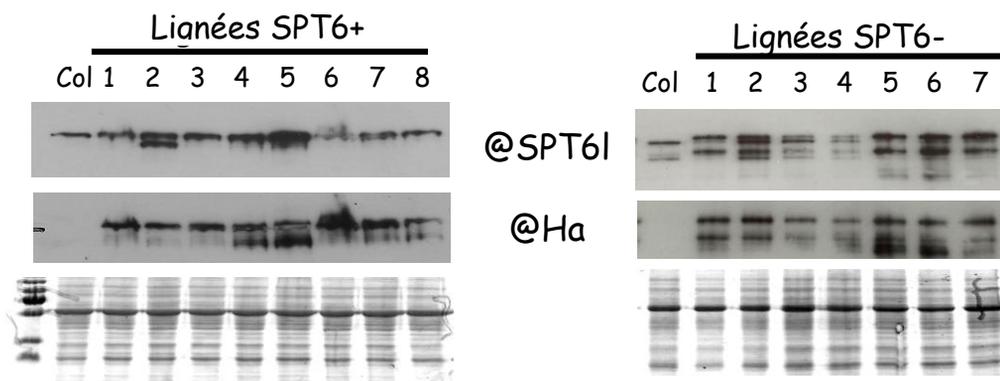


Figure 49. Caractérisation de lignées *spt6like* complémentées. La teneur en protéine SPT6like a été analysée par western blot à partir d'extraits protéiques de fleurs de plantes sauvages ou de mutants *spt6like* complémentés avec la forme SPT6+ ou SPT6-. Les anticorps antiSPT6like et HA ont indépendamment été utilisés. Une coloration au bleu de Coomassie est utilisée comme contrôle de charge.

l'expression des gènes *PHB* et *PHV*. De manière intéressante, les gènes *PHB* et *PHV* ne sont actifs, dans des embryons sauvages, que dans le domaine apical et périphérique de l'embryon au stade globulaire et au stade coeur, alors que chez le mutant *spt6like*, ces deux gènes sont ectopiquement actifs dans tout le domaine basal. Une situation inverse est observée dans les tissus provasculaires d'embryons sauvages au stade "torpille" dans lesquels les gènes *PHB* et *PHV* sont actifs, alors qu'aucune expression de ces gènes, dans ces tissus, n'est observée chez le mutant *spt6like*. L'ensemble de ces travaux démontre que la protéine SPT6like est à la fois impliquée dans la bonne coordination de l'expression de gènes et dans la mise en place des axes de symétries de l'embryon. Les auteurs de ces travaux proposent que cet effet pourrait impliquer les microARN sans en faire la moindre démonstration.

- Quels sont les effets résultant d'une diminution de la quantité de protéines SPT6like dans des plantes adultes ?

Afin de progresser dans la caractérisation de la fonction de la protéine SPT6like, nous avons cherché à diminuer la quantité de cette protéine chez la plante adulte. Pour cela nous avons mis en place une stratégie basée sur l'utilisation de microARN artificiels ciblant les transcrits *SPT6like*. Trois constructions différentes, réalisées à partir du squelette du pri-miR319a, ont été produites par le biais d'une collaboration avec Rebecca Schawb (figure 48A). Chaque construction (6-1, 6-2 et 6-3) permet la production d'un pre-miARN dont le produit va cibler une partie différente du transcrit *SPT6like*. Ces constructions ont été placées sous le contrôle d'un promoteur inductible en présence d'oestradiol. Après criblage de différentes lignées, au moins une par construction a été identifiée avec une teneur en protéine SPT6like réduite dans des plantules de 7 jours en présence d'oestradiol (figure 48B). Les phénotypes alors observés indiquent un important retard de croissance (figure 48C). Le fait que la stratégie microARN artificiel fonctionne permet de supposer que SPT6like n'est peut être pas impliquée dans cette voie, sauf si un équilibre est atteint. Mais se posait alors une importante question: le retard de croissance observé résulte-t-il de la fonction classique des protéines SPT6 ou du domaine AGO hook de cette protéine ?

- Complémentation de mutants *spt6like* avec une version SPT6like dont la plateforme est mutée.

Dans le but de déterminer le rôle physiologique de la plateforme AGO hook de la protéine SPT6like, des mutants *spt6like* ont été complétés soit avec une version SPT6like native (SPT6+) soit avec une version mutée (SPT6-), possédant des substitutions des résidus Tryptophane du domaine AGO hook par des résidus Phénylalanine et donc incapable d'interagir avec des protéines AGO. Les deux versions ont été placées sous le contrôle de la séquence promotrice (1456 pb) du gène *SPT6like*. En première génération, des transformants pour les deux types de versions, SPT6+ et SPT6- ont pu être isolés. Une analyse du niveau d'accumulation des protéines étiquetées, par l'utilisation d'un anticorps anti-HA, révèle que toutes les lignées sélectionnées produisent les protéines SPT6+ ou SPT6- (figure 49).

Le génotypage des lignées transformées indique que, pour au moins deux lignées indépendantes pour chaque version (SPT6+ et SPT6-), il est possible d'identifier des plantes homozygotes pour la mutation *spt6like* et que ces plantes ne présentent aucun phénotype anormal évident. Cela signifie que la mutation *spt6like* peut être complétée par l'un ou l'autre des transgènes et que donc la présence de résidus Tryptophane dans la partie C-terminale de la protéine SPT6like n'est pas indispensable à cette complémentation. Nous pouvons par conséquent conclure que la létalité des mutants *spt6like* est plutôt liée à la fonction classique des protéines SPT6 et donc,

Figure 50. Analyse du compartiment d'accumulation de la protéine SPT6like. Un fractionnement cellulaire à été réalisé à partir d'extraits protéiques de plantes produisant les protéines SPT6+ ou SPT6-. T, extrait initial, C, compartiment cytoplasmique, N, compartiment nucléaire. Les révélations ont été réalisées avec les anticorps antiSPT6like, anti histone et anti UGPase.

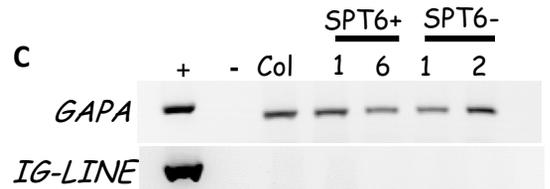
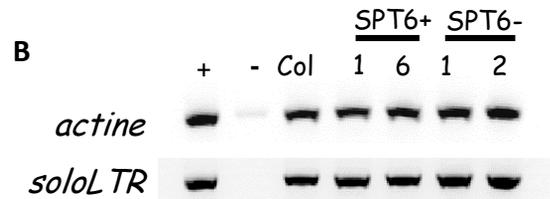
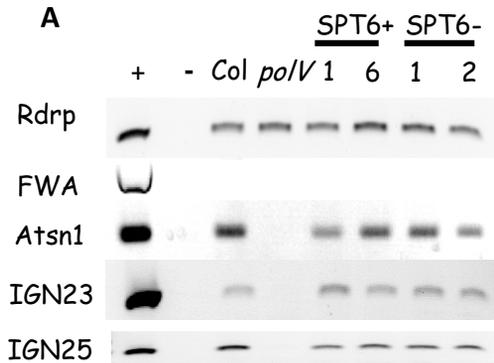
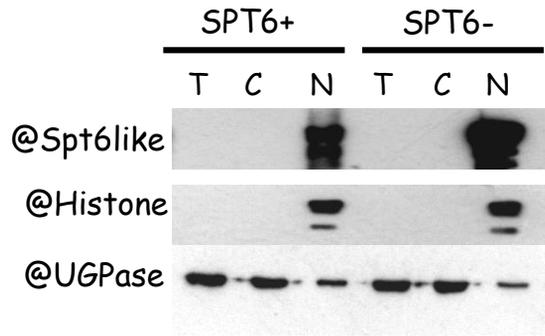
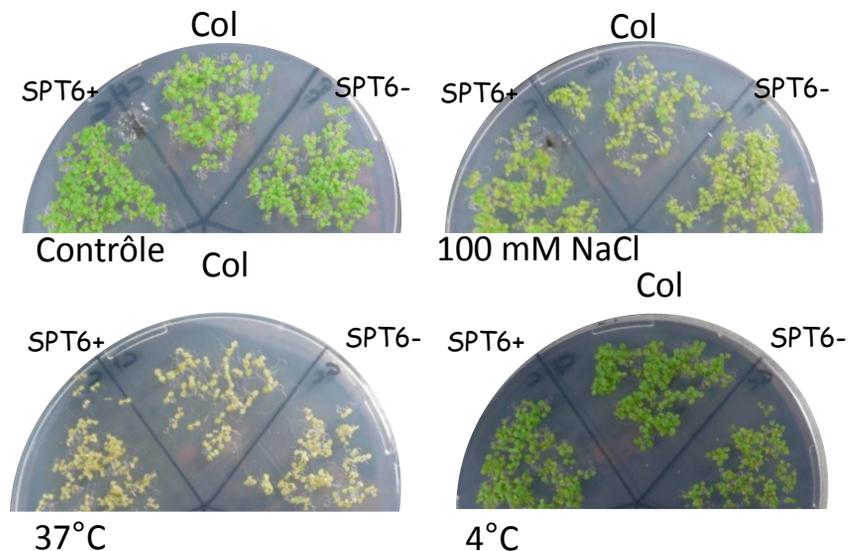


Figure 51. Méthylation et expression de cibles du RdDM dans les plantes SPT6+ et SPT6-. A, Méthylation des cibles *IGN23* et *IGN25*. Les ADN de plantes sauvages (Col), de mutants *POIV* et de plantes SPT6+ et SPT6- ont été digérés ou non (+) par l'enzyme HaeIII. Les amorces *Rdrp* sont utilisées comme contrôle d'équilibrage et les amorces *FWA* comme contrôle de digestion. **B, Méthylation de la cible *SoloLTR*.** Les ADN des différentes plantes ont été digérés par l'enzyme AluI. Les amorces *actine* sont utilisées comme contrôle d'équilibrage. **C, Expression du locus *IG/Line*.** RT-PCR réalisée à partir d'ARN de fleurs. De l'ADN génomique de plantes sauvages (+) a été utilisé comme contrôle positif. L'expression est révélée par amplification avec des amorces *IG-Line*, et des amorces *GAPA* qui permettent de contrôler l'équilibrage.

Figure 52. Analyse de la réponse des plantes SPT6+ et SPT6- à des stress environnementaux. Des graines sauvages (Col) et celles de plantes SPT6+ et SPT6- ont été placées à germer *in vitro* sur milieu MS en présence ou non de 100 mM de NaCl, ou incubées ou non après 3 jours à 37 °C ou 4 °C pendant 15 jours.



de la partie N terminale de cette protéine plutôt qu'à sa plateforme à protéines AGO. Ce résultat indique que, comme nous l'avons observé pour la grande sous-unité de la polymérase V et pour la protéine SPT5like, les résidus Tryptophane de l'extension C-terminale de la protéine SPT6like ne semblent pas indispensables pour certaines fonctions de cette protéine. Le fait que les plantes SPT6like- ne présentent apparemment pas les phénotypes typiques de mutants affectés dans la voie des microARN ou celle des tasiARN suggère que la plateforme SPT6like doit probablement être impliquée dans un autre mécanisme. L'association de la protéine SPT6like à la polymérase II peut aussi être un argument sur le fait qu'elle puisse difficilement être impliquée dans la dégradation des messagers qui se déroule en général dans le cytoplasme. Nous avons d'ailleurs analysé la localisation subcellulaire des deux versions SPT6+ et SPT6- dans les plantes transgéniques générées. L'analyse montre que les deux protéines s'accumulent préférentiellement dans la fraction nucléaire (figure 50). Cela est cohérent avec le rôle potentiel de facteur d'élongation de la transcription de la protéine SPT6like et démontre que les résidus Tryptophane de l'extension C-terminale ne sont pas indispensables pour sa rétention dans le noyau.

Recherche d'un phénotype lié à la mutation de la plateforme AGO de SPT6like.

- Implication dans l'initiation de la voie du RdDM ?

En caractérisant des mutants de la grande sous unité de la polymérase II, Zeng et ses collaborateurs ont observés que des mutants faibles de *nrbp2* (gène qui code la grande sous unité de la polymérase II) montrent un défaut sur la répression de certaines cibles du RdDM (Zheng et al., 2009). Les auteurs ont alors proposé que la Polymérase II permettrait la production de transcrits qui recruteraient des protéines AGO et qui permettraient d'initier la voie du RdDM (Zheng et al., 2009). Etant donné que SPT6like se lie au complexe de l'ARN polymérase II il semblait intéressant de tester une des cibles identifiées, comme par exemple le locus *soloLTR* (responsable de la répression de transcrits *Igline*) dans les plantes SPT6+ et SPT6-. L'analyse montre que dans les plantes SPT6- le degré de méthylation de l'ADN du locus *soloLTR*, tout comme celui des cibles AtSN1, IGN23, IGN25 qui sont aussi des cibles du RdDM, est similaire à celui observé dans des plantes sauvages (figure 51). Nos analyses montrent également que comme pour les plantes sauvages aucun transcrit *Igline* (correspondant au locus *soloLTR*) n'est détecté dans les plantes SPT6- (figure 51C). Une analyse plus globale des cibles du RdDM, réalisée en collaboration avec Steve Jacobsen, démontre par ailleurs un niveau de méthylation des cibles du RdDM des plantes SPT6- similaire à celui des plantes SPT6+ et différent de celui des mutants *polV* (*nrpe1*). L'ensemble de ces observations permet de conclure que la protéine SPT6like n'est pas être impliquée dans l'initiation de la voie du RdDM.

- Implication dans la réponse à des stress environnementaux ?

En raison de la relation entre AGO1 et la réponse à des stress environnementaux nous avons réalisé des tests exploratoires pour révéler un éventuel phénotype particulier des plantes SPT6+ et SPT6- en présence de différentes contraintes (figure 52). Les essais réalisées *in vitro*, en présence de 100 mM de NaCl ou une incubation de 15 jours à 37°C ou 4°C ne montrent, dans les conditions utilisées, aucune différence entre les plantules sauvages et les mutants *spt6like* complémentés avec la forme SPT6+ ou SPT6-. Les réponses à la présence de pathogènes n'ont pour le moment pas encore été analysées.

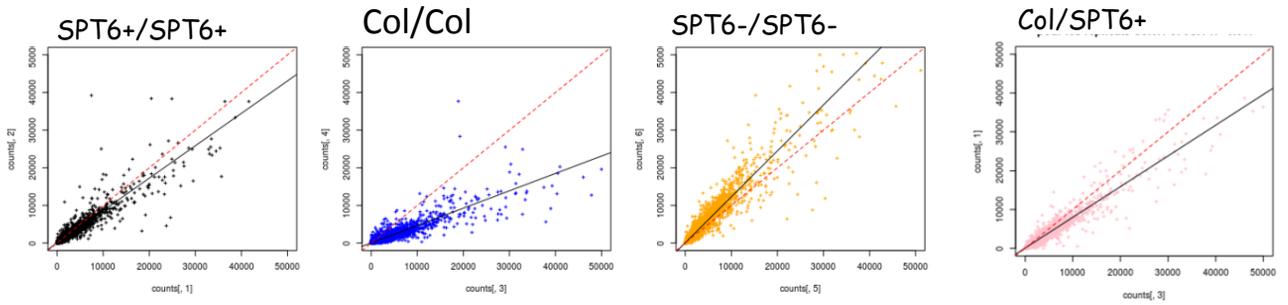


Figure 53. Analyse de la corrélation entre les réplicats. Comparaison des résultats obtenus entre différents réplicats

- Implication sur la transcription générale des gènes.

Du fait de l'association physique de SPT6like avec l'ARN Polymérase II il était nécessaire d'observer un éventuel effet de la mutation de la plateforme sur la transcription globale des gènes. Cette analyse devait également permettre de vérifier que les plantes SPT6-, qui ne présentent pas de phénotype caractéristique de la voie des microARN, ne sont réellement pas affectées dans cette voie. Des ARN polyA de boutons floraux de plantes sauvages, de plantes SPT6-, SPT6+ et sauvages ont été séquencés par le service HiSeq de Fasteris. Les séquences, obtenues à partir de deux réplicats pour chaque échantillon, ont été analysées par Marie Christine Charpentier, bio-informaticienne au laboratoire. Plus de 95% des séquences mappent des régions du génome et la corrélation entre les réplicats est correcte excepté pour un des réplicats obtenu à partir des ARN de fleurs sauvages (figure 53). La comparaison des différentes courbes de régression a conduit à exclure ce réplicat des futures analyses. Les comparaisons entre les échantillons ont permis de conclure que très peu de différences apparaissent entre les plantes SPT6+ et SPT6-. En effet seuls 12 gènes semblent être significativement différentiellement exprimés entre les plantes SPT6+ et SPT6-. Deux de ces gènes (*At1g12010* et *At3g44300*), ceux qui montraient les plus fortes différences, ont été testés et n'ont cependant pas été validés par RT-PCR et qRT-PCR. Le fait que la majorité des transcrits montre des quantités similaires entre les échantillons confirme que la voie classique de PTGS, celle qui implique des microARN, n'est probablement pas altérée. L'expression de transposons semble, elle aussi, ne pas être affectée par les mutations de la plateforme. Cette observation est en adéquation avec les analyses des cibles du RdDM précédemment décrites.

Finalement les seules différences qui ont été observées entre les échantillons concernent l'accumulation d'environ 100 transcrits alternatifs. Pour 48 d'entre eux, l'analyse indique une absence d'accumulation dans les plantes SPT6-. Ces transcrits sont pourtant présents en quantité similaires dans les plantes sauvages et les plantes SPT6+. Pour les 60 autres transcrits c'est une différence de quantité qui est observée. L'analyse plus approfondie de ces différences est actuellement en cours.

La conclusion sur le travail relatif à la protéine SPT6like est présentée dans la partie conclusion et projets.

6/ Conclusion et projets

6-1/ Conclusion

Du fait que les conclusions concernant les protéines LEA ou la protéine SPT5like ont déjà été publiées ou devraient l'être prochainement et ont été brièvement présentées dans les chapitres précédents ma conclusion portera uniquement sur la caractérisation de la protéine SPT6like.

Nos analyses préliminaires indiquent une interaction de SPT6like à la fois avec AGO1 et la grande sous-unité du complexe de l'ARN polymérase II. Ces informations permettent de définir SPT6like comme capable de faire le lien, probablement par un mécanisme co-transcriptionnel entre la machinerie de transcription Polymérase II et une éventuelle voie de régulation impliquant AGO1 et des petits ARN. Les analyses indiquent aussi que indépendamment de la présence de sa plateforme AGO hook, SPT6like est probablement la protéine qui assure chez *Arabidopsis* la fonction conservée de facteur d'élongation lié à l'ARN Polymérase II. Le premier argument en faveur de cette proposition est que le gène *Spt6like* est exprimé de manière constitutive et que des mutants d'insertion sont létaux. Le second argument est que le gène *SPT6* classique, sans plateforme, n'est lui exprimé que dans les fleurs et des mutants d'insertion dans ce gène ne présentent pas de phénotype particulier (Gu et al., 2012). Une approche par ARNi (microARN artificiel) a permis, en parallèle, de montrer qu'une diminution, même faible, de la quantité de protéine SPT6like dans des jeunes plantes entraîne un retard de croissance qui indique un rôle global dans le développement de la plante compatible avec les fonctions classiques des protéines SPT6.

Afin de découpler l'activité globale de SPT6like d'une fonction plus spécifique qui serait liée à son extension AGO hook, les mutants *spt6like* ont été complétés avec une version SPT6like sauvage (SPT6+) ou une version SPT6like pour laquelle les résidus Tryptophane de la plateforme sont remplacés par des résidus Phénylalanine (SPT6-). Le fait que les deux versions se sont révélées capables de compléter la létalité du mutant *spt6like* permet de conclure que la/les fonction(s) essentielle(s) de SPT6like semblent être parfaitement remplies par le domaine conservé de type SPT6. Il est intéressant à ce sujet d'ajouter que, tout comme les protéines SPT6 classique, SPT6like est capable de lier une protéine de type IWS (Yoh et al., 2008; Li et al., 2010). L'analyse plus approfondie des mutants complétés par les deux versions indique que la méthylation des cibles du RdDM et donc la voie du RdDM classique (TGS) n'est pas affectée. Le phénotype apparemment normal des plantes SPT6- suggère que la voie des microARN (PTGS) n'est pas non plus affectée. Hypothèse en partie confirmée par l'analyse des transcrits (RNAseq). Aucun effet n'a de plus été observé en réponse à certains stress environnementaux (température, stress osmotique).

L'analyse transcriptomique des plantes SPT6- suggère cependant une possible implication de la plateforme AGO de SPT6like dans l'épissage alternatifs ou AS (Alternative Splicing) de certains messagers. Mon premier projet va donc consister à poursuivre la caractérisation de la protéine SPT6like afin d'identifier la fonction de sa plateforme AGO hook en approfondissant d'une part le phénotype de modification de l'épissage alternatif et en testant également toutes les voies connues impliquant AGO1 et des petits ARN.

Deux projets portant sur la caractérisation d'autres protéines possédant des plateformes à AGO seront également décrits.

6-2/ Projet 1. Etude de la fonction de la plateforme AGO hook de la protéine SPT6like.

La piste dont nous disposons pour le moment est celle d'un rôle potentiel de SPT6like dans l'épissage alternatif. Ce mécanisme est important car il concerne près de 60 % des gènes d'*Arabidopsis* et plus de 90 % chez l'homme (Syed et al., 2012). L'épissage alternatif assure en fait une double fonction il permet de moduler la nature des transcrits mais aussi d'agir sur la production d'ARN aberrants capables d'initier des voies de répression. La régulation de ce mécanisme dépend en grande partie de la présence du complexe spliceosome, composé de hnRNP (heterogenous nuclear RNP) et de plus de 100 protéines souvent riches en Sérines et Arginine (protéines SR) mais aussi de séquences enhanceurs ou silencers d'épissage (Long and Caceres, 2009; Zhou et al., 2014). Il est intéressant de noter que les végétaux possèdent deux fois plus de gènes codant des protéines SR que les organismes non photosynthétiques et que leur épissage alternatif est particulièrement sensible aux conditions environnementales. Observations qui indiquent que des mécanismes particuliers de régulation pourraient être spécifiques des végétaux. Récemment différentes protéines végétales atypiques ont été identifiées comme ayant un effet à la fois sur l'épissage alternatif et sur des voies de répression. Il s'agit par exemple de la protéine SERRATE impliquée dans la voie des microARN, qui se lie à la fois au pre-ARNm et aux protéines DCL1 et les protéines SMD1 qui agissent sur certaines voies du PTGS, celle du S-PTGS (répression de transgènes), en augmentant la stabilité d'ARN aberrants qui initient le silencing, l'effet est alors dépendant de l'intensité de la répression (Iwata et al., 2013; Raczynska et al., 2014; Elvira-Matelot et al., 2016). Il est important de noter que les mutants *hyl1* et *dcl1* qui sont affectés dans la voie des microARN présentent également des défauts d'épissage (Raczynska et al., 2014). Par ailleurs différents marqueurs des spliceosomes co-localisent avec les sites de maturation des microARN (Bologna and Voinnet, 2014). La protéine SR45 est, elle, à la fois impliquée dans l'épissage et le RdDM (Ausin et al., 2012). Tout cela indique qu'il existe un lien entre l'épissage et certaines voies de répression.

L'étude de l'épissage alternatif chez les levures et les animaux a mis en évidence des caractéristiques importantes. Un premier point est le lien entre l'épissage alternatif et les marques d'histones, principalement la tri-méthylation H3K4 et H3K36, contrôlée par SET2, la mono-méthylation H3K9 et les acétylations d'histones qui auraient un effet sur le recrutement de protéines impliquées dans l'épissage alternatif (Nogues et al., 2002; Sims et al., 2007; Allo et al., 2009; Schor et al., 2009; Piacentini et al., 2009; Gunderson and Johnson, 2009; Luco et al., 2010; Dhami et al., 2010; Luco et al., 2011 a et b; Braunschweig et al., 2013). Il est démontré également que l'épissage alternatif est en partie modulé par la vitesse de progression du complexe de l'ARN Polymérase II et par la densité en nucléosomes (Zhou et al., 2014). Une corrélation existe aussi entre l'épissage et la méthylation de l'ADN (Li-Byarlay et al., 2013; Zhou et al., 2014). Enfin assez récemment il a été montré qu'un apport ectopique de siARN spécifique d'un transcrit affecte son épissage et que la protéine AGO1 serait impliquée dans ce contrôle chez l'homme (Schor et al., 2013; Braunschweig et al., 2013).

L'ensemble de ces observations montre une relation étroite entre l'épissage alternatif et les marques d'histones, la densité en nucléosomes, la vitesse d'élongation de la transcription, potentiellement de petits ARN et certaines voies de répression. Le fait que la protéine SPT6like puisse agir sur certains des facteurs cités et qu'elle possède une plateforme capable de recruter des protéines AGO et des petits ARN en fait par conséquent un bon candidat.

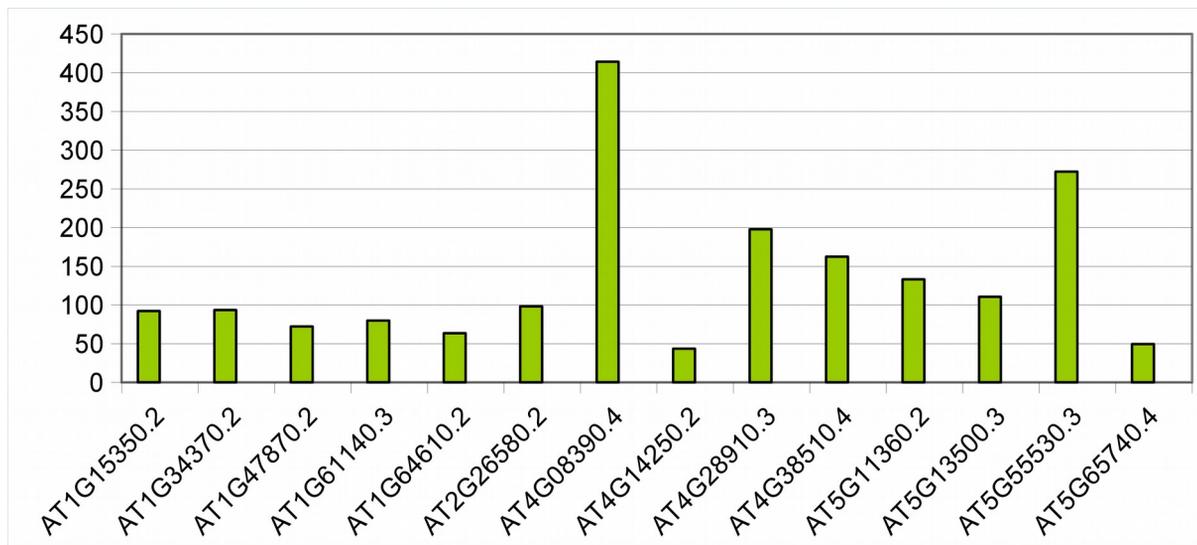


Figure 54. Transcrits alternatifs les plus fortement accumulés dans les plantes sauvages et SPT6+ et pas détectés dans les plantes SPT6-.

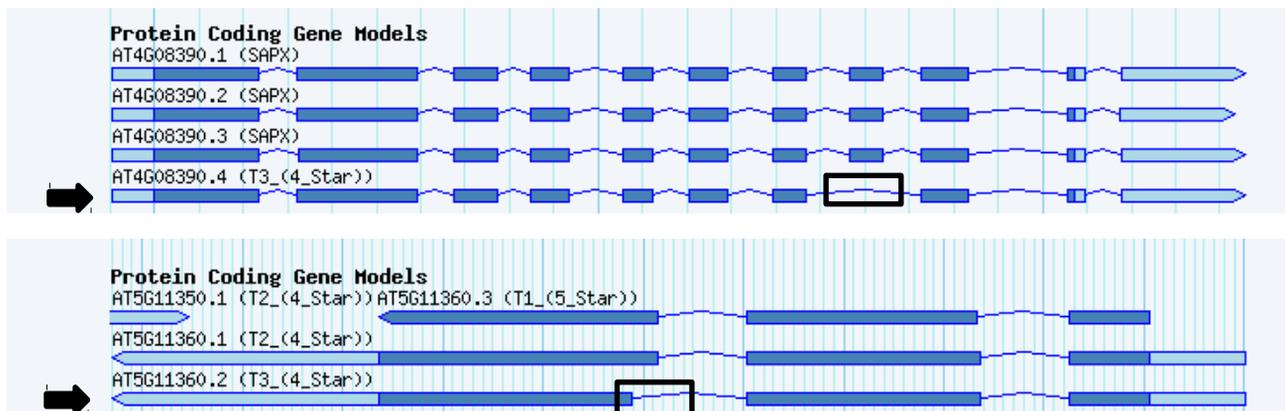


Figure 55. Structure de transcrits alternatif (At4g08390.4 et At5g11360.2) non produits dans les plantes SPT6-.

1. Analyse du défaut d'épissage.

Il est nécessaire dans un premier temps d'analyser à partir de l'ensemble des séquences d'ARN toutes les cibles potentielles du défaut d'épissage afin de déterminer si une seule catégorie d'épissage ou plusieurs pourraient être affectées. Chez les végétaux il semble que la majorité des événements d'épissage concerne les rétentions d'introns (40 % contre 5 % chez l'homme). La figure 54 indique les transcrits alternatifs qui sont le plus fortement détectés parmi les ARNs de plantes sauvages et SPT6+ et absents dans les ARNs de plantes SPT6-. La figure 55 donne deux exemples de la structure des transcrits affectés. Il est intéressant de noter que dans le second cas le transcrit AT5G11360.1 n'est présent que dans les plantes SPT6- et AT5G11360.2 que dans les plantes sauvages ou SPT6+. L'analyse complète de tous les transcrits affectés permettra d'identifier les cibles les plus fortement impactées qui feront alors l'objet d'une validation par comparaison de la quantité de leurs transcrits dans les plantes SPT6+ et SPT6- par analyse de type RT-PCR. Des cibles d'épissages alternatifs pourront également être analysées dans d'autres tissus en utilisant par exemple la base de donnée établie lors du travail réalisé avec le mutant *serrate* (Iwata et al., 2013; Raczynska et al., 2014).

Selon les résultats obtenus des analyses plus fines pourront alors être envisagées pour mieux appréhender la fonction de SPT6like. Une première analyse portera sur la comparaison des marques d'histones présentes au niveau des sites liés au défaut d'épissage par des expériences d'immuno-précipitation de chromatine. Une attention particulière sera portée aux marques H3K4 et H3K36 triméthylées, à la mono-méthylation H3K9 et aux histones acétylées qui ont, chez d'autres organismes, un effet sur l'épissage alternatif (Nogues et al., 2002; Allo et al., 2009; Schor et al., 2009; Luco et al., 2010; Braunschweig et al., 2013). L'épissage alternatif pouvant être impacté par la méthylation de l'ADN, on pourra comparer par des approches de CHOP-PCR ou par traitement bisulfite la méthylation des sites correspondant aux événements d'épissage défectueux entre les plantes SPT6+ et SPT6-. Enfin, l'identification des petits ARN potentiellement impliqués dans cette régulation se fera tout d'abord par une recherche de la présence de petits ARN correspondant aux sites d'épissages alternatifs dans les bases de données de petits ARNs ou en comparant les populations de petits ARN dans les plantes SPT6- et SPT6+. Une alternative consisterait à réaliser un séquençage de petits ARN après immuno-précipitation de SPT6like. Enfin il pourrait aussi être envisagé de comparer dans les plantes SPT6+ versus SPT6- le recrutement de protéines connues comme étant impliquées dans l'épissage alternatif.

L'ensemble de ces approches devrait permettre de confirmer ou non l'implication de SPT6like dans l'épissage alternatif et mieux comprendre son rôle dans ce mécanisme.

2. Poursuite de la caractérisation du rôle de la plateforme AGO de la protéine SPT6like.

En parallèle des travaux proposés ci-dessus la caractérisation de la plateforme SPT6like sera poursuivie par des analyses ciblées en se concentrant spécifiquement sur les voies qui impliquent AGO1.

1) Rôle dans le silencing des transgènes.

Il est établi que l'insertion d'un transgène peut entraîner par une voie similaire à celle du PTGS une répression initiée par la production de transcrits aberrants sens (on parle alors S-PTGS pour "sense transgene post transcriptional gene silencing") ou de produits résultant de séquences répétées inversées du transgène (on parle d'IR-PTGS "Inverted Repeat-PTGS") (figure 56) (Parent et al., 2015). Les ARN double brin, produits par RDR6 (RNA Dependant RNA polymérase)

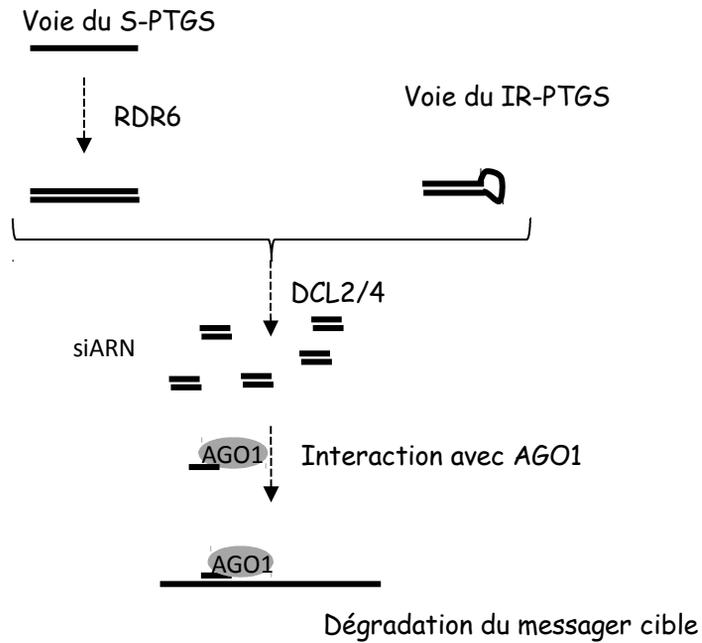


Figure 56. Modèle d'action du S-PTGS et IR-PTGS. Dans la voie IR-PTGS ce sont des structures tige boucle qui sont à l'origine des duplex d'ARNs.

ou provenant des régions répétées inversées sont coupés par DCL2 et/ou DCL4 en siARN, chargés par AGO1, et entraînent la dégradation des transcrits du transgène. De manière intéressante parallèlement à cette dégradation ciblée un mécanisme peu connu, probablement co-transcriptionnel, entraîne la méthylation du transgène et la maintenance du PTGS en impliquant AGO1 (Elmayan et al., 1998; Morel et al., 2000; Jauvion et al., 2012; Bologna and Voinnet, 2014). D'autres travaux comme celui sur la protéine JMJ14 (histone déméthylase) démontrent aussi un lien entre l'état de la chromatine ou des modifications d'histones et le S-PTGS (Le Masson et al., 2012). Au regard de ces informations il pourrait être envisagé que SPT6like puisse par le biais de petits ARN être impliquée dans la mise en place des modifications ciblées de la chromatine liées à la répression de transgènes. Afin de tester cette hypothèse, il est nécessaire de disposer d'un transgène initiant le S-PTGS, pour cela nous avons décidé d'utiliser la lignée L1 caractérisée par une forte répression du transgène *p35S::GUS* (Elmayan et al., 1998). L'objectif étant d'analyser l'effet de la mutation de la plateforme SPT6like sur la répression du transgène (contexte SPT6-). Dans cette optique et parce que les mutants d'insertion Salk, contenant un promoteur 35S, semblent peu appropriés pour étudier le S-PTGS les constructions SPT6+ et SPT6- ont été introduites dans un mutant *spt6like* provenant de la banque SAIL (569-E05) préalablement croisé avec la lignée L1 (Daxinger et al., 2008). Il s'agira alors d'étudier l'extinction du transgène, à la fois au niveau de l'accumulation des transcrits, de l'activité GUS, de la méthylation de l'ADN, de la production de siARN et des marques d'histones.

2) Rôle dans les mécanismes de réponses aux pathogènes

De nombreux pathogènes affectent les plantes en perturbant leurs voies de répression, souvent celle du PTGS ou du NMD (Nonsense Mediated Decay), pour pouvoir se multiplier et se propager. En fonction du pathogène, ce sont différentes étapes des voies de répression qui peuvent être ciblées (Pumplin and Voinnet, 2013 ; Garcia et al., 2014 ; Sun et al., 2015; Halter and Navarro, 2015). Afin de tester si le mécanisme impliquant la plateforme de SPT6like pourrait être une cible de pathogènes il est envisagé d'étudier la sensibilité des plantes SPT6+ versus SPT6- en présence de virus. Parce qu'il est établi que l'infection virale peut moduler la répression de transgènes cette analyse pourrait aussi être réalisée avec les plantes générées à partir de la lignée L1.

3) Rôle dans les voies *tasiARN* et *natsiARN*

Deux autres voies qui implique AGO1 et qui n'ont pas encore été testées sont celles des *tasiARN* et des *natsiARN*. Afin de tester si SPT6like peut agir via sa plateforme sur la voie des *tasiARN* il est nécessaire d'analyser l'accumulation *tasiARN* ou celle de leurs cibles, les transcrits ARF par exemple (Hunter et al., 2006). L'analyse de la voie des *natsiARN* peut elle être réalisée en quantifiant par exemple les *natsiARN SRO5/PCDH* (Borsani et al., 2005).

La caractérisation du rôle de la plateforme SPT6like en est encore à ses débuts mais les caractéristiques déjà identifiées permettent d'envisager un rôle important, probablement spécifique des végétaux, permettant de recruter, au niveau du complexe de la polymérase II, des protéines AGO1 et de petits ARNs.

6-3/ Projets 2 et 3: Caractérisation de nouvelles protéines à plateforme AGO hook.

La caractérisation de protéines possédant de potentielles plateformes à AGO a pour le moment été relativement productive. C'est ainsi que des protéines telles que SPT5like, SDE3 et NERD chez les plantes et la protéine humaine (prion PrPc) ont récemment fait l'objet de publications (Bies-Etheve et al., 2009; Garcia et al., 2012; Pontier et al., 2012; Gibbings et al., 2012). Cette approche, jugée originale lors de la publication du travail sur SPT5like par un littérature report dans EMBO reports (2009) a aussi été reconnue deux fois par le Faculty of 1000 Biology. Contrairement à d'autres approches plus dirigées la stratégie adoptée ne permet pas de supposer dans quelles voies les protéines identifiées vont être impliquées et permet ainsi d'identifier potentiellement de nouveaux mécanismes.

Parce que des petits ARN et des protéines AGO sont impliqués dans de nombreux mécanismes et qu'il est justifié de penser que toutes les voies de régulation ne sont peut être pas encore toutes identifiées il semble particulièrement intéressant de caractériser de nouvelles protéines à AGO hook. La liste des 20 meilleurs candidats issus du crible bio-informatique, contient des protéines riches en Glycine ou en Glutamine, certaines sont prédites comme étant nucléaires, certaines possèdent des domaines de liaison à des ARN (domaine RRM) et certaines enfin ne présentent pas de caractéristiques particulières permettant de supposer leur fonction biologique. J'envisage à la suite ou en parallèle du travail sur SPT6like de me concentrer sur l'étude de deux groupes particuliers de candidats. Le premier est composé de protéines possédant en plus de leur potentielle plateforme à AGO un domaine RRM (RNA recognition motif) et le second est constitué de protéines appartenant à la famille des AGD.

Contexte scientifique concernant les protéines possédant un domaine RRM.

196 protéines possédant un domaine RRM sont présentes chez Arabidopsis ces protéines sont en général impliquées dans la régulation du devenir des messagers auxquels elles se lient. Elles forment des hnRNP (heterogenous nuclear ribonucleoprotein) ou des snRNP (small nuclear ribonucleoprotein) et certaines comme par exemple, les protéines SR, les protéines U2AF2, les snRNP U1 et U1 ou la protéine RBM25 sont impliquées dans le contrôle de l'épissage, d'autres comme les protéines ORRM1, ORRM2, ORRM3 et ORRM4 sont impliquées dans l'édition des messagers (Sun et al., 2013; Jang et al., 2014; Jang et al., 2014; Zhan et al., 2015 ; Shi et al., 2016). Les protéines PABP sont elles principalement impliquées dans le contrôle de la stabilité et la traduction des messagers (Zuber et al., 2016). La protéine LARP1c qui appartient à la famille des LARP, qui possèdent un domaine LA et un domaine RRM, agit elle sur l'initiation de la senescence des feuilles (Zhang et al., 2012a). D'autres, comme la protéine MEL2 du riz ont un rôle dans le contrôle de la méiose (Nonomura et al., 2011; Miyazaki et al., 2015). L'ensemble de ces observations montre que les fonctions des protéines à domaines RRM sont divers et que par conséquent la caractérisation d'une protéine se liant à des messagers et recrutant des protéines AGO liées à de petits ARN représente un réel intérêt. De telles protéines intervenant à l'interface entre des ARN messagers cibles et les protéines AGO liées à de petits ARN pourraient moduler la stabilité, l'épissage, la dégradation de messagers, l'inhibition de leur traduction ou tout autre nouveau mécanisme.

Le groupe de protéines à domaine RRM qui possèdent également les caractéristiques d'une plateforme à AGO forme une famille composée de 4 membres qui n'ont pour le moment fait l'objet d'aucune publication et que nous envisageons donc de caractériser dans le cadre du projet 2.

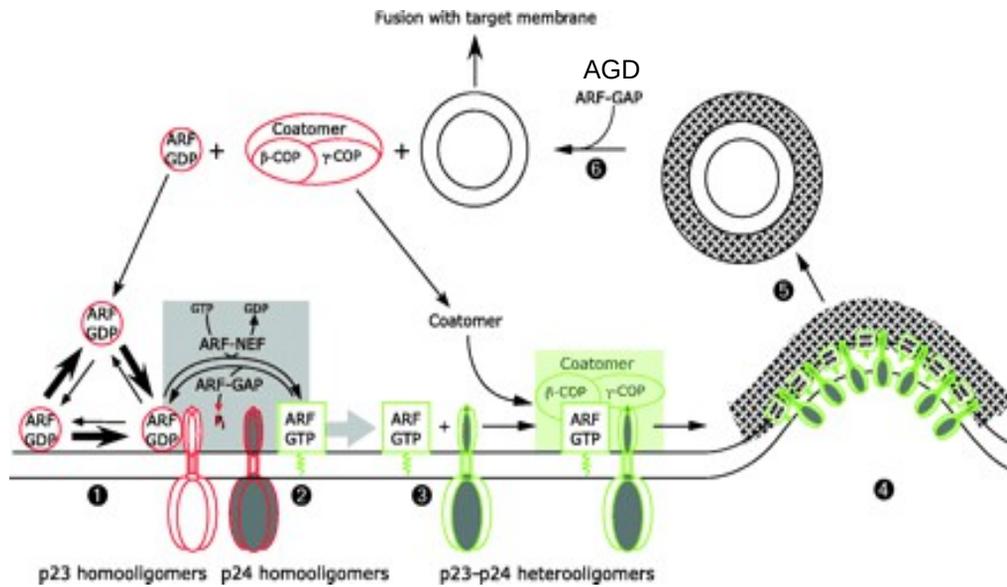


Figure 57. Modèle d'action des protéines AGD. Après formation des vésicules les AGD permettent de décrocher les protéines de l'enveloppe permettant ainsi la fusion de la vésicule avec la membrane réceptrice. Issu de Nickel et al, 2002

Contexte scientifique concernant les protéines appartenant à la famille des AGD.

Les protéines AGD ou ArfGAP (Arf GAP Domain, GAP; GTPases activating protein) ont pour fonction de cibler des GTPases (les Arf GTPases ou ADP-ribosylation factor GTPases) et stimuler leur activité hydrolytique permettant leur dissociation des membranes (figure 58) (Vernoud et al., 2003). Les Arf GTPases sont principalement impliquées dans le contrôle du transport de vésicules dans la cellule et les protéines AGD ont par conséquent un effet sur la formation et le devenir des vésicules (Vernoud et al., 2003; Memon, 2004). Au regard de ces informations, il semble intéressant d'appréhender le rôle de protéines AGD susceptibles de recruter des protéines Argonautes chargées avec des petits ARNs. Ces ARN auraient-ils un rôle de régulateur de la formation des vésicules ? Les protéines AGO et/ou leurs petits ARNs pourraient-ils être intégrés aux vésicules et être ainsi transportés et éventuellement exportés ? On sait par exemple que chez la drosophile le complexe RISC, qui recrute de nombreux petits ARNs, est souvent associé aux membranes de l'appareil de Golgi ou du réticulum endoplasmique (Lee et al., 2009). Chez Arabidopsis des protéines AGO1 sont associées à l'endomembrane du réticulum endoplasmique et pourraient être responsable de la répression traductionnelle (Li et al., 2013a; Bologna and Voinnet, 2014). Par ailleurs de nombreux ARNs, dont des siARN, sont sécrétés dans le milieu intercellulaire (Chen, 2010; Hoy and Buck, 2012; Tosar et al., 2015). C'est le cas par exemple chez les végétaux des petits ARN impliqués dans les réponses systémiques (Halter and Navarro, 2015).

Il y a 15 protéines AGD chez Arabidopsis et seules, deux d'entre elles, AGD6 et AGD7 possèdent une potentielle plateforme à AGO. Il est établi que la sur-accumulation de AGD7 entraîne à la fois un défaut de transport intra-celulaire, mis en évidence par l'étude de protéines marqueurs (sporamine et phaséoline) normalement transportées de l'appareil de Golgi vers la vacuole centrale et un défaut de sécrétion observé à l'aide d'une fusion invertase::GFP (Min et al., 2007). Les travaux réalisés démontrent aussi que AGD7 assure un rôle classique de protéine AGD en ciblant comme AGD5 et AGD8, autres AGD de la même classe, la Arf GTPase ARF1 (Min et al., 2007; Stefano et al., 2010). La caractérisation de ces deux protéines fera l'objet du projet 3.

Stratégie d'étude des projets 2 et 3.

Dans le but de caractériser ces nouvelles protéines à plateforme à AGO nous allons utiliser notre expérience dans ce domaine et aborder l'analyse par des approches systématiques qui vont constituer une sorte de pipeline d'étude. Il sera ensuite nécessaire d'affiner par des approches plus spécifiques qui dériveront des premiers résultats obtenus.

1) Confirmation de la fonction AGO hook.

Une première analyse consistera à confirmer la fonction AGO hook. Cela pourra être réalisé *in vitro* en testant la capacité de protéines recombinantes à lier des protéines AGO par GST pull down. La spécificité d'interaction pourra être aussi testée par l'utilisation de versions mutées sur les résidus Tryptophane. Si une interaction est observée, l'interaction pourra être confirmée *in vivo* à partir d'extraits de plantes transgéniques produisant des versions étiquetées permettant de réaliser des co-immuno-précipitation et tester la présence de protéines AGO. L'identité des protéines AGO retenues représentera un premier indice des voies pouvant être impactées. La suite des caractérisations, réalisées par différentes approches complémentaires, ne sera envisagée que lorsqu'une interaction avec des protéines AGO aura été observée.

Pour le moment seules les protéines à RRM ont été soumises à cette analyse qui a montré une interaction avec AGO4. Cette information indique que ces protéines RRM présentent certaines similarités avec la protéine Smd3 qui co-localise avec AGO4 dans les "cajals bodies" et est impliquée dans l'épissage (Swaraz et al., 2011).

2) Caractérisation phénotypique de mutants d'insertion.

Une analyse fonctionnelle sera initiée par la caractérisation de mutants d'insertions. Ces analyses seront guidées par les caractéristiques particulières des différentes protéines. Pour les protéines RRM par exemple, il sera important de tester d'une part les cibles de la voie du RdDM impliquant AGO4 et responsable d'une répression transcriptionnelle, et d'autre part un éventuel défaut d'épissage. Pour les mutants AGD une attention particulière sera portée à la structure et à l'organisation de l'appareil de Golgi et du réticulum mais aussi au transport et la sécrétion de protéines. Dans le cas où aucun phénotype anormal ne serait observé dans des conditions contrôle ou en présence de stress il pourra être envisagé d'initier l'étude de mutants multiples permettant de contourner d'éventuels problèmes de redondance de fonction des protéines étudiées. Les différentes voies connues pour impliquer de petits ARN pourront également être analysées. Si ces différentes approches ne fournissent pas de pistes fonctionnelles il pourra alors être envisagé de réaliser des analyses de transcrits (RNAseq) ou de comparer les populations de petits ARN.

3) Analyse de la localisation cellulaire des protéines.

La localisation cellulaire des différentes protéines sera étudiée par fractionnement cellulaire qui pourra être complété par immuno-localisation ou par utilisation de plantes transgéniques produisant des protéines de fusion GFP et microscopie confocale.

4) Identification de protéines partenaires.

La recherche de protéines associées sera réalisée par la purification, par des approches de biochimie classique, de complexes protéiques. Les protéines natives seront immuno-précipitées ou une alternative basée sur l'utilisation d'extraits de plantes transgéniques produisant des protéines étiquetées sera utilisée. Les complexes ainsi purifiés permettront l'identification de partenaires potentiels par spectrométrie de masse. Les équipements et services nécessaires nous seront disponibles grâce à des collaborations de longue date avec deux laboratoires localisés à Grenoble : le LAPM (équipe "Epigénétique et Parasites" dirigée par le Dr M.A Hakimi) et la plateforme de spectrométrie de masse du LBGE (équipe EDyP).

5) Identification des ARN associés.

Il pourra aussi être envisagé si nécessaire d'identifier les petits ARN liés aux protéines AGO recrutées. Pour cela une approche de type RIP sera mise en place. Pour les protéines à domaine RRM l'interaction avec des messagers constitue une étape importante.

L'ensemble de ces analyses, éventuellement complétées par d'autres qui pourraient dépendre des résultats obtenus, devraient fournir suffisamment d'informations pour pouvoir identifier le rôle biologique des protéines étudiées. Une analyse préliminaire par puces à ADN a déjà révélé, chez les mutants multiples RRM1,RRM2,RRM3,RRM4 une réponse de type SAR. Cette observation pourrait indiquer un rôle potentiel dans la réponse immunitaire des plantes aux agents pathogènes. Il va par conséquent être important de confirmer cette observation par l'analyse

de l'accumulation de transcrits caractéristiques et l'analyse d'éventuelles modifications de la chromatine au niveau des gènes affectés. La sensibilité des mutants à la présence de pathogènes sera également testée. Il sera ensuite probablement nécessaire d'affiner l'étude par des approches plus ciblées. La poursuite de ce projet fait partie des priorités et nous sommes actuellement à la recherche d'un étudiant en thèse susceptible de réaliser ce travail.

A terme les deux projets présentés pourraient permettre d'identifier de nouveaux mécanismes biologiques impliquant des protéines AGO et de petits ARN.

REFERENCES BIBLIOGRAPHIQUES

- Allo, M., Buggiano, V., Fededa, J.P., Petrillo, E., Schor, I., de la Mata, M., Agirre, E., Plass, M., Eyra, E., Elela, S.A., et al. (2009). Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.* 16, 717–724.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- Amente, S., Napolitano, G., Licciardo, P., Monti, M., Pucci, P., Lania, L., and Majello, B. (2005). Identification of proteins interacting with the RNAPII FCP1 phosphatase: FCP1 forms a complex with arginine methyltransferase PRMT5 and it is a substrate for. *FEBS Lett.* 579, 683–689.
- Ausin, I., Greenberg, M.V.C., Li, C.F., and Jacobsen, S.E. (2012). The splicing factor SR45 affects the RNA-directed DNA methylation pathway in Arabidopsis. *Epigenetics* 7, 29–33.
- Bahrndorff, S., Tunnacliffe, A., Wise, M.J., McGee, B., Holmstrup, M., and Loeschcke, V. (2009). Bioinformatics and protein expression analyses implicate LEA proteins in the drought response of Collembola. *J. Insect Physiol.* 55, 210–217.
- Bies, N., Aspart, L., Carles, C., Gallois, P., and Delseny, M. (1998). Accumulation and degradation of Em proteins in Arabidopsis thaliana: evidence for post-transcriptional controls. *J. Exp. Bot.* 49, 1925–1933.
- Bies-Etheve, N., da Silva Conceicao, A., Giraudat, J., Koornneef, M., Leon-Kloosterziel, K., Valon, C., and Delseny, M. (1999). Importance of the B2 domain of the Arabidopsis ABI3 protein for Em and 2S albumin gene regulation. *Plant Mol. Biol.* 40, 1045–1054.
- Bies-Etheve, N., Gaubier-Comella, P., Debures, A., Lasserre, E., Jobet, E., Raynal, M., Cooke, R., and Delseny, M. (2008). Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in Arabidopsis thaliana. *Plant Mol. Biol.* 67, 107–124.
- Bies-Etheve, N., Pontier, D., Lahmy, S., Picart, C., Vega, D., Cooke, R., and Lagrange, T. (2009). RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. *EMBO Rep.* 10, 649–654.
- Bologna, N.G., and Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. *Annu. Rev. Plant Biol.* 65, 473–503.
- 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, 1279–1291.
- Braunschweig, U., Gueroussov, S., Plocik, A.M., Graveley, B.R., and Blencowe, B.J. (2013). Dynamic integration of splicing within gene regulatory pathways. *Cell* 152, 1252–1269.
- Brocard, I.M., Lynch, T.J., and Finkelstein, R.R. (2002). Regulation and role of the Arabidopsis

- abscisic acid-insensitive 5 gene in abscisic acid, sugar, and stress response. *Plant Physiol.* *129*, 1533–1543.
- Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., and Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science* *320*, 1185–1190.
- Burugula, B.B., Jeronimo, C., Pathak, R., Jones, J.W., Robert, F., and Govind, C.K. (2014). Histone deacetylases and phosphorylated polymerase II C-terminal domain recruit Spt6 for cotranscriptional histone reassembly. *Mol. Cell. Biol.* *34*, 4115–4129.
- Li-Byarlay, H., Li, Y., Stroud, H., Feng, S., Newman, T.C., Kaneda, M., Hou, K.K., Worley, K.C., Elsik, C.G., Wickline, S.A., et al. (2013). RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 12750–12755.
- Cao, J., and Li, X. (2015). Identification and phylogenetic analysis of late embryogenesis abundant proteins family in tomato (*Solanum lycopersicum*). *Planta* *241*, 757–772.
- Carles, C., Bies-Etheve, N., Aspart, L., Leon-Kloosterziel, K.M., Koornneef, M., Echeverria, M., and Delseny, M. (2002). Regulation of *Arabidopsis thaliana* Em genes: role of ABI5. *Plant J. Cell Mol. Biol.* *30*, 373–383.
- Cavel, E., Pillot, M., Pontier, D., Lahmy, S., Bies-Etheve, N., Vega, D., Grimanelli, D., and Lagrange, T. (2011). A plant-specific transcription factor IIB-related protein, pBRP2, is involved in endosperm growth control. *PLoS One* *6*, e17216.
- Charfeddine, S., Saidi, M.N., Charfeddine, M., and Gargouri-Bouزيد, R. (2015). Genome-wide identification and expression profiling of the late embryogenesis abundant genes in potato with emphasis on dehydrins. *Mol. Biol. Rep.* *42*, 1163–1174.
- Chen, X. (2010). Small RNAs - secrets and surprises of the genome. *Plant J. Cell Mol. Biol.* *61*, 941–958.
- Chen, H., and Xiong, L. (2008). Role of HY5 in abscisic acid response in seeds and seedlings. *Plant Signal. Behav.* *3*, 986–988.
- Chen, R., Jiang, H., Li, L., Zhai, Q., Qi, L., Zhou, W., Liu, X., Li, H., Zheng, W., Sun, J., et al. (2012). The *Arabidopsis* mediator subunit MED25 differentially regulates jasmonate and abscisic acid signaling through interacting with the MYC2 and ABI5 transcription factors. *Plant Cell* *24*, 2898–2916.
- Chodosh, L.A., Fire, A., Samuels, M., and Sharp, P.A. (1989). 5,6-Dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits transcription elongation by RNA polymerase II in vitro. *J. Biol. Chem.* *264*, 2250–2257.
- Clarke, M.W., Boddington, K.F., Warnica, J.M., Atkinson, J., McKenna, S., Madge, J., Barker, C.H., and Graether, S.P. (2015). Structural and Functional Insights into the Cryoprotection of Membranes by the Intrinsically Disordered Dehydrins. *J. Biol. Chem.* *290*, 26900–26913.

- Colmenero-Flores, J.M., Moreno, L.P., Smith, C.E., and Covarrubias, A.A. (1999). Pvlea-18, a member of a new late-embryogenesis-abundant protein family that accumulates during water stress and in the growing regions of well-irrigated bean seedlings. *Plant Physiol.* *120*, 93–104.
- Cuming, A.C., and Lane, B.G. (1979). Protein synthesis in imbibing wheat embryos. *Eur. J. Biochem. FEBS* *99*, 217–224.
- Daxinger, L., Hunter, B., Sheikh, M., Jauvion, V., Gascioli, V., Vaucheret, H., Matzke, M., and Furner, I. (2008). Unexpected silencing effects from T-DNA tags in Arabidopsis. *Trends Plant Sci.* *13*, 4–6.
- DeGennaro, C.M., Alver, B.H., Marguerat, S., Stepanova, E., Davis, C.P., Bahler, J., Park, P.J., and Winston, F. (2013). Spt6 regulates intragenic and antisense transcription, nucleosome positioning, and histone modifications genome-wide in fission yeast. *Mol. Cell. Biol.* *33*, 4779–4792.
- Delseny, M., Bies-Etheve, N., Carles, C., Hull, G., Vicient, C., Raynal, M., Grellet, F., and Aspart, L. (2001). Late Embryogenesis Abundant (LEA) protein gene regulation during Arabidopsis seed maturation. *J. Plant Physiol.* *158*, 419–427.
- Dengl, S., Mayer, A., Sun, M., and Cramer, P. (2009). Structure and in vivo requirement of the yeast Spt6 SH2 domain. *J. Mol. Biol.* *389*, 211–225.
- Dhami, P., Saffrey, P., Bruce, A.W., Dillon, S.C., Chiang, K., Bonhoure, N., Koch, C.M., Bye, J., James, K., Foad, N.S., et al. (2010). Complex exon-intron marking by histone modifications is not determined solely by nucleosome distribution. *PLoS One* *5*, e12339.
- Ding, Z.J., Yan, J.Y., Li, G.X., Wu, Z.C., Zhang, S.Q., and Zheng, S.J. (2014). WRKY41 controls Arabidopsis seed dormancy via direct regulation of ABI3 transcript levels not downstream of ABA. *Plant J. Cell Mol. Biol.* *79*, 810–823.
- Dronamraju, R., and Strahl, B.D. (2014). A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. *Nucleic Acids Res.* *42*, 870–881.
- Durr, J., Lolas, I.B., Sorensen, B.B., Schubert, V., Houben, A., Melzer, M., Deutzmann, R., Grasser, M., and Grasser, K.D. (2014). The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis. *Nucleic Acids Res.* *42*, 4332–4347.
- 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.
- El-Shami, M., Pontier, D., Lahmy, S., Braun, L., Picart, C., Vega, D., Hakimi, M.-A., Jacobsen, S.E., Cooke, R., and Lagrange, T. (2007). Reiterated WG/GW motifs form functionally and evolutionarily conserved. *Genes Dev.* *21*, 2539–2544.
- Elvira-Matlot, E., Bardou, F., Ariel, F., Jauvion, V., Bouteiller, N., Le Masson, I., Cao, J., Crespi, M.D., and Vaucheret, H. (2016). The Nuclear Ribonucleoprotein SmD1 Interplays with Splicing, RNA Quality Control, and Posttranscriptional Gene Silencing in Arabidopsis. *Plant Cell* *28*, 426–438.

- Ezcurra, I., Wycliffe, P., Nehlin, L., Ellerstrom, M., and Rask, L. (2000a). Transactivation of the *Brassica napus* napin promoter by ABI3 requires interaction of the conserved B2 and B3 domains of ABI3 with different cis-elements: B2 mediates activation through an ABRE, whereas B3 interacts with an RY/G-box. *Plant J. Cell Mol. Biol.* *24*, 57–66.
- Ezcurra, I., Wycliffe, P., Nehlin, L., Ellerstrom, M., and Rask, L. (2000b). Transactivation of the *Brassica napus* napin promoter by ABI3 requires interaction of the conserved B2 and B3 domains of ABI3 with different cis-elements: B2 mediates activation through an ABRE, whereas B3 interacts with an RY/G-box. *Plant J. Cell Mol. Biol.* *24*, 57–66.
- Feng, C.-Z., Chen, Y., Wang, C., Kong, Y.-H., Wu, W.-H., and Chen, Y.-F. (2014). Arabidopsis RAV1 transcription factor, phosphorylated by SnRK2 kinases, regulates the expression of ABI3, ABI4, and ABI5 during seed germination and early seedling development. *Plant J. Cell Mol. Biol.* *80*, 654–668.
- Finkelstein, R.R., and Lynch, T.J. (2000). The Arabidopsis abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. *Plant Cell* *12*, 599–609.
- Florea, L., Hartzell, G., Zhang, Z., Rubin, G.M., and Miller, W. (1998). A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome Res.* *8*, 967–974.
- Gao, J., and Lan, T. (2016). Functional characterization of the late embryogenesis abundant (LEA) protein gene family from *Pinus tabulaeformis* (Pinaceae) in *Escherichia coli*. *Sci. Rep.* *6*, 19467.
- Garcia, D. (2008). A miRacle in plant development: role of microRNAs in cell differentiation and patterning. *Semin. Cell Dev. Biol.* *19*, 586–595.
- Garcia, D., Garcia, S., Pontier, D., Marchais, A., Renou, J.P., Lagrange, T., and Voinnet, O. (2012). Ago hook and RNA helicase motifs underpin dual roles for SDE3 in antiviral defense and silencing of nonconserved intergenic regions. *Mol. Cell* *48*, 109–120.
- Garcia, D., Garcia, S., and Voinnet, O. (2014). Nonsense-mediated decay serves as a general viral restriction mechanism in plants. *Cell Host Microbe* *16*, 391–402.
- Garcia-Ruiz, H., Carbonell, A., Hoyer, J.S., Fahlgren, N., Gilbert, K.B., Takeda, A., Giampetruzzi, A., Garcia Ruiz, M.T., McGinn, M.G., Lowery, N., et al. (2015). Roles and programming of Arabidopsis ARGONAUTE proteins during Turnip mosaic virus infection. *PLoS Pathog.* *11*, e1004755.
- Gaubier, P., Raynal, M., Hull, G., Huestis, G.M., Grellet, F., Arenas, C., Pages, M., and Delseny, M. (1993). Two different Em-like genes are expressed in *Arabidopsis thaliana* seeds during maturation. *Mol. Gen. Genet. MGG* *238*, 409–418.
- Gibbings, D., Leblanc, P., Jay, F., Pontier, D., Michel, F., Schwab, Y., Alais, S., Lagrange, T., and Voinnet, O. (2012). Human prion protein binds Argonaute and promotes accumulation of microRNA effector complexes. *Nat. Struct. Mol. Biol.* *19*, 517–524, S1.
- Gilmour, S.J., Artus, N.N., and Thomashow, M.F. (1992). cDNA sequence analysis and expression of two cold-regulated genes of *Arabidopsis thaliana*. *Plant Mol. Biol.* *18*, 13–21.

- Glover-Cutter, K., Kim, S., Espinosa, J., and Bentley, D.L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat. Struct. Mol. Biol.* *15*, 71–78.
- Graether, S.P., and Boddington, K.F. (2014). Disorder and function: a review of the dehydrin protein family. *Front. Plant Sci.* *5*, 576.
- Gu, X.-L., Wang, H., Huang, H., and Cui, X.-F. (2012). SPT6L encoding a putative WG/GW-repeat protein regulates apical-basal polarity of embryo in Arabidopsis. *Mol. Plant* *5*, 249–259.
- Gunderson, F.Q., and Johnson, T.L. (2009). Acetylation by the transcriptional coactivator Gcn5 plays a novel role in co-transcriptional spliceosome assembly. *PLoS Genet.* *5*, e1000682.
- Guo, M., Xu, F., Yamada, J., Egelhofer, T., Gao, Y., Hartzog, G.A., Teng, M., and Niu, L. (2008). Core structure of the yeast spt4-spt5 complex: a conserved module for regulation of transcription elongation. *Struct. Lond. Engl.* *1993 16*, 1649–1658.
- Guo, S., Yamaguchi, Y., Schilbach, S., Wada, T., Lee, J., Goddard, A., French, D., Handa, H., and Rosenthal, A. (2000). A regulator of transcriptional elongation controls vertebrate neuronal development. *Nature* *408*, 366–369.
- Halter, T., and Navarro, L. (2015). Multilayer and interconnected post-transcriptional and co-transcriptional control of plant NLRs. *Curr. Opin. Plant Biol.* *26*, 127–134.
- Hara, M., Terashima, S., Fukaya, T., and Kuboi, T. (2003). Enhancement of cold tolerance and inhibition of lipid peroxidation by citrus dehydrin in transgenic tobacco. *Planta* *217*, 290–298.
- Hartzog, G.A. (2003). Transcription elongation by RNA polymerase II. *Curr. Opin. Genet. Dev.* *13*, 119–126.
- Hartzog, G.A., and Fu, J. (2013). The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation. *Biochim. Biophys. Acta* *1829*, 105–115.
- Harvey, J.J.W., Lewsey, M.G., Patel, K., Westwood, J., Heimstadt, S., Carr, J.P., and Baulcombe, D.C. (2011). An antiviral defense role of AGO2 in plants. *PloS One* *6*, e14639.
- He, X.-J., Hsu, Y.-F., Zhu, S., Wierzbicki, A.T., Pontes, O., Pikaard, C.S., Liu, H.-L., Wang, C.-S., Jin, H., and Zhu, J.-K. (2009). An effector of RNA-directed DNA methylation in arabidopsis is an ARGONAUTE 4- and. *Cell* *137*, 498–508.
- Herr, A.J., Jensen, M.B., Dalmay, T., and Baulcombe, D.C. (2005). RNA polymerase IV directs silencing of endogenous DNA. *Science* *308*, 118–120.
- Holzinger, A., Kaplan, F., Blaas, K., Zechmann, B., Komsic-Buchmann, K., and Becker, B. (2014). Transcriptomics of desiccation tolerance in the streptophyte green alga *Klebsormidium* reveal a land plant-like defense reaction. *PloS One* *9*, e110630.
- Hoy, A.M., and Buck, A.H. (2012). Extracellular small RNAs: what, where, why? *Biochem. Soc. Trans.* *40*, 886–890.
- Hu, Y., and Yu, D. (2014). BRASSINOSTEROID INSENSITIVE2 interacts with ABSCISIC ACID

- INSENSITIVE5 to mediate the antagonism of brassinosteroids to abscisic acid during seed germination in *Arabidopsis*. *Plant Cell* 26, 4394–4408.
- Hull, G.A., Bies, N., Twell, D., and Delseny, M. (1996). Analysis of the promoter of an abscisic acid responsive late embryogenesis abundant gene of *Arabidopsis thaliana*. *Plant Sci.* 114, 181–192.
- Hundertmark, M., and Hinch, D.K. (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* 9, 118.
- Hunter, C., Willmann, M.R., Wu, G., Yoshikawa, M., de la Luz Gutierrez-Nava, M., and Poethig, S.R. (2006). Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in *Arabidopsis*. *Dev. Camb. Engl.* 133, 2973–2981.
- Iwata, Y., Takahashi, M., Fedoroff, N.V., and Hamdan, S.M. (2013). Dissecting the interactions of SERRATE with RNA and DICER-LIKE 1 in *Arabidopsis* microRNA precursor processing. *Nucleic Acids Res.* 41, 9129–9140.
- Jaehning, J.A. (2010). The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochim. Biophys. Acta* 1799, 379–388.
- Jakymiw, A., Lian, S., Eystathiou, T., Li, S., Satoh, M., Hamel, J.C., Fritzler, M.J., and Chan, E.K.L. (2005). Disruption of GW bodies impairs mammalian RNA interference. *Nat. Cell Biol.* 7, 1267–1274.
- Jang, Y.H., Park, H.-Y., Lee, K.C., Thu, M.P., Kim, S.-K., Suh, M.C., Kang, H., and Kim, J.-K. (2014). A homolog of splicing factor SF1 is essential for development and is involved in the alternative splicing of pre-mRNA in *Arabidopsis thaliana*. *Plant J. Cell Mol. Biol.* 78, 591–603.
- Jauvion, V., Rivard, M., Bouteiller, N., Elmayan, T., and Vaucheret, H. (2012). RDR2 partially antagonizes the production of RDR6-dependent siRNA in sense transgene-mediated PTGS. *PLoS One* 7, e29785.
- Ji, L., Liu, X., Yan, J., Wang, W., Yumul, R.E., Kim, Y.J., Dinh, T.T., Liu, J., Cui, X., Zheng, B., et al. (2011). ARGONAUTE10 and ARGONAUTE1 regulate the termination of floral stem cells through two microRNAs in *Arabidopsis*. *PLoS Genet.* 7, e1001358.
- Jia, F., Qi, S., Li, H., Liu, P., Li, P., Wu, C., Zheng, C., and Huang, J. (2014). Overexpression of Late Embryogenesis Abundant 14 enhances *Arabidopsis* salt stress tolerance. *Biochem. Biophys. Res. Commun.* 454, 505–511.
- Jiang, W., and Yu, D. (2009). *Arabidopsis* WRKY2 transcription factor mediates seed germination and postgermination arrest of development by abscisic acid. *BMC Plant Biol.* 9, 96.
- Kaplan, C.D., Morris, J.R., Wu, C., and Winston, F. (2000). Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in *D. melanogaster*. *Genes Dev.* 14, 2623–2634.
- Karlowski, W.M., Zielezinski, A., Carrere, J., Pontier, D., Lagrange, T., and Cooke, R. (2010). Genome-wide computational identification of WG/GW Argonaute-binding proteins in *Arabidopsis*. *Nucleic Acids Res.* 38, 4231–4245.

- Kato, H., Okazaki, K., and Urano, T. (2013). Spt6: two fundamentally distinct functions in the regulation of histone modification. *Epigenetics* 8, 1249–1253.
- Keith, K., Kraml, M., Dengler, N.G., and McCourt, P. (1994). *fusca3*: A Heterochronic Mutation Affecting Late Embryo Development in Arabidopsis. *Plant Cell* 6, 589–600.
- Kim, V.N. (2008). Sorting out small RNAs. *Cell* 133, 25–26.
- Kim, H., Kim, S.-H., Seo, D.H., Chung, S., Kim, S.-W., Lee, J.-S., Kim, W.T., and Lee, J.-H. (2016). ABA-HYPERSENSITIVE BTB/POZ PROTEIN 1 functions as a negative regulator in. *Plant Mol. Biol.* 90, 303–315.
- Kollen, K., Dietz, L., Bies-Etheve, N., Lagrange, T., Grasser, M., and Grasser, K.D. (2015). The zinc-finger protein SPT4 interacts with SPT5L/KTF1 and modulates transcriptional silencing in Arabidopsis. *FEBS Lett.* 589, 3254–3257.
- Komori, T., Inukai, N., Yamada, T., Yamaguchi, Y., and Handa, H. (2009). Role of human transcription elongation factor DSIF in the suppression of senescence and apoptosis. *Genes Cells Devoted Mol. Cell. Mech.* 14, 343–354.
- Koornneef, M., Jorna, M.L., Brinkhorst-van der Swan, D.L., and Karssen, C.M. (1982). The isolation of abscisic acid (ABA) deficient mutants by selection of induced revertants in non-germinating gibberellin sensitive lines of Arabidopsis thaliana (L.) heynh. *TAG Theor. Appl. Genet. Theor. Angew. Genet.* 61, 385–393.
- Kurkela, S., and Franck, M. (1990). Cloning and characterization of a cold- and ABA-inducible Arabidopsis gene. *Plant Mol. Biol.* 15, 137–144.
- Kwak, H., and Lis, J.T. (2013). Control of transcriptional elongation. *Annu. Rev. Genet.* 47, 483–508.
- Lahmy, S., Bies-Etheve, N., and Lagrange, T. (2010). Plant-specific multisubunit RNA polymerase in gene silencing. *Epigenetics* 5, 4–8.
- Larochelle, S., Batliner, J., Gamble, M.J., Barboza, N.M., Kraybill, B.C., Blethrow, J.D., Shokat, K.M., and Fisher, R.P. (2006). Dichotomous but stringent substrate selection by the dual-function Cdk7 complex revealed by chemical genetics. *Nat. Struct. Mol. Biol.* 13, 55–62.
- Lee, Y.S., Pressman, S., Andress, A.P., Kim, K., White, J.L., Cassidy, J.J., Li, X., Lubell, K., Lim, D.H., Cho, I.S., et al. (2009). Silencing by small RNAs is linked to endosomal trafficking. *Nat. Cell Biol.* 11, 1150–1156.
- 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.
- Li, L., Ye, H., Guo, H., and Yin, Y. (2010). Arabidopsis IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 107, 3918–3923.

- Li, L., Deng, D., Chen, X., Wu, B., Hu, K., Qiu, T., Cui, S., and Huang, F. (2015). Expression of the moss PpLEA4-20 gene in rice enhances membrane protection and client proteins stability. *Biochem. Biophys. Res. Commun.* *460*, 386–391.
- Li, S., Liu, L., Zhuang, X., Yu, Y., Liu, X., Cui, X., Ji, L., Pan, Z., Cao, X., Mo, B., et al. (2013a). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. *Cell* *153*, 562–574.
- Li, W., Cui, X., Meng, Z., Huang, X., Xie, Q., Wu, H., Jin, H., Zhang, D., and Liang, W. (2012). Transcriptional regulation of Arabidopsis MIR168a and argonaute1 homeostasis in abscisic acid and abiotic stress responses. *Plant Physiol.* *158*, 1279–1292.
- Li, Y.-J., Fang, Y., Fu, Y.-R., Huang, J.-G., Wu, C.-A., and Zheng, C.-C. (2013b). NFYA1 is involved in regulation of postgermination growth arrest under salt stress in Arabidopsis. *PloS One* *8*, e61289.
- Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R. 3rd, Parker, R., and Hannon, G.J. (2005). A role for the P-body component GW182 in microRNA function. *Nat. Cell Biol.* *7*, 1261–1266.
- Liu, Y., Xie, L., Liang, X., and Zhang, S. (2015). CpLEA5, the Late Embryogenesis Abundant Protein Gene from *Chimonanthus praecox*, Possesses Low Temperature and Osmotic Resistances in Prokaryote and Eukaryotes. *Int. J. Mol. Sci.* *16*, 26978–26990.
- Liu, Z.-Q., Yan, L., Wu, Z., Mei, C., Lu, K., Yu, Y.-T., Liang, S., Zhang, X.-F., Wang, X.-F., and Zhang, D.-P. (2012). Cooperation of three WRKY-domain transcription factors WRKY18, WRKY40, and WRKY60 in repressing two ABA-responsive genes ABI4 and ABI5 in Arabidopsis. *J. Exp. Bot.* *63*, 6371–6392.
- Long, J.C., and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.* *417*, 15–27.
- Lopez-Molina, L., Mongrand, S., Kinoshita, N., and Chua, N.-H. (2003). AFP is a novel negative regulator of ABA signaling that promotes ABI5 protein degradation. *Genes Dev.* *17*, 410–418.
- 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* *327*, 996–1000.
- Luerssen, H., Kirik, V., Herrmann, P., and Misera, S. (1998). FUSCA3 encodes a protein with a conserved VP1/AB13-like B3 domain which is of functional importance for the regulation of seed maturation in Arabidopsis thaliana. *Plant J. Cell Mol. Biol.* *15*, 755–764.
- Mallory, A.C., and Bouche, N. (2008). MicroRNA-directed regulation: to cleave or not to cleave. *Trends Plant Sci.* *13*, 359–367.
- Mallory, A., and Vaucheret, H. (2010). Form, function, and regulation of ARGONAUTE proteins. *Plant Cell* *22*, 3879–3889.
- 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.

- Manfre, A.J., Lanni, L.M., and Marcotte, W.R.J. (2006). The Arabidopsis group 1 LATE EMBRYOGENESIS ABUNDANT protein ATEM6 is required for normal seed development. *Plant Physiol.* *140*, 140–149.
- Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C.M., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* *30*, 1302–1310.
- 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.
- Matsui, A., Mizunashi, K., Tanaka, M., Kaminuma, E., Nguyen, A.H., Nakajima, M., Kim, J.-M., Nguyen, D.V., Toyoda, T., and Seki, M. (2014). tasiRNA-ARF pathway moderates floral architecture in arabidopsis plants subjected to drought stress. *BioMed Res. Int.* *2014*, 303451.
- Matzke, M., Kanno, T., Huettel, B., Daxinger, L., and Matzke, A.J.M. (2006). RNA-directed DNA methylation and Pol IVb in Arabidopsis. *Cold Spring Harb. Symp. Quant. Biol.* *71*, 449–459.
- Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Soding, J., and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. *Nat. Struct. Mol. Biol.* *17*, 1272–1278.
- McCullough, L., Connell, Z., Petersen, C., and Formosa, T. (2015). The Abundant Histone Chaperones Spt6 and FACT Collaborate to Assemble, Inspect, and Maintain Chromatin Structure in *Saccharomyces cerevisiae*. *Genetics* *201*, 1031–1045.
- Meinke, D.W. (1992). A Homoeotic Mutant of Arabidopsis thaliana with Leafy Cotyledons. *Science* *258*, 1647–1650.
- Memon, A.R. (2004). The role of ADP-ribosylation factor and SAR1 in vesicular trafficking in plants. *Biochim. Biophys. Acta* *1664*, 9–30.
- Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., et al. (2008). Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* *133*, 116–127.
- Min, M.K., Kim, S.J., Miao, Y., Shin, J., Jiang, L., and Hwang, I. (2007). Overexpression of Arabidopsis AGD7 causes relocation of Golgi-localized proteins to the endoplasmic reticulum and inhibits protein trafficking in plant cells. *Plant Physiol.* *143*, 1601–1614.
- Misera, S., Muller, A.J., Weiland-Heidecker, U., and Jurgens, G. (1994). The FUSCA genes of Arabidopsis: negative regulators of light responses. *Mol. Gen. Genet.* *MGG 244*, 242–252.
- Miyazaki, S., Sato, Y., Asano, T., Nagamura, Y., and Nonomura, K.-I. (2015). Rice MEL2, the RNA recognition motif (RRM) protein, binds in vitro to meiosis-expressed genes containing U-rich RNA consensus sequences in the 3'-UTR. *Plant Mol. Biol.* *89*, 293–307.
- Monke, G., Seifert, M., Keilwagen, J., Mohr, M., Grosse, I., Hahnel, U., Junker, A., Weisshaar, B., Conrad, U., Baumlein, H., et al. (2012). Toward the identification and regulation of the Arabidopsis

- thaliana ABI3 regulon. *Nucleic Acids Res.* *40*, 8240–8254.
- Nambara, E., Hayama, R., Tsuchiya, Y., Nishimura, M., Kawaide, H., Kamiya, Y., and Naito, S. (2000). The role of ABI3 and FUS3 loci in *Arabidopsis thaliana* on phase transition from late embryo development to germination. *Dev. Biol.* *220*, 412–423.
- Ni, Z., Saunders, A., Fuda, N.J., Yao, J., Suarez, J.-R., Webb, W.W., and Lis, J.T. (2008). P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. *Mol. Cell. Biol.* *28*, 1161–1170.
- Nogues, G., Kadener, S., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2002). Transcriptional activators differ in their abilities to control alternative splicing. *J. Biol. Chem.* *277*, 43110–43114.
- Nonomura, K.-I., Eiguchi, M., Nakano, M., Takashima, K., Komeda, N., Fukuchi, S., Miyazaki, S., Miyao, A., Hirochika, H., and Kurata, N. (2011). A novel RNA-recognition-motif protein is required for premeiotic G1/S-phase transition in rice (*Oryza sativa* L.). *PLoS Genet.* *7*, e1001265.
- Onodera, Y., Haag, J.R., Ream, T., Costa Nunes, P., Pontes, O., and Pikaard, C.S. (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* *120*, 613–622.
- Palangat, M., Renner, D.B., Price, D.H., and Landick, R. (2005). A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIS. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 15036–15041.
- Parent, J.-S., Jauvion, V., Bouche, N., Beclin, C., Hachet, M., Zytnicki, M., and Vaucheret, H. (2015). Post-transcriptional gene silencing triggered by sense transgenes involves uncapped antisense RNA and differs from silencing intentionally triggered by antisense transgenes. *Nucleic Acids Res.* *43*, 8464–8475.
- Pedrosa, A.M., Martins, C. de P.S., Goncalves, L.P., and Costa, M.G.C. (2015). Late Embryogenesis Abundant (LEA) Constitutes a Large and Diverse Family of Proteins Involved in Development and Abiotic Stress Responses in Sweet Orange (*Citrus sinensis* L. Osb.). *PloS One* *10*, e0145785.
- Perruc, E., Kinoshita, N., and Lopez-Molina, L. (2007). The role of chromatin-remodeling factor PKL in balancing osmotic stress responses during *Arabidopsis* seed germination. *Plant J. Cell Mol. Biol.* *52*, 927–936.
- Piacentini, L., Fanti, L., Negri, R., Del Vescovo, V., Fatica, A., Altieri, F., and Pimpinelli, S. (2009). Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in *Drosophila*. *PLoS Genet.* *5*, e1000670.
- Piskurewicz, U., Jikumaru, Y., Kinoshita, N., Nambara, E., Kamiya, Y., and Lopez-Molina, L. (2008). The gibberellic acid signaling repressor RGL2 inhibits *Arabidopsis* seed germination by stimulating abscisic acid synthesis and ABI5 activity. *Plant Cell* *20*, 2729–2745.
- 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.

- Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M.-A., Lerbs-Mache, S., Colot, V., and Lagrange, T. (2005). Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* *19*, 2030–2040.
- Pontier, D., Picart, C., Roudier, F., Garcia, D., Lahmy, S., Azevedo, J., Alart, E., Laudie, M., Karlowski, W.M., Cooke, R., et al. (2012). NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in *Arabidopsis*. *Mol. Cell* *48*, 121–132.
- Ponting, C.P. (2002). Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucleic Acids Res.* *30*, 3643–3652.
- Puhakainen, T., Hess, M.W., Makela, P., Svensson, J., Heino, P., and Palva, E.T. (2004). Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. *Plant Mol. Biol.* *54*, 743–753.
- Pumplin, N., and Voinnet, O. (2013). RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. *Nat. Rev. Microbiol.* *11*, 745–760.
- Raczynska, K.D., Stepien, A., Kierzkowski, D., Kalak, M., Bajczyk, M., McNicol, J., Simpson, C.G., Szweykowska-Kulinska, Z., Brown, J.W.S., and Jarmolowski, A. (2014). The SERRATE protein is involved in alternative splicing in *Arabidopsis thaliana*. *Nucleic Acids Res.* *42*, 1224–1244.
- Ramachandran, V., and Chen, X. (2008). Small RNA metabolism in *Arabidopsis*. *Trends Plant Sci.* *13*, 368–374.
- Raynal, M., Guillemot, J., Gueguen, C., Cooke, R., Delseny, M., and Gruber, V. (1999). Structure, organization and expression of two closely related novel *Lea* (late-embryogenesis-abundant) genes in *Arabidopsis thaliana*. *Plant Mol. Biol.* *40*, 153–165.
- Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A.D., Zhu, J.-K., Hagen, G., Guilfoyle, T.J., Pasa-Tolic, L., and Pikaard, C.S. (2009). Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol. Cell* *33*, 192–203.
- Rowley, M.J., Avrutsky, M.I., Sifuentes, C.J., Pereira, L., and Wierzbicki, A.T. (2011). Independent chromatin binding of ARGONAUTE4 and SPT5L/KTF1 mediates transcriptional gene silencing. *PLoS Genet.* *7*, e1002120.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., and Barrell, B. (2000). Artemis: sequence visualization and annotation. *Bioinforma. Oxf. Engl.* *16*, 944–945.
- Ryu, H., Cho, H., Bae, W., and Hwang, I. (2014). Control of early seedling development by BES1/TPL/HDA19-mediated epigenetic regulation of ABI3. *Nat. Commun.* *5*, 4138.
- Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., and Lepiniec, L. (2008). Deciphering gene regulatory networks that control seed development and maturation in *Arabidopsis*. *Plant J. Cell Mol. Biol.* *54*, 608–620.

- Schneider, D.A., French, S.L., Osheim, Y.N., Bailey, A.O., Vu, L., Dodd, J., Yates, J.R., Beyer, A.L., and Nomura, M. (2006). RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 12707–12712.
- Schor, I.E., Rascovan, N., Pelisch, F., Allo, M., and Kornblihtt, A.R. (2009). Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 4325–4330.
- Schor, I.E., Gomez Acuna, L.I., and Kornblihtt, A.R. (2013). Coupling between transcription and alternative splicing. *Cancer Treat. Res.* *158*, 1–24.
- Shekhawat, U.K.S., Srinivas, L., and Ganapathi, T.R. (2011). MusaDHN-1, a novel multiple stress-inducible SK(3)-type dehydrin gene, contributes affirmatively to drought- and salt-stress tolerance in banana. *Planta* *234*, 915–932.
- Shi, H., Chen, Y., Qian, Y., and Chan, Z. (2015). Low Temperature-Induced 30 (LTI30) positively regulates drought stress resistance in Arabidopsis: effect on abscisic acid sensitivity and hydrogen peroxide accumulation. *Front. Plant Sci.* *6*, 893.
- Shi, X., Bentolila, S., and Hanson, M.R. (2016). Organelle RNA recognition motif-containing (ORRM) proteins are plastid and mitochondrial editing factors in Arabidopsis. *Plant Signal. Behav.* *0*.
- Sims, R.J. 3rd, Millhouse, S., Chen, C.-F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol. Cell* *28*, 665–676.
- Stefano, G., Renna, L., Rossi, M., Azzarello, E., Pollastri, S., Brandizzi, F., Baluska, F., and Mancuso, S. (2010). AGD5 is a GTPase-activating protein at the trans-Golgi network. *Plant J. Cell Mol. Biol.* *64*, 790–799.
- Su, L., Zhao, C.-Z., Bi, Y.-P., Wan, S.-B., Xia, H., and Wang, X.-J. (2011). Isolation and expression analysis of LEA genes in peanut (*Arachis hypogaea* L.). *J. Biosci.* *36*, 223–228.
- Sun, T., Germain, A., Giloteaux, L., Hammani, K., Barkan, A., Hanson, M.R., and Bentolila, S. (2013). An RNA recognition motif-containing protein is required for plastid RNA editing in Arabidopsis and maize. *Proc. Natl. Acad. Sci. U. S. A.* *110*, E1169–E1178.
- Sun, Y.-W., Tee, C.-S., Ma, Y.-H., Wang, G., Yao, X.-M., and Ye, J. (2015). Attenuation of Histone Methyltransferase KRYPTONITE-mediated transcriptional gene silencing by Geminivirus. *Sci. Rep.* *5*, 16476.
- Svetlov, V., and Nudler, E. (2011). Clamping the clamp of RNA polymerase. *EMBO J.* *30*, 1190–1191.
- Swaraz, A.M., Park, Y.-D., and Hur, Y. (2011). Knock-out mutations of Arabidopsis SmD3-b induce pleiotropic phenotypes through altered transcript splicing. *Plant Sci. Int. J. Exp. Plant Biol.* *180*, 661–671.

- Swire-Clark, G.A., and Marcotte, W.R.J. (1999). The wheat LEA protein Em functions as an osmoprotective molecule in *Saccharomyces cerevisiae*. *Plant Mol. Biol.* *39*, 117–128.
- Syed, N.H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J.W.S. (2012). Alternative splicing in plants--coming of age. *Trends Plant Sci.* *17*, 616–623.
- Tanaka, M., Kikuchi, A., and Kamada, H. (2008). The Arabidopsis histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant Physiol.* *146*, 149–161.
- Taylor, R.M., and Cuming, A.C. (1993). Purification of an endoproteinase that digests the wheat “Em” protein in vitro, and determination of its cleavage sites. *FEBS Lett.* *331*, 76–80.
- Tezuka, K., Taji, T., Hayashi, T., and Sakata, Y. (2013). A novel *abi5* allele reveals the importance of the conserved Ala in the C3 domain for regulation of downstream genes and salt tolerance during germination in Arabidopsis. *Plant Signal. Behav.* *8*, e23455.
- Till, S., Lejeune, E., Thermann, R., Bortfeld, M., Hothorn, M., Enderle, D., Heinrich, C., Hentze, M.W., and Ladurner, A.G. (2007). A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. *Nat. Struct. Mol. Biol.* *14*, 897–903.
- Tosar, J.P., Gambaro, F., Sanguinetti, J., Bonilla, B., Witwer, K.W., and Cayota, A. (2015). Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines. *Nucleic Acids Res.* *43*, 5601–5616.
- Vaucheret, H. (2008). Plant ARGONAUTES. *Trends Plant Sci.* *13*, 350–358.
- Vernoud, V., Horton, A.C., Yang, Z., and Nielsen, E. (2003). Analysis of the small GTPase gene superfamily of Arabidopsis. *Plant Physiol.* *131*, 1191–1208.
- Vicient, C.M., Bies-Etheve, N., and Delseny, M. (2000). Changes in gene expression in the leafy cotyledon1 (*lec1*) and *fusca3* (*fus3*) mutants of Arabidopsis thaliana L. *J. Exp. Bot.* *51*, 995–1003.
- Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F., et al. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev.* *12*, 343–356.
- Wang, M.-B., and Dennis, E.S. (2009). SPT5-like, a new component in plant RdDM. *EMBO Rep.* *10*, 573–575.
- Wang, A.H., Zare, H., Mousavi, K., Wang, C., Moravec, C.E., Sirotkin, H.I., Ge, K., Gutierrez-Cruz, G., and Sartorelli, V. (2013a). The histone chaperone Spt6 coordinates histone H3K27 demethylation and myogenesis. *EMBO J.* *32*, 1075–1086.
- Wang, H., Zhang, X., Liu, J., Kiba, T., Woo, J., Ojo, T., Hafner, M., Tuschl, T., Chua, N.-H., and Wang, X.-J. (2011). Deep sequencing of small RNAs specifically associated with Arabidopsis AGO1 and AGO4 uncovers new AGO functions. *Plant J. Cell Mol. Biol.* *67*, 292–304.
- Wang, M., Li, P., Li, C., Pan, Y., Jiang, X., Zhu, D., Zhao, Q., and Yu, J. (2014). SiLEA14, a novel

- atypical LEA protein, confers abiotic stress resistance in foxtail millet. *BMC Plant Biol.* *14*, 290.
- Wang, Y., Li, L., Ye, T., Lu, Y., Chen, X., and Wu, Y. (2013b). The inhibitory effect of ABA on floral transition is mediated by ABI5 in Arabidopsis. *J. Exp. Bot.* *64*, 675–684.
- Wen, Y., and Shatkin, A.J. (1999). Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev.* *13*, 1774–1779.
- Wenzel, S., Schweimer, K., Rosch, P., and Wohrl, B.M. (2008). The small hSpt4 subunit of the human transcription elongation factor DSIF is a Zn-finger protein with alpha/beta type topology. *Biochem. Biophys. Res. Commun.* *370*, 414–418.
- Wenzel, S., Martins, B.M., Rosch, P., and Wohrl, B.M. (2010). Crystal structure of the human transcription elongation factor DSIF hSpt4 subunit in complex with the hSpt5 dimerization interface. *Biochem. J.* *425*, 373–380.
- Werner, F. (2012). A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J. Mol. Biol.* *417*, 13–27.
- Winston, F. (2001). Control of eukaryotic transcription elongation. *Genome Biol.* *2*, REVIEWS1006.
- Winston, F., Chaleff, D.T., Valent, B., and Fink, G.R. (1984). Mutations affecting Ty-mediated expression of the HIS4 gene of *Saccharomyces cerevisiae*. *Genetics* *107*, 179–197.
- Wise, M.J., and Tunnacliffe, A. (2004). POPP the question: what do LEA proteins do? *Trends Plant Sci.* *9*, 13–17.
- Wu, J., Seng, S., Sui, J., Vonapartis, E., Luo, X., Gong, B., Liu, C., Wu, C., Liu, C., Zhang, F., et al. (2015). *Gladiolus hybridus* ABSCISIC ACID INSENSITIVE 5 (GhABI5) is an important transcription factor in ABA signaling that can enhance *Gladiolus* corm dormancy and Arabidopsis seed dormancy. *Front. Plant Sci.* *6*, 960.
- Xiao, Y., Yang, Y.H., Burckin, T.A., Shiue, L., Hartzog, G.A., and Segal, M.R. (2005). Analysis of a splice array experiment elucidates roles of chromatin elongation factor Spt4-5 in splicing. *PLoS Comput. Biol.* *1*, e39.
- Xu, D., Li, J., Gangappa, S.N., Hettiarachchi, C., Lin, F., Andersson, M.X., Jiang, Y., Deng, X.W., and Holm, M. (2014). Convergence of Light and ABA signaling on the ABI5 promoter. *PLoS Genet.* *10*, e1004197.
- Yamaguchi, Y., Wada, T., Watanabe, D., Takagi, T., Hasegawa, J., and Handa, H. (1999). Structure and function of the human transcription elongation factor DSIF. *J. Biol. Chem.* *274*, 8085–8092.
- Yoh, S.M., Lucas, J.S., and Jones, K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev.* *22*, 3422–3434.
- Yokota, H., Iehisa, J.C.M., Shimosaka, E., and Takumi, S. (2015). Line differences in Cor/Lea and fructan biosynthesis-related gene transcript accumulation are related to distinct freezing tolerance

- levels in synthetic wheat hexaploids. *J. Plant Physiol.* *176*, 78–88.
- Yuan, C., Wang, J., Harrison, A.P., Meng, X., Chen, D., and Chen, M. (2015). Genome-wide view of natural antisense transcripts in *Arabidopsis thaliana*. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* *22*, 233–243.
- Zhan, X., Qian, B., Cao, F., Wu, W., Yang, L., Guan, Q., Gu, X., Wang, P., Okusolubo, T.A., Dunn, S.L., et al. (2015). An *Arabidopsis* PWI and RRM motif-containing protein is critical for pre-mRNA splicing and ABA responses. *Nat. Commun.* *6*, 8139.
- Zhang, B., Jia, J., Yang, M., Yan, C., and Han, Y. (2012a). Overexpression of a LAM domain containing RNA-binding protein LARP1c induces precocious leaf senescence in *Arabidopsis*. *Mol. Cells* *34*, 367–374.
- Zhang, Y., Li, Y., Lai, J., Zhang, H., Liu, Y., Liang, L., and Xie, Q. (2012b). Ectopic expression of a LEA protein gene TsLEA1 from *Thellungiella salsuginea* confers salt-tolerance in yeast and *Arabidopsis*. *Mol. Biol. Rep.* *39*, 4627–4633.
- Zheng, B., Wang, Z., Li, S., Yu, B., Liu, J.-Y., and Chen, X. (2009). Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes Dev.* *23*, 2850–2860.
- Zhou, H.-L., Luo, G., Wise, J.A., and Lou, H. (2014). Regulation of alternative splicing by local histone modifications: potential roles for RNA-guided mechanisms. *Nucleic Acids Res.* *42*, 701–713.
- Zhou, K., Kuo, W.H.W., Fillingham, J., and Greenblatt, J.F. (2009). Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 6956–6961.
- Zhou, X., Hao, H., Zhang, Y., Bai, Y., Zhu, W., Qin, Y., Yuan, F., Zhao, F., Wang, M., Hu, J., et al. (2015). SOS2-LIKE PROTEIN KINASE5, an SNF1-RELATED PROTEIN KINASE3-Type Protein Kinase, Is Important for Abscisic Acid Responses in *Arabidopsis* through Phosphorylation of ABSCISIC ACID-INSENSITIVE5. *Plant Physiol.* *168*, 659–676.
- Zorio, D.A., and Bentley, D.L. (2001). Transcription elongation: the “Foggy” is lifting ellipsis. *Curr. Biol. CB* *11*, R144–R146.
- Zuber, H., Scheer, H., Ferrier, E., Sement, F.M., Mercier, P., Stupfler, B., and Gagliardi, D. (2016). Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in *Arabidopsis*. *Cell Rep.* *14*, 2707–2717.