



**HAL**  
open science

# Analyse génétique et moléculaire du développement de la graine d'*Arabidopsis thaliana*: étude de la régulation de l'expression du gène **LEAFY COTYLEDON 2**

Nathalie Berger

► **To cite this version:**

Nathalie Berger. Analyse génétique et moléculaire du développement de la graine d'*Arabidopsis thaliana*: étude de la régulation de l'expression du gène **LEAFY COTYLEDON 2**. Sciences agricoles. Université Paris Sud - Paris XI, 2012. Français. NNT: 2012PA112033 . tel-00671310

**HAL Id: tel-00671310**

**<https://theses.hal.science/tel-00671310>**

Submitted on 17 Feb 2012

**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.

**THESE DE DOCTORAT DE L'UNIVERSITÉ PARIS-SUD 11**

Présentée pour obtenir le grade de

**DOCTEUR EN SCIENCES DE L'UNIVERSITE PARIS11, ORSAY**

Discipline : Sciences de la Vie  
Ecole doctorale Sciences du végétal

par Nathalie BERGER

Soutenance prévue le : 07 février 2012

**Analyse génétique et moléculaire du développement de la graine d'*Arabidopsis thaliana*: étude de la régulation de l'expression du gène *LEAFY COTYLEDON 2***

Mr Pascal Genschik	Directeur de recherche CNRS	Rapporteur
Mr Richard Thompson	Directeur de recherche INRA	Rapporteur
Mme Gwyneth Ingram	Chargée de recherche CNRS	Examineur
Mr Antoine Martin	Chargé de recherche CNRS	Examineur
Mr Michel Dron	Professeur, université de PARIS XI, Orsay	Examineur
Mr Bertrand Dubreucq	Chargé de recherche INRA	Dir. de thèse



## Remerciements :

Tout d'abord merci à Pascal Genschik, Richard Thompson, Michel Dron, Gwyneth Ingram et Antoine Martin d'avoir accepté d'examiner cette thèse.

Je tiens ensuite à remercier Bertrand Dubreucq et Loïc Lepiniec pour leur soutien et leurs encouragements dans la réalisation de cette thèse mais également depuis le début de ma carrière à l'INRA (ça fait quelques années). Cette thèse a été extrêmement stimulante et motivante grâce à vous. Travailler avec vous a été un vrai bonheur et je n'oublierai pas la chance que j'ai eu de vous avoir comme chefs (moi faillote ???).

Je remercie également Michel Dron, pour avoir énormément facilité mes démarches lors de mon inscription, notamment par la validation de mes acquis professionnels.

Maryelle Fradin pour m'avoir sauvé la vie bien des fois lors des inscriptions annuelles à la FAC et Jacqui shykoff pour sa constante disponibilité et sa gentillesse.

Martine.M et Christian pour avoir relu le manuscrit hyper rapidement et super sérieusement, merci à vous deux.

Je fais enfin un grand merci à Christian qui m'a communiqué une partie de sa passion de la recherche et de son sens du détail (pour ne pas dire maniaquerie) qui m'ont été très utiles dans les dernières étapes de rédaction.

Beaucoup de personnes m'ont aidés dans les manip comme Erwana pour les anticorps ou encore Anne, Béatrice et Johanne pour les commandes. Merci à tous les membres de l'ex-biologie des semences qui forment une vraie petite famille dont j'ai apprécié l'ambiance toutes ces années.

Je n'oublie pas non plus les « soigneurs » de nos plantes avec une spéciale dédicace à Mickael Anjuere et à Bruno Talbotec qui ont une telle conscience professionnelle qu'ils vous tuteurent vos plantes quand vous les zappez allègrement, absorbé par quelques autres manip. Merci d'avoir été un support inestimable toutes ces années. La même chose aux personnes du secrétariat qui par leur efficacité m'ont réconcilié avec la paperasse (enfin presque). Par ordre chronologique, tout d'abord Danielle puis Stéphane et enfin Andrée et Vanessa. Biensûr, l'équipe de la laverie indispensable à notre travail au laboratoire, surtout quand on prépare nos milieux levures au dernier moment et qu'ils doivent « à tout prix » partir à l'autoclave dans la minute...

Une thèse ça n'est biensûr pas que des manip, et voilà le temps des remerciements aux collègues dont certain sont devenus des amis. La première est biensûr mon éternelle collègue de bureau, collègue de travail puis copine et enfin amie, Erwana Harscoët qui a rendu la vie au labo bien plus rigolote et dont le départ a participé au mien! Les autres gais lurons comme Alex, Toto, Lolo, Sébounet, Kianou, Roulta, Ciboulette, Cabelo qui ont permis d'égayer la vie au labo et les jeudi soir au O'Paris et les super copines qui resteront dans mes tablettes Cricri, Clacla, Carinette, Sandra, Virginie, Aurélia. Les copines de pause thè ne seront pas oubliées bien que certaines ne puissent pas assister à la soutenance (elle se reconnaitra) j'ai nommé, Alex biensûr mais aussi Helen, Annie, zaz et plus récemment Elodie, n'oublions pas Lise ou Herman quand il y avait des gâteaux, hi !hi !

Enfin un grand merci à Christian et à ma famille pour leur soutien sans faille.



<b>I</b>	<b>INTRODUCTION.....</b>	<b>1</b>
<b>I.1</b>	<b>Le développement de la graine et les gènes AFL.....</b>	<b>2</b>
I.1.1	Le développement de la graine.....	2
I.1.2	Les mutants des gènes <i>AFL</i> .....	4
I.1.3	Facteurs de transcription et initiation de la transcription.....	6
<b>I.2</b>	<b>Description des protéines AFL et de leur fonction .....</b>	<b>7</b>
I.2.1	LEC1 et LEC1L.....	7
I.2.2	Structure protéique des AFL de type B3.....	9
I.2.3	Evolution des AFL.....	10
I.2.4	Profils d'expression des gènes AFL chez <i>Arabidopsis</i> .....	10
I.2.5	Les AFL : mécanismes moléculaires de la régulation de l'accumulation des réserves de la graine d' <i>Arabidopsis</i> .....	11
<b>I.3</b>	<b>Régulation des gènes AFL .....</b>	<b>14</b>
I.3.1	Boucles de régulation sur <i>ABI3</i> et <i>FUS3</i> .....	14
I.3.2	AGL15, AGL18 : uniques activateurs de l'expression de <i>LEC2</i> .....	15
I.3.3	Répresseurs de type facteurs de transcription.....	17
I.3.4	Répresseurs transcriptionnels de type chromatinien .....	18
<b>I.4</b>	<b>Objectifs de la thèse .....</b>	<b>30</b>
<b>II</b>	<b>RESULTATS.....</b>	<b>32</b>
<b>II.1</b>	<b>Etude fonctionnelle du promoteur de LEC2 .....</b>	<b>32</b>
II.1.1	La régulation transcriptionnelle de l'expression de <i>LEC2</i> implique RLE, un élément cis-régulateur qui conditionne la déposition de la marque H3K27me3.....	32
II.1.2	Résultats complémentaires : relations d'épistasie entre les trois éléments cis-régulateurs isolés.....	33
<b>II.2</b>	<b>Facteurs positifs, données supplémentaires.....</b>	<b>34</b>
II.2.1	AGL15, AGL18 et le promoteur de <i>LEC2</i> .....	34
II.2.2	Les BPC .....	35
<b>II.3</b>	<b>Facteurs négatifs, données supplémentaires .....</b>	<b>48</b>
II.3.1	PICKLE et le promoteur de <i>LEC2</i> .....	48
II.3.2	HDAC.....	49
<b>II.4</b>	<b>L'activité de LEC2 est elle régulée post-traductionnellement ?.....</b>	<b>50</b>
II.4.1	La sumoylation.....	50
II.4.2	LEC2 et la sumoylation .....	51
<b>II.5</b>	<b>Outils pour la détection de la protéine LEC2.....</b>	<b>52</b>
II.5.1	Anticorps anti-LEC2.....	52
II.5.2	Obtention d'une protéine étiquetée fonctionnelle .....	52
<b>III</b>	<b>DISCUSSION GENERALE/CONCLUSION .....</b>	<b>54</b>
<b>III.1</b>	<b>Avancées dans la régulation de LEC2 :.....</b>	<b>54</b>
<b>III.2</b>	<b>LEC2 est vraiment un très bon modèle pour l'étude de systèmes complexes de régulation .....</b>	<b>56</b>
<b>III.3</b>	<b>Régulation de LEC2 et des AFLs pendant le développement de la graine et projet CERES57</b>	
<b>III.4</b>	<b>Critique de la démarche de délétion de promoteurs et alternatives .....</b>	<b>58</b>
<b>IV</b>	<b>MATERIEL ET METHODES .....</b>	<b>60</b>
<b>IV.1</b>	<b>Matériel .....</b>	<b>60</b>
IV.1.1	Mutants d' <i>Arabidopsis thaliana</i> .....	60
IV.1.2	Souches de levures et de bactéries.....	60
<b>IV.2</b>	<b>Méthodes.....</b>	<b>61</b>
IV.2.1	Stérilisation des graines d' <i>Arabidopsis</i> et conditions de culture.....	61
IV.2.2	Transformation génétique d' <i>Arabidopsis thaliana</i> .....	61
IV.2.3	Constructions génétiques.....	62
IV.2.4	Colorations histochimiques GUS.....	62



IV.2.5	Extraction d'acides nucléiques .....	62
IV.2.6	PCR en temps réel ou PCR quantitative .....	64
IV.2.7	Immuno-précipitation de la chromatine (Le Page et al).....	65
IV.2.8	Manipulations de simple hybride .....	69
<b>V.</b>	<b>BIBLIOGRAPHIE .....</b>	<b>80</b>





## Table des illustrations :

Figure 1 : La graine d' <i>Arabidopsis thaliana</i> .....	2
Figure 2 : Développement de la graine et phénotypes des mutants <i>afls</i> .....	3
Figure 3 : Phénotypes du mutant <i>lec2</i> .....	4
Figure 4 : Mécanisme de base de la transcription. ....	6
Figure 5 : Structure des protéines AFLs.....	8
Figure 6 : Evolution des 3 AFL de type B3.....	10
Figure 7 : Profils d'expression des gènes <i>AFL</i> .....	10
Figure 8 : Fonctions des protéines AFL dans le développement de la plante....	11
Figure 9 : Modèle de recrutement à l'ADN pour les protéines AFL.....	12
Figure 10 : Régulation des gènes AFL dans la graine en développement.....	15
Tableau 1 : Protéines à domaine EAR et leur implication dans le développement d' <i>Arabidopsis</i> .....	16
Figure 11 : Modèle d'action des protéines à domaine EAR.....	16
Figure 12 : Modèle de recrutement d'ASIL1.....	17
Figure 13 : Régulation des gènes AFL dans la plantule en germination.....	18
Figure 14 : Structure de la chromatine.....	19
Tableau 2 : Exemple de code multifactoriel trouvé dans la chromatine, décrit par Steensel et al., 2011.....	20
Tableau 3 : Protéines impliquées dans la méthylation des histones chez <i>Arabidopsis thaliana</i> .....	23
Figure 15 : Cibles des complexes PRC2.....	26
Figure 16 : Analyse par la méthode de simple hybride de l'interaction entre les protéines AGL15 et AGL18 avec le promoteur de <i>LEC2</i> .....	34
Figure 17 : Expression de <i>BPC3</i> et <i>LEC2</i> dans des graines à 10 JAF.....	46
Figure 18 : Expression de <i>LEC2</i> dans le mutant <i>bpc1bpc2bpc3bpc4bpc6</i> .....	47
Figure 19 : Profils d'expression des promoteurs de la classe I des BPC.....	48
Figure 20 : Activité de PLEC2 en conditions de dérépression.....	49
Figure 21 : Effet de la mutation des sites potentiels de sumoylation sur la protéine <i>LEC2</i> .....	51
Tableau 5 : Construits réalisés pour la production d'une protéine <i>LEC2</i> étiquetée.....	52



### Liste des abréviations :

ABA	Abscisic Acid
ABI3	ABA Insensitive 3
ABRE	ABA Responsive Element
ADN	Acide DesoxyriboNucléique
AFL	ABI3-FUS3-LECs
AG	AGAMOUS
AP2	APETALA 2
ARF	Auxin Response Factor
ARN	Acide RiboNucléique
ARNpol	ARN polymérase
ASH1	ABSENT, SMALL, or HOMEOTIC DISCS 1
AtGAox2	Gibberellin 2-Oxidase
ATXR5,6	Arabidopsis Trithorax Related protein
BRM	BRAHMA
CBF	CCAAT Binding Factor
Chd	Chromodomain helicase DNA-binding protein 1
ChIP	Chromatin ImmunoPrecipitation
EAR	ERF-associated amphiphilic repression motif
EREBP	Ethylene-Responsive Element Binding Proteins
ERF	Ethylene-Responsive factors
E(Z)	ENHANCER OF ZESTE
ES	Embryogenèse Somatique
FLC	FLOWERING LOCUS C
FLD	FLOWERING LOCUS D
FIS2	FERTILIZATION-INDEPENDENT SEED 2
FUS3	FUSCA 3
GA	Gibberellic Acid
HAT	Histone Acetyl Transferase
HDA19	Histone deacetylase 9
HDAC	Histone DeACetylase
HSI	High level expression Sucrose Inducible gene
JAF	Jours Après Fécondation
JMJ	JumonjiC domain containing protein
INO80	Inositol auxotroph 80
iSWI	Imitation switch
LEA	Late Embryogenesis Abundant
LEC1	LEAFY COTYLEDON 1
LEC2	LEAFY COTYLEDON 2
LSD1	Lysine specific demethylase
LSDL1, 2	<i>LSD1-LIKE 1, 2</i>
MADS	MCM1 ( <i>Saccharomyces cerevisiae</i> ) AGAMOUS ( <i>Arabidopsis thaliana</i> ) DEFICIENS ( <i>Antirrhinum majus</i> ) SRF ( <i>homo sapiens</i> )
MET1	METHYTRANSFERASE 1
Mi-2	Dermatomyositis-specific Mi-2 autoantigen
MSI1	Multicopy suppressor of IRA1



NuRD	Nucleosome Remodelling and histone Deacetylation
NaBut	Butyrate de sodium
NAD	Nicotinamide Adenine Dinucleotide
PIC	PreInitiation Complex
PKL	PICKLE
PWWP	Pro-Trp-Trp-Pro motif
RAV	Related to ABI3/VP1
RBR	RetinoBlastoma Related protein
REM	Reproductive Meristem
RPD3	Reduced Potassium Deficiency 3
SA	Salicylic Acid
SAP18	Sin3A-Associated Protein 18
SET	Su(var)3-9,E(Z) and Trithorax
SU(VAR)	Suppressor of Variegation
SUVH	SU(VAR)3-9
SUVR	SU(VAR)3-9 related proteins
Swi2/Snf2	Switch 2/sucrose non-fermenting 2
TAFs	TBP Associated Factors
TBP	TATA Binding Protein
TSA	TrichoStatine A
TSS	Transcription Start
VRN1	VERNALISATION 1



## I INTRODUCTION

Le sujet de cette thèse est l'étude de la régulation de *LEAFY COTYLEDON 2*, gène régulateur clé du développement de la graine chez *Arabidopsis thaliana*. La graine est un organe crucial pour la plante car elle permet la dissémination de l'espèce et la résistance de l'embryon à des conditions environnementales défavorables. La graine est également l'élément de base de l'agriculture car l'efficacité de la levée des plantes cultivées est principalement déterminée par la qualité de la graine. De plus, l'alimentation humaine est largement basée sur les graines entières comme le maïs ou le riz ou sur les produits extraits de ces graines comme la farine et les huiles végétales entre autres. Enfin, la graine est à la base de nombreuses utilisations industrielles. Dans un monde où la population risque d'augmenter de 3,5 milliards d'ici 50 ans (Borlaug, 2007), la question de l'alimentation humaine et donc de l'agriculture en général est un point crucial nécessitant une recherche approfondie et l'étude du développement et du remplissage du grain est d'une grande importance cognitive, agronomique et industrielle.

La majorité des études menées lors de cette thèse concernent la régulation transcriptionnelle de l'expression du facteur de transcription *LEC2*, mais quelques pistes ont été abordées quant à sa régulation post-traductionnelle. Les facteurs de transcription représentent environ 5% des gènes chez *Arabidopsis* et jouent un rôle majeur dans l'expression coordonnée des gènes concourant au développement de la plante. Des changements dans les séquences régulatrices de ces facteurs, mais également dans les séquences codantes sont des systèmes très puissants permettant la sélection naturelle mais également la sélection de nombreuses espèces par l'homme, en représentant 70 % des gènes majeurs de domestication (Chen et al, 2005; Doebley et al, 2006; Martin et al, 2010). *LEC2* agit de façon concomitante et parfois redondante avec les gènes que l'on appelle AFL pour *ABI3* (*ABA insensitive 3*), *FUS3* (*FUSCA 3*), *LEC1* (*LEAFY COTYLEDON 1*) et *LEC2* (*LEAFY COTYLEDON 2*). Durant le développement de la graine, l'expression coordonnée des gènes AFL permet une accumulation de réserves nutritives dans l'embryon puis un passage de celui-ci à un état quiescent. Cet état lui permettra d'attendre des conditions pour germer si l'environnement, comme les conditions optimales de températures ou d'humidité, lui est hostile au moment de la dispersion de la graine. *LEC2* et *LEC1* sont exprimés plus tôt que *FUS3* et *ABI3* et sont considérés comme les agents déclenchant tous les phénomènes qui suivront jusqu'à la fin du développement de la graine.





Outre leur rôle de régulateurs majeurs de la maturation de la graine, *LEC2* et *LEC1* ont un rôle très important dans l'embryogenèse somatique (ES) (Lotan et al, 1998; Stone et al, 2001). En effet, le mutant *lec2* voit sa capacité à produire des embryons somatiques drastiquement diminuée, par rapport à des plantes de type sauvage. De plus, une expression ectopique de *LEC1* ou *LEC2* dans les parties végétatives permet d'induire l'ES. Cette caractéristique souligne l'importance d'une régulation stricte de l'expression de *LEC1* et *LEC2* car une embryogenèse déclenchée pendant le développement de la plante peut altérer sa croissance. D'un autre côté, le déclenchement de l'ES est aussi un challenge biotechnologique de poids pour améliorer la multiplication végétative de nombreuses espèces.

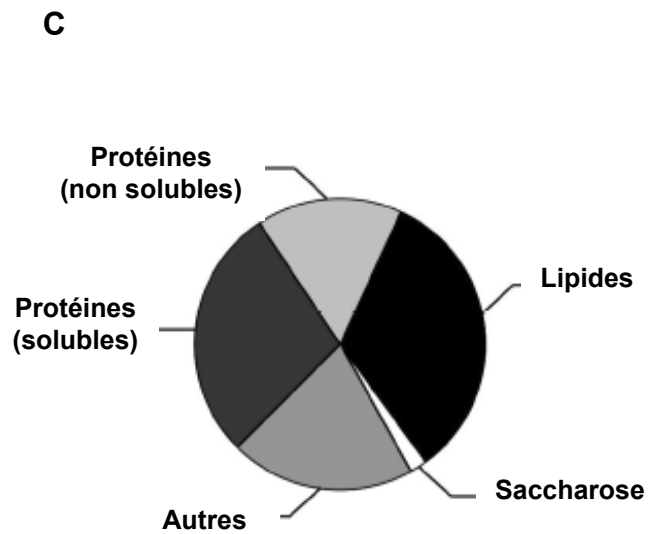
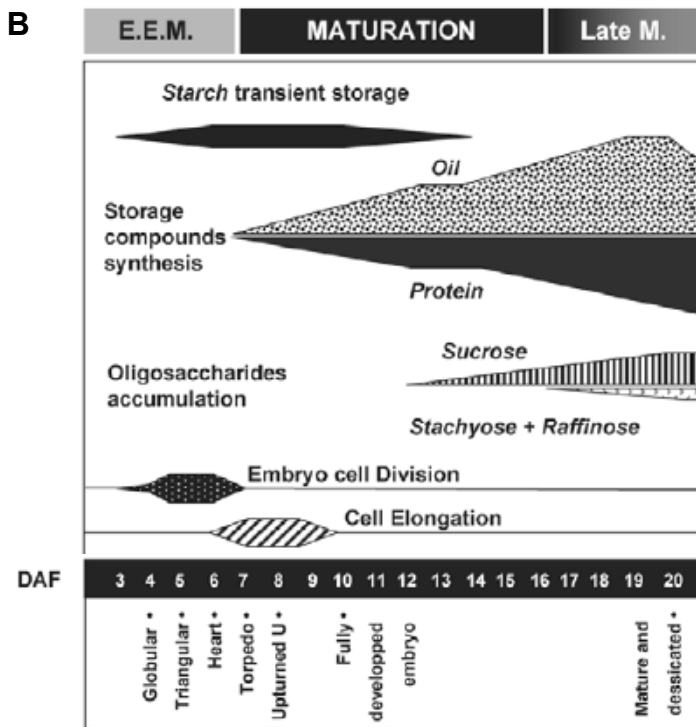
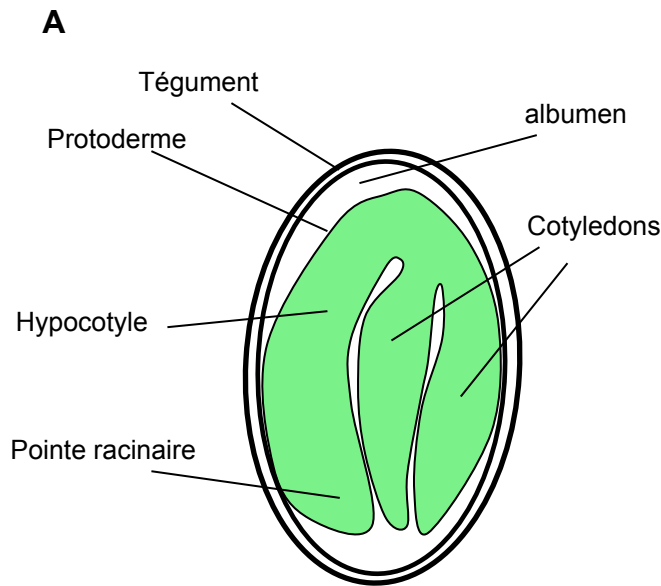
Enfin, l'expression des quatre gènes *AFL* est très strictement contrôlée au niveau spatial et temporel. Ils fournissent donc un excellent modèle pour l'étude de la régulation de l'expression des gènes. Pour toutes ces raisons, la régulation de l'expression des gènes *AFL* est étudiée par plusieurs équipes mais, malgré cela, les mécanismes moléculaires impliqués sont peu connus.

Après un bref rappel sur le développement de la graine d'*Arabidopsis*, ce chapitre s'attachera donc à décrire les mutants nuls de ces gènes, les fonctions des protéines correspondantes ainsi que ce qui est connu de leur régulation transcriptionnelle, grâce aux nombreux travaux réalisés ces dernières années. Ceci nous amènera à aborder la régulation de la transcription des gènes en allant des facteurs de transcription dits « classiques » à la régulation et l'organisation chromatinienne aux loci considérés.

## ***1.1 Le développement de la graine et les gènes AFL***

### **I.1.1 Le développement de la graine**

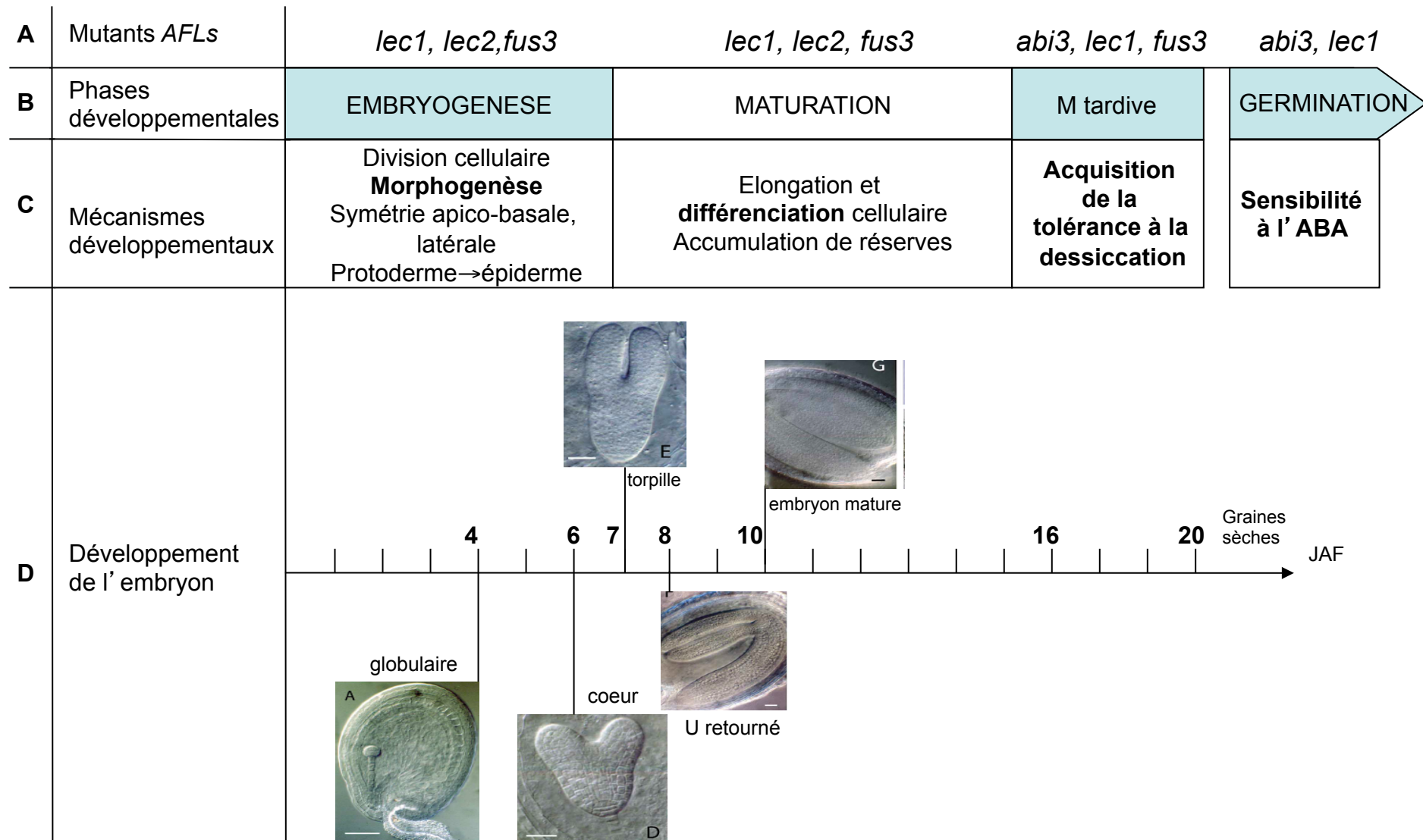
La graine est un organe très efficace pour la propagation d'une espèce végétale. Elle constitue une structure résistante aux stress biotiques et abiotiques, tout en fournissant à la jeune plantule l'énergie pour germer et se développer jusqu'à son indépendance énergétique vis-à-vis de ses réserves, c'est-à-dire jusqu'à ce qu'elle puisse réaliser sa propre photosynthèse (North et al, 2010; Vicente-Carbajosa & Carbonero, 2005). La graine est constituée d'un embryon entouré d'un albumen, plus ou moins persistant en fin de développement selon les espèces, et d'un tégument protecteur (Figure 1A). L'une ou l'autre de ces parties peuvent être utilisées par l'Homme et font de la graine un fort enjeu économique que ce soit dans l'alimentation, l'industrie ou la production d'énergie.



**Figure 1: La graine d' *Arabidopsis thaliana*.** (A) Structure de la graine en fin de développement. L'embryon qui est représenté en vert, avec son épiderme non différencié ou protoderme, occupe tout l'espace de la graine et est entouré d'une couche résiduelle d'albumen et du tégument qui vient de la plante mère. (B) Accumulation des réserves au cours du développement de la graine d' *Arabidopsis thaliana* (extrait de Baud et al., 2002, Plant Physiol. Biochem. 40 (2002) 151–160). (C) Composition globale de la graine mature (extrait de Baud et al., 2008, The Arabidopsis Book 6: e0113. 2008) .

Chez les plantes à graines ou angiospermes, la reproduction se fait par le mécanisme de la double fécondation entre un grain de pollen contenant deux noyaux spermatiques haploïdes (n) et l'ovule. L'un des noyaux spermatiques du grain de pollen (n) fusionne avec l'oosphère (n) pour former l'embryon ou zygote puis la future plantule. L'autre noyau spermatique (n) fusionne avec les deux noyaux de la cellule centrale de l'ovule pour donner un tissu triploïde appelé albumen qui servira à nourrir l'embryon. Au cours du développement, l'embryon accumule parfois transitoirement des réserves de type protéiques, lipidiques ou polysaccharidiques (Figure 1B). En fin de développement, les réserves de l'embryon sont globalement constituées de lipides et de protéines de réserves à parts égales (Figure 1C) (Baud et al, 2002) .

A partir de la fécondation, le développement de la graine peut être divisé en trois parties (Figure 2). Tout d'abord l'embryogenèse, qui débute juste après la fécondation et pendant laquelle ont lieu les divisions et différenciations cellulaires menant à la mise en place de la symétrie bilatérale de l'embryon (Jenik et al, 2007; Laux et al, 2004; Weijers & Jurgens, 2005). Un rôle primordial de l'auxine a été montré pour cette phase (Jenik & Barton, 2005; Weijers & Jurgens, 2005). La deuxième phase est la maturation de la graine qui consiste principalement en un arrêt de la croissance couplé à l'accumulation de réserves (Baud et al, 2002) stockées dans des structures appelées corps lipidiques et protéiques. La troisième étape, appelée maturation tardive, permet l'acquisition de la tolérance à la dessiccation puis l'entrée en dormance dans l'attente de conditions favorables à la germination (Angelovici et al, 2010). Il est important de noter l'implication des deux hormones ABA et GA. Un rapport ABA/GA important oriente le développement vers l'embryon puis vers la dormance alors qu'un rapport faible oriente vers la plantule puis vers la germination (Finkelstein et al, 2008; Koornneef et al, 2002; Koornneef et al, 1989; Ogawa et al, 2003). Les étapes de maturation ne sont pas indispensables au développement de la future plantule dans des conditions de croissance idéales, en effet, des embryons extraits de l'enveloppe de la graine juste après l'embryogenèse peuvent donner des plantules normalement développées (Vicente-Carbajosa & Carbonero, 2005). La maturation peut donc être considérée comme l'introduction d'une nouvelle étape durant l'évolution, qui permet d'améliorer la résistance de l'organe de dissémination et de multiplication de l'espèce qu'est la graine. Ceci est soutenu par le fait que certaines espèces ont des plantules qui se développent dans la nature sans l'étape de maturation des graines. Dans cette introduction, nous nous attacherons particulièrement à la description de l'étape de maturation et plus particulièrement aux quatre acteurs majeurs régulant ce mécanisme que



**Figure 2: Développement de la graine et phénotypes des mutants *AFL*.** Les mécanismes (C) ayant lieu durant les 4 étapes majeures du développement de la graine (B) sont décrits avec, en gras, les mécanismes affectés dans les mutants *AFL* mentionnés (A). Le développement embryonnaire sur une échelle de temps de 0 à 20 jours après fécondation (JAF) est représenté (D) (Photos de Baud et al., 2002).

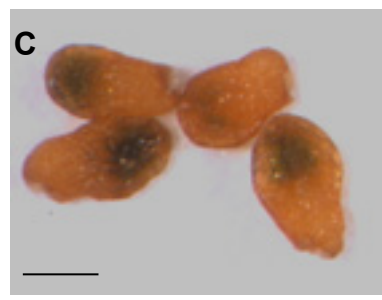
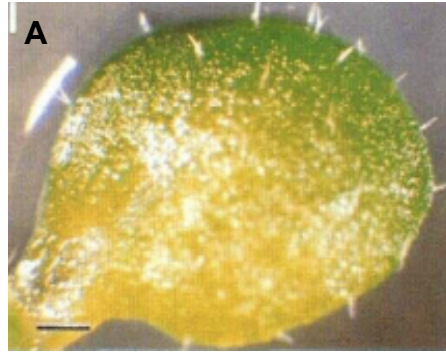
sont les gènes *ABI3*, *FUS3*, *LEC1* et *LEC2* appelés *AFL* (*ABI3-FUS3-LECs*) (pour revues : Gutierrez et al., 2007; Santos-Mendoza et al., 2008).

### **I.1.2 Les mutants des gènes *AFL***

Le développement de la graine d'*Arabidopsis* est étudié depuis plus de 50 ans et dans ce but, différentes équipes ont criblé des banques de mutants sur la base du phénotype de leurs graines. Un des types de mutants les plus faciles à isoler sous une simple loupe binoculaire a été le type *leafy cotyledon* (*lec*). Comme son nom l'indique, ce type de mutants possède des cotylédons présentant des caractéristiques de feuilles, comme la présence de trichomes (Figure 3A), mais, surtout, présente une coloration pourpre des cotylédons due à une accumulation d'anthocyanes, visible au travers des téguments de la graine sèche et facilitant l'identification (Meinke et al, 1994) (Figures 3B, 3C). Ces phénotypes visuels des mutants *lec* ont ensuite été étudiés de façon approfondie grâce à des analyses microscopiques et physiologiques plus fines qui ont permis d'appréhender l'étendue des mécanismes physiologiques touchés (Figure 2) (Meinke et al, 1994). On peut noter également qu'un gène nommé *TANMEI* (*TAN*) paraît avoir des fonctions chevauchantes avec *LEC1* et *LEC2* à un stade précoce de la maturation car le phénotype des simples mutants est similaire et les doubles mutants *tan lec1* et *tan lec2* ont des phénotypes similaires à un mutant *lec1 lec2* (Yamagishi et al, 2005), mais il ne fera pas partie de l'étude.

#### ***I.1.2.1 Le mutant *lec1****

Le mutant *lec1* a le phénotype anormal le plus fort (Lotan et al, 1998; Meinke, 1992; West et al, 1994). Comme tous les mutants *lec*, les cellules de ses cotylédons ressemblent à des cellules de feuilles avec un système vasculaire qui peut être déjà différencié au stade xylème mature contrairement au procambium non différencié que l'on trouve en observant une graine de type sauvage au même âge. Les caractéristiques de feuilles s'expriment également par la présence de stomates, qui n'apparaissent à la surface des cotylédons qu'à la fin de la germination dans une plante dont le gène *LEC1* est intact. On peut noter également un élargissement du méristème apical ainsi qu'une viviparité associée à une diminution quasi-complète de la tolérance à la dessiccation. Les réserves sont également affectées, en effet, les



**Figure 3: Phénotypes du mutant *lec2*.** (A) Cotyledon (extrait de Meinke et al., 1994) (Echelle=300 $\mu$ m), graines sèches de plantes sauvages Col0 (B) et de plantes mutantes *lec2-4* (C)

corps protéiques sont quasiment absents des cotylédons comme les corps lipidiques que l'on peut par contre encore trouver en petit nombre dans l'hypocotyle (Meinke et al, 1994).

### ***1.1.2.2 Le mutant fus3***

Les phénotypes au niveau de l'accumulation des réserves du mutant *fus3* sont similaires à ceux de *lec1* mais plus faibles (Baumlein et al, 1994; Gazzarrini et al, 2004; Keith et al, 1994; Kirik et al, 1996; Luerksen et al, 1998). Les cotylédons de *fus3* montrent également des caractéristiques foliaires via l'accumulation d'acide gibbérélique qui est responsable de l'apparition des trichomes, caractéristiques d'un embryon en phase de germination (Curaba et al, 2004).

### ***1.1.2.3 Le mutant lec2***

Les phénotypes de *lec2* sont assez différents et également plus faibles que ceux de *lec1* (Meinke et al, 1994; Stone et al, 2001). *lec2* partage avec les autres *lec* des cotylédons ressemblant à des feuilles (présence de trichomes, stomates, et système vasculaire différencié). L'accumulation de réserves dans *lec2* est très dissymétrique, le profil d'accumulation ressemble à celui d'une plante sauvage dans l'hypocotyle et à celui de *lec1* dans les cotylédons (Meinke et al, 1994).

### ***1.1.2.4 Le mutant abi3***

*abi3*, qui a été isolé au départ pour son insensibilité à l'ABA lors de la germination, fait partie des gènes AFL mais ne partage pas les caractéristiques foliaires des cotylédons et est touché plus tardivement (Giraudat et al, 1992; Koornneef et al, 1989; Koornneef et al, 1984; Nambara et al, 2000). Il est donc moins affecté dans l'accumulation en composants de réserve, il contient 2/3 des protéines de réserve et 1/3 des acides gras par rapport à une graine de type sauvage et est par contre totalement intolérant à la dessiccation (Finkelstein & Somerville, 1990; Haslekas et al, 1998; Meinke, 1992; Wehmeyer & Vierling, 2000).





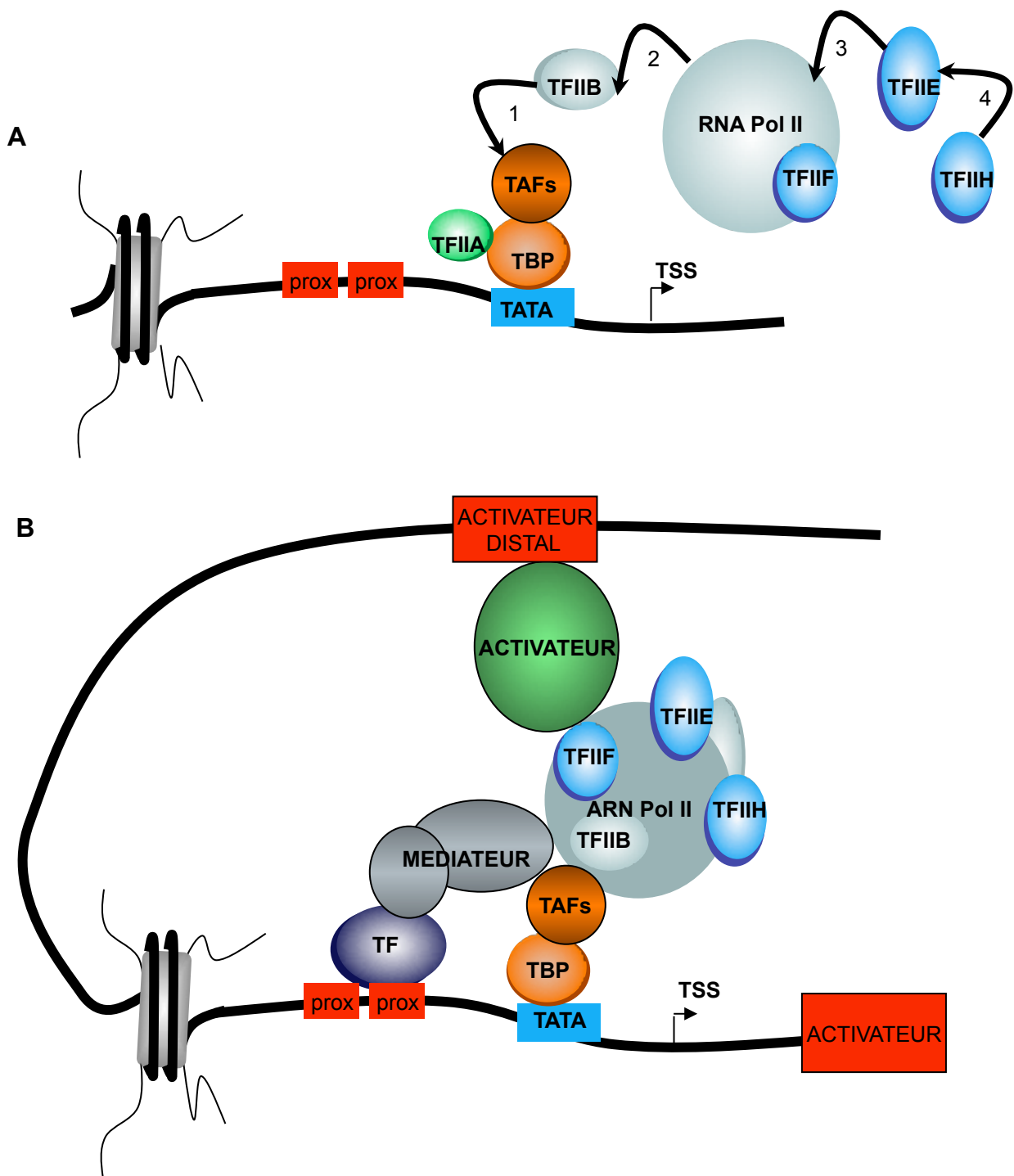
En résumé, les mutants affectés dans les gènes *AFL* ont comme points majeurs en commun un défaut dans l'accumulation des réserves et une identité cellulaire modifiée qui évoquent une transition directe entre l'embryogénèse et la germination sans passer par l'étape de maturation. Il faut noter qu'aucun de ces mutants n'est létal. Les gènes affectés dans ces mutants ont pu être clonés et correspondent tous les quatre à des facteurs de transcription.

### **I.1.3 Facteurs de transcription et initiation de la transcription**

Les facteurs de transcription sont en général caractérisés par un signal d'adressage nucléaire, un site de fixation à l'ADN et un site d'activation ou de répression de la transcription (Wray et al, 2003). Chez *Arabidopsis*, environ 2000 gènes codent pour des facteurs de transcription (Riechmann & Ratcliffe, 2000) et il apparaît que le nombre de membres dans certaines familles est beaucoup plus important chez les plantes que dans les autres règnes, ce qui laisserait penser à une spécialisation de ces facteurs dans des fonctions végétales (Dubos et al, 2010; Shiu et al, 2005). Ces protéines participent à la régulation de la transcription des gènes qui est un des principaux mécanismes permettant la vie et le développement des organismes.

La transcription est un mécanisme très complexe qui consiste en un passage d'un ADN double-brin composé d'acides désoxyribonucléiques à un ARN simple brin constitué d'acides ribonucléiques grâce à l'ARN polymérase (ARNpol). Dans le cadre de ce travail, l'initiation de la transcription qui est le lieu de nombreuses régulations, sera plus particulièrement étudiée. Les gènes ou unités de transcription sont souvent divisés en trois classes selon le type de transcrits pour lequel ils codent et selon l'ARNpol impliquée : la classe I est composée des ARN ribosomaux et est transcrite par l'ARNpolII. L'ARNpolIII transcrit quasi-exclusivement les ARNm qui constituent les gènes de classe II et l'ARNpolIII synthétise les ARN de transfert, l'ARN ribosomique 5S et d'autres petits ARN qui composent la classe III. Dans la suite du manuscrit nous ne considérerons que les gènes de la classe II et le fonctionnement de l'ARNpolIII.

Les TATA binding protéines (TBP) se lient à la TATA box présente en règle général dans les gènes de type II, dans les 100pb en amont de l'ATG et recrutent alors des TAFs (TBP associated factors) pour former le complexe TFIID (Figure 4). Ce complexe TBP se lie ensuite à TFIIB qui recrute à son tour l'ARNpolIII assisté par TFIIIF qui permet la fixation de TFIIE qui s'associe à TFIIH dont l'activité hélicase permet l'ouverture de la double hélice d'ADN (Figure 4A) (Roeder, 1996). Ainsi se met en place le complexe de pré-initiation



**Figure 4: Mécanisme de base de la transcription.** (A) Les TATA binding protein (TBP) se lient à la TATAbox (TATA) et recrutent les TBP associated factor (TAF) qui permettent l'assemblage du complexe de pré-initiation contenant l'ARNpolIII. (B) Les facteurs de transcription (TF), activateurs ou represseurs et avec ou sans médiateur peuvent alors se fixer aux éléments cis-régulateurs proximaux (prox) ou distaux (distal) pour permettre le démarrage de la transcription ou modifier le niveau de transcription.

(PIC), constitué de 44 polypeptides qui suffisent à assurer un niveau de base de transcription très faible et qui doit être ensuite renforcé ou réprimé par des facteurs de transcription ou des médiateurs (Roeder, 2005). Ces facteurs de transcription se fixent à l'ADN par l'intermédiaire de séquences activatrices ou répressives présentes dans la zone promotrice du gène considéré (Figure 4B) que l'on appelle éléments cis-régulateurs. Chez les animaux, des séquences de ce type peuvent être trouvées jusqu'à 100kb de la zone à transcrire, comme pour les deux clusters majeurs des gènes Hox chez la drosophile (Gorman & Kaufman, 1995; Martin et al, 1995). Chez les plantes, les régions nécessaires à une expression correcte sont généralement localisées dans une zone entre l'ATG et trois kilobases en amont de celui-ci, comme le prouvent la plupart des analyses de promoteurs fusionnés à des gènes rapporteurs chez *Arabidopsis thaliana* (Hartmann et al, 2005; Takada & Jurgens, 2007). On trouve également dans la littérature, nombre d'exemples d'éléments de régulation situés dans un intron ou dans la partie 3' UTR du gène, comme pour les gènes *FLOWERING LOCUS C (FLC)* ou *AGAMOUS (AG)* (Sheldon et al, 2002; Sieburth & Meyerowitz, 1997). D'autres éléments, appelés paramutations, montrent l'influence de séquences se trouvant parfois à 100kb du gène régulé chez le maïs et chez *Arabidopsis*, impliquant l'intervention de SiRNA (Small interfering RNA) (pour revue : (Arteaga-Vazquez & Chandler, 2010; Erhard & Hollick, 2011).

La régulation de la transcription des gènes *AFL* semble quant à elle, déterminée entièrement par les séquences contenues dans les 3kb en amont de l'ATG. En effet, l'expression d'un gène rapporteur sous contrôle de ces zones promotrices suffit à reproduire le profil des ARNm et à compléter les mutations (sauf pour *LEC1* dont le promoteur n'a jamais été cloné à notre connaissance). Le clonage des gènes affectés dans les mutants *AFL*, codant pour des facteurs de transcription a permis l'étude fonctionnelle de ces protéines et la détermination de leurs cibles, bien qu'il subsiste des zones d'ombre quant à leurs cibles ou leur mode d'action.

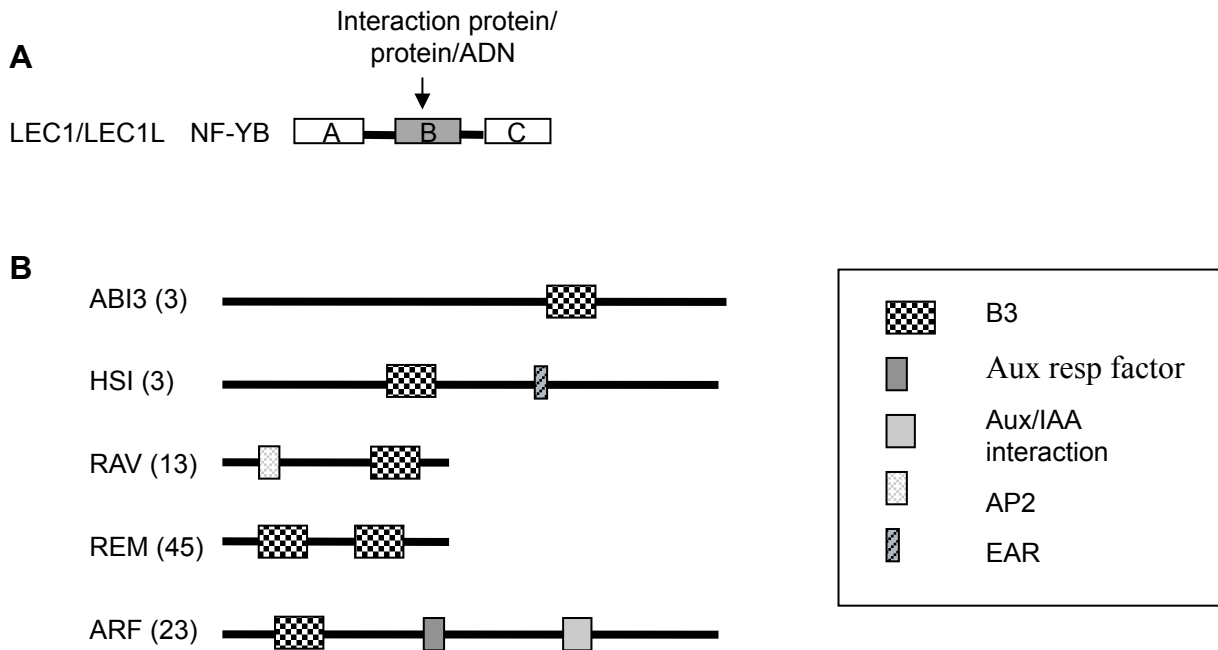
## ***1.2 Description des protéines AFL et de leur fonction***

### **1.2.1 LEC1 et LEC1L**

LEC1 présente de fortes homologues avec les sous-unités HAP3 de facteurs de levures se liant à des séquences CCAAT appelés CCAAT binding factor (CBF) et connus également sous le nom de NF-Y (Lotan et al, 1998). Ces structures oligomériques sont constituées de trois sous-unités différentes chez les mammifères (HAP2 ou CBF-B ou NF-YA, HAP3 ou CBF-A



ou NF-YB et HAP5 ou CBF-C ou NF-YC) et de quatre chez la levure (avec HAP4 en plus). Ces trois ou quatre sous-unités sont indispensables à la fixation du complexe sur des séquences CCAAT, fixation qui se fera par l'intermédiaire de la sous-unité HAP2 (Kim et al, 1996; Mantovani, 1999; McNabb et al, 1995; Sinha et al, 1996; Sinha et al, 1995). Plusieurs indices concordants quant à l'existence de ce complexe chez les plantes ont été révélés, comme l'identification d'une activité de fixation aux séquences CCAAT dans des extraits de noyaux de plantes (Gusmaroli et al, 2001) sans que le complexe ait pu être purifié et, à fortiori, sans que LEC1 ait pu y être associé. De plus, l'identification d'homologues de HAP2, HAP3 et HAP5 chez le maïs, *Medicago truncatula* et *Arabidopsis* (Albani & Robert, 1995; Combiér et al, 2006; Li et al, 1992; Newman et al, 1994) permet également de supposer une similarité fonctionnelle entre les protéines de plantes et de mammifères. La protéine LEC1 n'a jamais été montrée comme se liant directement à l'ADN seule, comme son homologue chez les mammifères et est composée de trois domaines : le domaine A, le domaine B qui permet des interactions avec les autres HAP et l'ADN chez les animaux et la levure et qui est extrêmement conservé chez les eucaryotes, et un domaine C (Figure 5A) (Li et al, 1992). Chez *Arabidopsis*, 10 homologues au NF-YB de levure ont été identifiés dont LEC1 et LEC1L qui sont similaires et fonctionnellement redondants. En effet, LEC1L complémente le mutant *lec1* lorsqu'on le place sous le contrôle du promoteur de *LEC1* et LEC1L peut induire, comme LEC1, la synthèse de lipides lorsqu'il est exprimé dans des parties végétatives (Tan et al, 2011). De plus, ils possèdent tous les deux un acide aminé dans le domaine B qui définit leur spécificité de fonction par rapport aux 8 autres NF-YB (Kwong et al, 2003; Lee et al, 2003). LEC1 et LEC1L se détachent des NF-YB des autres eucaryotes et des 8 autres NF-YB d'*Arabidopsis* lorsque l'on compare leurs séquences protéiques, ce qui laisse penser que les NF-YB de type LEC1 représentent un nouveau type de NF-YB, spécifique des végétaux. Comme d'autres facteurs NF-Y d'*Arabidopsis* (Liu and Howell 2010), LEC1 et LEC1L ont récemment été montrés se liant à des facteurs NF-YC comme chez les animaux et à un facteur liant les éléments cis-régulateurs ABRE appelé bZIP63. Cette interaction en présence d'ABA, permet la transcription de la CRUCIFERINE C et de la SUCROSE SYNTHASE 2 (SUS2) (Yamamoto et al, 2009). Il est à noter également que, chez le riz (*Oryza sativa*), OsNF-YB1, le facteur le plus proche d'AtLEC1, interagit avec OsMADS18 en double hybride mais le complexe ne semble pas se fixer aux CCAAT box (Masiero et al, 2002),



**Figure 5: Structure des protéines AFLs.** (A) Représentation schématique des protéines LEC1 et LEC1L (B) Représentation des 5 familles composant les facteurs de transcription de type B3. Le chiffre entre parenthèses correspond au nombre de représentants de la famille dans le génome d'*Arabidopsis*, d'après Romanel et al., 2009.

## **I.2.2 Structure protéique des AFL de type B3**

Les trois autres protéines AFL sont des facteurs de transcription de la famille des B3, famille qui est classiquement divisée en cinq sous-groupes (Figure 5B) (Romanel et al, 2009). Tout d'abord, le type ARF (Auxin Response Factor) est caractérisé par un domaine B3 N-terminal qui reconnaît les AuxREs (Auxin Response Element : TGTCTC), un domaine central qui semble avoir un rôle dans l'activation ou la répression de la transcription et, enfin, un domaine C-terminal qui permet la dimérisation, avec des motifs similaires aux protéines Aux/IAA (Ulmasov et al, 1997). Le deuxième groupe est appelé le type REM (REproductive Meristem). Il est composé de facteurs qui ont apparemment perdu leur capacité à se fixer à l'ADN de façon spécifique comme VRN1 (vernalisation 1) et contiennent deux domaines B3 (Franco-Zorrilla et al, 2002). La troisième classe de facteurs B3 est appelée RAV (Related to ABI3/VP1), le domaine B3 de RAV1 a été cristallisé et cela a permis de déterminer la présence de 7 feuillets  $\beta$  suivis de 2 hélices  $\alpha$  (Yamasaki et al, 2004). L'interface entre le domaine B3 et l'ADN a également pu être modélisée, permettant de souligner les acides aminés potentiellement importants pour la fixation aux éléments cis-régulateurs. Cette famille est caractérisée par un domaine AP2/EREBP qui reconnaît les séquences CAACA et un domaine B3 en position C-terminale qui reconnaît les boîtes CACCTG (Kagaya et al, 1999). Le groupe des HSI (High level expression of Sucrose Inducible gene) contient un domaine B3 en C-terminal et un domaine de répression de la transcription (Suzuki et al, 2007; Tsukagoshi et al, 2005). La dernière classe de protéines est le type ABI3, qui reconnaît les éléments Sph/RV (CATGCA) et dont font partie LEC2, FUS3 et ABI3. Comme on peut le voir dans les annexes 1 et 2, les facteurs B3 se classent selon les familles décrites ci-dessus, que l'on considère la séquence protéique complète ou uniquement le domaine B3. Les protéines de type HSI se retrouvent quant à elles proches des AFL avec leur séquence protéique totale ou proches des REM, en ne considérant que le domaine B3.

La protéine ABI3 est la plus divergente des B3/AFL et contient un domaine acide A1 en position N-terminale suivi de trois domaines basiques, comme son homologue VP1 chez le maïs : B1, B2 et B3. B1 et B2 sont impliqués dans des interactions avec d'autres protéines comme ABI5 (Finkelstein & Lynch, 2000) et, par là même, dans l'induction de cibles par le biais de séquences ABRE (Ezcurra et al, 2000a; Ezcurra et al, 2000b; Nakamura et al, 2001). La protéine FUS3, en plus du domaine B3, contient un domaine présentant des similarités avec le domaine B2 d'ABI3 (Luerssen et al, 1998). LEC2 contient un domaine B3 sans





domaine B1 ni B2 ou aucun autre domaine connu à l'heure actuelle. Aucune protéine n'a encore pu être associée physiquement à LEC2 et FUS3. Il est intéressant de noter que le domaine B3 d'ABI3 doit être produit seul pour que l'on observe une fixation à la séquence Sph/R<sub>Y</sub> *in vitro* ou en levure (Suzuki et al, 1997). On peut alors se demander si la protéine ABI3 native est capable de se fixer aux séquences B3 *in planta*.

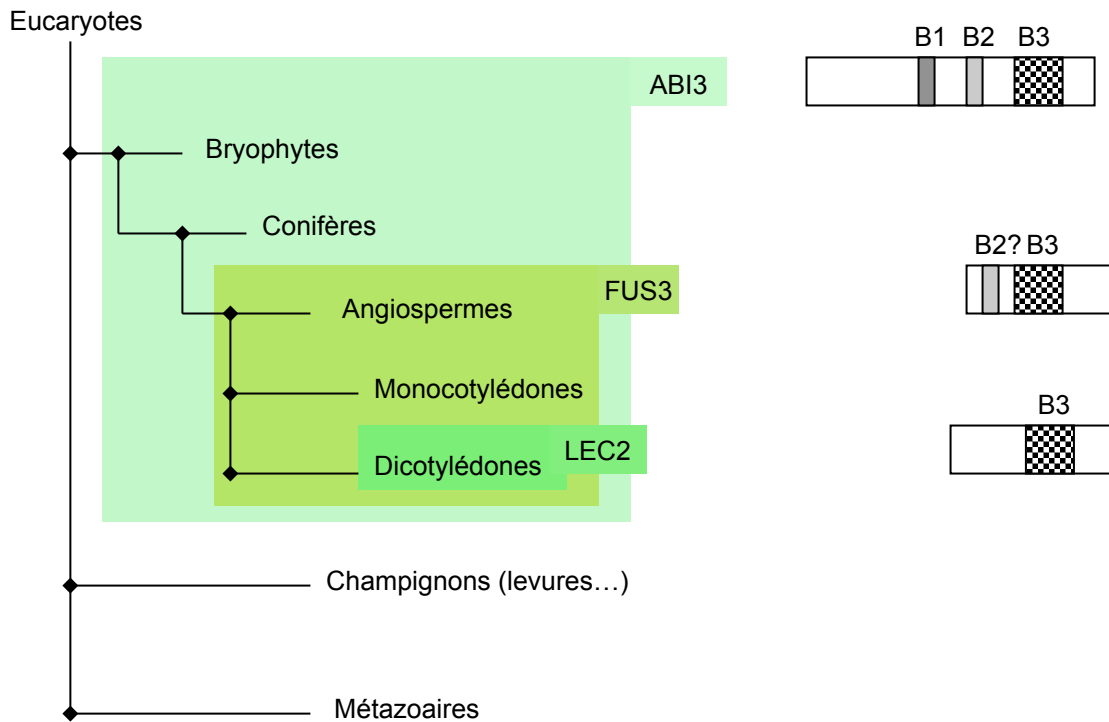
### **I.2.3 Evolution des AFL**

Des orthologues de LEC1 ont été identifiés dans tous les règnes (voir chapitre I.2.1) et, chez les plantes, des conifères aux angiospermes monocotylédones et dicotylédones (Uddenberg et al, 2011). La situation n'est pas si claire pour les AFL de type B3. *ABI3* est trouvé dans tout le règne végétal, de la mousse *Physcomitrella patens* (Sakata et al, 2010) à une monocotylédone comme le maïs avec VP1 (McCarty et al, 1991). *FUS3* semble être limité aux plantes à graines et les B3 de type *LEC2* ne sont retrouvés que chez les dicotylédones (Li et al, 2010). Toutes ces données tendent à prouver que l'ancêtre des AFL de type B3 est ABI3 et que sont apparus ensuite FUS3 et LEC2 chez les angiospermes (Figure 6).

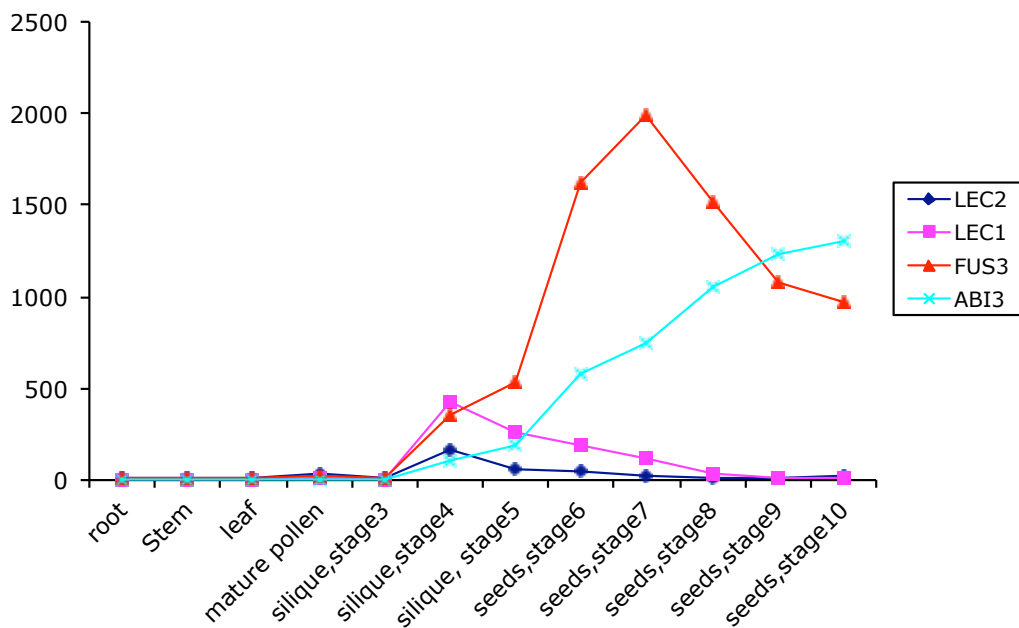
### **I.2.4 Profils d'expression des gènes AFL chez *Arabidopsis***

Excepté *ABI3*, les gènes *AFL* sont exprimés spécifiquement dans la graine en cours de développement. *LEC1* présente une expression transitoire de 4 à 14 jours après floraison avec un pic à 10 jours après floraison (Figure 7) (Baumbusch et al, 2004). Il est détecté dans l'embryon et l'albumen en hybridation *in situ* (Lotan et al, 1998). *LEC1L* est exprimé exactement au même moment que *LEC1* dans l'embryon et dans l'albumen d'après les bases de données Atgenexpress (<http://jsp.weigelworld.org/expviz/expviz.jsp?experiment=development&normalization>) et eFPbrowser (<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>). En raison de leur similarité d'expression et de fonction, il serait intéressant d'étudier le double mutant *lec1 lec1l* pour analyser les relations génétiques entre ces 2 gènes, ce qui n'a jamais été réalisé. *LEC2* est co-exprimé avec *LEC1* dans l'embryon (Kroj et al, 2003; Stone et al, 2001) mais n'est pas exprimé dans l'albumen. Il convient d'insister sur le fait que toutes ces études d'expression ont été menées au niveau de l'accumulation des ARNm, et non des protéines.

*FUS3* est exprimé de façon transitoire entre 8 et 14 jours après floraison mais reste détectable jusqu'à la fin de la maturation (Kroj et al, 2003). *FUS3*, dès le stade globulaire, est



**Figure 6: Evolution des 3 AFL de type B3.** A gauche, arbre phylogénétique montrant les grandes étapes évolutives et l'apparition des AFL de type B3. A droite, les domaines connus de ABI3, FUS3 et LEC2 sont schématisés.



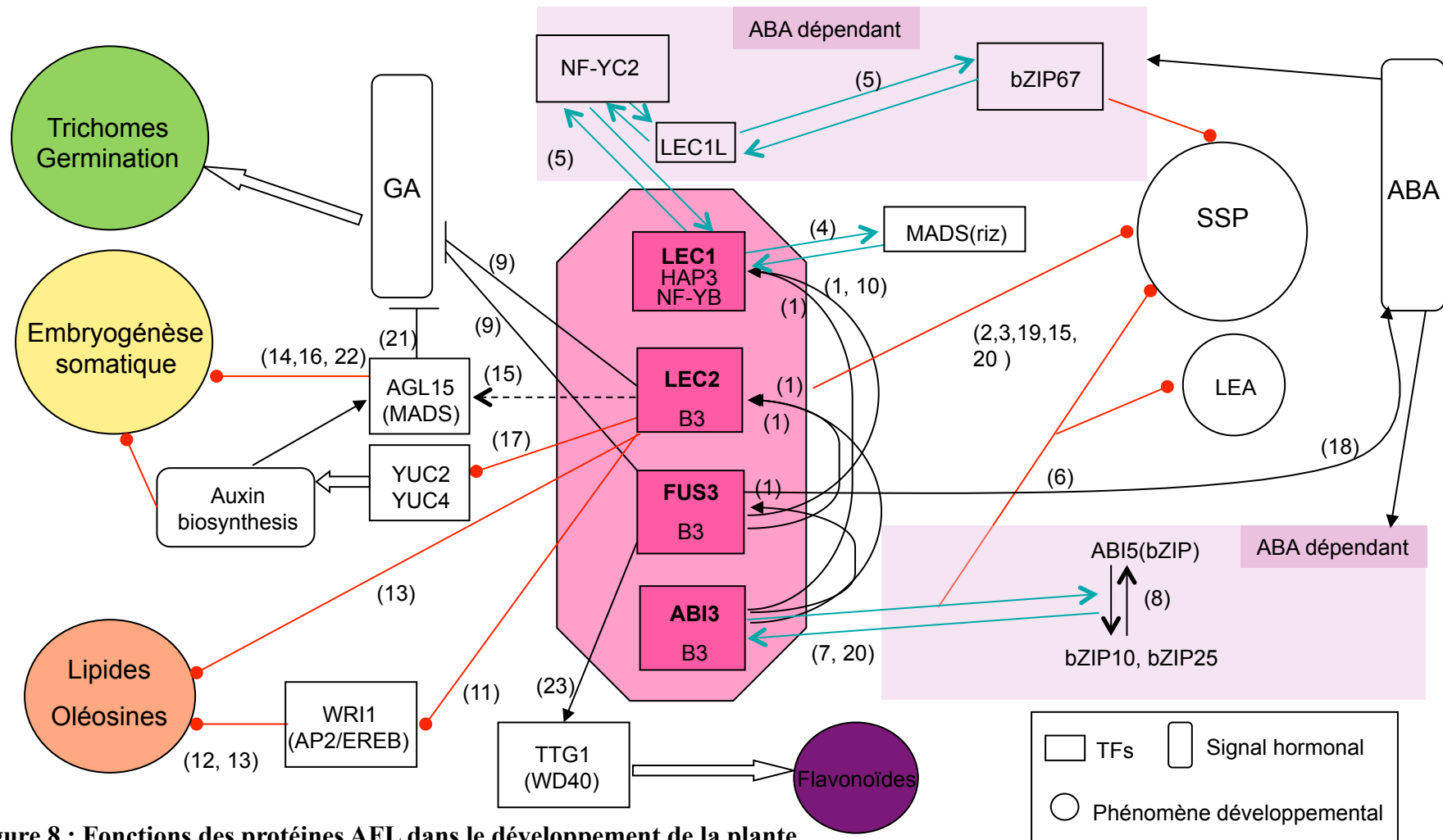
**Figure 7 : Profils d'expression des gènes AFL** (Atgenexpress: <http://jsp.weigelworld.org/expviz/>)

exprimé principalement dans le protoderme (Gazzarrini et al, 2004) et une expression uniquement protodermique (sous contrôle du promoteur d'*ATML1*) suffit à la complémentation du mutant *fus3* (Tsuchiya et al, 2004), ce qui suggère que l'expression de *FUS3* est limitée à cette couche cellulaire.

*ABI3* n'est pas spécifique de la maturation car son expression est détectable et stable de 10 jours après floraison jusqu'à 2 jours après semis lorsque l'on quantifie les ARNm (Baumbusch et al, 2004) et jusqu'à 6 jours après semis lorsque l'on analyse des plantes contenant le construit  $P_{ABI3} :GUS$  (Parcy et al, 1994). Il est également détecté dans les méristèmes. En accord avec son profil d'expression, *ABI3* est en effet impliqué non seulement dans la maturation tardive, les phénomènes de dormance et de germination, mais également dans le développement des racines latérales en réponse à l'auxine (Brady et al, 2003).

### **I.2.5 Les AFL : mécanismes moléculaires de la régulation de l'accumulation des réserves de la graine d'*Arabidopsis***

Comme nous allons le voir dans ce chapitre, les AFL sont impliqués dans la régulation de l'expression des protéines de réserves. La régulation des gènes codant pour les enzymes du métabolisme des lipides peuvent également être régulées directement ou indirectement par les AFLs. La transcription des oléosines, qui entrent dans la composition des oléosomes, lieu de stockage des lipides de la graine est également sous le contrôle des protéines AFL. Comme décrit précédemment, les mutants affectés dans les gènes *AFL* présentent des quantités de réserves diminuées par rapport à une graine de type sauvage. De plus, la fonction des AFL pendant la maturation de la graine a été confirmée par des manipulations d'expression ectopique dans les parties végétatives ou dans des cultures cellulaires (Braybrook et al, 2006; Crowe et al, 2000; Haslekas et al, 2003; Kagaya et al, 2005b; Santos Mendoza et al, 2005). Par ailleurs, la capacité à se fixer aux séquences cis-régulatrices Sph/R<sub>Y</sub> présentes dans les zones promotrices de leur cibles, *in vitro* et en système hétérologue est un argument très fort (Monke et al, 2004; Reidt et al, 2000) qui, combiné aux autres observations citées ci-dessus et à des études intégrant toutes ces stratégies, permet d'identifier des gènes cibles des AFL pendant la maturation de façon fiable (Figure 8) (Baud et al, 2009; Baud et al, 2007; Che et al, 2009; Kotak et al, 2007; Kroj et al, 2003; Nakashima et al, 2006; Parcy et al, 1994; Roschztardt et al, 2009; Stone et al, 2008). Il est assez étonnant de voir la diversité de



**Figure 8 : Fonctions des protéines AFL dans le développement de la plante**

Dans cette figure sont représentés les interactions physiques démontrées entre 2 protéines (  $\rightleftarrows$  ), les régulations positives prouvées uniquement génétiquement (  $\longrightarrow$  ) ou génétiquement et moléculairement (  $\longrightarrow$  ● ). (-----) représente des inductions uniquement observables en expression ectopique et (—|) représente une répression révélée de façon génétique. (1) To et al., 2006; (2) Meinke et al., 1994; (3) Santos Mendoza et al., 2008 (revue); (4) Masiero et al., 2002; (5) Yamamoto et al., 2009; (6) Finkelstein and Lynch 2000; (7) Nakamura et al., 2001; (8) Alonso et al., 2009; (9) Curaba et al., 2004; (10) Parcy et al., 1997; (11) Baud et al., 2007; (12) Cernac and Benning 2004; (13) Baud et al., 2009; (14) Perry et al., 1999; (15) Braybrook et al., 2006; (16) Harding et al., 2003; (17) Stone et al., 2008; (18) Gazzarrini et al., 2004; (19) Kroj et al., 2003; (20) Ezcurra et al., 2000; (21) Wang et al., 2004; (22) Thakare et al., 2008; (23) Tsuchiya *et al.*, 2004.

mécanismes que peuvent utiliser ces quatre protéines régulatrices pour aboutir à l'activation des mêmes cibles.

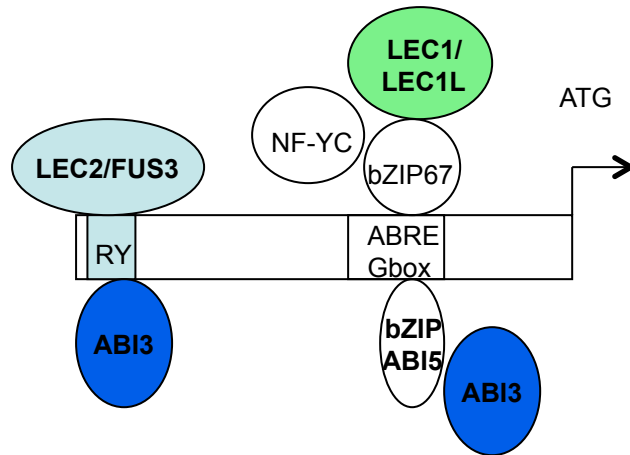
#### ***1.2.5.1 Les facteurs précoces : LEC1 et LEC2***

LEC1, comme on a pu le voir précédemment, peut induire l'expression des gènes codant pour les protéines de réserve en interagissant avec un facteur bZIP qui se lie, en présence d'ABA, aux promoteurs de ces gènes (Yamamoto et al, 2009) (Figure 9), ou en induisant l'expression de *FUS3* et d'*ABI3*, peut-être par le même moyen (Kagaya et al, 2005a). Il est à noter que cette activation est dans l'état actuel des connaissances CCAAT indépendante, soulignant les différences avec les levures et les animaux. *LEC1* est exprimé trop tôt lors du développement pour être responsable de la synthèse totale des protéines de réserve, à moins que la protéine LEC1 soit très stable. *FUS3* et *ABI3* prennent probablement le relais, compte tenu de leur patron d'expression, même si l'expression de *FUS3* dans la couche L1 de l'épiderme ne va pas dans le sens d'un rôle moteur dans l'accumulation des réserves.

LEC2, grâce à son domaine B3, peut induire ses cibles en se fixant directement aux séquences cis-régulatrices de type Sph/Ry de gènes codant pour des protéines de réserve (Figure 9) (Braybrook et al, 2006; Kroj et al, 2003). Il peut également se lier à des gènes codant pour des oléosines et des facteurs de régulation de l'accumulation des huiles comme WRI1 (Baud et al, 2009; Baud et al, 2007; Cernac & Benning, 2004) l'impliquant directement dans la régulation des réserves de type lipidique également (Figure 8). Le mutant *lec2* montre également une quantité de GA supérieure au sauvage (Curaba et al, 2004), l'impliquant dans la répression du métabolisme de cette hormone, qui comme nous l'avons vu au chapitre (I.1.1), oriente le développement vers un état plantule. Un autre système d'action n'est pas exclu pour LEC2 car des manipulations d'expression ectopique ont mis en avant des cibles directes qui ne contiennent cependant pas de séquences RY dans leur zone promotrice (Santos Mendoza et al, 2005).

#### ***1.2.5.2 Les facteurs tardifs : ABI3 et FUS3***

*FUS3* se lie aux boîtes Sph/Ry des promoteurs de ses cibles (Figure 9) que sont les protéines de réserve (Kroj et al, 2003). *FUS3* agit, comme *LEC2* sur le métabolisme des GA.



**Figure 9: Modèle de recrutement à l'ADN pour les protéines AFL.** Un promoteur schématisé et pouvant représenter un promoteur de protéines de réserve est pris comme modèle pour indiquer les boîtes essentielles au recrutement des AFL (RY, ABRE, Gbox).

L'expression de *FUS3* dans une feuille induit une augmentation de la quantité d'ABA et une diminution de la quantité de GA présentes dans les tissus. Ceci est cohérent avec le fait que dans un mutant *fus3*, la quantité d'ABA est réduite par rapport au sauvage à partir de 8 jours après pollinisation (Gazzarrini et al, 2004). FUS3 peut également réprimer l'expression du gène *AtGA3ox2* qui est responsable du passage de GA inactif en GA actif (Figure 8) (Curaba et al, 2004).

ABI3 peut potentiellement se lier directement aux promoteurs grâce à son domaine B3 (Figure 9). Il possède également un système d'action par le biais de son interaction avec des bZIP comme EEL ou ABI5 (Lara et al, 2003) par l'intermédiaire d'éléments cis-régulateurs de type ABRE et ceci de façon ABA dépendante (Alonso et al, 2009). Il est important de noter qu'ABI3 peut agir indépendamment de son domaine B3, par exemple chez *Arabidopsis* pour l'induction des LEA (Kagaya et al, 2005a; Sakata et al, 2010), ou VP1 l'homologue d'ABI3 chez le maïs (Carson et al, 1997; Vasil et al, 1995). Ceci n'a jamais été montré chez les autres AFL de type B3 (Figure 8).

### ***1.2.5.3 Fonctions potentielles***

Bien qu'étudié depuis de nombreuses années, le mode d'action de ces facteurs n'est cependant pas si clair. En effet, l'introduction d'un motif Sph/RV dans une séquence promotrice ne suffit pas à en faire un gène exprimé spécifiquement dans la graine et l'environnement de ces motifs semble être crucial (Ezcurra et al, 2000b; Nakashima et al, 2006).

Par ailleurs, l'activation de la transcription par les AFL, notamment ABI3 et FUS3, a été montrée corrélée à une modification de la répartition nucléosomale aux loci considérés chez *Phaseolus vulgaris* (Li et al, 1999; Li et al, 1998; Ng & Hall, 2008) et le maïs (Busk et al, 1999), sans qu'aucune liaison directe n'ait été établie. ABI3 a également été montré chez le riz comme interagissant avec rad23, une enzyme impliquée dans la réparation de l'ADN (Schultz & Quatrano, 1997). Il reste donc encore peut-être de nombreuses voies à explorer quant à la fonction et au mode de fonctionnement des AFLs.

En conclusion, les AFLs, en induisant la formation de composés de réserve et en réprimant la synthèse de GA dans l'embryon, permettent un allongement de la phase de transition quiescente entre l'état embryonnaire et l'état de plantule.





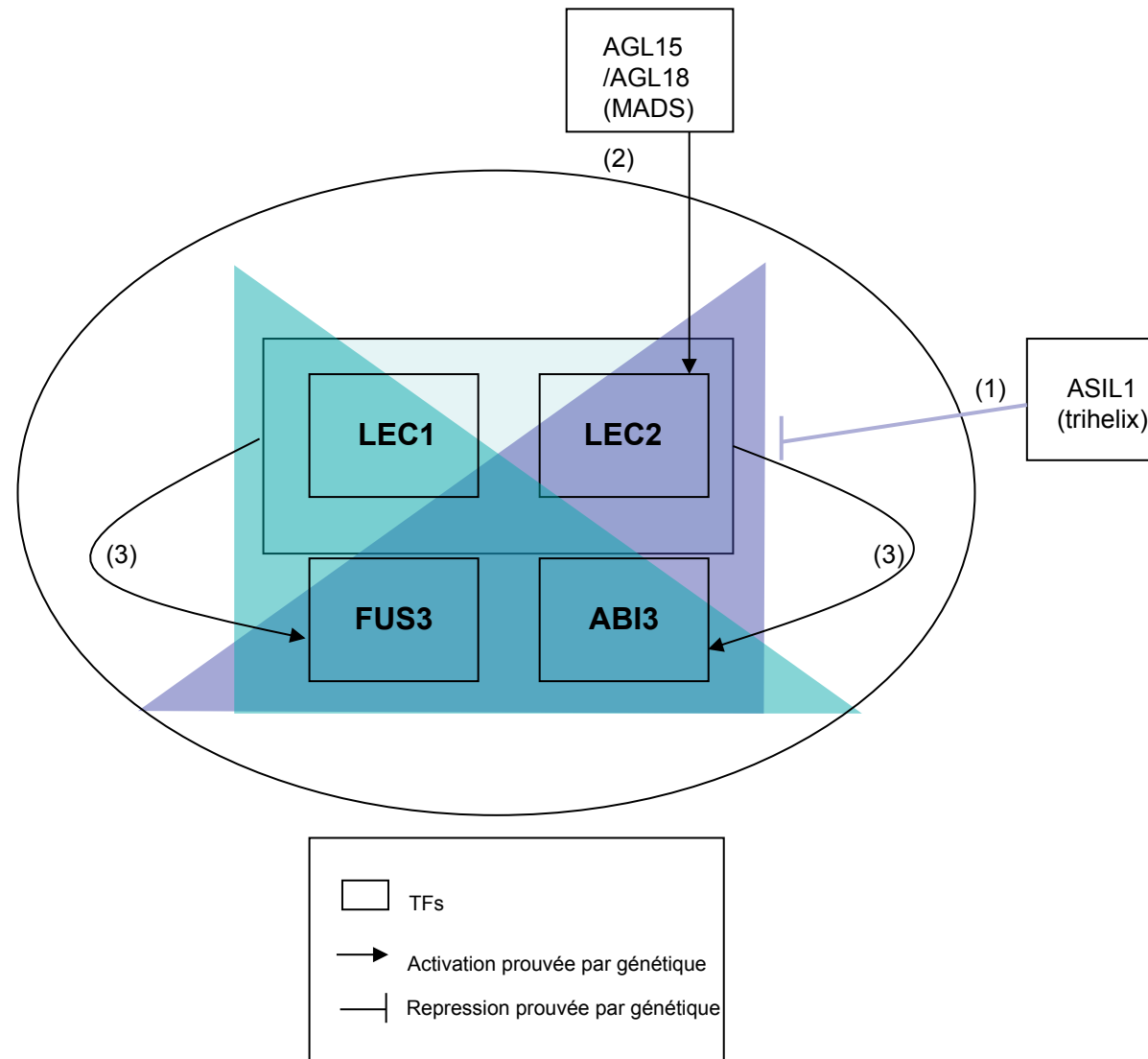
### ***1.3 Régulation des gènes AFL***

#### **1.3.1 Boucles de régulation sur *ABI3* et *FUS3***

Les études de mutants ont également montré des interactions génétiques entre les AFLs comme par exemple *LEC1* qui impacte l'accumulation des réserves par l'intermédiaire d'*ABI3* et de *FUS3* (Parcy et al, 1997), sans qu'une interaction physique entre ces trois protéines puisse être montrée en système hétérologue comme la levure (Brocard-Gifford et al, 2003). Il est important de noter que, dans ces expériences de double hybride en levures, les protéines ont été testées deux à deux et non les trois ensemble, un complexe tripartite n'est donc pas exclu en l'état actuel des connaissances. De la même façon, la nécessité de la présence d'autres partenaires ne peut être exclue.

La régulation de l'expression des *AFL*, et de *LEC2* en particulier, est le sujet de cette thèse mais d'autres équipes s'y sont déjà attachées. Considérant les liens génétiques étroits observés entre les gènes *AFL*, To et collaborateurs ont voulu déterminer s'ils se régulaient entre eux et si l'on pouvait déterminer une hiérarchie entre eux au niveau génétique. La stratégie utilisée a été d'introduire des construits contenant les promoteurs des gènes *AFL* en fusion avec un gène rapporteur dans les différents fonds génétiques mutants *afl*. Ceci a permis d'appréhender la grande complexité de ce réseau de régulation (Figure 8). *LEC1* et *LEC2* semblent agir en amont du réseau de régulation, étant capables d'induire *FUS3* et *ABI3* dans les cotylédons, l'hypocotyle et la pointe racinaire. *ABI3* et *FUS3* s'autorégulent dans les cotylédons, *FUS3* s'autoactive dans la pointe racinaire et *ABI3* active *FUS3* dans l'hypocotyle (To et al, 2006). Ces manipulations ont permis de considérer *LEC1* et *LEC2* comme des gènes maîtres de ce réseau, mais la régulation de ces deux gènes restait cependant largement méconnue.

A l'heure actuelle, le double mutant *agl15 agl18* est le seul génotype à montrer une expression diminuée de *LEC2* pendant le développement de la graine (Zheng et al, 2009), aucun autre mutant montrant une diminution de l'expression des gènes *LEC1* ou *LEC2* n'a été trouvé, nous laissant dans l'expectative quant à l'identité d'un régulateur positif de ce réseau. Cependant, ces dernières années, de nombreux mutants présentant une expression ectopique des *AFL* dans la plantule en germination ont pu être identifiés et ont permis l'identification de répresseurs, soit transcriptionnels, soit chromatinien.



**Figure 10 : Régulation des gènes *AFL* dans la graine en développement.** Les rectangles sont utilisés pour définir un ou deux gènes régulés de la même façon, pour plus de clarté, les triangles colorés définissent des groupes contenant 3 gènes régulés par le même facteur. (1) Gao et al., 2009; (2) Zheng et al., 2009; (3) To et al., 2006

### **I.3.2 AGL15, AGL18 : uniques activateurs de l'expression de LEC2**

Le double mutant *agl15 agl18* présente une diminution de l'expression de *LEC2* et *ABI3* dans des cultures d'embryons somatiques et dans des graines en développement alors que l'expression de *LEC2* n'est pas affectée dans les simples mutants (Zheng et al, 2009). Il faut noter que le mutant *agl15* n'a pas de phénotype notable alors que le double mutant *agl15 agl18* présente quant à lui des phénotypes de floraison mais surtout une capacité moindre à induire l'ES, même en présence d'auxine, soulignant ainsi une redondance fonctionnelle au moins partielle entre ces deux gènes existant également pour d'autres mécanismes (Adamczyk et al, 2007; Thakare et al, 2008). Une surexpression d'*AGL15* induit une augmentation de la quantité de transcrits de *LEC2* et de *FUS3* dans des graines en développement et dans des cultures d'embryons somatiques. Récemment, des manipulations de CHIP ont révélé une liaison de la protéine AGL15 avec les loci *LEC2*, *FUS3* et *ABI3* dans des cultures d'embryons somatiques (Zheng et al, 2009). Il apparaît donc que AGL15, de façon redondante avec AGL18 (facteurs de transcription de type MADS appartenant au sous-groupe des MICK) (Kaufmann et al, 2005), activent l'expression de *LEC2* dans la graine en développement et dans des embryons somatiques (Figure 10). *AGL15* est fortement lié à l'identité embryonnaire, en effet, comme *LEC1* et *LEC2*, il est préférentiellement exprimé dans les tissus ayant des propriétés d'embryons (microspore, méristème apical...) avec une expression 10 fois plus importante dans l'embryon lui-même, où il est exprimé tout au long du développement de la graine (Perry et al, 1999; Perry et al, 1996). Lorsqu'*AGL15* est exprimé, les cellules deviennent compétentes à une induction de l'embryogénèse somatique (ES) par l'auxine. Une surexpression d'*AGL15* augmente la capacité de tissus comme le méristème apical à réaliser des embryons somatiques (Thakare et al, 2008) et augmente la capacité de tissus déjà compétents à acquérir une identité d'embryon (Harding et al, 2003). Comme nous l'avons vu pour *FUS3* et *LEC2*, AGL15 peut également réprimer la synthèse de GA en activant le gène *AtGAox2* qui permet la transformation de GA actif en GA inactif (Wang et al, 2004a).

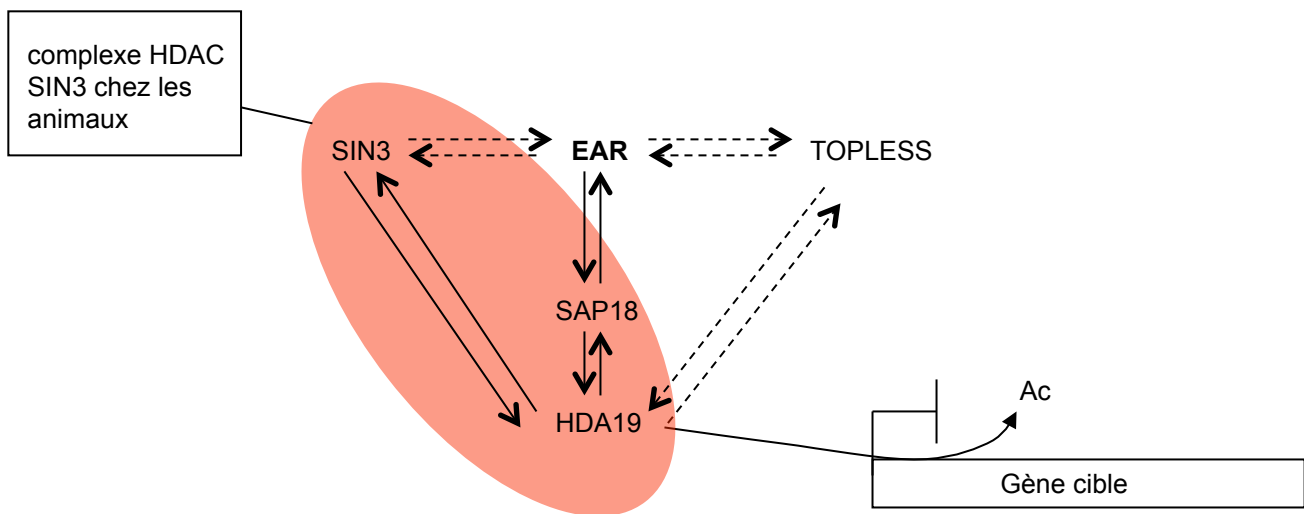
Les relations entre les AFL et *AGL15* sont assez compliquées, non seulement AGL15 peut activer leur transcription mais *AGL15* est également induit par *LEC2* lors de manipulations d'expression inductible (Braybrook et al, 2006). L'expression d'*AGL15* n'est cependant pas modifiée dans le mutant *lec2* (Perry et al, 1996). *AGL15* n'est pas dérégulé en fond *abi3* ni *fus3*, mais dans un allèle *lec1*, la protéine AGL15 disparaît de l'embryon à partir de la moitié

**Tableau 1: Protéines à domaine EAR et leur implication dans le développement d' *Arabidopsis***

SA= acide salicylique; GA= acide gibbéréllique

Protéines	Signalisation/développement	Signature peptidique
SUP (1)	Fleurs	LxLxL, DLNxxP
Aux/IAA (2)	Auxine/	
HSI2 (3, 4, 5)	Saccharose/graine	
AGL15 (6)	Auxine,GA/graine, méristème	
NIMIN (7)	SA, résistance aux pathogènes	

(1) Hiratsu et al., 2002; (2) Tiwari et al., 2004; (3) Tsukagoshi et al., 2005; (4) Tsukagoshi et al., 2007; (5) Suzuki et al., 2007; (6) Hill et al., 2008; (7) Weigel et al., 2005

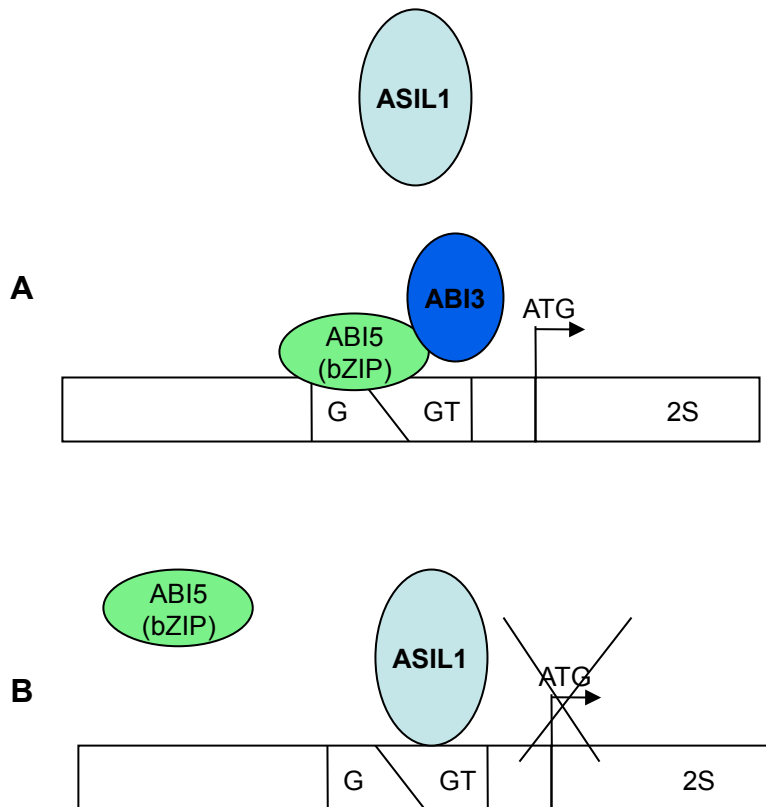


**Figure 11: Modèle d' action des protéines à domaine EAR.** Le domaine EAR a récemment été associé à des éléments du complexe HDAC (  $\leftrightarrow$  ) indique des interactions génétiques et moléculaires et (  $\dashrightarrow$  ) indique des interactions génétiques uniquement (Ac= groupe acétyle)

du développement de la graine (Perry et al, 1996). *AGL15* apparaît donc comme une cible probable des AFLs, en plus d'en être un régulateur, probablement au sommet d'une boucle complexe de régulation mêlant développement et contrôle hormonal. Les MADS se fixent aux éléments cis-régulateurs CARG qui correspondent à la séquence consensus CC(A/T)<sub>6</sub>GG (Riechmann & Meyerowitz, 1997) avec une petite variante pour *AGL15* qui se fixe préférentiellement à C(A/T)<sub>8</sub>G (Tang & Perry, 2003). Même si *AGL15* se lie aux loci *LEC2*, *FUS3* et *ABI3* lorsqu'il est surexprimé dans des embryons somatiques (Zheng et al, 2009), les séquences nucléotidiques nécessaires à sa fixation n'ont pas été déterminées à l'heure actuelle dans les promoteurs des *AFL*.

Les facteurs MADS sont connus pour pouvoir interagir entre eux en formant des homo- ou des hétérocomplexes ce qui est le cas pour *AGL15* (de Folter et al, 2005; Hill et al, 2008). *AGL15*, comme de nombreux MADS, a aussi été identifié comme interagissant avec des protéines autres que des MADS (Causier et al, 2003; Fujita et al, 2003; Gamboa et al, 2001; Masiero et al, 2002; Pelaz et al, 2001; Sridhar et al, 2006), comme HDA19 ou SAP18 qui sont supposés faire partie du complexe SIN3 HDAC (Hill et al, 2008), un des deux types de complexes impliquant des HDAC, comme nous le verrons dans les chapitres suivants. *AGL15* contient en outre un domaine EAR en position C-terminale (Hill et al, 2008). Ce domaine Ethylene-responsive element binding factor-associated Amphiphilic Repression, est présent dans des protéines impliquées dans divers mécanismes biologiques (Tableau 1) (Hiratsu et al, 2002; Kagale & Rozwadowski, 2011; Tiwari et al, 2004; Weigel et al, 2005) et est caractérisé par deux types de signatures peptidiques : LxLxL, DLNxxP. Il est appelé domaine de répression active car, lorsqu'il est fusionné à un facteur de transcription ayant un domaine d'activation de la transcription, un phénotype dominant négatif est induit (Hiratsu et al, 2003). Un modèle récent postule que les protéines à domaine EAR telles qu'*AGL15* réprimeraient leurs cibles en recrutant au locus le complexe HDAC (Figure 11) qui, en déacétylant les histones, participerait à la répression du gène cible (Hill et al, 2008). Les HDAC sont également impliquées dans la régulation des gènes *AFL* et font partie des facteurs de modification de la chromatine, participant ainsi à la régulation de l'expression des gènes. En ce qui concerne les *AFL*, *AGL15* apparaîtrait plutôt comme un activateur de la transcription, il n'aurait donc pas de rôle conjoint avec les HDAC dans ce contexte (Figure 12).

Différents type de mécanismes ont pu être identifiés comme réprimant transcriptionnellement l'expression des gènes *AFL*, en dehors de la graine en développement. Nous nous intéresserons tout d'abord aux facteurs de transcription puis aux facteurs chromatinien.



**Figure 12: Modèle de recrutement d' ASIL1** sur les promoteurs de ses cibles en se basant sur le promoteur du gène codant pour la protéine de réserve 2S (d' après Gao et al., 2008). (A) Activation de la transcription grâce au complexe contenant au moins ABI5 et ABI3. (B) ASIL1 empêche ABI5 de se fixer et par conséquent ABI3 d'être recruté, causant une répression de la transcription.

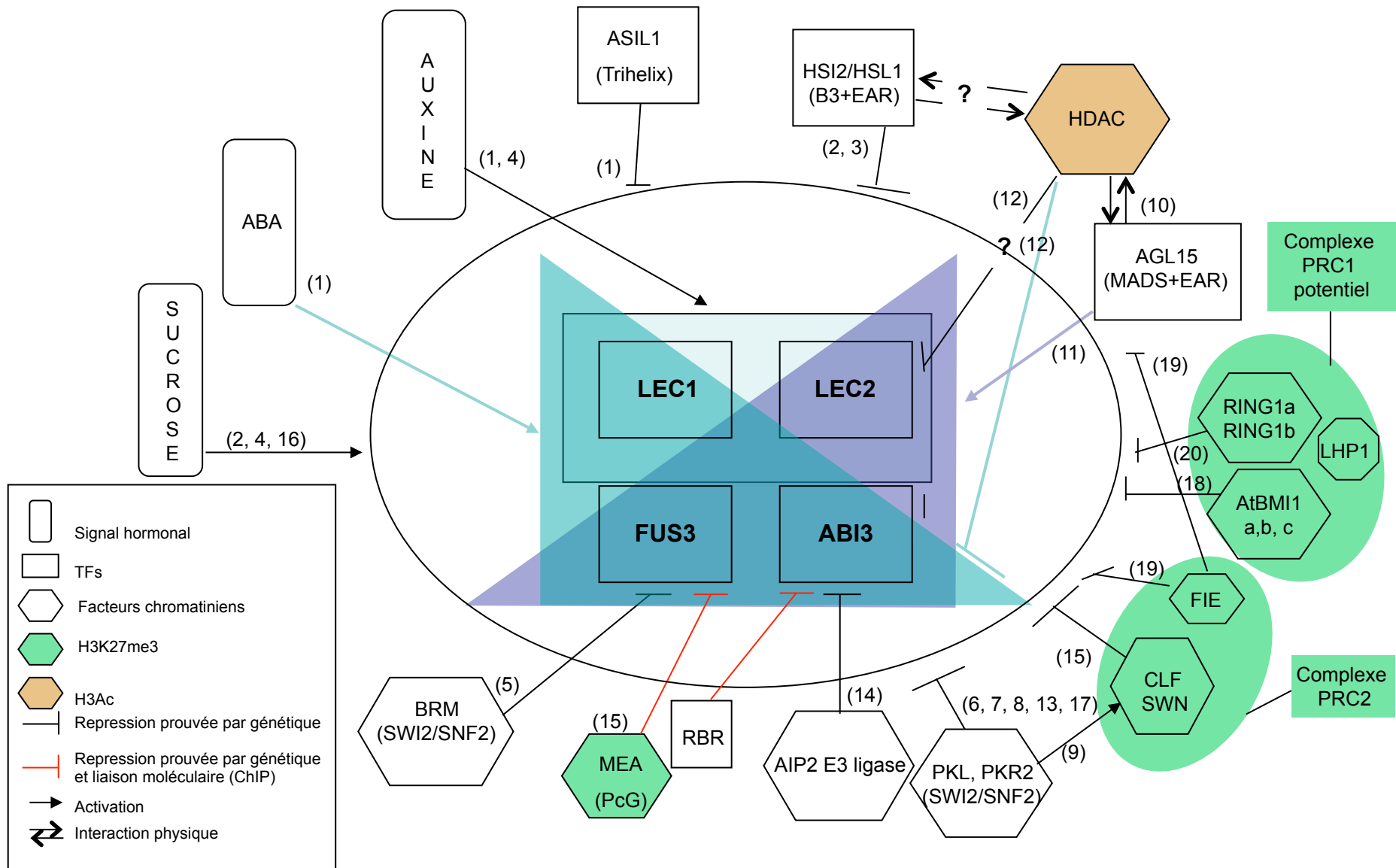
### **I.3.3 Répresseurs de type facteurs de transcription**

#### ***I.3.3.1 ASIL1***

ASIL1 (Arabidopsis 6b-interacting protein 1-like 1) est un facteur de transcription de type trihelix, famille spécifique des plantes, qui a été montré comme se fixant aux éléments GT (GTGATT) (Ayadi et al, 2004). Dans le mutant *asil1*, les gènes *AFL* ainsi que certains gènes codant des protéines de réserve majoritaires comme *At2S3* sont déréprimés dans la plantule en début de germination (Gao et al, 2009). Il faut noter également que des signaux hormonaux tels que l'application d'ABA et d'auxine exogènes augmente l'induction des gènes *AFL* dans le mutant *asil1* dans des plantules au stade 10 jours après semis, sauf pour *LEC2* dont l'expression paraît même diminuée dans le mutant *asil1* en présence d'ABA exogène (Gao et al, 2009). Les auteurs ont étudiés la fixation du facteur ASIL1 sur le promoteur d'*At2S3*. L'arrangement positionnel entre la boîte G (ACGT, fixation par des bZIP), l'élément GT (adjacentes dans le promoteur d'*At2S3*) et les boîtes Sph/RV semble être important pour la fixation du facteur ainsi que la répression de l'expression du gène cible. Le modèle, en ce qui concerne les gènes de réserve contenant des séquences consensus Sph/RV et G/GTbox, suppose une activité de répression d'ASIL1 par compétition avec des bZIP activateurs (par exemple ABI5) au niveau des boîtes GT (Figure 12). Aucune étude n'indique qu'ASIL1 se fixe directement aux régions cis-régulatrices des gènes *AFL*, même si des GT-box potentielles ont pu être déterminées *in silico* (Gao et al, 2009). ASIL1, en plus de réprimer l'expression de *LEC2* dans les parties végétatives, participe à sa faible intensité d'expression en début de maturation. En effet, Gao et collaborateurs ont montré que le niveau d'expression de *LEC2* jusqu'à sept jours après floraison est deux fois plus élevé dans le mutant *asil1*. L'expression de *FUS3* et *ABI3* est décalée vers le début de la maturation, celle de *LEC1* restant inchangée. Ceci est très intéressant car rares sont les mutants connus à l'heure actuelle qui présentent une mauvaise expression des gènes *AFL* pendant le développement de la graine (Figure 10). ASIL1 est donc le premier facteur identifié comme réprimant l'accumulation de transcrits des gènes *LEC2*, *FUS3* et *ABI3* dans l'embryon en début de développement.

#### ***I.3.3.2 HSI2/HSL1***





**Figure 12: Régulation des gènes *AFL* dans la plantule en germination**

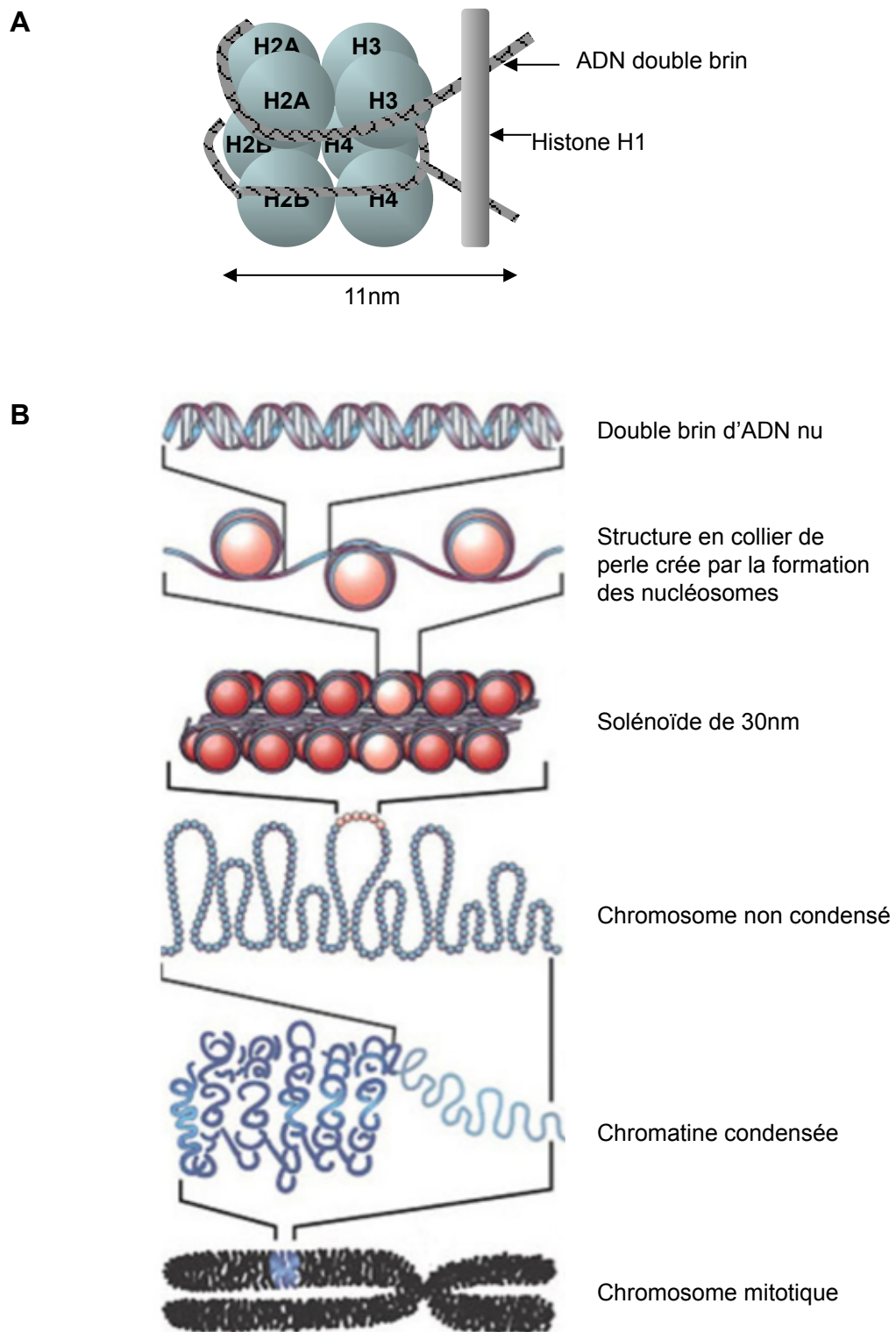
(1) Gao et al., 2009; (2) Tsukagoshi et al., 2007; (3) Suzuki et al., 2007; (4) Casson et al., 2006; (5) Tang et al., 2008; (6) Ogas et al., 1999; (7) Perruc et al., 2007; (8) Rider et al., 2003; (9) Aichinger et al., 2009; (10) Hill et al., 2008; (11) Zheng et al., 2009; (12) Tanaka et al., 2008; (13) Li et al., 2005; (14) Zhang et al., 2005; (15) Makarevich et al., 2006; (16) Gutzat et al., 2011; (17) Zhang et al., 2008; (18) Chen et al., 2010; (19) Bouyer et al., 2011; (20) Bradzel et al., 2010.

Un deuxième type de facteurs de transcription appartenant à la famille des B3 et à la sous-famille des HSI (Figure 5B), a pu être identifié comme réprimant l'expression des *AFL* dans la plantule en début de germination. Le type HSI est composé, en plus d'un domaine B3, d'une zone de répression EAR. Ces facteurs sont au nombre de trois dans le génome d'*Arabidopsis* : *HSI2* (High level expression of Sugar Inducible gene 2), *HSL1* (HSI2-like 1) et *HSL2* (HSI2-like 2) et ont été initialement montrés comme réprimant des gènes induits par le sucrose (Tsukagoshi et al, 2005). Dans le double mutant *hsi2 hsl1* on observe, comme dans le mutant *asill*, une expression ectopique des gènes *AFL* et de *LEC1L* dans la plantule en tout début de germination (Figure 13) (Suzuki et al, 2007), ainsi qu'une dérégulation quasi-totale des caractéristiques embryonnaires, comme, par exemple, une accumulation anormale de lipides spécifiques de la graine dans la plantule jusqu'à 12 jours après semis (Tsukagoshi et al, 2007). L'ajout de saccharose dans le milieu de semis du double mutant *hsi2hsl1* permet une expression encore plus forte des caractéristiques embryonnaires et, en particulier, de l'expression des gènes *AFL*. L'impact du saccharose avait déjà été observé, en effet, il augmente la pénétrance du phénotype d'un mutant surexprimeur de *LEC1* (Casson & Lindsey, 2006) et les *AFL* sont exprimés après la fin de la maturation dans des mutants *rbr* (RetinoBlastoma-Related protein) mis à germer sur saccharose (Gutzat et al, 2011). L'hypothèse proposée est donc que les HSI se fixeraient sur les promoteurs de leur cible via les boîtes Sph/R<sub>Y</sub> grâce à leur domaine B3, pour les réprimer par l'intermédiaire de leur domaine EAR. Encore une fois, aucune étude n'a été menée à l'heure actuelle sur l'interaction directe de ces facteurs avec les zones cis-régulatrices des *AFL*. Cependant, des analyses *in silico* ont permis d'identifier des boîtes R<sub>Y</sub> dans les premiers kb en amont de l'ATG ou dans les premiers introns de tous les *AFL* sauf *LEC2* (Suzuki et al, 2007). Un autre point important concernant *HSI2* est sa similarité de séquence avec des domaines PHD et CW qui sont souvent trouvés dans des facteurs de modification de la chromatine (Aasland & Stewart, 1995; De Lucia et al, 2008; Perry & Zhao, 2003), ce qui suggérerait un mode de répression plus complexe.

### **I.3.4 Répresseurs transcriptionnels de type chromatinien**

#### **II.2.2.3 La chromatine**

L'ADN est présent dans les noyaux sous une forme compactée appelée chromatine, structure complexe dont la base est constituée de nucléosomes. Un nucléosome comprend 146pb



**Figure 14: Structure de la chromatine.** (A) Représentation d'un nucléosome (B) Compaction de l'ADN d'un ADN double brin à un chromosome (image de <http://themedicalbiochemistrypage.org/dna.html>)

d'ADN enroulées autour d'un complexe protéique de 8 protéines histone [2x (H2A, H2B, H3, H4)] et contient également une histone H1 comme liant dans certains cas (Woodcock et al, 2006) (Figure 14A). L'ADN est ensuite compacté pour former un chromosome qui contient deux grands types de structures que sont l'euchromatine et l'hétérochromatine (Figure 14B). L'hétérochromatine est généralement située dans une zone péricentromérique, c'est-à-dire dans la zone du chromosome correspondant à la fixation sur le fuseau achromatique pendant la division cellulaire. On y trouve des séquences répétées de type transposon. L'euchromatine est généralement plutôt associée à la présence de gènes transcrits. Depuis quelques années l'existence d'un code d'histones a vu le jour, en considérant les modifications post-traductionnelles sur les extrémités des protéines histones. De plus, des études de plus en plus nombreuses couvrant le génome entier des animaux et des plantes tendent à souligner la richesse d'un code constitué non seulement des modifications des histones mais également des protéines associées à la chromatine (Filion et al, 2010; Roudier et al, 2011; van Steensel, 2011). En effet, chez la drosophile, il semble qu'il y ait cinq types de chromatine, distribués par les auteurs en code couleur, en fonction des protéines associées et des modifications post-traductionnelles des histones : deux correspondant à des zones de transcription active (jaune et rouge) et trois zones où la transcription semble globalement réprimée (bleue, noire et verte) (Tableau 2). Des études réalisées sur le génome d'*Arabidopsis* ne concernant que les modifications post-traductionnelles des histones et la méthylation de l'ADN ont déjà permis de distinguer deux types de zones actives correspondant aux zones jaunes et rouges et deux zones réprimées (Roudier et al, 2011).

Ces éléments sont essentiels pour structurer et compacter l'ADN dans le noyau, mais également pour réguler l'expression des gènes. Pour que la machinerie de transcription puisse atteindre l'ADN au travers de cette structure dense, des mécanismes ont été mis au point par les eucaryotes pour l'ouvrir, en déstabilisant l'organisation des nucléosomes et/ou en modifiant d'une autre façon la structure de la chromatine. Cette accessibilité peut être atteinte par deux stratégies majeures. La première est de modifier la structure chromatinienne grâce à des enzymes utilisant l'hydrolyse de l'ATP que l'on appelle facteurs de modification de la chromatine ATP-dépendant. La deuxième est de modifier post-traductionnellement la zone N-terminale des histones par phosphorylation, ubiquitination, sumoylation, acétylation et méthylation. L'acétylation et la méthylation des histones sont les deux modifications les plus impliquées dans la régulation des gènes à l'heure actuelle, les chapitres suivants traiteront donc de ces deux types de modifications.

**Tableau 2: Exemple de code multifactoriel trouvé dans la chromatine, décrit par Steensel et al., 2011**

Code couleur	Type de gènes	Modifications histones	Protéines associées
JAUNE	Gènes de ménage exprimés ubiquitairement	Présence de la modification H3K36me3	HAT, HDAC
ROUGE	Gènes tissus spécifiques	Absence totale d' H3K36me3	Caractérisé par un grand nombre de protéines s' y fixant. Présence du variant d' histone H3.3 et diminution de la densité des nucléosomes
BLEUE	Gènes régulant le développement	Caractérisé par la marque H3K27me3	PcG
NOIRE	Gènes tissus spécifiques	Aucune des modifications testées	Hitone H1, AT-hook protein D1, SuUR
VERTE	Correspond à ce que l' on appelle hétérochromatine	H3K9me2, me3	HP1 et protéines associées

↑ GENES ACTIFS  
 ↓ GENES INACTIFS

### ***1.3.4.3 Acétylation des histones et régulation des AFL***

#### ***1.3.4.2.1 Acétylation/déacétylation des histones***

L'acétylation des histones consiste en un transfert d'un groupement acétyle à partir de l'acétyl-coA sur un résidu lysine d'une histone. De façon schématique, une acétylation importante relâche la structure de la chromatine (Mutskov et al, 1998; Puig et al, 1998) et est presque toujours associée à une forte activité de transcription alors qu'un faible degré d'acétylation correspond à une chromatine compactée et à des gènes réprimés chez la levure (Flaus et al, 2006), le poulet (Hebbes et al, 1988) ou les plantes (Chua et al, 2001; Chua et al, 2003). De toutes les modifications, l'acétylation est la plus susceptible de modifier la structure de la chromatine, car elle neutralise les groupements basiques de la lysine ((CH<sub>2</sub>)<sub>4</sub>-NH<sub>2</sub>). L'acétylation de la lysine 9 de l'histone H3 (H3K9) est associée à des gènes transcriptionnellement actifs et influence un grand nombre d'étapes du développement et de la biologie des plantes supérieures (Ausin et al, 2004; Benhamed et al, 2006; Ng et al, 2006). De plus, les résidus lysine des histones H3 (K9, K14, K18 and K23) et H4 (K5, K8, K12, K16 and K20) ont été identifiés comme des cibles d'acétylation/déacétylation chez *Arabidopsis* (Earley et al, 2007; Zhang et al, 2007a; Zhou et al, 2010). L'équilibre entre histones, acétylées ou non, est maintenue par l'action antagoniste des histones déacétylases (HDAC) et des histones acétyl transférases (HAT) (Peserico & Simone, 2011). Quinze HAT ont été trouvées chez *Arabidopsis* (Pandey et al, 2002) dont GCN5, qui est impliquée dans l'acétylation des résidus H3K9 et qui, chez la levure, acétyle des histones différents selon le complexe protéique auquel il est associé. Le mutant *gcn5* d'*Arabidopsis* présente des phénotypes pléiotropiques développementaux et de réponses aux stress ou à différents facteurs environnementaux (des réponses à court et moyen termes). Quatre pourcent des gènes sont affectés dans les plantules de *gcn5* (Benhamed et al, 2006; Benhamed et al, 2008).

Les HDAC, qui par leur activité déacétylase ont un rôle répressif sur l'expression des gènes, ont été récemment associées à des gènes transcriptionnellement actifs dans des cellules humaines (Wang et al, 2009) nous imposant de reconsidérer l'étendue du rôle des HDAC. Les HDAC sont organisées en trois classes. La première est constituée des homologues du facteur de levure RDP3 (Reduced Potassium Deficiency 3). Présents chez tous les eucaryotes, ils sont les plus étudiés, HDA6 et HDA19, les deux HDAC impliquées dans la répression de



l'expression des *AFL*, font partie de ce groupe. Les deux autres groupes sont les homologues du facteur de levure SIR2 appelés HDA1/SIR2, ou Sirtuin, et le groupe des HDtuin (HD2) qui est spécifique des plantes (Hollender & Liu, 2008; Zhou et al, 2004) (Annexe 1). Chez les animaux, il existe deux principaux types de complexes HDAC, le complexe NuRD et le complexe SIN3. Le complexe SIN3 est composé, entre autres, des protéines SIN3, SAP18 et d'HDAC de type RPD3. Chez *Arabidopsis*, les homologues de ces protéines existent et des interactions entre SIN3, SAP18 et HDA19 d'*Arabidopsis* (qui est une HDAC de type RPD3) ont déjà été identifiées (Figure 11). Le complexe NuRD, quant à lui, en plus de contenir une activité HDAC contient une protéine de la famille des SWI/SNF, qui sont des facteurs de modifications de la chromatine ATP-dépendants (voir chapitre suivant) (Ahringer, 2000).

Les HAT sont associées à des co-activateurs qui permettent leur recrutement aux nucléosomes et/ou orientent leur activité (Sternier & Berger, 2000; Utley et al, 1998). Les HDAC sont également associées à des co-répresseurs, orientant ainsi leur activité vers la répression de la transcription (Gonzalez et al, 2007; Sternier & Berger, 2000; Utley et al, 1998) ou vers d'autres modifications épigénétiques en considérant, par exemple, l'interaction d'HDA6 avec une histone déméthylase (Yu et al, 2011). Les HDAC semblent également nécessaires dans des modifications épigénétiques comme la méthylation de l'ADN et des histones à certains loci (Aufsatz et al, 2002; Lawrence et al, 2004; Lippman et al, 2003; Probst et al, 2004).

Les HDAC et les HAT, qui ont donc, par l'identité de leurs interactants, des rôles variés et sont essentiels pour le développement des organismes, sont en cours d'étude chez les plantes. On suppose que la spécificité de ces facteurs est régulée par des interactions avec des séquences spécifiques d'ADN, ceci incluant donc certainement des facteurs de transcription ou des régulateurs de l'état chromatique. HDA6 et HDA19, qui font partie de la classe des RPD3, sont les HDAC les plus étudiées et sont impliquées dans de nombreux mécanismes développementaux comme la floraison (He et al, 2003; Wu et al, 2000; Wu et al, 2008; Yu et al, 2011), l'embryogenèse (Long et al, 2006), la germination (Chen et al, 2010b; Tai et al, 2005; Tanaka et al, 2008), la réponse à la lumière (Benhamed et al, 2006) et la réponse aux pathogènes (Devoto et al, 2002; Zhou et al, 2005) (annexe1). Le nombre de mécanismes contrôlés par HDA6 ou HDA19 est confirmé par le phénotype pléiotropique du mutant *hda19* (Tian & Chen, 2001; Tian et al, 2003) et par les modifications globales dans l'acétylation des histones observées dans ce mutant (Tian et al, 2005). HDA6 et HDA19 ont souvent un rôle redondant et l'étude de lignées déficientes dans les deux transcrits est nécessaire pour observer certains phénotypes. Des outils pharmacologiques sont également disponibles pour étudier le rôle de ces protéines. En effet, il existe des inhibiteurs des HDAC de type RPD3 et





des HD2, comme la TSA (TrichoStatine A) ou le butyrate de sodium (NaBut) qui permettent de lever les problèmes de redondance (Jung et al, 1999; Yoshida et al, 1995). La dernière classe d'HDAC, les Sirtuin, peut être inhibée spécifiquement par le sirtinol (Grozinger et al, 2001). L'usage de ces inhibiteurs spécifiques a permis de révéler des rôles dans la prolifération du méristème racinaire, la dominance nucléaire, et le cycle cellulaire dans des cellules de tabac (Chua et al, 2004; Xu et al, 2005).

#### ***1.3.4.2.2 Les HDAC dans la maturation de la graine***

L'usage d'inhibiteurs a également permis d'identifier une augmentation de l'acétylation des gènes spécifiques de la maturation tardive, comme les protéines LEA (Late Embryogenesis Abundant), à un jour après semis. Cette acétylation conduit à réaliser correctement la transition entre la phase de maturation et la phase de germination (Tai et al, 2005). Les AFL sont également réprimés durant la germination par les HDAC. Quand des graines de plantes sauvages sont mises à germer sur un milieu contenant de la TSA, une expression anormale des AFL est observée en début de germination chez *Arabidopsis* (Tanaka et al, 2008). Il a également été récemment montré que, en présence de TSA, les transcrits des *LEC1* et *LEC1L* de *Picea abies* et *Pinus sylvestris* sont dérèprimés en fin de maturation de l'embryon, provoquant un arrêt dans le développement (Uddenberg et al., 2011). Chez *Arabidopsis*, *LEC1*, *FUS3* et *ABI3* sont également anormalement exprimés dans la lignée double RNAi HDA6/HDA19, au contraire de *LEC2* (Figure 13), suggérant qu'une autre HDAC doit être impliquée dans la répression de *LEC2*, à la place, ou de façon redondante, avec HDA6 et HDA19 (Tanaka et al, 2008). La TSA peut aussi agir de façon indirecte, en activant un gène cible qui active lui-même l'expression de *LEC2*, en effet, l'acétylation des histones au locus *LEC2* n'a jamais été étudiée avec TSA, dans la plantule ou dans la graine. L'acétylation des histones au locus *LEC2* dans des plantules en germination a cependant déjà été détectée (Zhang et al, 2008).

#### ***1.3.4.3 Méthylation des histones et régulation des AFL***

La méthylation des histones est un des systèmes utilisés par la cellule pour moduler la structure de la chromatine et, ainsi, réguler l'expression de certains gènes. Il existe trois types d'acteurs pour le "code" de méthylation des histones: ceux qui réalisent la ou les méthylations, les HKMT (Histones Lysine Méthyl Transferase), ceux qui les effacent et ceux



qui les lisent (les effecteurs) (Tableau 3). La méthylation ne change pas la charge globale de la protéine histone mais peut élever son hydrophobicité, ce qui peut modifier des interactions intra- ou inter-moléculaires et/ou permettre la fixation de nouveaux partenaires.

#### ***1.3.4.3.1 La méthylation des histones : les clés d'un code***

Il existe quatre types majeurs de méthylation des histones, se situant toutes sur des lysines (Tableau 3). Globalement, les mono, di et triméthylations des lysines 4 et 36 de l'histone H3 (H3K4me1, me2, me3; H3K36me1, me2, me3) sont associées à une transcription active. Les mono, di et triméthylation des lysines 9 et 27 de l'histone H3 (H3K9me1, me2, me3; H3K27me1, me2, me3) sont quant à elles considérées comme des marques de répression (Hublitz et al, 2009). La méthylation en H3K9 est conditionnée à la méthylation de l'ADN. En effet le mutant *met1*, qui code pour une CG méthyltransférase chez *Arabidopsis*, présente une diminution de la marque H3K9me dans l'hétérochromatine (Mathieu et al, 2005; Soppe et al, 2002; Tariq et al, 2003). Inversement, la méthylation en H3K9 est nécessaire à la méthylation de l'ADN (Jackson et al, 2002) et à sa maintenance (Malagnac et al, 2002).

La marque H3K27, comme H3K9, est une marque préférentielle de l'inactivation des gènes présente dans l'euchromatine bien que l'on n'observe pas de réel chevauchement entre ces deux modifications. H3K9me2, H3K27me1/me2 peuvent être associées à l'hétérochromatine mais H3K27 n'est pas du tout associée à la méthylation de l'ADN. Son fonctionnement est appelé "ADN méthylé indépendant", ce qui implique certainement des mécanismes de déposition différents. ATXR5 et 6 sont connues pour réaliser la modification H3K27me1 (Jacob et al, 2009). Les protéines de type SET de la famille E(Z) (MEDEA, SWINGER et CURLY LEAF) sont elles connues comme réalisant la modification H3K27me3, même si aucune preuve biochimique ne le démontre à l'heure actuelle. Les H3K27me3 sont reconnues par la protéine LHP1 (Like Heterochromatin Protein1) (Lippman et al., 2004; Turck et al., 2007; Zhang et al., 2007a; Zhang et al., 2007b; Zilberman et al., 2008) qui, par le recrutement d'autres partenaires, permet la répression du locus. Les protéines de type SET et LHP1 seront décrites plus largement dans la suite du manuscrit (Chapitres I.3.4.3.3 et I.3.4.3.5.b). La méthylation sur la lysine 36 de l'histone H3 est généralement associée à une transcription active. Chez l'humain, H3K36me3 est associée à l'élongation de la transcription (Li et al, 2008).

**Tableau 3: Protéines impliquées dans la méthylation des histones chez *Arabidopsis thaliana***

Modifications	Protéines qui déposent la marque	Protéines qui retirent la marque	Protéines qui lisent la marque
H3K9me1/2/3	SUVH4 (KYP) (1,3,17,4), SUVH5 (5), SUVH6 (1,2), SUVR4 (6)	JMJ25 (7), REF6? (8)	Inconnues
H3K27me1	Inconnues	Inconnues	Inconnues
H3K27me2	Inconnues	JMJ12 (24)	Inconnues
H3K27me3	CLF, SWN, MEA	JMJ12 (24)	Inconnues
H3K4me1	Inconnues	JMJ14 (10,9); JMJ15 (10); LDL1 (11); LDL2 (9); FLD? (9,12)	ORC1/PHD (13), AtING (14), WDR5a (15), AL(16)
H3K4me2	ATX2 (23)	Inconnues	Inconnues
H3K4me3	ATX1(22,19)	Inconnues	Inconnues
H3K36me1/2/3	SDG8 (18,19), SDG4 (20), SDG26	Inconnues	Inconnues

(1) Ebbs et al., 2005; (2) Naumann et al., 2005; (3) Jackson et al., 2004; (4) Malagnac et al., 2002; (5) Ebbs et al., 2006; (6) Thorstensen et al., 2006; (7) Saze et al., 2008; (8) Yu et al., 2008; (9) Jiang et al., 2007; (10) Lu et al., 2010; (11) Spedaletti et al., 2008; (12) Liu et al., 2007; (13) de la Paz Sanchez et al., 2009; (14) Lee et al., 2009; (15) Jiang et al., 2009; (16) Grafi et al., 2007; (17) Jackson et al., 2002; (18) Dong et al., 2008; (19) Xu et al., 2008; (20) Cartagena et al., 2008; (21) Zhao et al., 2005; (22) Pien et al., 2008; (23) Saleh et al., 2008; (24) Lu et al., 2011

#### ***1.3.4.3.2 Histone méthyl transférases (HMT) diverses***

La plupart des méthylation des lysines des histones est réalisée par des protéines contenant des domaines SET, qui sont au nombre de 39 dans le génome d'*Arabidopsis*. On peut les classer en quatre catégories: les SU(VAR)3-9 (SUVH et SUVR), les E(Z) (Enhancer of Zeste), les TRX (TRithoraX) et les ASH1 (Absent, Small, or Homeotic discs 1) (Baumbusch et al, 2001; Springer et al, 2003; Thorstensen et al, 2011; Zhao & Shen, 2004). Les protéines homologues à E(Z) chez *Arabidopsis* seront décrites dans un chapitre spécifique (I.3.4.3.3) car elles sont les seules à avoir un rôle sur l'expression des gènes AFL en l'état actuel des connaissances.

Les méthyl transférases de type SU(VAR)3-9 sont responsables de la déposition de résidus méthyl sur la lysine H3K9. KRYPTONITE (SUVH4) est la première H3K9 méthylase à avoir été identifiée mais SUVH2 et SUVH6 ont également été observées plus tard comme agissant de façon partiellement redondante (Ebbs et al, 2005; Naumann et al, 2005).

Dans le génome d'*Arabidopsis*, on peut trouver au moins quatre homologues du gène *ASH1* de drosophile et trois protéines proches nommées SDG (Set Domain Group) (Baumbusch et al, 2001). Parmi elles, SDG8 semble être responsable de la déposition de la majorité des marques H3K36me<sub>2/3</sub> (Dong et al, 2008; Xu et al, 2008).

Les protéines de type TRX réalisent la méthylation des résidus H3K4 de façon antagoniste aux protéines de type PcG (Polycomb Group) qui déposent la marque H3K27me<sub>3</sub> chez les animaux. Chez les plantes, ces fonctions sont conservées et cinq homologues de TRX et sept protéines similaires ont été identifiées (Pien et al, 2008; Saleh et al, 2007; Xu et al, 2008). ATX2 est le paralogue d'ATX1 à la suite d'une duplication d'un chromosome mais leurs profils d'expression ne sont pas similaires et ATX1 semble effectuer les modifications H3K4me<sub>3</sub> alors que les H3K4me<sub>2</sub> sont sous contrôle d'ATX2 (Saleh et al, 2008).

#### ***1.3.4.3.3 HMT de type E(Z) et leur implication dans la régulation des AFL***

La dernière classe d'HMT est la classe des homologues à E(Z), qui permettent de réaliser la modification H3K27me<sub>3</sub> et ainsi de réprimer des gènes devant être finement régulés. Ces facteurs font partie de la famille des PcG qui agissent de façon antagoniste aux TRX chez les animaux et les plantes (Pien & Grossniklaus, 2007; Schwartz & Pirrotta, 2007). Ils sont au nombre de trois chez *Arabidopsis* : CURLY LEAF (CLF), SWINGER (SWN) et MEDEA



(*MEA*) (Grossniklaus et al, 1998). Ils sont impliqués dans un grand nombre de mécanismes développementaux, de la transition florale au développement végétatif en passant par le développement du gamétophyte femelle et de la graine (pour revue (Butenko & Ohad, 2011; Kohler & Villar, 2008; Pien & Grossniklaus, 2007).

#### **I.3.4.3.3.1 Description des HMT de type E(Z)**

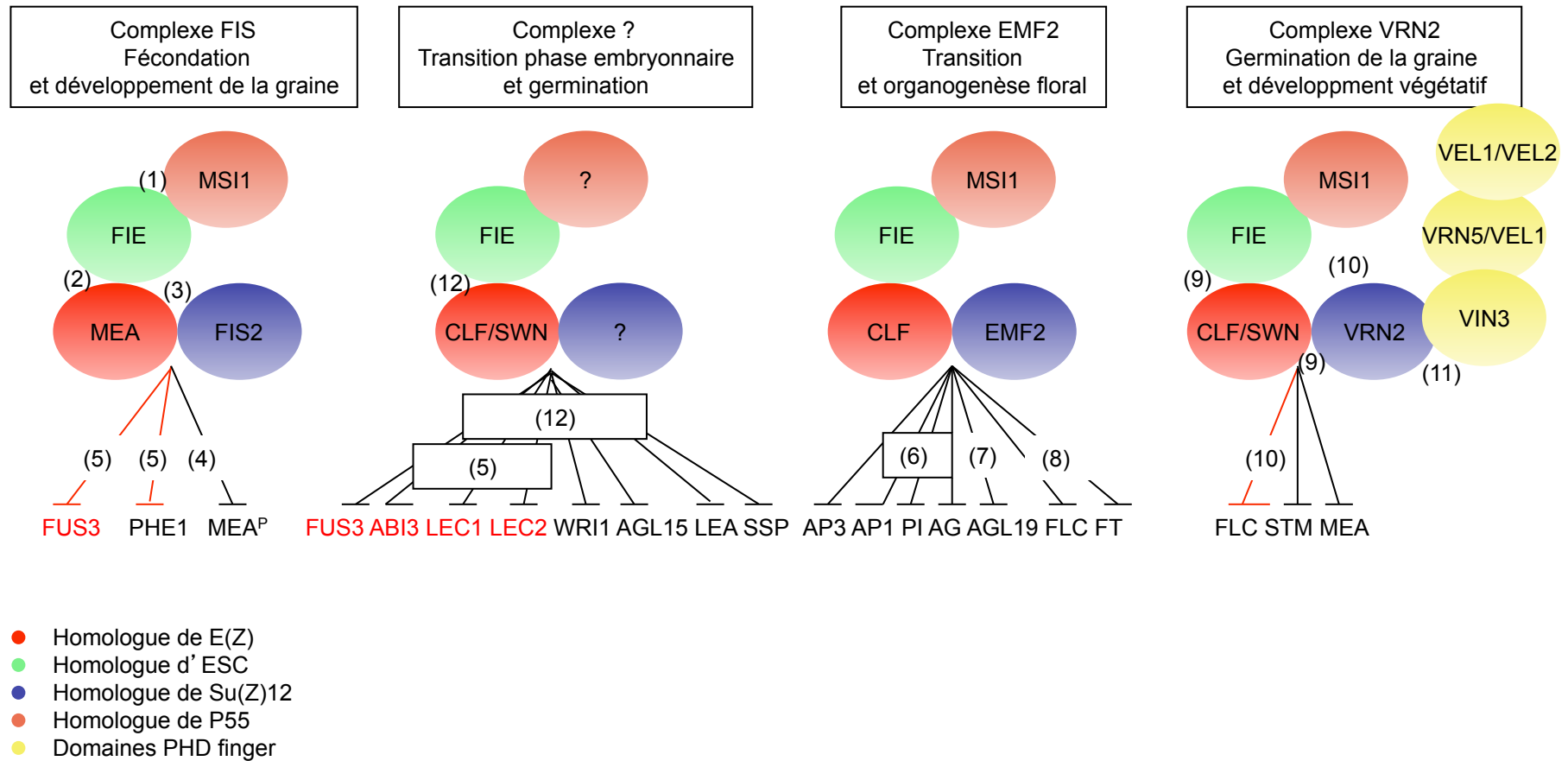
*MEA* est préférentiellement exprimé dans les fleurs en développement et les graines (Grossniklaus et al, 1998). *SWN* et *CLF* sont exprimés de façon plus ubiquitaire (Chanvivattana et al, 2004a; Chanvivattana et al, 2004b; Goodrich et al, 1997; Wang et al, 2006).

Le mutant *mea* peut, sans fécondation, voir son albumen se diviser jusqu'à former des structures ressemblant à des graines. Quand la fécondation a lieu, on observe une prolifération cellulaire dans l'albumen et l'embryon qui peut conduire à une graine avortée. Le mutant *clf* présente des phénotypes pléiotropiques sans gravité pour le développement de la plante (feuilles enroulées, floraison précoce et transformation homéotique partielle des organes floraux) (Goodrich et al, 1997; Schubert et al, 2006). Le mutant *swn* n'est pas différentiable d'une plante sauvage. Les faibles phénotypes des deux derniers mutants peuvent s'expliquer par leur redondance fonctionnelle observée lors de l'étude de double mutants (Makarevich et al, 2006).

#### **I.3.4.3.3.2 Les HMT de type E(Z) dans les complexes PRC2**

Chez la drosophile, le complexe PRC2 est responsable de la modification H3K27me3 grâce à l'activité méthyl transférase des protéines E(Z), composantes du complexe (Muller et al, 2002; Tolhuis et al, 2006). PRC2 est composé de quatre protéines : ESC (Extra Sex Comb), E(Z) (Enhancer of Zeste), SU(Z)12 (Supressor of Zeste 12), et p55 (Czermin et al, 2002). Chez les plantes, on trouve des homologues à ces trois familles de protéines en plus des E(Z) vues précédemment (*MEA*, *CLF*, *SWN*). On trouve des homologues d'ESC comme FIE (Chaudhury et al, 1997; Ohad et al, 1999), des homologues à Su(Z)12 comme FIS2, EMF2 ou VRN2 (Luo et al, 1999) et des homologues à P55 comme MSI1 (Kohler et al, 2003). Trois principaux types de complexes PRC2 ont été isolés chez *Arabidopsis* (Figure 15). Le premier





**Figure 15: Cibles des complexes PRC2.** Adapté de Pien et Grossniklaus 2006. Les cercles placés de façon contiguë correspondent à des interactions prouvées génétiquement et moléculairement au moins *in vitro*; une barre noire représente une répression directe ou indirecte (dérépression dans les mutants), une barre rouge signifie que l'interaction a été vue moléculairement. Les AFL sont écrits en rouge parmi les cibles des complexes. (1) Kohler et al., 2003; (2) Spillane et al., 2000; (3) Wang et al., 2006; (4) Jullien et al., 2006; (5) Makarevich et al., 2006; (6) Moon et al., 2003; (7) Schonrock et al., 2006; (8) jiang et al., 2008; (9) Chanvivattana et al., 2004; (10) De lucia et al., 2008; (11) Wood et al., 2006; (12) Bouyer et al., 2011.

est le complexe contenant FIS, composé de FIS2, MEA (ou FIS1), FIE (ou FIS3) et MSI1 (Figure 15) (Kohler et al, 2003; Spillane et al, 2000; Wang et al, 2006). En l'état actuel des connaissances, il n'a été impliqué que lors du développement du gamétophyte femelle et de la graine (Jullien et al, 2006; Makarevich et al, 2006) (Figure 15). Le deuxième type de complexe est EMF2-associé composé de EMF2, CLF, FIE et MSI1. Il a été identifié pour son rôle dans la répression de la transition florale et des gènes homéotiques floraux (Jiang et al, 2008; Moon et al, 2003; Schonrock et al, 2006). Le troisième complexe contient VRN2, CLF ou SWN, FIE et MSI1, il est également associé à trois protéines à domaine PHD finger VIN3, VRN5 (VIL1) (liu et al., 2010) et VEL1/VEL2 (Chanvivattana et al, 2004a; De Lucia et al, 2008; Wood et al, 2006).

#### **I.3.4.3.3 Recrutement de PRC2 au locus**

Grâce aux études chez la drosophile et chez l'Homme, des régions cis-régulatrices de quelques centaines de paires de bases, essentielles à la répression de gènes majeurs, ont pu être isolées et associées au complexe PRC2. Elles ont été nommées: PREs pour Polycomb Responsive Element (Muller & Kassis, 2006; Ringrose & Paro, 2007; Schuettengruber et al, 2007). Ces zones sont assez mal définies et peu prédictibles, mais contiennent généralement les sites de reconnaissance du facteur GAGA (Mahmoudi et al, 2003) de protéines polycomb comme PHO et PHOL (Brown et al, 2003; Brown et al, 1998; Mihaly et al, 1998) et de la protéine zeste (Mahmoudi et al, 2003; Mulholland et al, 2003; Saurin et al, 2001). Toutes ces caractéristiques sont essentielles à l'activité de répression de PRC2 (Brown et al, 1998; Hagstrom et al, 1997; Hur et al, 2002). Chez les plantes, aucun élément de type PRE n'a pu être déterminé clairement, même si des séquences nécessaires à la répression de cibles connues de PRC2 comme *FLC* ou *AG* ont déjà été isolées. De plus, chez les plantes, aucun homologues de PHO ou de zeste n'ont pu être détectés, mais des homologues fonctionnels du facteur GAGA de drosophile ont pu être identifiés avec les BPC. Les BPC reconnaissent le même type de séquences consensus que le facteur GAGA mais n'ont aucune similarité de séquence. Ces protéines seront abondamment et spécifiquement décrites au chapitre (II.2.2).

#### **I.3.4.3.4 PRC2 et les AFL**



En ce qui concerne les AFL, *FUS3* est le seul gène à avoir été démontré comme réprimé par PRC2 pendant le développement de la graine. Lorsque le gène rapporteur *UIDA* est placé sous le contrôle du promoteur de *FUS3* en fond mutant *mea*, on observe une expression du gène rapporteur dans l'embryon et l'albumen (Makarevich et al, 2006) alors que l'expression reste spécifique de l'embryon en fond sauvage. Il est important de noter que la même manipulation avec le promoteur de *LEC2* ne provoque absolument pas d'expression ectopique. De plus, des manipulations de ChIP-PCR ont pu montrer une liaison de MEA au locus *FUS3* dans des fleurs encore fermées (Makarevich et al, 2006). Makarevich et collaborateurs ont également démontrés une expression ectopique de *FUS3*, *LEC2* et *LEC1* en fond double mutant *clf swn* dans des plantules. Une expression ectopique de *FUS3*, *ABI3*, *LEC2* et *LEC1* et de leurs gènes cibles est également observée dans des plantules de 7 jours du mutant *fie* (Bouyer et al, 2011). Toutes ces études ont été menées sur des plantules, nous n'avons aucun renseignement actuellement sur une possible répression des AFLs, hors *FUS3*, pendant le développement de la graine qui pourrait expliquer les patrons d'expression très spécifiques de ces gènes. Il faut également noter que même s'il est prouvé que PRC2 réprime les AFL lors de la germination, nous ne connaissons pas la composition exacte de ce complexe car seulement CLF/SWN et FIE ont pu y être associés (Figure 15).

#### ***1.3.4.3.4 Histones déméthylases***

Il existe deux types de déméthylases, les LSD1 qui déméthylent par oxydation d'une amine (Jiang et al, 2007) et les protéines contenant un domaine JmjC ou JMJ qui déméthylent par hydroxylation (Tsukada et al, 2006). Chez *Arabidopsis*, quatre homologues d'histone déméthylases de type LSD1 ont été identifiées : FLD (Jiang et al, 2007; Liu et al, 2007), LDL1 (Spedaletti et al, 2008), LDL2 (Jiang et al, 2007) et LDL3. Les trois premiers sont impliqués dans la floraison. Les mutants affectés dans ces trois gènes apparaissent avec une floraison retardée, liée à la surexpression du gène répresseur de la floraison *FLC*, dans lequel une hyperacétylation H4 et une diminution de l'H3K27me2 et de la H3K9me2 sont détectées dans la région 5' de *FLC* (Jiang et al, 2007).

Les protéines JUMONJI (JMJ) (pour revue (Chen et al, 2011) représentent une seconde classe de déméthylases qui, chez les animaux, sont impliquées dans la soustraction d'un groupement methyl des lysines des histones H3 et réalisent des modifications telles que H3K9me3, H3K36me2 et H3K36me3, H3K27me2 et H3K27me3, H3K4me2 et H3K4me3. Chez *Arabidopsis*, il existe 20 protéines à domaine JmjC (Liu et al, 2010; Lu et al, 2008).



Chez les plantes, le gène de riz *JMJ706* et le gène d'*Arabidopsis* *JMJ25* sont impliqués dans la déméthylation des résidus H3K9 (Saze et al, 2008; Yu et al, 2008). OsJMJ706 est le seul à avoir été montré comme ayant une activité de déméthylation des lysines de type H3K9me2 et H3K9me3 *in vitro* et *in vivo* (Sun & Zhou, 2008). JMJ14 et JMJ15 sont quant à eux impliqués dans la déméthylation des résidus H3K4 (Jiang et al, 2007; Lu et al, 2010).

Des protéines impliquées dans la déméthylation de type H3K27 chez les plantes viennent seulement d'être connues chez *Arabidopsis*. La protéine REF6 ou JMJ12 permet la déméthylation des histones portant les modifications H3K27me3 et H3K27me2 (Lu et al, 2011).

Une fois la marque déposée, des protéines sont capables de "lire" ces modifications permettant ainsi le recrutement d'acteurs plus tardifs. Ces effecteurs sont décrits dans le chapitre suivant.

#### ***1.3.4.3.5 Effecteurs de la méthylation des histones***

Les effecteurs de la méthylation des histones sont classés en trois grands groupes: les protéines contenant un domaine chromolike, de la superfamille "royale" (chromodomain, tudor domain, malignant brain tumor (MBT), PWWP et plant Agenet module), la superfamille des facteurs à homéodomains (PHD finger) et, enfin, les effecteurs contenant une répétition WD40 comme WDR5 (Ruthenburg et al, 2007; Taverna et al, 2007). Chez *Arabidopsis*, quelques protéines sont connues pour lire les marques de méthylation des histones comme ORC1, AtING, AL et WDR5a qui lisent les méthylations en H3K4, et LHP1 qui lit les méthylations en H3K27.

##### **1.3.4.3.5.1 Effecteurs divers**

La protéine ORC1 d'*Arabidopsis* contient un domaine PHD finger et peut interagir avec la marque H3K4me3 pour activer la transcription de ses gènes cibles (de la Paz Sanchez & Gutierrez, 2009). ING (Lee et al, 2009) et AL (Grafi et al, 2007) contiennent un domaine PHD finger et peuvent interagir avec la marque H3K4me2/3. WDR5a est quant à lui présent au locus *FLC* et peut se lier *in vitro* à un peptide portant la marque H3K4me2 (Jiang et al, 2009). La protéine LHP1 a le même type de fonction qu'un membre du complexe PRC1 animal qui a un rôle effecteur sur les modifications H3K27me3, plus précisément, PRC1, recruté par les marques H3K27me3 déposées par le complexe PRC2, permet l'ubiquitination



des histones à proximité et ainsi, la répression de la transcription au locus.

#### **I.3.4.3.5.2 Le complexe PRC1**

Chez les animaux, PRC1 est composé de polycomb (Pc), polyhomeotic (Ph), Posterior sex combs (Psc), Zeste, dSbf1, une histone H2A ubiquitine transférase dRing1 (appelée également Sex combs extra, (Sce)) (Fritsch et al, 2003), HSC4, et de cinq facteurs généraux de la transcription. Chez *Arabidopsis*, la protéine LHP1 fonctionne comme Pc, sans pour autant avoir la moindre similarité de séquence, en se liant par l'intermédiaire de son chromodomaine aux marques H3K27me3 (Exner et al, 2009). Ce lien est renforcé par une colocalisation entre LHP1 et la marque H3K27me3 (Turck et al, 2007; Zhang et al, 2007b). LHP1 permet la répression de la transcription de gènes tels que *FLC* (Mylne et al, 2006) mais n'a pas encore été associée à un des gènes *AFL*.

Par contre, AtRING1a et AtRING1b, homologues de l'H2A ubiquitine transférase dRing1 de drosophile et qui interagissent avec LHP1 (Xu & Shen, 2008), permettent la répression des quatre *AFL* et d'*AGL15* pendant la germination (Chen et al, 2010a). Trois autres protéines de type ring (AtBMI1a,1b et 1c) (Qin et al, 2008) permettent une monoubiquitination de l'histone 2A chez *Arabidopsis*. Conjointement avec la marque H3K27me3 médiée par PRC2 elles jouent un rôle déterminant dans la répression des gènes *AFL* ainsi que d'autres gènes impliqués dans le développement de l'embryon durant la phase végétative (Bratzel et al, 2010).

#### ***1.3.4.4 Les complexes ATP-dépendant et les AFL***

Les complexes ATP-dépendant peuvent être classés en quatre familles selon le type d'ATPase SNF2 qu'ils contiennent : les Swi/Snf, les Iswi, les Chd/Mi2 et les INO ATPases (Conaway & Conaway, 2009; Flaus et al, 2006; Kunert & Brehm, 2009; Racki & Narlikar, 2008). Ces complexes ont été montrés comme « menant à une zone » transcriptionnellement active et pauvre en nucléosomes chez la drosophile et les mammifères et ont des rôles cruciaux dans le fonctionnement cellulaire (Johnson et al, 2005; Maier et al, 2008). Chez *Arabidopsis*, deux protéines de la famille des SWI2/SNF2 ont été identifiées comme réprimant les *AFL*, BRAHMA (BRM) et PICKLE (PKL) (Ogas et al, 1997; Ogas et al, 1999). Les facteurs SNF2 reconnaissent les H3K4me grâce à leurs chromodomains (Flanagan 2006).





Dans le mutant *brm*, les transcrits codant pour *FUS3* et les protéines de réserve de la graine d'*Arabidopsis* sont détectés dans des plantules en cours de germination (Tang et al, 2008). Des expériences de ChIP ont pu montrer le recrutement de BRM aux loci des protéines de réserve dans des plantules de type sauvage ainsi qu'une sensibilité à la DnaseI différente aux loci, indiquant une structure chromatinienne modifiée. BRM est donc impliqué dans la répression de *FUS3* et des gènes codant certaines protéines de réserve, mais son rôle apparaît comme inexistant au niveau de *LEC2*, *LEC1* et *ABI3* (Figure 13). En ce qui concerne *FUS3*, les auteurs n'ont pas testé d'interactions physiques entre le locus *FUS3* et BRM. L'homologue de BRM est associé, chez l'Homme, à la méthylation de l'ADN (Harikrishnan et al, 2005), par contre BRM n'a jamais été détecté aux loci AFL, suggérant des fonctions différentes entre les BRM d'animaux et de plantes.

PKL, le deuxième facteur de type SWI/SNF2, est impliqué directement dans une répression bien plus générale des facteurs déterminants l'identité embryonnaire, mais uniquement dans la racine primaire de la plantule en tout début de germination. PKL réprime les AFL et leurs cibles (Li et al, 2005; Perruc et al, 2007; Rider et al, 2003) et, par exemple, *AGL15* (Aichinger et al, 2011). Récemment, cette répression a été liée à une modification des histones de type H3K27me3 des cibles et non à une déacétylation (Zhang et al, 2008) comme chez ses homologues de mammifères qui sont associés à des HDAC dans des complexes de type NuRD (Bouazoune & Brehm, 2006; Hall & Georgel, 2007). Une étude récente a montré que PKL et son homologue PKR2 agissent de façon redondante (Figure 13). Grâce à des expériences de ChIP-PCR avec des anticorps anti-PKL, PKL a été identifié comme se liant directement aux cibles sous-exprimées dans le mutant *pkl* et qu'il aurait donc un rôle activateur et non répresseur (Aichinger et al, 2009). À l'heure actuelle, PKL semble agir par deux voies : 1) d'un côté, il peut activer EMF2 et SWN, deux protéines de type PcG et, par la même, réprimer ses cibles via l'activité méthyl-transférase de ces dernières. On peut donc dire que, les cibles étant le siège de la marque H3K27me3, sont des cibles indirectes dont la répression est médiée par l'activation des PcG. 2) d'un autre côté, PKL peut agir de façon PcG indépendante et activer d'autres cibles.

#### **1.4 Objectifs de la thèse**

Comme nous l'avons vu dans le chapitre précédent, la régulation transcriptionnelle est extrêmement complexe. Beaucoup de facteurs protéiques sont connus comme réprimant



l'expression des gènes *AFL*, mais on ne connaît pas leur mode de recrutement au locus considéré, ni même si leur action est directe. Comme nous l'avons aussi vu précédemment, *LEC2* est l'élément déclencheur, avec *LEC1*, de la cascade complexe de régulation régissant la maturation de la graine. L'étude du promoteur de *LEC1* n'a fait l'objet d'aucune publication, celle du promoteur de *FUS3* a été limitée à un fragment de 1,5kb en amont de l'ATG (Gazzarrini et al, 2004), celle du promoteur d'*ABI3* à un fragment de 6,4kb (Parcy et al, 1994) et, enfin, celle du promoteur de *LEC2* à 2kb en amont de l'ATG (Kroj et al, 2003). Il est étonnant que les promoteurs de ces gènes cruciaux, dont l'expression ectopique permet la mise en place complète du programme d'embryogenèse et d'accumulation des réserves, n'aient jamais été étudiés en détails et qu'on ne connaisse donc pas la mise en place de la mécanistique de la répression.

Les objectifs de la thèse étaient donc :

- 1) En premier lieu, pour tenter de répondre à toutes ces interrogations, de réaliser une étude fonctionnelle du promoteur de *LEC2* afin de mieux comprendre les mécanismes moléculaires impliqués dans sa répression par ces facteurs déjà connus.
- 2) En second lieu, de rechercher de nouveaux acteurs, positifs ou négatifs, de la régulation de l'expression de ce gène grâce à des cribles en hôtes hétérologues avec, comme appât, le promoteur de *LEC2*. Le tout dans le but de comprendre comment l'expression de *LEC2* est si finement régulée pendant le développement de la graine.
- 3) Enfin, le troisième objectif de la thèse était de tenter d'obtenir les outils permettant l'étude de la protéine *LEC2* pour la localiser au niveau tissulaire et sub-cellulaire, et détecter d'éventuelles modifications post-traductionnelles.



## II RESULTATS

### ***II.1 Etude fonctionnelle du promoteur de LEC2***

#### **II.1.1 La régulation transcriptionnelle de l'expression de *LEC2* implique RLE, un élément cis-régulateur qui conditionne la déposition de la marque H3K27me3**

Pour répondre au premier objectif de la thèse, qui portait sur le déterminisme moléculaire de la régulation de l'expression du gène *LEC2* j'ai entrepris des expériences de délétions du promoteur de *LEC2* et ai pu caractériser trois boîtes essentielles à une expression correcte. Ce travail a été mené à son terme et a fait l'objet d'une publication dans « The Plant Cell » présentée ci-après.



# Transcriptional Regulation of *Arabidopsis* *LEAFY COTYLEDON2* Involves *RLE*, a *cis*-Element That Regulates Trimethylation of Histone H3 at Lysine-27 <sup>W</sup>

Nathalie Berger,<sup>a</sup> Bertrand Dubreucq,<sup>a</sup> François Roudier,<sup>b</sup> Christian Dubos,<sup>a</sup> and Loïc Lepiniec<sup>a,1</sup>

<sup>a</sup> Institut Jean-Pierre Bourgin, Unité Mixte de Recherche 1318 Institut National de la Recherche Agronomique–Agro-ParisTech, Saclay Plant Sciences, 78026 Versailles cedex, France

<sup>b</sup> Institut de Biologie de l'École Normale Supérieure, Centre National de la Recherche Scientifique Unité Mixte de Recherche 8197–Institut National de la Santé et de la Recherche Médicale U1024, 75230 Paris cedex 05, France

***LEAFY COTYLEDON2 (LEC2)* is a master regulator of seed development in *Arabidopsis thaliana*. In vegetative organs, *LEC2* expression is negatively regulated by Polycomb Repressive Complex2 (PRC2) that catalyzes histone H3 Lys 27 trimethylation (H3K27me3) and plays a crucial role in developmental phase transitions. To characterize the *cis*-regulatory elements involved in the transcriptional regulation of *LEC2*, molecular dissections and functional analyses of the promoter region were performed in vitro, both in yeast and in planta. Two *cis*-activating elements and a *cis*-repressing element (*RLE*) that is required for H3K27me3 marking were characterized. Remarkably, insertion of the *RLE cis*-element into *pF3H*, an unrelated promoter, is sufficient for repressing its transcriptional activity in different tissues. Besides improving our understanding of *LEC2* regulation, this study provides important new insights into the mechanisms underlying H3K27me3 deposition and PRC2 recruitment at a specific locus in plants.**

## INTRODUCTION

In eukaryotes, complex developmental programs require tightly regulated expression of a specific set of genes at both spatial and temporal levels. Chromatin state dynamics and transcription factors are involved in the underlying molecular mechanisms (Beisel and Paro, 2011). Recent research has emphasized the importance of histone posttranslational modifications, such as acetylation, methylation, ubiquitination, or phosphorylation, in these regulatory processes (Berger, 2007; Kouzarides, 2007; Zhang et al., 2007a; Lee et al., 2010). In plants, genome-wide chromatin and transcriptomic analyses have confirmed the correlation existing between histone posttranslational modifications and transcriptional activity (Pfluger and Wagner, 2007; Roudier et al., 2009, 2011; Lauria and Rossi, 2011). For instance, Histone H3 Lys (residues 4, 9, 27, and 36) methylations are strongly associated with the transcriptional regulation of plant genes (Li et al., 2008; Liu et al., 2010; Lafos et al., 2011; Zheng and Chen, 2011). H3K4me3 and H3K36me2/me3 are predominantly linked with active chromatin, whereas H3K9me1/me2 and H3K27me3 modifications are considered to be repressive. More specifically, dynamic regulation of H3K27me3 marking has been shown to play a crucial role in the developmental regulation of gene

expression in eukaryotes, including plants (Kouzarides, 2007; Zhang et al., 2007a; Schatlowksi et al., 2008; Lafos et al., 2011; Zheng and Chen, 2011). H3K27me3 is extensively distributed over the transcribed and promoter regions of genes. It extends over large intergenic regions in animals, whereas the distribution seems to be more gene centered in plants (Turck et al., 2007; Zhang et al., 2007b; Schatlowksi et al., 2008).

H3K27me3 deposition is catalyzed by Polycomb-group (PcG) proteins with a SET [for Su(var)3-9, Enhancer-of-zeste, Trithorax] domain that belong to Polycomb Repressive Complex2 (PRC2) (Schuettengruber and Cavalli, 2009; Simon and Kingston, 2009; Morey and Helin, 2010; Margueron and Reinberg, 2011). On this line, it has been recently shown that PRC2 activity is necessary for H3K27me3 deposition in plants (Schubert et al., 2006; Bouyer et al., 2011). PcG proteins were first identified in *Drosophila melanogaster* and are evolutionary conserved among higher eukaryotes, including plants (Ringrose et al., 2004; Hennig and Derkacheva, 2009; Chen et al., 2010; Sawarkar and Paro, 2010; Margueron and Reinberg, 2011; Zheng and Chen, 2011). The PRC2 complex is composed of four core proteins, including ENHANCER OF ZESTE [E(Z); a PcG protein with a SET domain and methyltransferase activity], EXTRA SEX COMBS (ESC; a WD40 domain protein), SUPPRESSOR OF ZESTE-12 (a zinc finger protein), and NUCLEOSOME REMODELING FACTOR55. PcG proteins act in establishing and maintaining inactive chromatin states at numerous target genes throughout development (Francis and Kingston, 2001; Schwartz et al., 2006; Schwartz and Pirrotta, 2008; Müller and Verrijzer, 2009; Köhler and Hennig, 2010; Margueron and Reinberg, 2011).

Recruitment of PRC2 at a specific locus likely relies on multiple mechanisms (Margueron and Reinberg, 2011). Genomic

<sup>1</sup> Address correspondence to loic.lepiniec@versailles.inra.fr.

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors ([www.plantcell.org](http://www.plantcell.org)) are: Bertrand Dubreucq ([bertrand.dubreucq@versailles.inra.fr](mailto:bertrand.dubreucq@versailles.inra.fr)) and Loïc Lepiniec ([loic.lepiniec@versailles.inra.fr](mailto:loic.lepiniec@versailles.inra.fr)).

<sup>W</sup>Online version contains Web-only data.  
[www.plantcell.org/cgi/doi/10.1105/tpc.111.087866](http://www.plantcell.org/cgi/doi/10.1105/tpc.111.087866)





domains named Polycomb Repressive Element (PRE) have been described mainly in *Drosophila* (Ringrose and Paro, 2007; Schuettengruber et al., 2007). PRE domains are characterized by short DNA motifs that include binding sites for regulatory proteins like PLEIOHOMEOTIC (PHO), ZESTE (Z), and GAGA factors (GAFs) (Schuettengruber and Cavalli, 2009). PHO and GAFs would be directly involved in recruiting PRC2. Recently, PRE-like elements have also been identified in mice and humans (Sing et al., 2009; Woo et al., 2010). Nevertheless, our understanding of PRC2 recruitment is still scarce, and no PRE elements have been identified in plants to date (Schatlowski et al., 2008; Morey and Helin, 2010; Sawarkar and Paro, 2010; Smith and Shilatifard, 2010; Enderle et al., 2011; Margueron and Reinberg, 2011).

In *Arabidopsis thaliana*, the three E(Z) homologs, namely MEDEA (MEA) and the two partially redundant proteins CURLY LEAF (CLF) and SWINGER (SWN), act throughout plant development (Kiyosue et al., 1999; Luo et al., 1999; Pien and Grossniklaus, 2007; Hennig and Derkacheva, 2009; Köhler and Hennig, 2010; Rodrigues et al., 2010; Zheng and Chen, 2011). MEA predominantly represses the MADS box gene *PHERES1* and *MEA* itself in the endosperm (Baroux et al., 2006; Gehring et al., 2006; Jullien et al., 2006). CLF and SWN are required for the repression of floral homeotic genes, such as *AGAMOUS* (*AG*) or *SHOOTMERISTEMLESS* (*STM*) (Goodrich et al., 1997; Katz et al., 2004; Schubert et al., 2006; Farrona et al., 2008). PRC2 activity also represses the transcription of the master regulators of seed development during vegetative and/or endosperm development (Makarevich et al., 2006; Bouyer et al., 2011). These master regulators include LEAFY COTYLEDON1 (*LEC1*), a protein homologous to the HAP3 subunits of the CAAT Box binding Factors, and three B3-domain transcription factors, namely, ABA INSENSITIVE3 (*ABI3*), FUSCA3 (*FUS3*), and *LEC2* (Braybrook and Harada, 2008; Santos-Mendoza et al., 2008). These AFL (for *ABI3*, *FUS3*, and *LEC2*) regulatory proteins control a network of local and redundant gene regulations (Santos Mendoza et al., 2005; To et al., 2006). Among these regulators, *LEC2* is specifically expressed in embryos within a short temporal window, from 4 d after pollination (DAP) to 14 DAP, peaking at 10 DAP, and is not expressed in any other tissues (Stone et al., 2001; Kroj et al., 2003). The ectopic expression of *LEC2* in transgenic plants triggers the formation of somatic embryos as well as other organ-like structures and confers embryonic characteristics to seedlings (Stone et al., 2001).

Consistent with the involvement of proteins with a SET domain, other reports have confirmed that the expression of *LEC2* is controlled, in a developmental manner, by PcG complexes. It has been shown recently that FERTILIZATION INDEPENDENT ENDOSPERM, the *Arabidopsis* ESC homolog of the PRC2 complex, negatively regulates the expression *LEC2* and other AFL genes (Bouyer et al., 2011). PRC1-like ring finger proteins also repressed the expression of the AFL genes (Chen et al., 2010). In addition, PICKLE (PKL), a chromatin remodeler of the CHD3 family, and a closely related homolog (PKR2) are also involved in this regulation (Ogas et al., 1999; Zhang et al., 2008; Aichinger et al., 2009). PKL promotes indirectly H3K27me3 deposition at different loci, including *LEC2*, in vegetative tissues, through the induction of PcG gene expression (e.g., *SWN* or

*EMBRYONIC FLOWER2*) (Aichinger et al., 2009, 2011). Other mechanisms involved in the regulation of *LEC2* expression have been identified (Zhang and Ogas, 2009). For example, the inhibition of histone deacetylase activity triggers *LEC2* expression upon germination (Tanaka et al., 2008). ASIL1 (for *Arabidopsis* 6b-interacting protein 1-like 1), a plant-specific trihelix factor, represses *LEC2* during both germination and early maturation phase (Gao et al., 2009).

Besides the large amount of data accumulated regarding the regulation of *LEC2* expression, the DNA *cis*-elements involved in the specific expression of *LEC2* during seed development are unknown. More specifically, the means by which the *LEC2* locus is marked by H3K27me3 during vegetative development remain to be determined. In this work, we identified and functionally characterized three *cis*-regulatory elements involved in the developmental regulation of *LEC2* expression. Two of them were found to be essential for the transcriptional activation of *LEC2*. The first one presents some similarities with *CArG* regulatory elements recognized by MADS box transcription factors. The second one is similar to a GAFs box and bound by plant BASIC PENTACYSSTEINE (BPC) proteins, both in yeast and in vitro. The third element, which we named *Repressive LEC2 Element* (*RLE*), is 50 bp long and is necessary for the repression and H3K27 trimethylation of *LEC2* in vegetative tissues. Remarkably, the *RLE cis*-element was also sufficient to regulate the activity of an unrelated promoter ( $P_{F3H}$ ) and to promote H3K27me3 mark deposition in transgenic plants. Therefore, *RLE* acts as a functional homolog of the PRE element found in *Drosophila*. Taken together, the data presented here shed new light on the regulation of *LEC2*, a master regulator of seed development and bring important new insights into the regulation of H3K27me3 deposition and, thus, PRC2 recruitment in plants.

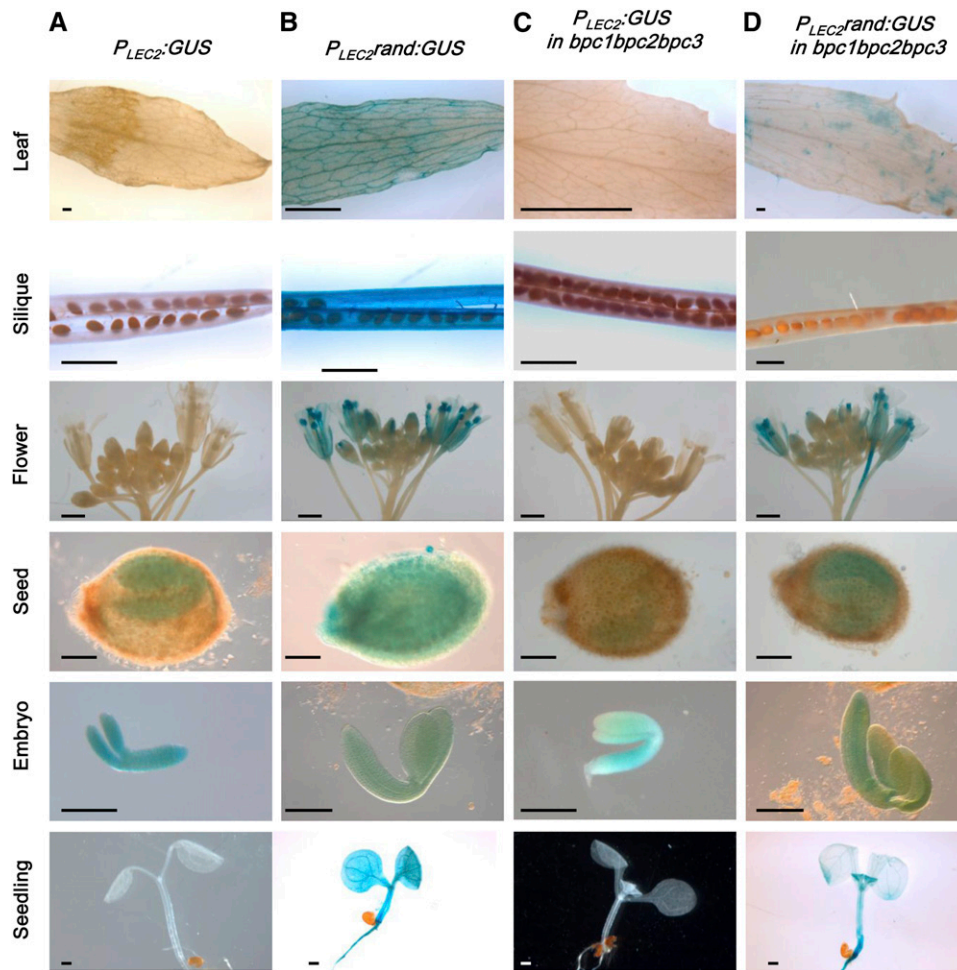
## RESULTS

### Identification of the *cis*-Regulatory Elements That Control the Specific Expression of *LEC2*

It has been previously shown that the 2.5-kb sequence located upstream of the *LEC2* open reading frame recapitulates the embryo-specific pattern of *LEC2* expression when fused to the  $\beta$ -glucuronidase (*GUS*) reporter gene (Stone et al., 2001; Kroj et al., 2003). Here, it was shown that a 500-bp-long fragment, hereafter denoted as  $P_{LEC2}$ , was sufficient to faithfully confer a similar expression pattern (Figure 1A). In addition, the expression of the *LEC2* cDNA under the control of  $P_{LEC2}$  fully complemented the *lec2-4* mutation (see Supplemental Figure 1 online).

To gain further insight into the *cis*-regulatory elements necessary for *LEC2* expression during embryogenesis, a series of 5'-end deletions was performed on  $P_{LEC2}$ , and the resulting DNA fragments were fused to the *GUS* reporter gene to test their respective transcriptional activity in transgenic plants (Figure 2A). The analysis of *GUS* activity identified a region between -447 and -400 bp upstream of the first codon that is necessary for the expression of *LEC2* in the embryo. Deletions every 5 bp narrowed this region down to a 5-bp motif (*CATAT*; between -447 and -442 bp). This 5-bp regulatory sequence showed





**Figure 1.** Expression Patterns Using the Wild Type and Mutated *LEC2* Promoter.

Representative GUS patterns are presented for *P<sub>LEC2</sub>::GUS* (A) and *P<sub>LEC2rand</sub>::GUS* (B) in the wild-type Columbia-0 background and for *P<sub>LEC2</sub>::GUS* (C) and *P<sub>LEC2rand</sub>::GUS* (D) in the *bpc1 bpc2 bpc3* triple mutant background, respectively. Bars = 1 mm in leaf, silique, and flower, and 0.1 mm in seed, embryo, and seedling. Independent transgenic lines ( $n > 12$  for each construct) were assayed for GUS activity in leaves, siliques, flowers, seeds, embryos, and seedlings (10 d after sowing).

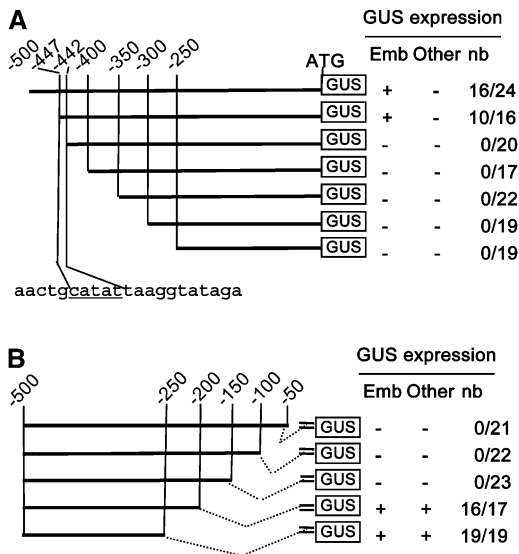
no similarity to any known consensus *cis*-regulatory element described in public databases.

In parallel, a series of 3'-end deletions of *P<sub>LEC2</sub>* was performed and fused to a minimal 35S promoter carrying a transcription start site derived from the cauliflower mosaic virus that was placed upstream of the *GUS* reporter gene (Debeaujon et al., 2003). Analysis of the GUS activity in transgenic plants indicated that the 50-bp region upstream of the first codon, which corresponds to the 5'-untranslated region of *LEC2*, was required for the activity of *P<sub>LEC2</sub>* in this context (Figure 2B). Surprisingly, further deletion of 150 bp resulted in the recovery of GUS expression. In fact, the promoter fragment extending from -500 to -250 bp was sufficient to confer ubiquitous GUS expression throughout plant and seed development. This analysis suggested that the 50-bp domain comprised between -200 and -150 bp contained sufficient information to repress the transcriptional activity of *P<sub>LEC2</sub>* (Figure 2B).

### ***LEC2* Expression Relies on a CT-Rich Element Bound by BPC Proteins**

In silico analysis of the 50-bp sequence located upstream of the first codon revealed a GAGA box motif (inverted CT repeats; Figure 3A) previously found in different plant promoters (Meister et al., 2004; Monfared et al., 2011; Wanke et al., 2011). The conversion by site-directed mutagenesis of CT to CA repeats (*mCA*) did not affect the transcriptional activity of *P<sub>LEC2</sub>* in planta, whereas changing CT to AT repeats (*mAT*) fully abolished the activity of *P<sub>LEC2</sub>* (Figure 3A). Previous analyses have shown the positive effect of adding CT repeats on promoter activity (Pauli et al., 2004). Here, the addition of three CT repeats to the GAGA box led to the ectopic activation of *P<sub>LEC2</sub>* in pollen grains (see Supplemental Figure 2 online). These data indicated that the GAGA box made of eight C-T (or A) dinucleotide repeats is essential for the correct expression of *LEC2*.





**Figure 2.** Molecular Dissection Analysis of the *LEC2* Promoter.

Schematic representation of the deletion series of the *LEC2* promoter ( $P_{LEC2}$ ) fused to the *GUS* reporter gene. The presence or lack of *GUS* expression is indicated by (+) or (-), respectively. The ratio (nb) indicates the number of transgenic lines expressing *GUS* (numerator) and the total number of transgenic lines analyzed (denominator). *GUS* expression was tested in embryos (emb), valves of siliques, flowers, rosettes, cauline leaves, and stems (other).

**(A)** The 5'-end deletions. The nucleotide sequence underlined is essential for  $P_{LEC2}$  activity.

**(B)** The 3'-end deletions were fused to a minimal promoter derived from the 35S of the cauliflower mosaic virus (=).

In plants, GAGA motifs are bound by BPC proteins (Meister et al., 2004; Monfared et al., 2011; Wanke et al., 2011). In *Arabidopsis*, the BPC family contains seven members grouped into three classes. The accumulation of BPC mRNA in embryos shown in the databases (<http://bbc.botany.utoronto.ca/efp/cgi-bin/efpWeb.cgi>) was confirmed by RT-PCR experiments (Figure 3B). All BPC genes but *BPC5* (which likely is a pseudogene) (Meister et al., 2004; Monfared et al., 2011) were expressed in 10-d-old embryos.

Then, the ability of BPC proteins to interact with  $P_{LEC2}$  was tested in yeast one-hybrid experiments (Figure 3C). A strong interaction was detected for the three BPCs of class I (BPC1, 2, and 3), whereas a weak interaction was detected for BPC proteins that belong to class II (BPC4) and III (BPC7). To define the ability of BPC to bind directly to the GAGA motif found in  $P_{LEC2}$ , in vitro electrophoretic mobility shift assays (EMSAs) were performed with recombinant BPC2 (class I) and BPC4 (class II) proteins. Both proteins BPC2 and BPC4 proteins were able to interact with the poly-CT double-stranded probe (Figure 3D). Competition experiments with wild-type and mutated (*mAT*) probes demonstrated the specificity of these interactions. Finally, the accumulation of *LEC2* mRNA was measured in seeds of available *bpc* mutants (i.e., *bpc1*, *bpc2*, and *bpc3* single or the corresponding multiple mutants). No significant differences were observed in comparison to wild-type seed mRNA level (see Supplemental Figure 3 online).

Taken together, these results showed that *LEC2* expression relies on a GAGA box and that BPC proteins are probably involved in the transcriptional regulation of *LEC2*. Further analyses of additional *bpc* mutants are required to determine which of the BPC proteins are involved. Interestingly, the functional analysis of a modified *LEC2* promoter, active in vegetative tissues (see the results presented below), strongly supported the role of the BPC proteins in the control of *LEC2* expression.

### Characterization of a *cis*-Regulatory Element (*RLE*) That Restricts *LEC2* Expression to the Embryo

To further characterize the domain located between 200 and 150 bp upstream of the first codon (Figure 2B), a new series of deletions fused to the 35S minimal promoter upstream of the *GUS* reporter gene was constructed and tested in transgenic plants (Figure 4). This experiment suggested that at least the sequences between -185/-178 and -168/-160bp were necessary to repress  $P_{LEC2}$  transcriptional activity in seed and vegetative tissues, respectively (Figure 4A). To supplement this deletion analysis, mutagenesis experiments were performed using the entire  $P_{LEC2}$  promoter. The 50-bp element associated with the repression activity was replaced by a random 50-bp nucleotide sequence, with no similarity within the *Arabidopsis* genome, providing the  $P_{LEC2}rand$  promoter. In comparison to the embryo-specific *GUS* staining observed with  $P_{LEC2}:GUS$  lines (Figure 1A), the  $P_{LEC2}rand:GUS$  lines showed ectopic expression in leaves, stems, flowers, pollen, siliques, and seed coat (Figures 1B and 4B). Similar results were obtained in the Columbia-0 background (see Supplemental Figure 4 online), and these results confirmed the role of the repressive element in the developmental regulation of *LEC2* expression. Thus, we named this element *RLE*. In addition, site-directed mutagenesis was performed to support and refine the structure of the repressive *cis*-regulatory element. Some of the mutations (*m1* and *m3*) did not modify the pattern of *GUS* activity, whereas other mutations (*m2* and *m4*) led to the ectopic activity of *GUS* in pollen (Figure 4B).

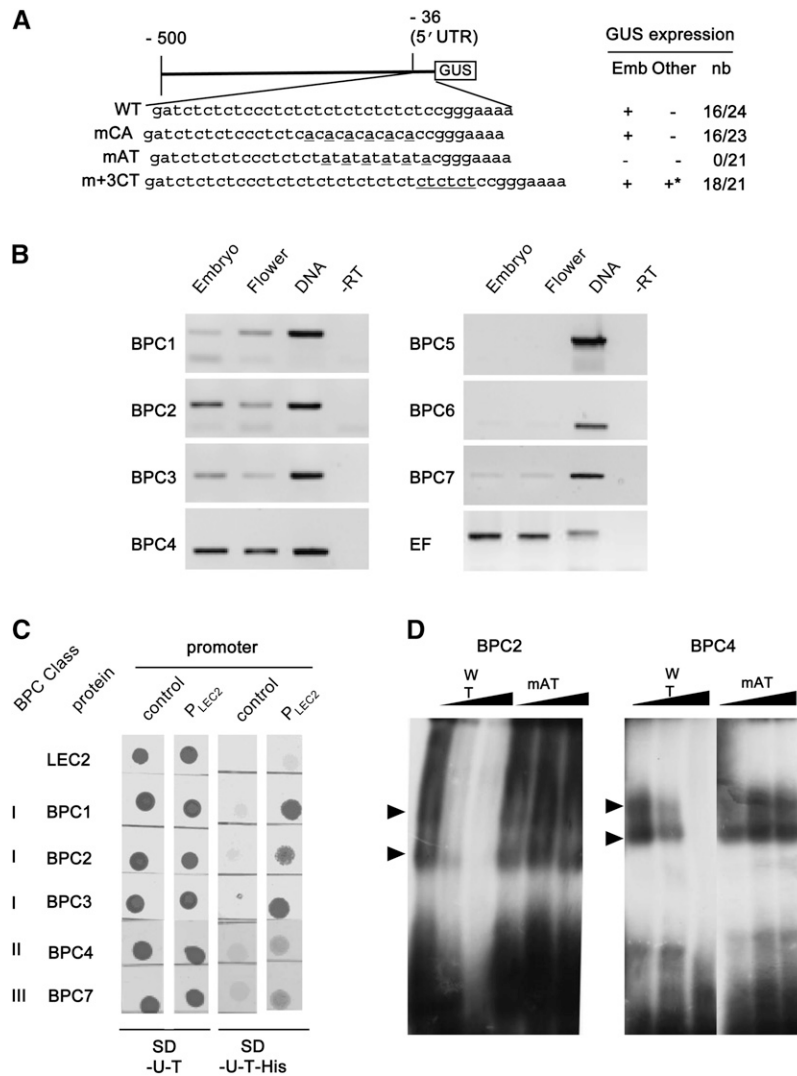
Taken together, these results indicated that the 50-bp element located between -150 and -200 bp upstream of the first codon has a negative effect on the transcriptional activity of  $P_{LEC2}$  in planta. In addition, they suggested that *RLE* likely contains different functional motifs that are necessary to confer this repressive activity.

Finally, to test if the BPC proteins are involved in the ectopic activity of  $P_{LEC2}$  when *RLE* is mutated,  $P_{LEC2}rand:GUS$  was introduced into the triple *bpc1 bpc2 bpc3* mutant by transformation. In this genetic background, *GUS* activity was completely lost in siliques and seed integument and was strongly down-regulated in leaves and cotyledons, in comparison to the wild type (Figure 1C). These results strongly suggested that class I BPCs were involved in the regulation of *LEC2* expression.

### The *RLE* Element Identified in $P_{LEC2}$ Is Necessary for H3K27me3 Deposition

Previous reports have shown that PRC2, which catalyzes H3K27 trimethylation, is involved in the transcriptional repression of





**Figure 3.** The Promoter of *LEC2* Contains a CT-Rich Repeat Bound by BPC Proteins.

**(A)** The wild-type (WT) *P<sub>LEC2</sub>* sequence is shown in the top row (WT). Mutated nucleotides are underlined in the next three rows (mCA, mAT, and m+3CT). GUS activity in pollen is indicated with an asterisk. Presence or lack of GUS expression is indicated by (+) or (-), respectively. The ratio (nb) indicates the number of transgenic lines expressing GUS (numerator) and the total number of transgenic lines analyzed (denominator). GUS expression was tested in embryos (emb), valves of siliques, flowers, rosettes, cauline leaves, and stems (other). UTR, untranslated region.

**(B)** Analysis of BPC expression in embryos and flowers. Genomic DNA, embryo RNA untreated by the superscript II enzyme (-RT), and *EF1 $\alpha$ A4* (EF) were used as control.

**(C)** One-hybrid analysis. The interactions of the different full-length BPC proteins, fused to GAL4, with *P<sub>LEC2</sub>* were tested. The ability to grow in the absence of His indicates an interaction.

**(D)** EMSA competition assays. BPC2 and BPC4 proteins were tested with labeled wild-type CT repeat probe, and competitions were performed with wild-type or mutated CT repeat (mAT) unlabeled probes. Shifts are indicated with arrowheads.

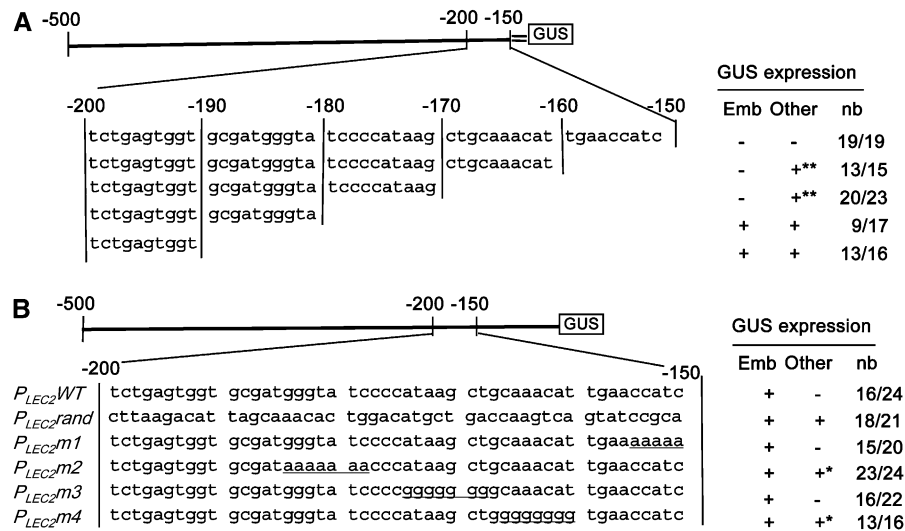
*LEC2* in seedlings (Makarevich et al., 2006; Chen et al., 2010; Bouyer et al., 2011). To determine if the *RLE* element identified in *P<sub>LEC2</sub>* was involved in this chromatin-based regulatory process, H3K27me3 enrichment was investigated by chromatin immunoprecipitation (ChIP) and subsequently analyzed by quantitative PCR. The H3K27me3 mark was detected over the *LEC2* locus, in the promoter and more strongly right downstream of the first

codon. These results are consistent with the data obtained by genome-wide analyses (Zhang et al., 2007b; Roudier et al., 2011).

Then, the level of H3K27me3 at the *P<sub>LEC2</sub>:GUS* and *P<sub>LEC2</sub>rand:GUS* loci was investigated in 10-d-old seedlings of transgenic plants. The distribution of H3K27me3 at the *P<sub>LEC2</sub>:GUS* loci was similar to the pattern found at the endogenous *LEC2* locus







**Figure 4.** Molecular Dissection Analysis of the *RLE* Domain.

Presence or absence of *GUS* activity is indicated by (+) or (–), respectively. The ratio to the right (nb) indicates the number of transgenic lines that display *GUS* activity (numerator) and the total number of transgenic lines analyzed (denominator). *GUS* activity was tested in embryos (emb), valves of siliques, flowers, rosettes, cauline leaves, and stems (other). The detection of *GUS* activity in pollen (+\*) or in all organs with the exception of the seed coat and embryo (+\*\*) is indicated.

**(A)** Analysis using 3′-end deletion constructs fused to the minimal 35S promoter (=) and to the *GUS* reporter gene.

**(B)** Analysis using modifications of the *RLE* domain by replacement with a random 50-bp sequence (*P<sub>LEC2</sub><sup>rand</sup>*) or by site-directed mutagenesis (*P<sub>LEC2</sub><sup>m1</sup>*, *m2*, *m3*, and *m4*).

(Figure 5A). Importantly, the level of H3K27me3 over the *P<sub>LEC2</sub><sup>rand</sup>:GUS* locus was significantly lower than over *P<sub>LEC2</sub>:GUS* locus (at  $P < 0.002$ , according to a Student's unpaired *t* test). The level was between 45- and 85-fold lower, depending on the probes considered (Figure 5B). The analyses were performed using several ( $n > 10$ ) independent transgenic lines for each construct. Therefore, the result indicated that the difference between *P<sub>LEC2</sub><sup>rand</sup>:GUS* and *P<sub>LEC2</sub>:GUS* loci was strictly dependent on the presence of *RLE cis*-element, independent of the transgene insertion sites within the genome. These observations therefore demonstrated that H3K27me3 enrichment over *LEC2* depends on the *RLE* sequence. Taken together, these data and the ectopic *GUS* activity detected in the *P<sub>LEC2</sub><sup>rand</sup>:GUS* transgenic plants (Figure 1B) strongly suggest that the *RLE cis*-element is a major determinant of PRC2-mediated repression of *LEC2* outside of embryogenesis (Figure 1B).

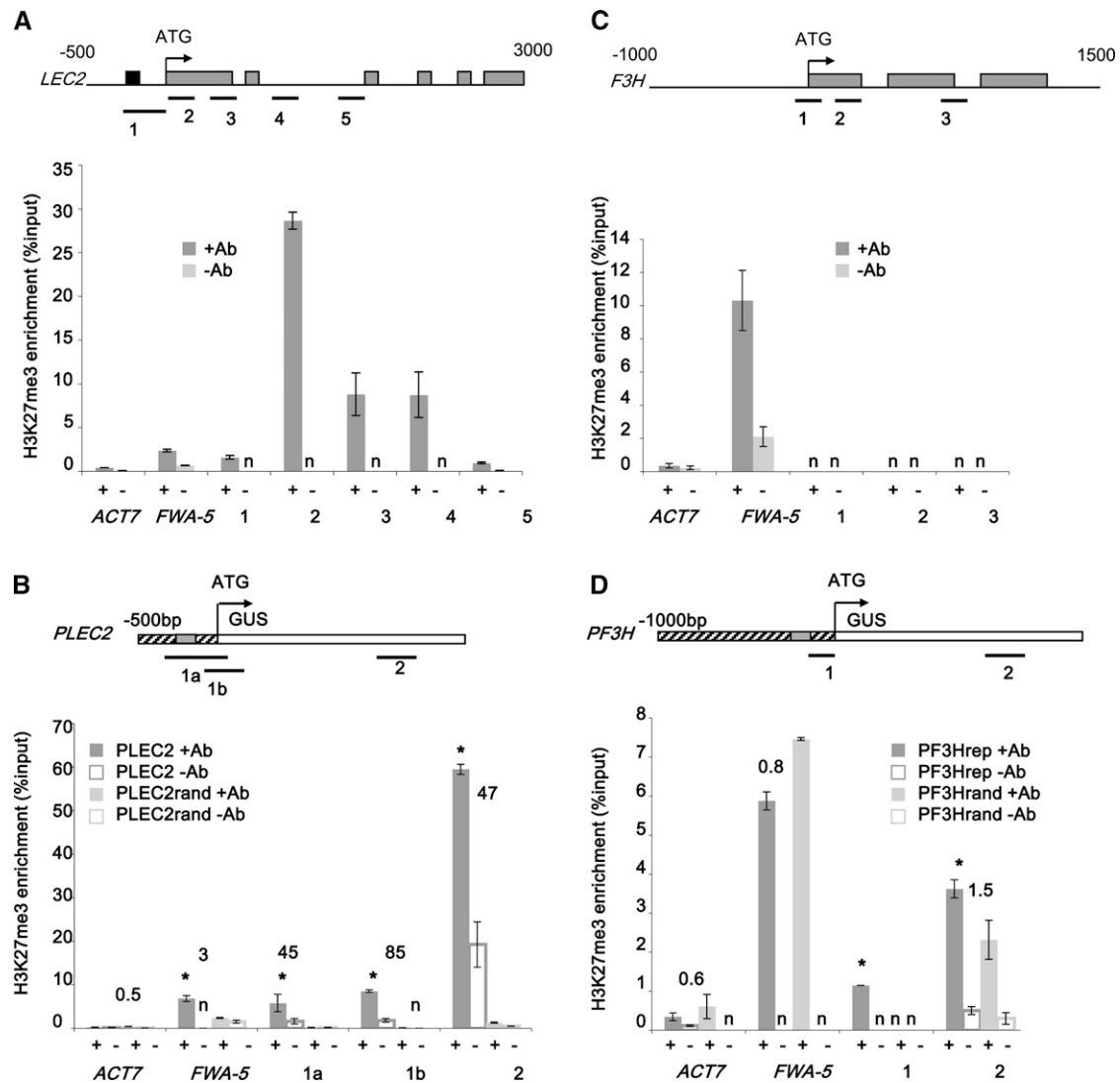
#### Insertion of *RLE* in pF3H, an Unrelated Promoter, Is Sufficient to Repress Its Transcriptional Activity in Seed and Can Trigger H3K27me3 Deposition

The repressive function and the role of *RLE* in the initiation of H3K27me3 marking were tested by a gain-of-function experiment with an unrelated promoter. The *RLE* sequence or a mutated version (random) were introduced at similar position (i.e., 150 bp upstream of the first codon) into the promoter of the *FLAVONONE 3-HYDROXYLASE (F3H)* gene to generate *P<sub>F3H</sub><sup>rle</sup>:GUS* and *P<sub>F3H</sub><sup>rand</sup>:GUS* constructs, respectively. The *F3H* promoter was chosen for two reasons. First the *cis*-regulating

elements that control its activity have been described, and no motifs were located at the insertion site (Hartmann et al., 2005). Furthermore, no H3K27me3 was detectable in ChIP-quantitative PCR analyses (Figure 5C), in agreement with previous analyses (Zhang et al., 2007b). Similar analyses were performed to establish the H3K27me3 landscape at the *P<sub>F3H</sub><sup>rle</sup>:GUS* and *P<sub>F3H</sub><sup>rand</sup>:GUS* loci in transgenic seedlings. It revealed a weak enrichment of H3K27me3 at *P<sub>F3H</sub><sup>rle</sup>:GUS* loci when compared with *P<sub>F3H</sub><sup>rand</sup>:GUS* (Figure 5D).

The effect of *RLE* insertion on *P<sub>F3H</sub>* transcriptional activity was investigated. The endogenous *F3H* promoter was fused to the *GUS* reporter gene (*P<sub>F3H</sub>:GUS*) and introduced into *Arabidopsis* as a control. *GUS* activity was analyzed in transgenic plants carrying one of the three different constructs, namely, *P<sub>F3H</sub>:GUS*, *P<sub>F3H</sub><sup>rle</sup>:GUS*, and *P<sub>F3H</sub><sup>rand</sup>:GUS*. For each construct, 24 independent transformants were analyzed. Whereas a strong *GUS* activity was detected in seeds of 19 transgenic lines carrying *P<sub>F3H</sub>:GUS* and 17 *P<sub>F3H</sub><sup>rand</sup>:GUS* lines, it was undetectable in seeds of the 14 *P<sub>F3H</sub><sup>rle</sup>:GUS* transgenic lines showing *GUS* activities in other tissues (Figure 6). These results demonstrated that the insertion of *RLE* is sufficient to repress *P<sub>F3H</sub>* transcriptional activity in the seed (embryo and seed coat). On the contrary, similar *GUS* activity was found in vegetative tissues for the three constructs, suggesting that H3K27me3 marking was not sufficient to fully inhibit gene expression in these tissues. Taken together, these observations demonstrate that the *RLE cis*-element is sufficient to trigger H3K27me3 deposition and control transcriptional activity in a developmentally regulated manner.





**Figure 5.** ChIP-PCR Analysis of H3K27me3 Enrichment over the Endogenous and Transgenic *LEC2* and *F3H* Loci.

The H3K27me3 enrichments in 10-d-old seedlings at *LEC2* (A), *P<sub>LEC2</sub>:GUS* (B), *F3H* (C), and *P<sub>F3H</sub>:GUS* (D) loci are presented. The enrichment is expressed relative to chromatin input (%input). All the experiments were repeated at least twice on a pool of five independent transgenic lines. *ACT7* and *FWA* genes were used as negative and positive controls, respectively. Bars denote SE, and n = not detectable. For (A) and (C), the positions of the primers used (thick lines), *RLE* element (black box), introns (thin line), and exons (dark boxes) are indicated on the top. For (B) and (D), the positions of the primers (thick lines), *LEC2* promoter (striped box), *RLE* element (dark box), and *GUS* coding sequence (white box) are indicated on the top. The numbers above the histogram bars represent the ratio of the H3K27me3 enrichment when comparing constructs with and without the *RLE* element (i.e., *P<sub>LEC2</sub>* with *P<sub>LEC2</sub>Rand* [B] and *P<sub>F3H</sub>* with *P<sub>F3H</sub>Rand* [D]). The asterisks indicate that the difference (with or without *RLE*) is significant at P < 0.002 according to Student's unpaired t test.

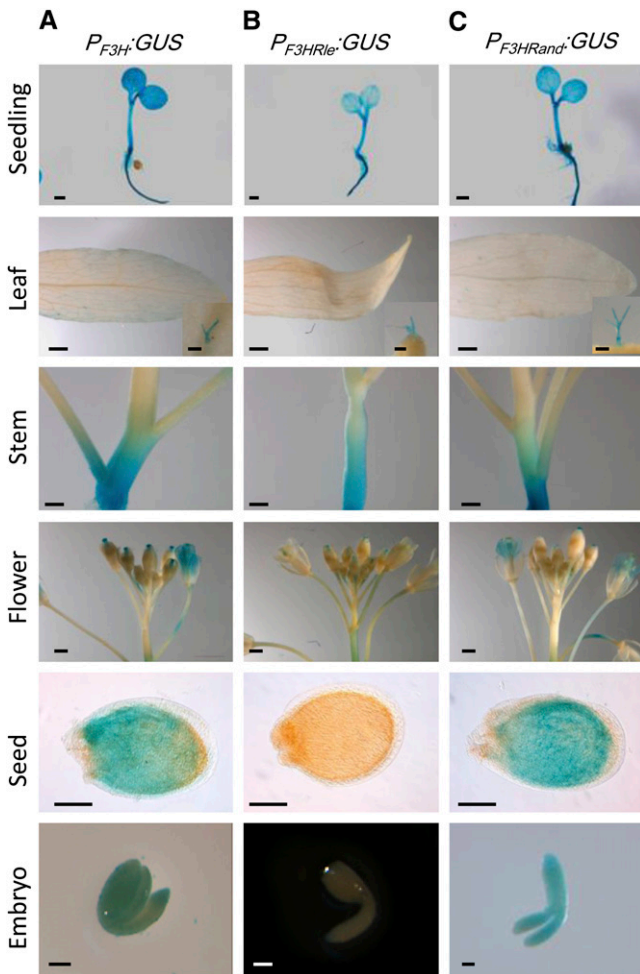
## DISCUSSION

### Both a *CArG* and a CT-Rich Activating *cis*-Element Are Essential for *P<sub>LEC2</sub>* Activity

The molecular and functional dissection of the *LEC2* promoter in planta revealed two essential activating sequences that are necessary for embryo-specific expression (Figure 7). The first *cis*-regulatory element is made of five nucleotides (5'-CATAT-3')

and located 447 bp upstream of the translational start codon. This very short motif did not present any obvious similarities with known *cis*-regulatory sequences. Nevertheless, when a wider frame is considered (i.e., GCATATTAAGG), a similarity was found with the consensus *CArG* motif [i.e., CC(A/T)<sub>6</sub>GG] (Riechmann and Meyerowitz, 1997). The *CArG* motifs are known to be recognized by MADS box proteins, a family of transcription factors that control multiple developmental processes, ranging from root to floral or embryo development (Becker and Theissen, 2003).





**Figure 6.** Insertion of the *RLE* Element Inhibits  $P_{F3H}$  Activity in Seed.

Representative patterns of GUS staining are presented for  $P_{F3H}$ :GUS (A),  $P_{F3H}Rle$ :GUS (B), and  $P_{F3H}Rand$ :GUS (C). Independent transgenic lines ( $n = 24$  for each construct) were analyzed for GUS activity in seedling, leaf, trichome, stem, flower, seed, and embryo. Bars = 1 mm for seedling, leaf, and flower and 0.1 mm for trichome, seed, and embryo.

Interestingly, the AGAMOUS-like15 (AGL15) MADS box protein preferentially binds *CArG* motifs with longer *A/T*-rich region, such as  $C(A/T-rich)_8G$  and  $C(A/T-rich)_7GG$  sequences (Tang and Perry, 2003; Hill et al., 2008; Zheng et al., 2009). This later motif is similar to the *CArG* sequence found in  $P_{LEC2}$  [i.e.,  $C(A/T)_7GG$ ]. In addition, genetic and genome-wide ChIP analyses have demonstrated that *LEC2* is a direct target of AGL15 (Zheng et al., 2009). Taken together, these results strongly suggested that the  $C(A/T-rich)_7gg$  motif found in  $P_{LEC2}$  is one of the AGL15 binding sites involved in the regulation of *LEC2* expression. Furthermore, AGL15 can interact with histone deacetylase HDA19 (Hill et al., 2008) and thus could be involved in histone deacetylase recruitment at the *LEC2* locus.

The second *cis*-regulatory element essential for *LEC2* expression is a *CT*-rich domain (containing eight *CT* repeats) located in

the 5'-untranslated region of *LEC2* (Figure 3A). This domain displays strong similarities with *GAGA* elements (inverted *CT* repeats). Consistent with functional data previously obtained (Meister et al., 2004; Kooiker et al., 2005), the results presented here demonstrated that the *CT*-rich element of  $P_{LEC2}$  is bound by BPC proteins (preferentially of class I), both in yeast and in vitro. The quantitative differences in DNA binding observed between the BPC in vitro and yeast (Figures 3C and 3D) may be related to the target sequences used [i.e., the entire 500-bp-long promoter for the one-hybrid experiment and a short repetition of three *CT* motifs,  $3 \times (CT)_8$ , in vitro] (see Supplemental Table 1 online). It has been shown that, in vitro, BPC1 binding was affected by modification of the stretch of purine residues (i.e., by introducing any pyrimidine residue) (Kooiker et al., 2005). Our results suggested that, in planta, the replacement of cytosine by adenine residues strongly affected  $P_{LEC2}$  activity (Figure 3A). This suggested that the strong base pairing between complementary C-G nucleotides is important for the function of the *CT* element. Interestingly, the addition of three extra *CT* repeats was sufficient to trigger ectopic  $P_{LEC2}$ :GUS activity in pollen grains (see Supplemental Figure 2 online). This result was consistent with evidence indicating that *CT* repeats have an enhancer effect on promoter activity (Pauli et al., 2004). The three additional *CT* repeats inserted in  $P_{LEC2}$  may increase transcriptional activation, up to the threshold of detection of GUS activity, in pollen. A second nonexclusive possibility is that the addition of these repeats increased the affinity for another BPC protein expressed in pollen. This later hypothesis is supported by the fact that the promoter of *BPC2*, *BPC3*, *BPC4*, and *BPC7* are all active in pollen (Monfared et al., 2011).

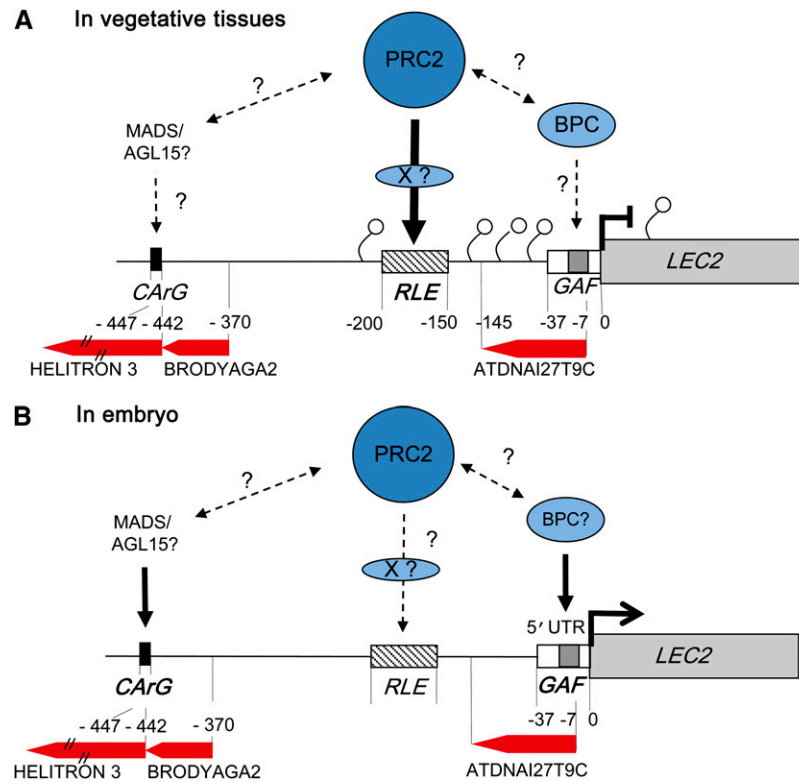
The *BPC* mRNAs of class I (*BPC1*, 2, and 3) together with *BPC4* (class II) were the most abundant in embryo (Figure 3). Other *BPC* mRNAs were also detected (with the exception of *BPC5*) but at very low levels. These expression profiles together with the results of DNA binding assays suggested that BPC factors are likely to regulate *LEC2* through binding to the *GAGA cis*-element found in  $P_{LEC2}$ . Consistent with this hypothesis, the ectopic GUS activity driven by  $P_{LEC2}rand$ :GUS was fully abolished in siliques and seeds integuments and was strongly affected in leaves and cotyledons of the triple class I mutant (Figure 1C).

Altogether, these results suggested that BPC proteins that belong to class II or III act redundantly with class I BPCs to activate *LEC2* expression through its specific *GAGA cis*-regulatory motif. This hypothesis is in agreement with the high genetic redundancy recently reported for BPCs (Monfared et al., 2011). Moreover, it has been shown that BPCs of group II can form homodimers or heterodimers with group I BPCs (Wanke et al., 2011). Therefore, obtaining new *bpc* single and multiple mutants of the different classes will be necessary to unravel which of the BPC proteins are involved in the regulation of *LEC2* expression.

#### ***RLE* Is Sufficient to Repress Transcriptional Activity and Is Necessary to Trigger H3K27me3 Deposition**

The molecular and functional dissection of  $P_{LEC2}$  presented here has identified a new and negative regulatory *cis*-element (*RLE*) involved in restricting the expression of *LEC2* to the seed (Figure 7). This 50-bp-long *cis*-acting sequence located 150 bp





**Figure 7.** Schematic Structure and Regulation of the *LEC2* Promoter.

Putative model of regulation in vegetative tissues (A) and in the embryo (B). The *CARG* box, *GAF* box, and *RLE* element are represented as well as the transposon marks (*MuDR* and *HEL:Helitron*) and putative functional interactions between the proteins. Small circles represent H3K27me3 marks.

upstream of the *LEC2* first codon is required for repressing *LEC2* expression in leaves, stems, flowers, pollen, seed coat, and siliques (Figures 1 and 2B). Additional experiments (i.e., deletions of 10 bp and site-directed mutagenesis approaches) confirmed these observations and suggested that *RLE* is made of different functional stretches of nucleotides that are necessary to confer this repressive activity (Figure 4B).

Remarkably, the presence of *RLE* in *P<sub>LEC2</sub>:GUS* is sufficient to trigger H3K27me3 deposition in transgenic seedlings (Figures 5C and 5D). In addition, the insertion of *RLE* into the *F3H* promoter (*P<sub>F3H</sub>RLE*) was sufficient to inhibit its transcriptional activity in both the seed coat and embryo but not in vegetative tissues after germination (Figure 6B). The detection of GUS activity in vegetative tissues is not totally unexpected. Indeed, it may be related to the environmental regulation of *F3H* (Pelletier and Shirley, 1996; Blödner et al., 2007; Dubos et al., 2008; Zhang et al., 2011) and to counteracting chromatin-based regulation (Liu et al., 2010; Schmitges et al., 2011). Consistent with this hypothesis, genome-wide analyses showed that in seedlings the *F3H* endogenous locus is largely enriched in H3K4me3, a mark usually associated with transcriptionally active chromatin (Roudier et al., 2011). In conclusion, the *RLE cis*-element is necessary to trigger H3K27me3 enrichment and concomitant inhibition of transcriptional activity, although additional factors are likely involved in maintaining or alleviating this repression.

### **RLE Functions as PREs Identified in *Drosophila***

It has been shown that PcG complexes that control the H3K27me3 level are involved in the transcriptional repression of *LEC2* in seedlings (Makarevich et al., 2006; Chen et al., 2010; Bouyer et al., 2011). Our results demonstrated that the *RLE* element found in *P<sub>LEC2</sub>* is sufficient to initiate H3K27me3 deposition at the *LEC2* locus. In addition, some PREs have been found in the proximal promoter of target genes in *Drosophila* (Morey and Helin, 2010; Enderle et al., 2011), as *RLE* in *LEC2*. Last, it has been demonstrated recently that H3K27me3 deposition is abolished in *fie* mutant lacking PRC2 activity (Bouyer et al., 2011). These results suggest that the *RLE* element functions as *Drosophila* PRE. Nevertheless, the weak effect of *RLE* on H3K27me3 deposition at *pF3H* suggests that it may not be sufficient in any genomic contexts.

PREs contain a few short transcription factor binding sites, including a GAGA box (Hagstrom et al., 1997; Horard et al., 2000). Similarly, *RLE* is associated with a CT-rich element in *P<sub>LEC2</sub>*. In addition, some other plant promoters that are regulated by PcG also display CT-rich elements and are regulated by BPC proteins. For instance, *SEEDSTICK* and *INNER NO OUTER (INO)* are two targets of BPC1 and BPC2, respectively (Villanueva et al., 1999; Meister et al., 2004; Kooiker et al., 2005). Both loci are associated with strong H3K27me3 enrichment (Zhang et al., 2007b; Roudier et al., 2011). GA repeats can be found in the





regulatory sequences of many other PcG targets, such as *AG*, *FLOWERING LOCUS C (FLC)*, and *STM* (Schatlowski et al., 2008). In *Drosophila*, the *GAGA* factors were shown to be involved in the regulation of gene expression in relation with chromatin packaging (Adkins et al., 2006; Nakayama et al., 2007). Therefore, it is very tempting to speculate that there is an interaction between BPC and PcG silencing in plants. The availability of new and multiple *bpc* mutants will give the opportunity to further investigate this hypothesis, by looking at the chromatin landscape, at different loci containing both type of elements, such as *LEC2*, but also *INO*, *STM*, *FLC*, or *STM*.

No sequences showing similarity with *RLE* were detected genome-wide. This is not surprising considering that *PREs* are made of different short transcription factor binding sites. These motifs are usually small, degenerate, and difficult to detect by simple DNA sequence comparison (Ward and Bussemaker, 2008; Badis et al., 2009; Moyroud et al., 2011). Additional functional analyses, similar to those presented in this report, will be necessary to establish a model allowing a more systematic identification of putative *RLE* elements in plant genomes. Sequence analysis of *P<sub>LEC2</sub>* revealed the presence of three transposable elements (Figure 7). The transcription start site and the *GAGA* box are included in a class II transposable element sequence of the MuDR superfamily (Lisch, 2002). The *CArG* motif is located at the beginning of a HELITRON element (Kapitonov and Jurka, 2007; Yang and Bennetzen, 2009), which is itself juxtaposed with another MuDR sequence. Finally, *RLE* is flanked by the two MuDR sequences. It can be reasonably speculated that insertion of these transposable elements likely shaped the specific expression pattern of *LEC2*. It has been shown that transposons can be involved in transcriptional regulation of nearby genes, including *PRE*, although the precise mechanisms remain unknown and seem to be rather gene specific (Valenzuela and Kamakaka, 2006; Weil and Martienssen, 2008; Comet et al., 2011). In addition, selected transposons are specifically targeted by the PCR2 complex in *Arabidopsis* endosperm (Weinhofer et al., 2010). However, the hypothesis that transposons have a general function in recruiting or regulating PcG activity in plants requires further experimental data.

To conclude, the data presented here demonstrate that the developmental regulation of *LEC2* expression results from a combination of activating and repressive mechanisms involving histone H3 methylation. This regulation involves at least three distinct *cis*-regulatory elements located within the proximal *P<sub>LEC2</sub>* promoter, including repressive *RLE* element. Remarkably, *RLE* is sufficient to initiate H3K27me<sub>3</sub> deposition and to repress transcription in a developmentally regulated manner. This study provides important new insights into the mechanism by which PRC2 might be recruited at specific target loci and paves the way for a more comprehensive understanding of PcG-mediated gene silencing in plants.

## METHODS

### Plant Material and Growth Conditions

*Arabidopsis thaliana*, accession Wassilewskija, was used for plant transformation. The mutant lines *bpc1* (SALK\_072966.43.30.x), *bpc2*

(SALK\_090810C), and *bpc3* (N89719) were obtained from the Nottingham Arabidopsis Stock Centre (Scholl et al., 2000) and have all been described by Monfared et al. (2011). Plants were grown, transformed, and selected as previously described (Harscoët et al., 2010).

### Constructs and Plant Transformation

All the primers used in this work are described in Supplemental Table 1 online. For 5'-end deletions of the *LEC2* promoter (*P<sub>LEC2</sub>*), genomic DNA was PCR amplified with the Phusion DNA polymerase (Finnzymes) using primers carrying B1 and B2 GATEWAY recombination sites (Invitrogen). PCR products were cloned by BP recombination (according to the manufacturer's recommendations) into the entry vector pDONR207 and then LR recombined into the destination vector pBI101R1R2 for plant transformation (Baudry et al., 2004). Modified CT repeat constructs were performed similarly using 3' primers as follows: AttB2mCA, AttB2mAT, and AttB2m+3CT.

The 3'-end deletion series constructs were PCR amplified using 5' and 3' oligonucleotides carrying *KpnI* and *XhoI* restriction sites, respectively. The 5' primer used was P500*KpnI*, and the 3' primers were P250*XhoI*, P200*XhoI*, P190*XhoI*, P180*XhoI*, P170*XhoI*, P160*XhoI*, P150*XhoI*, P100*XhoI*, and P50*XhoI*. PCR products were digested with *KpnI-XhoI* and cloned into *KpnI-XhoI*-digested pBS35sminGUS (Debeaujon et al., 2003). *P<sub>LEC2</sub>* fragments fused to the minimum 35S promoter and the *UidA* reporter gene were transferred into pCAMBIA1300 linearized by *HindIII-KpnI* digestion and ligated with T4 DNA ligase (New England Biolabs).

For the *P<sub>LEC2</sub>rand* construct, the fragment P500-200 was amplified with the Phusion DNA polymerase using the 5' primer P500*KpnI* and the 3' primer P200rand*XhoI*. The sequence (5'-CTTAAGACATTAGCAAA-CACTGGACATGCTGACCAAGTCAGTATCCGCAT-3') was randomly chosen, and in silico analysis was performed to ensure that no known *cis*-regulatory elements were found in the sequence. The PCR product was digested with *KpnI-XhoI* and cloned into *KpnI-XhoI*-digested pBS35sminGUS (Debeaujon et al., 2003). The minimum 35S promoter was discarded after *XhoI-NcoI* digestion of the p200rand-pBS35sminGUS and replaced by the PCR amplified fragment between -150 and the translational start site, with P150*XhoI* and PATGNcoI 5' and 3' primers, respectively. The *F3H* promoter sequence was PCR amplified with Phusion DNA polymerase and primers AttB1PF3H and AttB2PF3H carrying B1 and B2 GATEWAY recombination sites. *P<sub>F3H</sub>ire* and *P<sub>F3H</sub>rand* were amplified with three successive PCRs using AttB1PF3H/PF3Hrev1, AttB1PF3H/PF3Hrev2, and AttB1PF3H/AttB2PF3H and were cloned using the approach described for wild-type *PF3H*. All these promoter fragments were recombined in the entry vector pDONR207 and then LR recombined into pBI101R1R2 for plant transformation, as described above.

### One-Hybrid Assays

Reporter constructs containing the different promoters were made in the pHIS-i plasmid (Clontech), which carries the *URA3* selection marker and the minimal *HIS3* promoter upstream of the *HIS3* reporter gene (Li and Herskowitz, 1993). The 500-bp-long *LEC2* promoter (*P<sub>LEC2</sub>*) was PCR amplified with the Phusion DNA polymerase with P500*KpnI* and PATGNcoI 5' and 3' primers, respectively, and blunt cloned into pHIS-i digested with *SmaI*. The control (-500 to -250 fragment of the *LEC2* promoter) was cloned using the same method with 5' primer P500*KpnI* and the 3' primer P250*LEC2* (see Supplemental Table 1 online). *LEC2* and *BPCs* cDNAs were amplified with the Phusion DNA polymerase using primers carrying B1 and B2 GATEWAY recombination sites. PCR products were BP recombined according to the manufacturer's (Invitrogen) recommendations into pDONR207 and then LR recombined into pDEST22 vector that allows fusion with the GAL4-activating domain. pHIS-i constructs



were stably transformed into the EGY 48 yeast (*Saccharomyces cerevisiae*) strain and selected on SD-Uracil medium prior to transformation with pDEST22 vectors for protein expression. The interactions were tested on SD-Uracil-Trp-His medium. To increase the stringency of the test, various concentrations of 3-aminotriazol were used, from 5 to 50 mM.

### EMSA Experiments

The cDNAs obtained for one-hybrid experiments were introduced into pDEST15 (Invitrogen), which produces an N-terminal glutathione S-transferase tag fusion. BPCs proteins were produced in *Escherichia coli* and purified with the Glutathione Sepharose 4B system (Amersham Biosciences) according to the manufacturer's instructions. One hundred nanograms of both forward and reverse primers (EMSACT/mAT forw and rev) were boiled for 5 min and annealed at room temperature. Double-stranded probes were <sup>32</sup>P labeled after Klenow treatment (Invitrogen). Binding conditions have been previously described (Kooiker et al., 2005).

### Histochemical GUS Assays

*Arabidopsis* tissues were immersed in GUS staining solution (100 mM sodium phosphate buffer, pH 7.2, 10 mM EDTA, 0.1% Triton, and 1 mM X-GlcA) and then placed under vacuum for 1 h. After incubation at 37°C overnight, the staining solution was removed and samples were cleared by sequential changes of 75 and 95% (v/v) ethanol.

### ChIP Experiments

The experiments were performed as previously described (Gendrel et al., 2002). Briefly, all the experiments have been duplicated and performed using pools of five independent transformants. Relative enrichment was measured by comparison between input and ChIP value. *ACT7* was used as a negative control (H3K27me3 modification is barely detectable over the *ACT7* locus), and *FWA* was used as a positive control for H3K27me3 enrichment (Roudier et al., 2011). Primers used in this analysis are described in Supplemental Table 1 online.

### Analysis of Gene Expression

RNA extractions were performed using the total mammalian RNA kit (Sigma-Aldrich) according to the manufacturer's instructions. cDNAs were synthesized using the SuperScriptII first-strand synthesis system for RT-PCR (Invitrogen) and dT oligonucleotides. The PCR experiments were conducted as previously described (Baud et al., 2005), with specific primers described in Supplemental Table 1 online. The data presented are the results of three biological and three technical repeats.

### Accession Numbers

Sequence and mutant data from this article can be found in the GenBank/EMBL data libraries or the Arabidopsis Genome Initiative database under the following accession numbers: *bpc1* (SALK\_072966.43.30), *bpc2* (SALK\_090810C), *bpc3* (N89719), *BPC1* (At2g01930), *BPC2* (At1g14685), *BPC3* (At1g68120), *BPC4* (At2g21240), *BPC5* (At4g38910), *BPC6* (At5g42520), *BPC7* (At2g35550), *LEC2* (At1g28300), *ACT7* (At5g09810), *FWA-5* (At4g25530), and *EF1α4* (AT5G60390).

### Supplemental Data

The following materials are available in the online version of this article.

**Supplemental Figure 1.** Complementation of the *lec2-4* Mutation by *P<sub>LEC2</sub>:LEC2*.

**Supplemental Figure 2.** Activation of *PLEC2* in Pollen by Adding 3CT Repeats in the GAF Box of *P<sub>LEC2</sub>*.

**Supplemental Figure 3.** *LEC2* Expression Is Not Modified in *bpc* Mutants.

**Supplemental Figure 4.** Expression Patterns of *PLEC2:GUS* and *PLEC2rand:GUS* in the Col-0 Ecotype Background.

**Supplemental Table 1.** List of Primers Used in This Study.

### ACKNOWLEDGMENTS

Part of this work was supported by grants from the Agence Nationale de la Recherche to L.L., namely, TF-code (BLAN-07-185566) and CERES (BLAN-10-123801). We thank Charles S. Gasser for providing *bpc* mutant seeds. Our laboratory belongs to the Laboratoire d'Excellence Saclay Plant Sciences (<https://www6.inra.fr/saclay-plant-sciences>).

### AUTHOR CONTRIBUTIONS

N.B., B.D., and C.D. performed the experiments. All the authors designed the research, analyzed data, and wrote the article.

Received May 31, 2011; revised October 11, 2011; accepted October 30, 2011; published November 11, 2011.

### REFERENCES

- Adkins, N.L., Hagerman, T.A., and Georgel, P. (2006). GAGA protein: A multi-faceted transcription factor. *Biochem. Cell Biol.* **84**: 559–567.
- Aichinger, E., Villar, C.B., Di Mambro, R., Sabatini, S., and Köhler, C. (2011). The CHD3 chromatin remodeler PICKLE and polycomb group proteins antagonistically regulate meristem activity in the *Arabidopsis* root. *Plant Cell* **23**: 1047–1060.
- Aichinger, E., Villar, C.B., Farrona, S., Reyes, J.C., Hennig, L., and Köhler, C. (2009). CHD3 proteins and polycomb group proteins antagonistically determine cell identity in *Arabidopsis*. *PLoS Genet.* **5**: e1000605.
- Badis, G., et al. (2009). Diversity and complexity in DNA recognition by transcription factors. *Science* **324**: 1720–1723.
- Baroux, C., Gagliardini, V., Page, D.R., and Grossniklaus, U. (2006). Dynamic regulatory interactions of Polycomb group genes: MEDEA autoregulation is required for imprinted gene expression in *Arabidopsis*. *Genes Dev.* **20**: 1081–1086.
- Baud, S., Wuilleme, S., Lemoine, R., Kronenberger, J., Caboche, M., Lepiniec, L., and Rochat, C. (2005). The AtSUC5 sucrose transporter specifically expressed in the endosperm is involved in early seed development in *Arabidopsis*. *Plant J.* **43**: 824–836.
- Baudry, A., Heim, M.A., Dubreucq, B., Caboche, M., Weishaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* **39**: 366–380.
- Becker, A., and Theissen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol. Phylogenet. Evol.* **29**: 464–489.
- Beisel, C., and Paro, R. (2011). Silencing chromatin: Comparing modes and mechanisms. *Nat. Rev. Genet.* **12**: 123–135.
- Berger, S.L. (2007). The complex language of chromatin regulation during transcription. *Nature* **447**: 407–412.
- Blödner, C., Goebel, C., Feussner, I., Gatz, C., and Polle, A. (2007). Warm and cold parental reproductive environments affect seed



- properties, fitness, and cold responsiveness in *Arabidopsis thaliana* progenies. *Plant Cell Environ.* **30**: 165–175.
- Bouyer, D., Roudier, F., Heese, M., Andersen, E.D., Gey, D., Nowack, M.K., Goodrich, J., Renou, J.P., Grini, P.E., Colot, V., and Schnittger, A.** (2011). Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. *PLoS Genet.* **7**: e1002014.
- Braybrook, S.A., and Harada, J.J.** (2008). LECs go crazy in embryo development. *Trends Plant Sci.* **13**: 624–630.
- Chen, D., Molitor, A., Liu, C., and Shen, W.H.** (2010). The *Arabidopsis* PRC1-like ring-finger proteins are necessary for repression of embryonic traits during vegetative growth. *Cell Res.* **20**: 1332–1344.
- Comet, I., Schuettengruber, B., Sexton, T., and Cavalli, G.** (2011). A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. *Proc. Natl. Acad. Sci. USA* **108**: 2294–2299.
- Debeaujon, I., Nesi, N., Perez, P., Devic, M., Grandjean, O., Caboche, M., and Lepiniec, L.** (2003). Proanthocyanidin-accumulating cells in *Arabidopsis* testa: Regulation of differentiation and role in seed development. *Plant Cell* **15**: 2514–2531.
- Dubos, C., Le Gourrierec, J., Baudry, A., Huep, G., Lanet, E., Debeaujon, I., Routaboul, J.M., Alboresi, A., Weisshaar, B., and Lepiniec, L.** (2008). MYB2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *Plant J.* **55**: 940–953.
- Enderle, D., Beisel, C., Stadler, M.B., Gerstung, M., Athri, P., and Paro, R.** (2011). Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. *Genome Res.* **21**: 216–226.
- Farrona, S., Coupland, G., and Turck, F.** (2008). The impact of chromatin regulation on the floral transition. *Semin. Cell Dev. Biol.* **19**: 560–573.
- Francis, N.J., and Kingston, R.E.** (2001). Mechanisms of transcriptional memory. *Nat. Rev. Mol. Cell Biol.* **2**: 409–421.
- Gao, M.J., Lydiate, D.J., Li, X., Lui, H., Gjetvaj, B., Hegedus, D.D., and Rozwadowski, K.** (2009). Repression of seed maturation genes by a trihelix transcriptional repressor in *Arabidopsis* seedlings. *Plant Cell* **21**: 54–71.
- Gehring, M., Huh, J.H., Hsieh, T.F., Penterman, J., Choi, Y., Harada, J.J., Goldberg, R.B., and Fischer, R.L.** (2006). DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124**: 495–506.
- Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R.A.** (2002). Dependence of heterochromatic histone H3 methylation patterns on the *Arabidopsis* gene DDM1. *Science* **297**: 1871–1873.
- Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G.** (1997). A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature* **386**: 44–51.
- Hagstrom, K., Muller, M., and Schedl, P.** (1997). A Polycomb and GAGA dependent silencer adjoins the Fab-7 boundary in the *Drosophila* bithorax complex. *Genetics* **146**: 1365–1380.
- Harscoët, E., Dubreucq, B., Palauqui, J.C., and Lepiniec, L.** (2010). NOF1 encodes an *Arabidopsis* protein involved in the control of rRNA expression. *PLoS ONE* **5**: e12829.
- Hartmann, U., Sagasser, M., Mehrtens, F., Stracke, R., and Weisshaar, B.** (2005). Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. *Plant Mol. Biol.* **57**: 155–171.
- Hennig, L., and Derkacheva, M.** (2009). Diversity of Polycomb group complexes in plants: Same rules, different players? *Trends Genet.* **25**: 414–423.
- Hill, K., Wang, H., and Perry, S.E.** (2008). A transcriptional repression motif in the MADS factor AGL15 is involved in recruitment of histone deacetylase complex components. *Plant J.* **53**: 172–185.
- Horard, B., Tatout, C., Poux, S., and Pirrotta, V.** (2000). Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. *Mol. Cell Biol.* **20**: 3187–3197.
- Jullien, P.E., Katz, A., Oliva, M., Ohad, N., and Berger, F.** (2006). Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in *Arabidopsis*. *Curr. Biol.* **16**: 486–492.
- Kapitonov, V.V., and Jurka, J.** (2007). Helitrons on a roll: Eukaryotic rolling-circle transposons. *Trends Genet.* **23**: 521–529.
- Katz, A., Oliva, M., Mosquna, A., Hakim, O., and Ohad, N.** (2004). FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. *Plant J.* **37**: 707–719.
- Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J., Wells, D., Katz, A., Margossian, L., Harada, J.J., Goldberg, R.B., and Fischer, R.L.** (1999). Control of fertilization-independent endosperm development by the MEDEA polycomb gene in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **96**: 4186–4191.
- Köhler, C., and Hennig, L.** (2010). Regulation of cell identity by plant Polycomb and trithorax group proteins. *Curr. Opin. Genet. Dev.* **20**: 541–547.
- Kooiker, M., Airoldi, C.A., Losa, A., Manzotti, P.S., Finzi, L., Kater, M.M., and Colombo, L.** (2005). BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic *Arabidopsis* gene SEEDSTICK. *Plant Cell* **17**: 722–729.
- Kouzarides, T.** (2007). Chromatin modifications and their function. *Cell* **128**: 693–705.
- Kroj, T., Savino, G., Valon, C., Giraudat, J., and Parcy, F.** (2003). Regulation of storage protein gene expression in *Arabidopsis*. *Development* **130**: 6065–6073.
- Lafos, M., Kroll, P., Hohenstatt, M.L., Thorpe, F.L., Clarenz, O., and Schubert, D.** (2011). Dynamic regulation of H3K27 trimethylation during *Arabidopsis* differentiation. *PLoS Genet.* **7**: e1002040.
- Lauria, M., and Rossi, V.** (2011). Epigenetic control of gene regulation in plants. *Biochim. Biophys. Acta* **1809**: 369–378.
- Lee, J.S., Smith, E., and Shilatifard, A.** (2010). The language of histone crosstalk. *Cell* **142**: 682–685.
- Li, J.J., and Herskowitz, I.** (1993). Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science* **262**: 1870–1874.
- Li, X., et al.** (2008). High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. *Plant Cell* **20**: 259–276.
- Lisch, D.** (2002). Mutator transposons. *Trends Plant Sci.* **7**: 498–504.
- Liu, C., Lu, F., Cui, X., and Cao, X.** (2010). Histone methylation in higher plants. *Annu. Rev. Plant Biol.* **61**: 395–420.
- Luo, M., Bilodeau, P., Koltunow, A., Dennis, E.S., Peacock, W.J., and Chaudhury, A.M.** (1999). Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **96**: 296–301.
- Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and Köhler, C.** (2006). Different Polycomb group complexes regulate common target genes in *Arabidopsis*. *EMBO Rep.* **7**: 947–952.
- Margueron, R., and Reinberg, D.** (2011). The Polycomb complex PRC2 and its mark in life. *Nature* **469**: 343–349.
- Meister, R.J., Williams, L.A., Monfared, M.M., Gallagher, T.L., Kraft, E.A., Nelson, C.G., and Gasser, C.S.** (2004). Definition and interactions of a positive regulatory element of the *Arabidopsis* INNER NO OUTER promoter. *Plant J.* **37**: 426–438.
- Monfared, M.M., Simon, M.K., Meister, R.J., Roig-Villanova, I., Kooiker, M., Colombo, L., Fletcher, J.C., and Gasser, C.S.** (2011). Overlapping and antagonistic activities of BASIC PENTACYSTEINE



- genes affect a range of developmental processes in *Arabidopsis*. *Plant J.* **66**: 1020–1031.
- Morey, L., and Helin, K.** (2010). Polycomb group protein-mediated repression of transcription. *Trends Biochem. Sci.* **35**: 323–332.
- Moyroud, E., Minguet, E.G., Ott, F., Yant, L., Pose, D., Monniaux, M., Blanchet, S., Bastien, O., Thevenon, E., Weigel, D., Schmid, M., and Parcy, F.** (2011). Prediction of regulatory interactions from genome sequences using a biophysical model for the *Arabidopsis* LEAFY transcription factor. *Plant Cell* **23**: 1293–1306.
- Müller, J., and Verrijzer, P.** (2009). Biochemical mechanisms of gene regulation by polycomb group protein complexes. *Curr. Opin. Genet. Dev.* **19**: 150–158.
- Nakayama, T., Nishioka, K., Dong, Y.X., Shimojima, T., and Hirose, S.** (2007). Drosophila GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading. *Genes Dev.* **21**: 552–561.
- Ogas, J., Kaufmann, S., Henderson, J., and Somerville, C.** (1999). PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **96**: 13839–13844.
- Pauli, S., Rothnie, H.M., Chen, G., He, X., and Hohn, T.** (2004). The cauliflower mosaic virus 35S promoter extends into the transcribed region. *J. Virol.* **78**: 12120–12128.
- Pelletier, M.K., and Shirley, B.W.** (1996). Analysis of flavanone 3-hydroxylase in *Arabidopsis* seedlings. Coordinate regulation with chalcone synthase and chalcone isomerase. *Plant Physiol.* **111**: 3393–345.
- Pfluger, J., and Wagner, D.** (2007). Histone modifications and dynamic regulation of genome accessibility in plants. *Curr. Opin. Plant Biol.* **10**: 645–652.
- Pien, S., and Grossniklaus, U.** (2007). Polycomb group and trithorax group proteins in *Arabidopsis*. *Biochim. Biophys. Acta* **1769**: 375–382.
- Riechmann, J.L., and Meyerowitz, E.M.** (1997). MADS domain proteins in plant development. *Biol. Chem.* **378**: 1079–1101.
- Ringrose, L., Ehret, H., and Paro, R.** (2004). Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. *Mol. Cell* **16**: 641–653.
- Ringrose, L., and Paro, R.** (2007). Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* **134**: 223–232.
- Rodrigues, J.C., Luo, M., Berger, F., and Koltunow, A.M.** (2010). Polycomb group gene function in sexual and asexual seed development in angiosperms. *Sex. Plant Reprod.* **23**: 123–133.
- Roudier, F., et al.** (2011). Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*. *EMBO J.* **30**: 1928–1938.
- Roudier, F., Teixeira, F.K., and Colot, V.** (2009). Chromatin indexing in *Arabidopsis*: An epigenomic tale of tails and more. *Trends Genet.* **25**: 511–517.
- 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.* **54**: 608–620.
- Santos-Mendoza, M., Dubreucq, B., Miquel, M., Caboche, M., and Lepiniec, L.** (2005). LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in *Arabidopsis* leaves. *FEBS Lett.* **579**: 4666–4670.
- Sawarkar, R., and Paro, R.** (2010). Interpretation of developmental signaling at chromatin: The Polycomb perspective. *Dev. Cell* **19**: 651–661.
- Schatlowski, N., Creasey, K., Goodrich, J., and Schubert, D.** (2008). Keeping plants in shape: polycomb-group genes and histone methylation. *Semin. Cell Dev. Biol.* **19**: 547–553.
- Schmitges, F.W., et al.** (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell* **42**: 330–341.
- Scholl, R.L., May, S.T., and Ware, D.H.** (2000). Seed and molecular resources for *Arabidopsis*. *Plant Physiol.* **124**: 1477–1480.
- Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., and Goodrich, J.** (2006). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J.* **25**: 4638–4649.
- Schuettengruber, B., and Cavalli, G.** (2009). Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development* **136**: 3531–3542.
- Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G.** (2007). Genome regulation by polycomb and trithorax proteins. *Cell* **128**: 735–745.
- Schwartz, Y.B., Kahn, T.G., Nix, D.A., Li, X.Y., Bourgon, R., Biggin, M., and Pirrotta, V.** (2006). Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nat. Genet.* **38**: 700–705.
- Schwartz, Y.B., and Pirrotta, V.** (2008). Polycomb complexes and epigenetic states. *Curr. Opin. Cell Biol.* **20**: 266–273.
- Simon, J.A., and Kingston, R.E.** (2009). Mechanisms of polycomb gene silencing: Knowns and unknowns. *Nat. Rev. Mol. Cell Biol.* **10**: 697–708.
- Sing, A., Pannell, D., Karaiskakis, A., Sturgeon, K., Djabali, M., Ellis, J., Lipshitz, H.D., and Cordes, S.P.** (2009). A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell* **138**: 885–897.
- Smith, E., and Shilatifard, A.** (2010). The chromatin signaling pathway: Diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes. *Mol. Cell* **40**: 689–701.
- Stone, S.L., Kwong, L.W., Yee, K.M., Pelletier, J., Lepiniec, L., Fischer, R.L., Goldberg, R.B., and Harada, J.J.** (2001). LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proc. Natl. Acad. Sci. USA* **98**: 11806–11811.
- 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.
- Tang, W., and Perry, S.E.** (2003). Binding site selection for the plant MADS domain protein AGL15: An in vitro and in vivo study. *J. Biol. Chem.* **278**: 28154–28159.
- To, A., Valon, C., Savino, G., Guillemot, J., Devic, M., Giraudat, J., and Parcy, F.** (2006). A network of local and redundant gene regulation governs *Arabidopsis* seed maturation. *Plant Cell* **18**: 1642–1651.
- Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V.** (2007). *Arabidopsis* TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet.* **3**: e86.
- Valenzuela, L., and Kamakaka, R.T.** (2006). Chromatin insulators. *Annu. Rev. Genet.* **40**: 107–138.
- Villanueva, J.M., Broadhvest, J., Hauser, B.A., Meister, R.J., Schneitz, K., and Gasser, C.S.** (1999). INNER NO OUTER regulates abaxial-adaxial patterning in *Arabidopsis* ovules. *Genes Dev.* **13**: 3160–3169.
- Wanke, D., Hohenstatt, M.L., Dynowski, M., Bloss, U., Hecker, A., Elgass, K., Hummel, S., Hahn, A., Caesar, K., Schleifenbaum, F., Harter, K., and Berendzen, K.W.** (2011). Alanine zipper-like coiled-coil domains are necessary for homotypic dimerization of plant GAGA-factors in the nucleus and nucleolus. *PLoS ONE* **6**: e16070.
- Ward, L.D., and Bussemaker, H.J.** (2008). Predicting functional transcription factor binding through alignment-free and affinity-based analysis of orthologous promoter sequences. *Bioinformatics* **24**: i165–i171.





- Weil, C., and Martienssen, R.** (2008). Epigenetic interactions between transposons and genes: Lessons from plants. *Curr. Opin. Genet. Dev.* **18**: 188–192.
- Weinhofer, I., Hehenberger, E., Roszak, P., Hennig, L., and Köhler, C.** (2010). H3K27me3 profiling of the endosperm implies exclusion of polycomb group protein targeting by DNA methylation. *PLoS Genet.* **6**: e1001152.
- Woo, C.J., Kharchenko, P.V., Daheron, L., Park, P.J., and Kingston, R.E.** (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* **140**: 99–110.
- Yang, L., and Bennetzen, J.L.** (2009). Structure-based discovery and description of plant and animal Helitrons. *Proc. Natl. Acad. Sci. USA* **106**: 12832–12837.
- Zhang, H., and Ogas, J.** (2009). An epigenetic perspective on developmental regulation of seed genes. *Mol. Plant* **2**: 610–627.
- Zhang, H., Rider, S.D., Jr., Henderson, J.T., Fountain, M., Chuang, K., Kandachar, V., Simons, A., Edenberg, H.J., Romero-Severson, J., Muir, W.M., and Ogas, J.** (2008). The CHD3 remodeler PICKLE promotes trimethylation of histone H3 lysine 27. *J. Biol. Chem.* **283**: 22637–22648.
- Zhang, K., Sridhar, V.V., Zhu, J., Kapoor, A., and Zhu, J.K.** (2007a). Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *PLoS ONE* **2**: e1210.
- Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y.V., Pellegrini, M., Goodrich, J., and Jacobsen, S.E.** (2007b). Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.* **5**: e129.
- Zhang, Y., Zheng, S., Liu, Z., Wang, L., and Bi, Y.** (2011). Both HY5 and HYH are necessary regulators for low temperature-induced anthocyanin accumulation in *Arabidopsis* seedlings. *J. Plant Physiol.* **168**: 367–374.
- Zheng, B., and Chen, X.** (2011). Dynamics of histone H3 lysine 27 trimethylation in plant development. *Curr. Opin. Plant Biol.* **14**: 123–129.
- Zheng, Y., Ren, N., Wang, H., Stromberg, A.J., and Perry, S.E.** (2009). Global identification of targets of the *Arabidopsis* MADS domain protein AGAMOUS-Like15. *Plant Cell* **21**: 2563–2577.



**Transcriptional Regulation of *Arabidopsis* *LEAFY COTYLEDON2* Involves *RLE*, a *cis*-Element That Regulates Trimethylation of Histone H3 at Lysine-27**

Nathalie Berger, Bertrand Dubreucq, François Roudier, Christian Dubos and Loïc Lepiniec  
*Plant Cell*; originally published online November 11, 2011;  
DOI 10.1105/tpc.111.087866

This information is current as of December 1, 2011

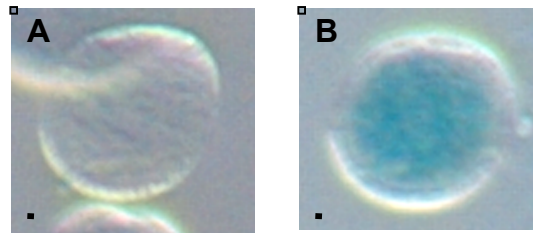
<b>Supplemental Data</b>	<a href="http://www.plantcell.org/content/suppl/2011/10/31/tpc.111.087866.DC1.html">http://www.plantcell.org/content/suppl/2011/10/31/tpc.111.087866.DC1.html</a>
<b>Permissions</b>	<a href="https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X">https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X</a>
<b>eTOCs</b>	Sign up for eTOCs at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a>
<b>CiteTrack Alerts</b>	Sign up for CiteTrack Alerts at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a>
<b>Subscription Information</b>	Subscription Information for <i>The Plant Cell</i> and <i>Plant Physiology</i> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a>





**Supplemental Figure 1: Complementation of the *lec2-4* mutation by  $P_{LEC2}:LEC2$**   
Dry seeds of the WT plant (A), *lec2-4* mutant (B) and *lec2-4* mutant transformed with  $P_{LEC2}:LEC2$  that contains  $P_{LEC2}$  fused to the LEC2 cDNA (ATG-STOP) (C).  
Scale bar = 0.5 mm



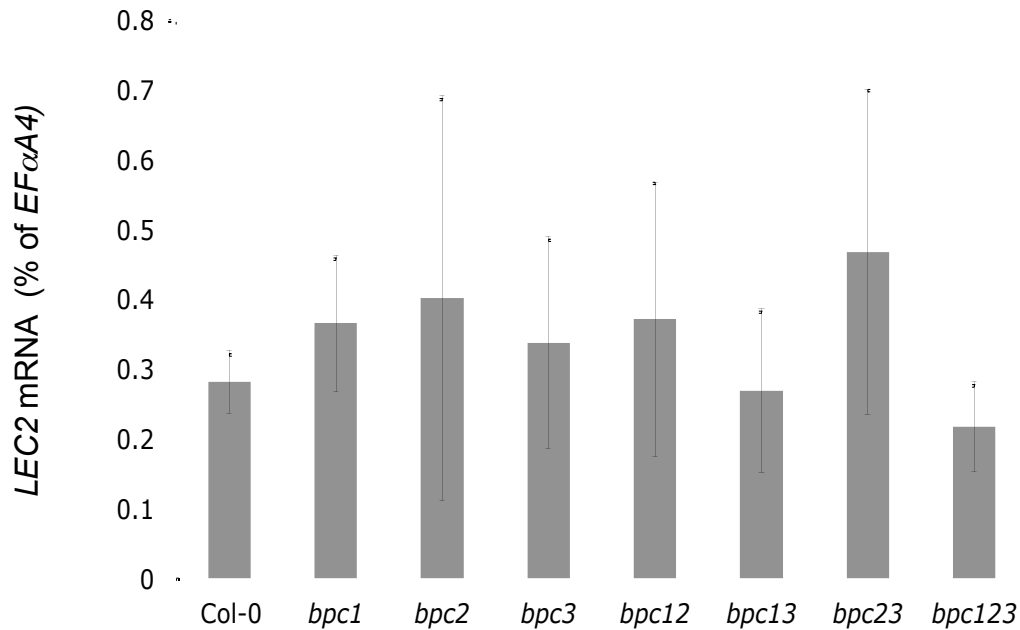


**Supplemental Figure 2: Activation of *PLEC2* in pollen by adding 3CT repeats in the GAF box of *P<sub>LEC2</sub>*.**

GUS staining of *P<sub>LEC2</sub>* (A) and *P<sub>LEC2</sub>*m+3CT (B) in pollen of transgenic plants. Pollen staining was observed in 18 lines versus 24 independent lines studied. Scale bar = 5  $\mu$ m



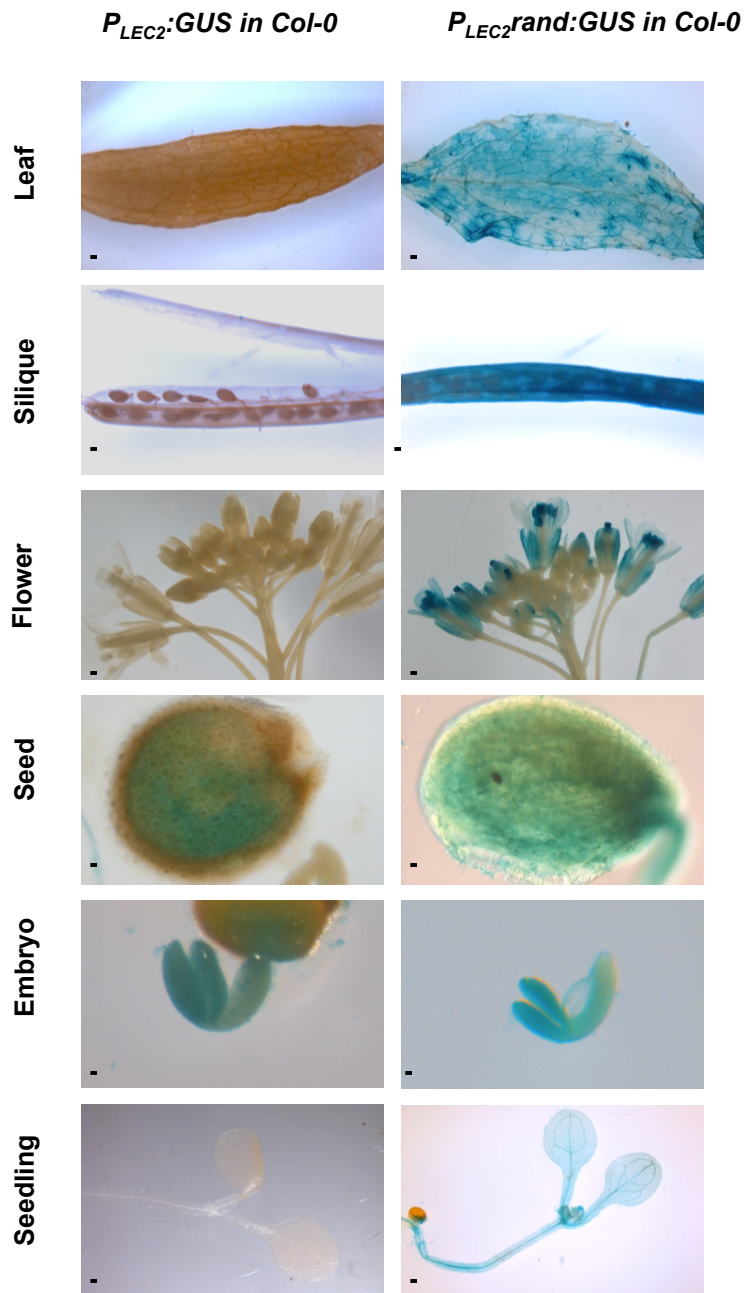




**Supplemental Figure 3: *LEC2* expression is not modified in *bpc* mutants.**

The accumulation of mRNA was investigated by RT PCR in siliques 10 days after pollination (DAP) in single, double and triple *bpc1*, *bpc2* and *bpc3* mutants. Three biological and three technical repeats have been performed.





**Supplemental Figure 4: Expression patterns of *P<sub>LEC2</sub>:GUS* and *P<sub>LEC2</sub>rand:GUS* in Col-0 ecotype background.**

Scale bars = 1 mm in leaf, silique and flower, and 0.1 mm in seed, embryo and seedlings. Independent transgenic lines ( $n > 12$  for each construct) were assayed for GUS activity in leaves, siliques, flowers, seeds, embryos and seedlings (10 days after sowing).













4 QLEC2GAR TCCGGTTTGC GTAACCGGGA  
QLEC2G3F TGAAATGCATTGATGAAAATGA  
QLEC2G3R CGCAACGATTTTTGTGCTT  
5 QLEC2G2F TGGTTTACCACTTTACCTCGAA  
QLEC2G2R GCATTCCAAAGCCATGATTC

PLEC2/PF3H:GUS

1A PLEQ1F GCCAAGGAACTCCAAA  
GUSQ-2R GTTTCTACAGGACGTAACAT  
1B PLEC2Q(AV2011)F GTCTCTTCTCACTCGCTCTCTCCT  
PLEC2Q(AV2011)REV GAATGCCACAGGCCGTCGA  
1 QPF3H-2 CATTACACATCTCTTCTTTC  
GUSQ-2R GTTTCTACAGGACGTAACAT  
2 GUSQ3F(1249) CGTCCGCAAGTGCACGGGAA  
GUSQ3R(1541) GAGTGCAGCCCGGCTAACGT

CONTROL

ACT7FWPCRQ TTCGCACATGTA CTCTGTTTCG  
ACT7REVPCRQ AGAGAGAAAGATAGAGAAATGGAGGAGAA  
FWA\_5F ATCTTGCCATTACGGCTTTG  
FWA\_5R AGCTCTGAAGCCTCCACAA

ANALYSIS OF GENE EXPRESSION

EF-1F CTGGAGGTTTTGAGGCTGGTA  
EF-1R GCAAGGGTGAAAGCAAGAAGA  
BPC1-F GATTAATCAACCAAAGGACA  
BPC1-R AGGTAAACCAGAAATGTCCA  
BPC2-F ACCATTGAYCGGAACACTAA  
BPC2-R CTTTTGCATTGGTCTTTCTC  
BPC3-F GAGATGGTGATAAATGGTGT  
BPC3-R GAAATCAAAACCATCAGAAG  
BPC4-F AGGAAGAAAGAGAACAAGCA  
BPC4-R TAGTGATGTAGCGGTTTGTC  
BPC5-F GCACTTGCACCGGGACTGCT  
BPC5-R TGGAAATCCGCTATGGTTCTTGCCG  
BPC6-F ACCACCTGCCGGATCCACCT  
BPC6-R CCTTCTGCAGCAAGCCGGCT  
BPC7-F AAACGCAAACAAGGACTCAG  
BPC7-R GTTTCTCCTTTGCGGTTTT



## II.1.2 Résultats complémentaires : relations d'épistasie entre les trois éléments cis-régulateurs isolés

Une des questions restant sans réponse après les études présentées au chapitre précédent (II.1.1) est la relation entre ces 3 boîtes de régulation. Sont-elles épistatiques les unes sur les autres et peut-on supposer une interaction entre les différents intervenants ? Pour répondre génétiquement à ce point particulier, j'ai réalisé des construits présentant des combinaisons de mutations des 3 boîtes étudiées (décrites dans le tableau ci-dessous).

Nom du construit	Boîte MADS	Boîte RLE	Boîte BPC	Oligos	Matrice ADN
PL---	Délétée	Remplacée par la boîte rand	Mutée	AttB1P442/AttB2mAT	PLEC2rand :GUS
PL+--	Sauvage	Remplacée par la boîte rand	Mutée	AttB1p500/AttB2mAT	PLEC2rand :GUS
PL--+	Délétée	Sauvage	Mutée	AttB1P442/AttB2mAT	PLEC2 :GUS
PL--+	Délétée	Remplacée par la boîte rand	Sauvage	AttB1P442/AttB2pATG	PLEC2rand :GUS
PL-++	Délétée	Sauvage	Sauvage	AttB1P442/ AttB2pATG	PLEC2 :GUS

Les plantes transgéniques sont en cours d'obtention. L'analyse de ces plantes devrait nous permettre d'élucider les relations existant ou pas entre ces trois boîtes de régulation. Par conséquent, nous pourrions nous intéresser aux relations entre les facteurs interagissant potentiellement avec elles et peut-être détecter l'existence d'autres séquences essentielles. Une fois tout cela étudié, il serait intéressant de tenter de créer des promoteurs synthétiques en introduisant ces trois boîtes dans d'autres séquences pour voir si elles suffisent à recréer le profil d'expression de *LEC2* et surtout à recruter la marque H3K27me3 ou d'autres.



## **II.2 Facteurs positifs, données supplémentaires**

### **II.2.1 AGL15, AGL18 et le promoteur de *LEC2***

Comme nous l'avons vu au chapitre II.1, le double mutant *agl15agl18* est le seul génotype connu, montrant une diminution de l'expression de *LEC2* dans la graine bien que le locus *LEC2* soit intact. Le locus *LEC2* a également été identifié comme cible d'AGL15 lors de manipulations de ChIP (Zheng et al, 2009). De plus, lors de l'étude fonctionnelle du promoteur de *LEC2*, un site présentant des homologies avec le site de fixation d'AGL15 a été identifié comme essentiel à l'activité de PLEC2 (chapitre II-1-1). Je me suis donc attachée à déterminer, en utilisant la technique du simple hybride, si AGL15 et/ou AGL18 sont capables de se fixer au promoteur de *LEC2* et, si c'est le cas, grâce à quelle séquence nucléotidique. Comme nous pouvons le voir figure 16, nous n'avons pas pu détecter d'interactions entre ces facteurs et le promoteur de *LEC2*.

Cependant, bien que ce premier résultat soit négatif, les facteurs MADS agissant souvent en hétérocomplexe, il est donc très intéressant de tester si une interaction entre AGL15 et AGL18 est nécessaire pour une fixation sur PLEC2. De Folter et collaborateurs (De folter et al., 2005) ont réalisé une matrice d'interaction grâce à des tests systématiques en double hybride sur plus de 100 facteurs MADS, mais ils n'avaient pas inclus AGL18 dans leurs cibles. Comme on peut le voir figure 16, que l'on considère les protéines indépendamment ou 2 à 2, aucune interaction n'est détectable. Les hypothèses pour cela sont très variées : la structure chromatinienne au locus peut jouer un rôle essentiel pour la fixation des MADS au locus, il conviendrait donc de réaliser des expériences de ChIP avec la protéine AGL15 pour déterminer précisément la région chromosomique du locus *LEC2* fixée. Une autre explication serait que l'interaction entre AGL15 et le locus *LEC2* nécessite une protéine ou un complexe protéique intermédiaire inconnu, absent de la levure. Pour tenter de pallier cette difficulté, nous allons réaliser le même type de manipulations, en remplaçant le système hétérologue levure par des protoplastes de mousse (Thevenin et al, 2011) qui sera plus proche d'un système végétal que la levure. Cette technique est utilisée en routine au laboratoire. Comme dans un système simple hybride, si une activation du promoteur est observée, il sera ensuite possible de mutagéniser les sites potentiels de fixations des facteurs MADS sur le promoteur de *LEC2* pour confirmer nos hypothèses. Enfin, AGL15 et AGL18 sont des facteurs positifs de l'expression de *LEC2*, comme les « Basic Penta Cystein » (BPC, cf chapitre II.1). Des

		pDEST22			pDEST32				
		vide	AGL15	AGL18				AGL15	AGL18
Sans sélection	PBAN								
	PLEC2								
Avec sélection	PBAN								
	PLEC2								

**Figure 16: Analyse par la méthode de simple hybride de l'interaction entre les protéines AGL15 et AGL18 avec le promoteur de *LEC2*. Les protéines AGL15 et AGL18 n'interagissent pas avec le promoteur de *LEC2* en levure.**

manipulations de double hybride sont donc en cours pour tester la liaison des MADS avec les BPC qui interviennent également dans l'activation et non uniquement entre MADS.

## II.2.2 Les BPC

Dans le chapitre II.1, la boîte de reconnaissance des BPC apparaît indispensable pour une activité du promoteur de *LEC2*, mais aucune liaison génétique ni physique n'a pu être établie *in planta* entre une protéine BPC et le promoteur de *LEC2* sauvage. Le seul fait établi est que la surexpression du gène *uidA* dans les plantes contenant le construit PLEC2rand :GUS est sous le contrôle des BPC et de la boîte CTn. On ne sait donc toujours pas si les BPC sont nécessaires à l'expression de *LEC2* dans la graine ou s'ils sont capables d'activer *LEC2* dans les plantules dans un contexte de promoteur sauvage, une fois la répression levée, comme un inducteur constitutif. Les BPC présentent des similarités de fonction avec le facteur GAGA de drosophile qui a de nombreuses fonctions, essentiellement en recrutant d'autres facteurs possédant une fonction plus spécifique. Ces fonctions vont de la régulation de l'élongation de la transcription au recrutement de PRC2 et elles sont décrites plus avant dans l'article de revue ci-après qui a été accepté pour publication à « Biochimica et Biophysica Acta-gene regulatory mechanisms ».





### ***II.2.2.1 Homologues des BPC dans les autres règne***

*Title : Evolution goes GAGA : GAGA binding proteins across kingdoms*

**Nathalie Berger, Bertrand Dubreucq\*.**

Institut Jean-Pierre Bourgin, UMR1318 INRA-Agro-ParisTech, 78026 Versailles Cedex, France.

\*Corresponding Author: B. dubreucq ([Bertrand.dubreucq@versailles.inra.fr](mailto:Bertrand.dubreucq@versailles.inra.fr))

IJPB, INRA, centre de Versailles, Bâtiment 2

Route de Saint Cyr

78026 versailles Cedex , France

#### **Highligths**

- Chromatin-associated proteins (CAP) play a key role in the regulation of gene expression.
- CAPs are relatively well conserved across the animal and plant kingdoms.
- The *Drosophila* GAGA factor (dGAF) interacts with CAP but is not conserved in plants.
- Plant BBR/BPC proteins functionally converge with dGAF despite divergent sequences.

**Keywords:** GAGA Factor, chromatin, plant, BPC proteins



**Summary** (170 words).

Chromatin-associated proteins (CAP) play a crucial role in the regulation of gene expression and development in higher organisms. They are involved in the control of chromatin structure and dynamics. CAP have been extensively studied over the past years and are classified into two major groups: enzymes that modify histone stability and organization by post-translational modification of their N-Terminal tails; and proteins that use ATP hydrolysis to modify chromatin structure. All of these proteins show a relatively high degree of sequence conservation across the animal and plant kingdoms. Surprisingly, the essential *Drosophila melanogaster* GAGA factor (dGAF), which interacts with these two types of CAP to regulate homeobox genes and thus contributes to a wide range of developmental events, is not conserved in plants. In this review, following an overview of fly GAF functions, we discuss the role of plant BBR/BPC proteins. These appear to functionally converge with dGAF despite a completely divergent amino acid sequence. Some suggestions are given for further investigation into the function of BPC proteins in plants.



**Glossary:**

ACP	associated chromatin protein
ATP	Adenosine-5'-triphosphate
BBR	barley b recombinant
BTB/POZ	broad complex tramtrack bric a brac/poxvirus zinc finger
BPC	basic pentacystein protein
CHD	chromodomain Helicase/ATPase DNA binding protein
Dam	DNA adenine methyl transferase
eve	even skipped
FACT	facilitates chromatin transcription
GBP	GAGA Binding Protein
HDAC	histone deacetylase
hsp70	heat shock protein 70
ISWI	imitation of switch
NELF	negative elongation factor
NURF	nucleosomal remodeling factor
PcG	Polycomb group
PIC	preinitiation complex
PRE	polycomb repressive element
psq	pipsqueak
RNA polII	RNA polymerase II
SNF	sucrose non-fermenting
SWI	homothalic switching deficient
TAF3	TBP-associated factor 3
TBP	TATA binding protein
Trl	trithorax like
TrxG	trithorax group





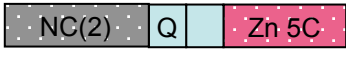
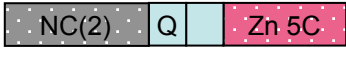
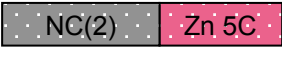
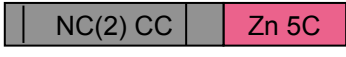
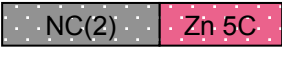
## 1. Chromatin structure determines transcriptional activity

Chromatin is a complex molecular ensemble of genomic DNA and directly or indirectly associated proteins. This dynamic structure is necessary for the correct recruitment of transcription factors and coordinated gene transcription. Nucleosomes are the basic unit of chromatin DNA packaging and consist of approximately 147bp of DNA wrapped around a protein histone core. A wide range of histone N-terminal modifications has been implicated in the regulation of gene expression. These modifications lead to the recruitment of specific factors, which confer a transcriptional repressive or permissive status to chromatin. Methylation and acetylation of histone N-terminal tails are the predominant post-translational modifications influencing gene regulation and chromatin structure (Kouzarides, 2007). Histone acetylation is usually associated with an active chromatin state and HDACs (Histone deacetylases) are consequently considered as repressors. In comparison, histone tail methylation appears to be a far more complex process. Mono, di or tri-methylation of the lysine 4 or 36 of histone H3 (H3K4me1/me2/me3; H3K36me1/me2/me3) is globally considered as an active mark of gene expression. On the other hand, mono, di or tri-methylation of the lysine 9 or 27 of histone H3 (H3K9me1/me2/me3; H3K27me1/me2/me3) is associated with repression of gene expression (Hublitz et al, 2009). TrxG and PcG are responsible for H3K4me3 and H3K27me3 mark deposition, respectively, and can therefore activate or repress homeobox genes in *Drosophila melanogaster* (Papp & Muller, 2006).

The second way altering chromatin structure and mobilizing nucleosomes is dependent on ATP hydrolysis. To date, four classes of ATP-dependent chromatin remodelling complexes are known (each containing one of the following ATPases belonging to the SNF2 ATPase family: SWI2/SNF2, ISWI, CHD/Mi-2, or INO80 ATPases) (Conaway & Conaway, 2009; Flaus et al, 2006; Kunert & Brehm, 2009; Racki & Narlikar, 2008) and were found conducive to a transcriptionally active and nucleosome free chromatin region (Johnson et al, 2005; Maier et al, 2008).



FIGURE 1

Family name	Protein structure	DNA specificity
Trl (GAF581)		GAGAG
Human <i>c-Krox/Th-POK</i>		(GA) <sub>8</sub>
<i>Zea mays</i>		?
Barley BBR		(CT) <sub>8</sub>
<i>Arabidopsis</i> BPC1,2,3		RGARAGRRA
<i>Arabidopsis</i> BPC4,5,6		(CT) <sub>6</sub> , ?
<i>Arabidopsis</i> BPC 7		(CT) <sub>6</sub> , ?

**Figure 1: Protein structure of dGAF orthologues.** Functionally proven domains are in color and putative functions are in dotted squares. dGAF are composed of a N-terminal protein-protein interaction domain (grey square) and a C-terminal DNA binding domain (pink square). BPB/POZ: broad complex tramtrack bric a brac/poxvirus zinc finger; NC(1), NC(2) and CC: alanine Zipper-Like Coiled-Coil Domain; C2H2 : Zinc finger C2H2 domain; BasicQ, Q and Zn5C: DNA binding domain.

Enzymes which modify histone N-terminal tails and ATP-dependent chromatin modifiers are known as chromatin-associated proteins (CAP) and contribute to DNA replication, mitosis, recombination and gene transcription. Many CAP were found to act in complexes with the *Drosophila* GAGA factor (dGAF) in a wide range of cellular functions. Furthermore, in keeping with their fundamental role in gene expression, most CAP are conserved across the animal and plant kingdoms (Lauria & Rossi, 2011a; Lauria & Rossi, 2011b; Zheng & Chen, 2011), but, surprisingly, dGAF orthologues were only found in animals. In this review we will first give an overview of the multiple functions for GAF proteins in animals. We will then discuss the functional similarities between dGAF and the plant protein family named BBR/BPC, despite the fact that their amino-acid sequences are fully divergent.

## **2. The *Drosophila* GAGA factor**

dGAF or *Trl* was first identified in *Drosophila* as a positive regulator of homeotic genes, a position-effect variegation modifier (Farkas et al, 1994) and a nucleosome disrupter at the *hsp70* locus (Tsukiyama et al, 1994). Later, it was found to act in boundary activity (Ohtsuki & Levine, 1998), chromosome segregation and nuclear division (Bhat et al, 1996). Genome-wide analysis showed that dGAF targets include up to 250 members involved in at least 28 cellular pathways (van Steensel et al, 2003).

dGAF exists in two isoforms (GAGA519 and GAGA581) produced from alternative splicing (Benyajati et al, 1997), which share a BTB/POZ protein-protein interaction domain and DNA binding region (C<sub>2</sub>H<sub>2</sub> zinc finger domain) but diverge in their C-terminal region that contains a glutamine rich domain (Q domain) essential for transcriptional activation (Espinass et al, 1999; Pedone et al, 1996; Vaquero et al, 2000) (Figure 1). As the BTB/POZ domain is crucial for homo and hetero-dimer formation (Benyajati et al, 1997) this certainly implies that dGAF acts through the recruitment of different interactors. Interestingly, some of the identified partners, such as Pipsqueak, Batman or Lolalike, which associate with dGAF to regulate homeobox genes expression, also contain a BTB/POZ domain (Faucheux et al, 2003; Mishra et al, 2003). Interaction of dGAF with other BTB/POZ domain proteins can also



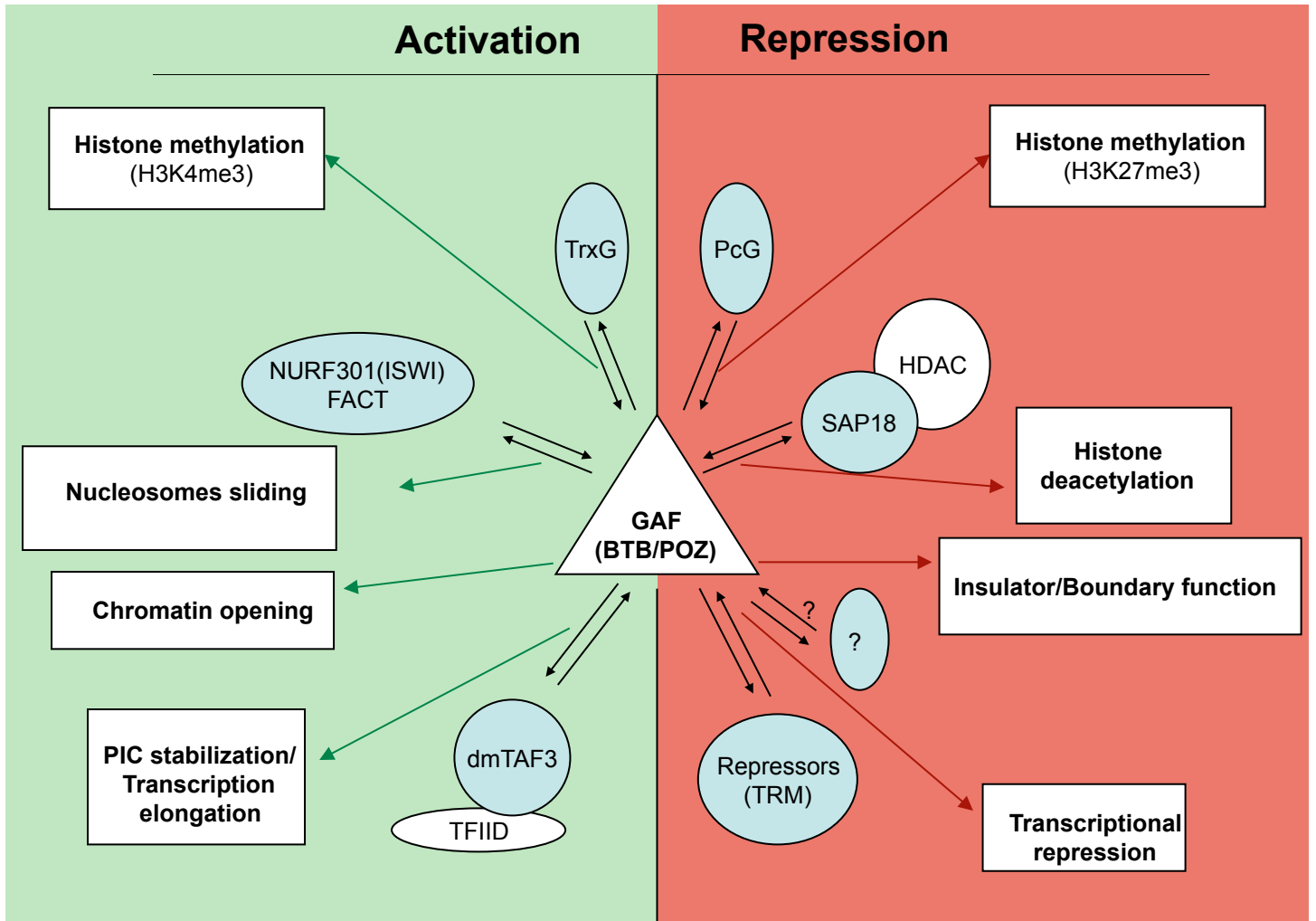
counteract the Tramtrack repressor activity at the *eve* locus (Pagans et al, 2002). In addition, oligomerization of BTB/POZ proteins was shown to form higher order complexes that contribute to open chromatin and increase the DNA binding specificity at promoters that contain multiple GAF binding sites (Katsani et al, 1999).

dGAF was named for its capacity to bind GAGAG stretches, which are fine-tuned by post-translational modification at the DNA binding site. For instance, acetylation of two lysines or phosphorylation of two serines decreases the affinity of the factor for DNA without abolishing binding (Bonet et al, 2005). In *Drosophila*, genome wide studies revealed that GAGAG motifs are dispersed evenly throughout the genome in active or inactive chromatin regions, introns and exons. Genome wide analysis of GAF binding sites *in vivo* with the Dam method (van Steensel et al, 2003) revealed binding to both transcriptionally active and inactive regions, confirming its dual role in activation and repression. Furthermore, in active regions, only closely spaced GAGAG elements were bound (van Steensel et al, 2003), consistent with the hypothesis that dGAF oligomerization plays a role in DNA binding specificity (Katsani et al, 1999). In addition, a lack of binding in exons was also observed by van Steensel and collaborators, which can be explained by the low representation of closely located GAGAG repeats. This robust consensus led to research into the functions of GAF through the presence and prevalence of its DNA recognition sites.

## **2. dGAF in transcriptional activation processes**

Mutations in dGAF recognition sites have provided important insights into dGAF functions. When GAGAG sequences located in the *Fab7* and *Evx2-Hoxd13* loci are mutated, they lose their insulator activity (Belozarov et al, 2003; Ohtsuki & Levine, 1998; Schweinsberg & Schedl, 2004; Vasanthi et al, 2010). Insulator, or boundary, activity is specific to animal genomes and depends on the clustered organization of some genes that need proper enhancers, sometimes located far away from the coding region, at a specific developmental stage. Insulators were first identified in the flanking regions of the *Drosophila hsp70* gene (Cai & Levine, 1995; Kellum & Schedl, 1991; Kellum & Schedl, 1992; Scott & Geyer, 1995)

FIGURE 2



**Figure 2: The function of the GAGA factor (GAF) in *Drosophila*.** Associations with proteins which interact directly (blue circles) and indirectly (white circles) with GAF (triangle) are shown with double arrows. Associated functions are shown in white squares.

and allow chromatin loops to isolate *cis*-regulatory elements from each other. Thus, dGAF can activate transcription by modifying chromatin structure to tether distant DNA sequences. dGAF binding sequences were found to co-localize with DNase I hypersensitive sites (Cartwright & Elgin, 1986; Georgel, 2005; Kerrigan et al, 1991; Lis & Wu, 1993; Lu et al, 1993; Tsukiyama et al, 1994). dGAF was also implicated in local modifications to chromatin structure, creating nucleosome free regions, consistent with its interaction with NURF301 (Xiao et al, 2001) and FACT (Orphanides et al, 1998; Orphanides et al, 1999; Shimojima et al, 2003) which act in nucleosome sliding.

FACT is involved in transcriptional elongation, modifying nucleosome structure, (Orphanides et al, 1998; Orphanides et al, 1999) and its cooperation with dGAF is fully consistent with the co-localization of dGAF with RNA polIII. An interaction between FACT and dGAF supports the hypothesis that dGAF can bind to its high affinity binding site and then spreads through the gene to be induced, opening up the chromatin structure to the RNA polymerase II enzyme (O'Brien et al, 1995). Indeed, ChIP experiments showed that dGAF binds to 3'-end gene regions (van Steensel et al, 2003). In addition, the GAF/FACT association was shown to participate in the replacement of K9 methylated histone H3 by H3.3, which is not a repressive mark, preventing heterochromatin spreading (Nakayama et al, 2007). Another partner, NURF301, is an ISWI family CAP that enhances nucleosome sliding, and was shown to interact *in vitro* and in co-IP experiments with GAF (Xiao et al, 2001). Furthermore, ISWI and GAF binding sites were found to partially overlap (about 30% of overlap) in double label immuno-fluorescence microscopy and ChIP-chip experiments (Sala et al, 2011).

In addition to its activity in nucleosome sliding, dGAF also plays a more direct role in the activation of target gene expression by interacting with the basal transcriptional machinery (Figure 2). Indeed, dGAF was identified as a positive regulator of the Ultrabithorax gene of the bithorax complex in *Drosophila*, which itself interacts with dmTAF3, a component of the TFIID complex (Chopra et al, 2008). dGAF is also required to enhance TFIID binding (Leibovich, 2002) and stabilize PIC formation through its Q-domain (Vaquero et al, 2008).



dGAF can also play a repressive role in transcriptional elongation, acting with NELF for promoter proximal RNAPolIII pausing (Gilchrist et al, 2008; Lee et al, 2008). This is not the only example of GAF duality; in the opposite mechanism involving the TrxG and PcG proteins, it can interact with each type of complex, and GAGAG sequences are essential for the recruitment of both (Figure 2).

### 3. Repressive activity of GAF

GAF recognition sites are often found in PREs, which are necessary for the recruitment of PcG complexes (Ringrose & Paro, 2007; Schuettengruber et al, 2007). In PREs, GAGAG stretches are predominantly associated with PHO and Zeste recognition sites (Mahmoudi et al, 2003; Mulholland et al, 2003) and all three are essential for many PcG-mediated repression events (Brown et al, 1998; Hagstrom et al, 1997; Hur et al, 2002). The PcG proteins are involved in repression of gene expression via the H3K27me3 deposition mark and are implicated in three major types of complexes: PhoRC, PRC1 and PRC2 (Katsani et al, 1999). GAF was found to associate with PRC1, co-precipitating with at least two of the PRC1 subunits, namely PC and PH (Hodgson et al, 2001; Horard et al, 2000) and is genetically bound to PHO.

*Pho* is a PcG gene that encodes a zinc finger DNA binding protein related to the mammalian transcription factor YY1 (Brown et al, 1998). It is essential in many PRE activities (Busturia et al, 2001; Fritsch et al, 1999; Mihaly et al, 1998; Mishra et al, 2001; Shimell et al, 2000), and interacts with E(z), ESC (Wang et al, 2004b), PC and PH (Mohd-Sarip et al, 2002). Genetic association between dGAF and *Pho* in PRE domains is reinforced by the fact that dGAF facilitates PHO binding to chromatin (Mahmoudi et al, 2003). dGAF can also participate in transcriptional repression, interacting with SAP18, which itself is associated with HDAC complexes (Espinosa et al, 2000).

Altogether, the data suggest that dGAF is a central actor in *Drosophila* gene regulation. Consequently, a search for homologues was made across the animal kingdom. The first dGAF homologue identified in vertebrates was the human *c-Krox/Th-POK* protein,





followed by the mouse and zebrafish orthologues, which all bind GAGAG sequence in hox genes clusters as in *Drosophila* (Matharu et al, 2010). In vertebrates, longer and less well defined GA stretches can be bound (Matharu et al, 2010). The protein structure is well conserved across the animal kingdom except for the C<sub>2</sub>H<sub>2</sub> zinc finger domain, which is repeated four times in vertebrates but only once in flies (Figure 1).

#### **4. GAF functional homologues in plants**

GAF homologues could not been identified in plants using basic protein sequence comparisons. Nevertheless GA or CT binding proteins were discovered based on their binding abilities. CT binding proteins were first identified in soybean where GBP was found to bind CT stretches essential for the expression of the chlorophyll synthesis gene *Gsa1* (Sangwan & O'Brian, 2002). This study led to the discovery of a new type of protein that binds CT repeats. BLAST searches then identified orthologues in rice (*Oryza sativa*), tomato (*Lycopersicon esculentum*), potato (*Solanum tuberosum*) (Sangwan & O'Brian, 2002), *Arabidopsis thaliana* and barley (*Hordeum vulgare*) (Santi et al, 2003). All these genes show conservation in their C-terminal DNA binding domain (Meister et al, 2004) (Figure 1).

CT repeats in the vicinity of the transcriptional start were found to influence plant gene transcription (Bao et al, 1997; Bolle et al, 1996; Burke & Kadonaga, 1997; de Boer et al, 1999; Enjuto et al, 1995; Sangwan & O'Brian, 2002). The addition of CT repeats to a GAF recognition site was shown to enhance the associated transcriptional activity (De Amicis et al, 2007; Pauli et al, 2004). The same phenomenon was observed in chicken, in the promoter of the malic enzyme gene, where CT repeats were demonstrated to be essential for promoter activity (Xu & Goodridge, 1998).

The barley BBR protein recognizes (CT)<sub>8</sub> stretches in intron IV of the homeobox gene *Bkn3* (Santi et al, 2003). In *Arabidopsis*, seven BPC proteins were found to bind CT elements and classified into three classes according to their amino-acid sequence: BPC1, BPC2 and BPC3 are in the class I, BPC4, BPC5 and BPC6 are in class II and BPC7 is the only member of class III (Meister et al, 2004; Monfared et al, 2011). All but one (BPC5 that is likely to be a



pseudogene) of these proteins is ubiquitously expressed (Monfared et al, 2011) like BBR in barley (Santi et al, 2003) or dGAF in *Drosophila*. As was studied for dGAF, the DNA selectivity of these plant GAGA binding proteins, was tested *in vitro*: BPC1 interacts with the RGARAGRRA consensus sequence where R is A or G; BPC7 can bind to at least the (CT)<sub>6</sub> motif (Meister et al, 2004).

CT microsatellites were analyzed in the rice and *Arabidopsis thaliana* genomes and are dispersed throughout the genome like in *Drosophila* or human but, nevertheless, an enrichment in 5' flanking coding regions was observed (Fujimori et al, 2003). Moreover, multiple BPC mutants in *Arabidopsis* were recently identified which appear to be impaired in several developmental processes such as cell growth and shape, and hormone perception (Monfared et al, 2011).

These data suggest that the plant BBR/BPC proteins, like animal GAF, are likely to be involved in many gene transcriptional regulation processes. In addition, BBR/BPC were not only implicated in the activation of gene transcription, such as *INO* (Meister et al, 2004), *BKN3* homeobox and *LEC2* gene activation (Berger et al, 2011) but also in repressing transcription, since BPC1 was shown to repress *STK* expression in ovules. This observation is of major importance because it provides the first evidence that BBR/BPC proteins can have a dual, activation or repressive, effect on transcriptional activity, as was shown in animals for CT binding proteins.

We do not know if BBR/BPC function is dependent on as yet unknown interacting partners, but the BPC6 N-terminal contains an Alanine Zipper-Like Coiled-Coil domain essential for BPC6 homo or hetero dimerization with other members of the BPC class II family (Wanke et al, 2011). It is thus tempting to speculate that this domain is functionally similar to the BTB/POZ domain in GAF, even if this domain was not found in other BBR/BPC proteins. To date, the oligomerization properties of other BBR/BPC have not been studied but, interestingly, multiple CT stretches were identified in the *INO* and *STK* promoters. Similarly, CT elements in the *STK* promoter are essential for the chromatin modification mediated by BPC1 *in vitro*, suggesting that BPC function likely depends on

Table 1 | GAGA factor function across kingdoms

	<b>Organisms</b>	
	<b>Animals</b>	<b>Plants</b>
insulator function	+	-
nuclear division	+	?
chromosome integrity	+	?
nucleosome sliding	+	?
Histone modification	+	?
PIC stabilisation	+	?
gene transcription activation	+	+
gene transcription repression	+	+
Chromatin remodeling	+	+
Homeobox genes regulation	+	+
hormone response	+	+

oligomerization as in *Drosophila*.

## 5. Concluding remarks

In animal cells, dGAF orthologues carry out an insulator function, a mechanism which remains unknown to date in plants. However, other mechanisms involving GAGA binding factors in animals are conserved across kingdoms. The function of dGAF was shown to be largely governed by interacting factors. Homo and hetero-dimerization were observed between class II BPC members, in the same manner as dGAF forms associations with itself and other BTB/POZ proteins. The link between BPC and PcG can certainly be investigated through the study of *INO*, *STK* and *LEC2* regulation. All three loci contain H3K27me3 mark deposition in plantlets, and *LEC2* was shown to be a target of PcG in vegetative tissues (Bouyer et al, 2011; Makarevich et al, 2006). It would thus be interesting to test the possible link between BPC and PcG complexes at these loci. Other plant orthologues of dGAF interactors are well characterized and it would be worthwhile investigating their capacity to interact with BBR/BPC proteins.

The occurrence and importance of CT repeats in plant genomes together with the pleiotropic phenotypes of multiple *Arabidopsis bpc* mutants (Monfared et al, 2011), provide crucial indications about the general role of BPC proteins in plant development (table1). We expect that future studies will demonstrate the overall importance of CT binding proteins in the development of life.

## Acknowledgments:

We are grateful to Christian Dubos and Loïc Lepiniec for discussions and helpful comments on the manuscript.



### II.2.2.2 Surexpression des BPC

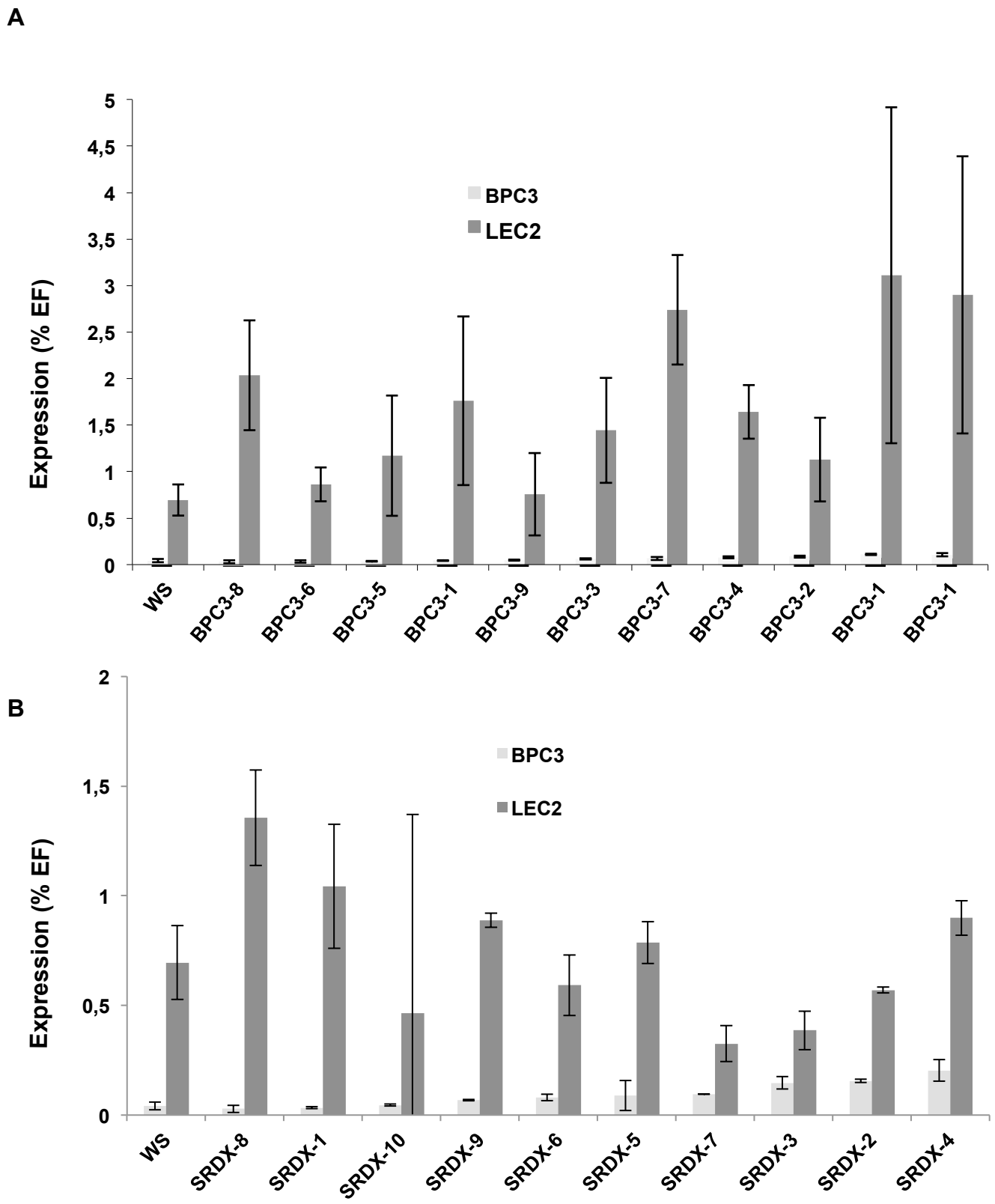
Pour tenter d'estimer l'effet des BPC sur l'expression de *LEC2 in planta*, l'effet sur l'expression du gène *LEC2* d'une surexpression des BPC dans différents organes a été étudié. Pour cela, j'ai réalisé une grande quantité de construits permettant d'exprimer tous les BPC sous contrôle de différents promoteurs. Ces constructions sont décrites dans le tableau ci-dessous.

Promoteurs	Gènes						
		BPC1	BPC2	BPC3	BPC4	BPC6	BPC7
35S	√	√	√	√	√	√	√
P <sub>LEC2</sub>			√				
P <sub>LEC2</sub> rand			√				

Les construits contenant le promoteur constitutif 35SCAMV ont été réalisés dans PGWB2 (voir annexes), les construits contenant les promoteurs de *LEC2* sauvage (PLEC2) et mutants (PLEC2rand) ont été réalisés dans pBIB-Hyg-GTW (Dubos et al, 2008). Aucune plante contenant les constructions avec toutes les BPCs n'a montré de surexpression des gènes *BPC*, avec le promoteur 35S ou les PLEC2 et PLEC2rand. Il semblerait qu'il existe des phénomènes de cosuppression qui ne permettent pas à ces protéines de s'exprimer au-delà de leur niveau endogène qui est très faible.

L'ADNc de *BPC3* a été cloné pour produire une protéine native (Figure 17A) ainsi qu'une fusion avec le domaine répresseur SRDX (Hiratsu et al, 2003) afin de créer un système d'expression dominant négatif (Figure 17B). L'expression du transcrit *BPC3* a été mesurée en PCR-quantitative dans 11 (*BPC3*) et 10 (*BPC3::SRDX*) transformants indépendants. *BPC3* est exprimé tellement faiblement qu'il est difficile d'établir la significativité de la surexpression dans ces plantes qui varient de 0,03 à 0,2 en pourcentage d'expression du gène de référence *EF1αA4*. En conclusion, dans nos conditions, il semble impossible de produire en grande quantité un transcrit codant pour une protéine BPC. Une alternative aurait pu être d'utiliser un système inductible pour tenter de trouver un organe où une surexpression est





**Figure 17: Expression de *BPC3* et *LEC2* dans des graines à 10 JAF dans des plantes contenant le construit *PLEC2:BPC3* (A) ou *PLEC2:BPC3::SRDX* (B).**

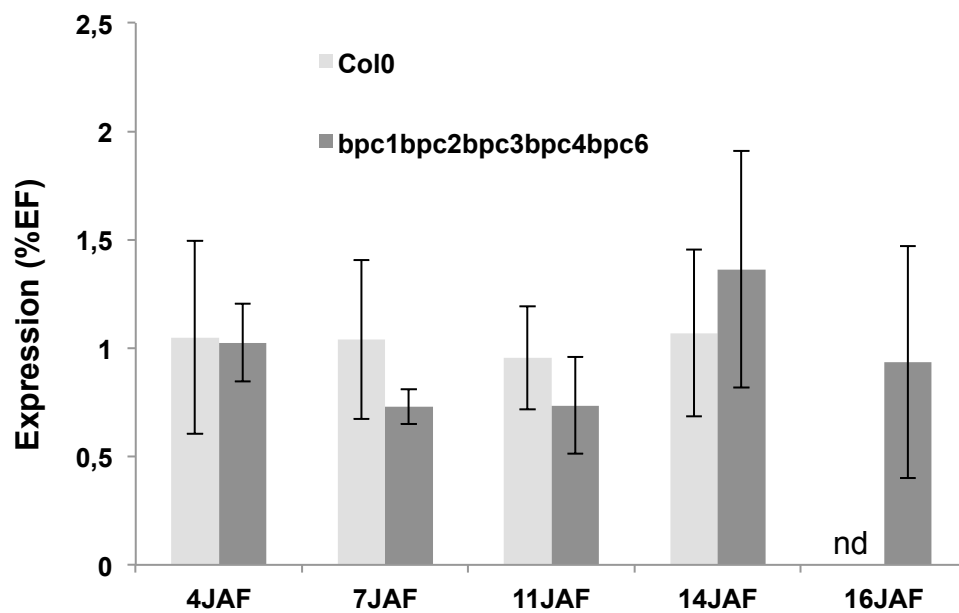
possible mais nous voulions rester dans le contexte graine. La deuxième solution s'offrant à nous a donc été d'étudier l'expression de *LEC2* en fond mutants *bpc*.

### ***II.2.2.3 LEC2 en fonds mutants bpcs***

Afin de déterminer le rôle des BPC dans l'expression de *LEC2 in planta*, j'ai commencé par tester les mutants nuls disponibles (à l'époque) et par réaliser des doubles et triples mutants entre *bpc1*, *bpc2* et *bpc3*. Comme on a pu le voir dans les données supplémentaires du chapitre II.1, l'expression de *LEC2* n'est absolument pas altérée dans ces mutants tout au long du développement.

Récemment, Monfared et collaborateurs ont publié des quadruples et quintuples mutants entre *bpc1*, 2, 3, 4, et 6 (Monfared et al, 2011). Ils nous ont aimablement fourni ces plantes et j'ai donc vérifié si l'expression de *LEC2* était affectée durant le développement de la graine de ces mutants. Comme on peut le voir dans la figure 18, le profil d'expression de *LEC2* ne semble pas différent entre une plante Col0 sauvage et le mutant *bpc1bpc2bpc3bpc4bpc6*. Le seul point qui paraît varier est à 16 jours après fécondation où *LEC2* est encore détectable dans le mutant multiple contrairement à la plante sauvage. Cette expérience n'a été réalisée qu'une fois, sur siliques entières, et de nouvelles plantes sont en serre pour répéter l'expérience. Un autre point surprenant dans ces résultats est le niveau d'expression de *LEC2* qui, même dans les trois plantes sauvages testées, apparaît quasi constant entre 4 et 14 jours après fécondation alors que le profil attendu est un pic d'expression à 10 jours. Toutes ces données demandent donc à être répétées, avant toute interprétation hasardeuse. L'hypothèse la plus probable est que BPC7, qui est le seul BPC à être fonctionnel dans le multiple mutant utilisé, agit de façon redondante avec d'autres BPC pour permettre une expression de *LEC2*. Il nous faudrait donc considérer un sextuple mutant *bpc*. Dans l'état actuel de nos résultats nous ne pouvons donc pas exclure que la boîte CT est, certes, indispensable à une expression correcte de *LEC2* mais que ce sont d'autres types de facteurs qui s'y fixent que ceux habituellement connus comme se fixant sur ces éléments.

### ***II.2.2.4 Analyse des promoteurs des BPC***



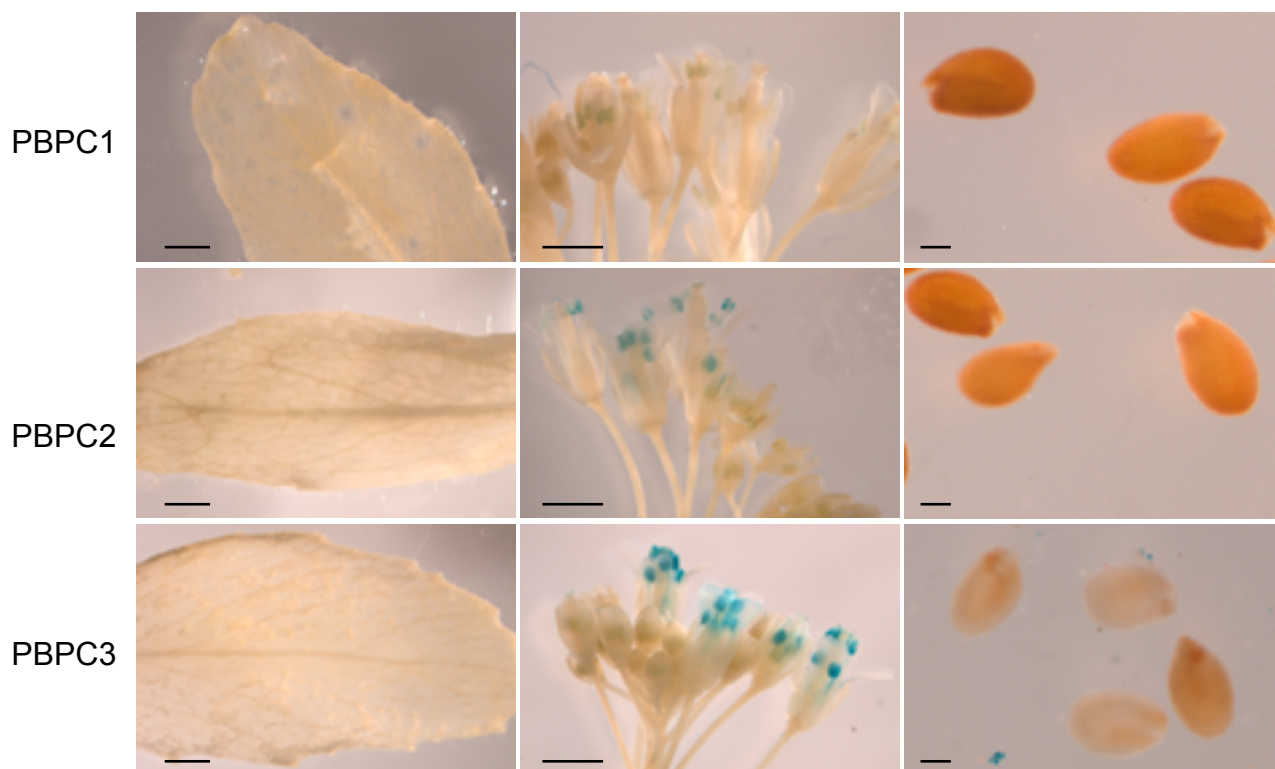
**Figure 18:** Expression de *LEC2* dans le mutant *bpc1bpc2bpc3bpc4bpc6* au cours du développement de la graine. (JAF:jour après floraison; nd: non détecté)

L'hypothèse de travail étant que les BPC permettent l'expression de *LEC2*, il fallait donc vérifier les profils d'expression des BPC au niveau tissulaire. Une RT-PCR au plateau sur des ARNm extraits de graines disséquées a permis de déterminer que tous les BPCs s'expriment dans l'embryon (cf chapitre II.1), mais sans indiquer dans quelle couche cellulaire exactement. J'ai donc réalisé le clonage des promoteurs des gènes de la classe I que nous pensions impliquée dans la régulation de *LEC2* afin d'en déterminer leur patron d'expression. De façon très surprenante, les lignées obtenues présentaient des profils inattendus avec une spécificité d'expression dans le pollen (figure 19). L'hypothèse la plus facile à envisager est que les fragments de 2kb de promoteurs utilisés ne contiennent pas toutes les informations nécessaires à une expression native. Peu de temps après l'obtention de nos résultats, Monfared et collaborateurs ont publié les promoteurs de tous les BPC fusionnés au gène rapporteur *GUS* et en utilisant 2,5kb de séquence nucléotidique en amont du site d'initiation de la traduction. Les profils d'expression obtenus correspondent bien aux données d'expression des bases de données, ainsi qu'à leurs propres résultats d'hybridation *in situ* (Monfared et al, 2011). Il apparaît donc clair que des éléments cis-régulateurs sont présents dans la zone comprise entre -2,5kb et -2kb par rapport à l'ATG dans ces promoteurs. Cette donnée est d'un grand intérêt pour une étude fonctionnelle complète de ces gènes.

### **II.3 Facteurs négatifs, données supplémentaires**

#### **II.3.1 PICKLE et le promoteur de *LEC2***

Le promoteur de 500pb de *LEC2* suffit à reproduire le profil d'expression d'un promoteur de 2,5kb et permet de compléter le mutant *lec2* quand il est placé en amont de l'ADNc de *LEC2* (cf chapitre II.1). Sachant que *LEC2* est exprimé de façon ectopique dans la pointe de la racine primaire juste après la germination dans un mutant *pkl* (Ogas et al, 1997; Ogas et al, 1999), j'ai introduit le construit PLEC2 :GUS dans ce fond génétique. De façon très surprenante, nous n'avons pu retrouver cette mauvaise expression. Comme on peut le voir figure 20A, aucune coloration GUS n'est obtenue dans les plantules contenant PLEC2 :GUS en fond *pkl1.1*, bien que le transcrit *LEC2* endogène soit bien détecté (Figure 20B). Ceci est très intéressant et permet alors de conclure à un comportement différent de la copie endogène de *LEC2* par rapport au transgène PLEC2 :GUS. Il semble donc y avoir des éléments manquants nécessaires à la répression par PKL dans la séquence clonée. Il serait très informatif d'introduire des promoteurs de *LEC2* de taille différente fusionnés au gène *uidA*

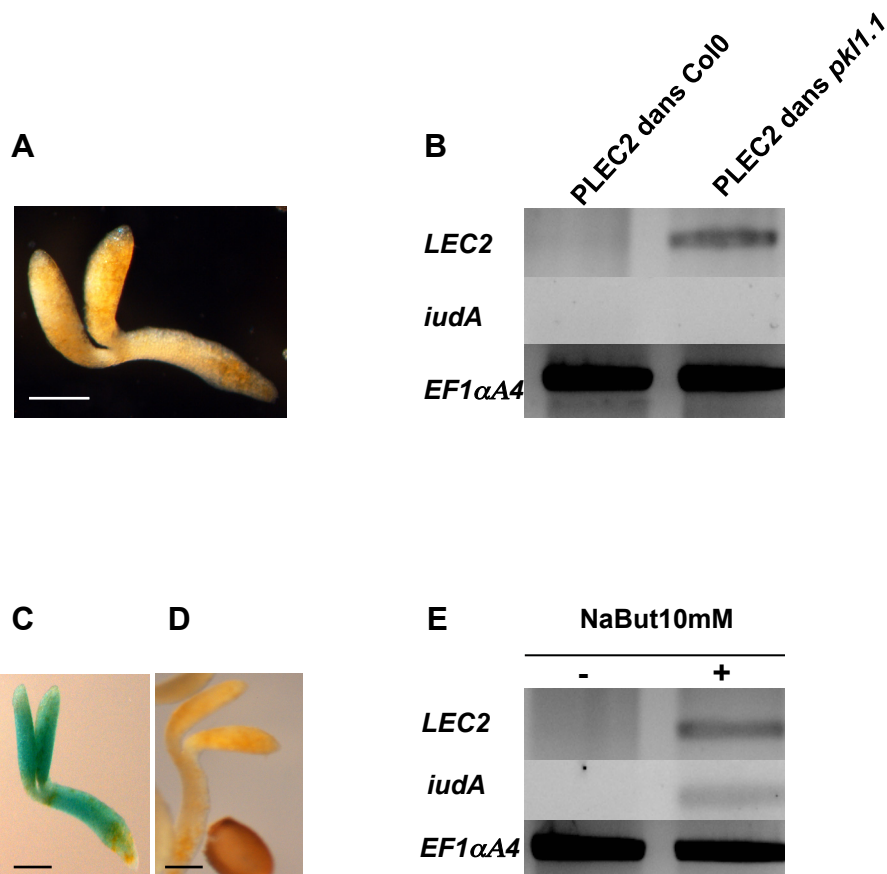


**Figure 19: Profils d'expression des promoteurs de la classe I des BPC.** Coloration GUS sur des plantes contenant les promoteurs de 2kb de *BPC1*, *BPC2* et *BPC3* fusionnés au gène rapporteur *iuidA* (scale bars= 0,2mm).

jusqu'à obtention d'une dérèpression dans les plantules. On peut noter également qu'aucune coloration GUS n'est observée dans les pointes de racine primaire des plantes contenant le construit PLEC2 :GUS, ce qui devrait cependant être le cas si on pose comme hypothèse que la séquence nécessaire à la répression par PKL dans PLEC2 est incomplète. Peut-être a-t-on perdu également une séquence activatrice en considérant uniquement un promoteur de *LEC2* de 500pb, séquence qui serait responsable de cette activation différente dans le mutant *pkl*. Cependant, dans cette dernière hypothèse, l'absence de cette séquence activatrice hypothétique ne semble pas perturber la restauration phénotypique de *lec2* par PLEC2500 :LEC2. Comme la séquence de PLEC2 est suffisante pour être réprimée par les PcG (cf, I.1), nous avons ici mis en évidence un outil très utile pour découpler l'action des PcG et de PKL et ainsi disséquer ces relations qui semblent très complexes (Figure 13).

### II.3.2 HDAC

Les histones acétylées transférases (HDAC) ont été montrées comme jouant un rôle dans le contrôle de l'expression des gènes (voir chapitre I-3-4-2-2). Pour aborder cette question vis-à-vis d'un rôle possible sur l'expression de *LEC2* nous avons fait germer des graines de plantes contenant le construit PLEC2 :GUS sur butyrate de sodium (NaBut) qui est un inhibiteur des HDAC (cf chapitre I.3.4.2). Comme on peut le voir figure 20C, 20D et 20E, le gène *iudA* est bien réactivé dans des plantules germant en présence de NaBut, suggérant une répression de *LEC2* par les HDAC. Contrairement au cas de PKL, nous pouvons dire, à partir de cette expérience, que les éléments nécessaires à la répression par les HDAC sont bien présents dans les 500pb du promoteur de *LEC2*. La délétion de la zone de -200 à -150 appelée *RLE* (cf II.1), dans le promoteur de *LEC2* permet la dérèpression de *LEC2*. On peut donc faire l'hypothèse qu'en plus de contenir les informations pour la répression par les PcG, elle contient peut-être les éléments cis-régulateurs nécessaires à l'action des HDAC. Les HDAC sont d'ailleurs peut-être recrutées par les mêmes complexes se fixant à l'ADN que les PcG dans le cas de *LEC2*. Il est important de noter que, à notre connaissance, l'acétylation des histones aux loci des AFL n'a pas été étudiée dans les mutants HDAC ou sur NaBut. On ne peut donc pas lier de façon directe la dé-répression des AFL dans des plantules germées sur inhibiteur avec une action directe des HDAC au locus. Les HDAC pourraient également réprimer l'expression d'un transcrit codant pour une protéine capable d'activer l'expression des AFL, de façon indirecte.



**Figure 20: Activité de PLEC2 en conditions de dérepression.** Coloration GUS de plantules contenant le construit PLEC2:GUS en fond *pkl1.1* (A) et test en RT-PCR au plateau de l'expression de *LEC2*, du gène *iudA* et de *EF1αA4* comme témoin positif dans des plantules contenant le construit PLEC2 en fond génétique *Col0* sauvage et en fond *pkl1.1*(B). Coloration GUS de plantules contenant le construit PLEC2 en fond sauvage sur NaBut (C) et sur milieu normal (D). RT-PCR sur les même plantules dans ces deux même conditions (E) (scale bars = 0,1mm).

## **II.4L 'activité de LEC2 est elle régulée post-traductionnellement ?**

L'ensemble des travaux réalisés dans le cadre de ce travail s'est attaché à la régulation transcriptionnelle de l'expression de *LEC2*. L'expression ectopique de *LEC2* conduit à des phénotypes sévères et la question d'une éventuelle régulation post-traductionnelle pour réguler l'abondance ou l'activité de la protéine *LEC2* semble légitime.

La façon la plus simple et directe de tester si *LEC2* est modifiée post-traductionnellement est d'obtenir des extraits protéiques d'embryons, de les immuno-précipiter avec des anticorps spécifiques anti-*LEC2* et de réaliser un western blot sur ces extraits avec des anticorps détectant les protéines phosphorylées, ubiquitinylées et/ou sumoylées par exemple. Le problème majeur pour ce genre d'étude est que nous ne disposons pas d'anticorps anti-*LEC2* (cf II.5.1) ni de construits permettant la production d'une protéine *LEC2* étiquetée et fonctionnelle jusqu'à récemment (cf chapitre II.4.2). La démarche que nous avons choisie a donc été indirecte. Nous avons tout d'abord recherché des sites potentiels de modifications post-traductionnelles, nous les avons mutagénisés et nous avons regardé l'effet produit sur l'activité de la protéine *LEC2* en testant la réversion phénotypique du mutant *lec2-4*. Les prédictions *in silico* pour la phosphorylation et l'ubiquitination fournissent soit trop d'acides aminés cibles pour notre démarche indirecte, soit ne sont pas assez fiables (comme la prédiction des sites d'ubiquitination). En attendant d'avoir des outils nous permettant de détecter si *LEC2* est phosphorylée ou ubiquitinée, je me suis donc attachée à étudier les sites potentiels de sumoylation dont la prédiction semble facilitée par l'existence d'une unique enzyme E2 responsable du transfert du peptide SUMO sur ses cibles (pour revue (Geiss-Friedlander & Melchior, 2007))

### **II.4.1 La sumoylation**

Ces dernières années, plusieurs équipes ont identifié, tout d'abord chez la levure *Saccharomyces cerevisiae* (Meluh & Koshland, 1995), des protéines similaires structurellement bien que présentant seulement 18% d'identité en acides aminés avec l'ubiquitine, appelées SUMO. Ces protéines peuvent se fixer de façon réversible et covalente à des protéines cibles. Contrairement à l'ubiquitine, elles ne semblent pas orienter les protéines vers la dégradation, mais peuvent, au contraire, entrer en compétition avec l'ubiquitination au niveau des mêmes résidus lysine et ainsi protéger cette protéine de la dégradation par le protéasome 26S (Ulrich, 2005). La sumoylation peut modifier la structure

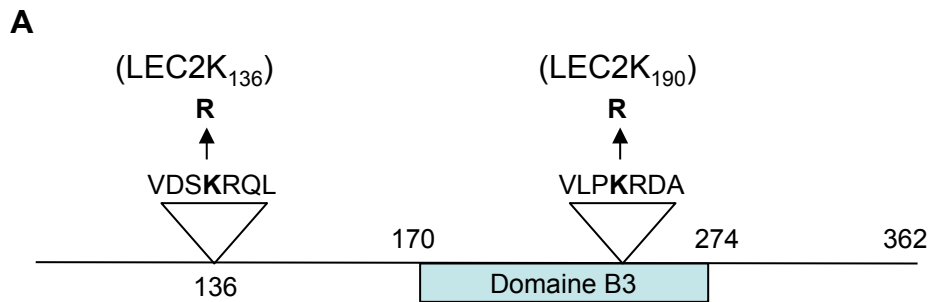










d'une protéine, ou ses interactions avec ses partenaires et, ainsi, en modifier sa fonction (Geiss-Friedlander & Melchior, 2007) ou sa localisation (Terui et al, 2004). La sumoylation est donc susceptible de modifier la stabilité, la localisation ou l'activité d'une protéine (Geiss-Friedlander & Melchior, 2007). La sumoylation a été impliquée dans un grand nombre de mécanismes chez la levure et les métazoaires, comme le cycle cellulaire, la maintenance de l'intégrité du génome, le transport sub-cellulaire et la répression ou l'activation de la transcription de certaines cibles (pour revue (Hay, 2005). Il existe une seule enzyme responsable du transfert des protéines SUMO, ce qui permet de déterminer un site consensus cible de sumoylation de façon plus fiable que pour l'ubiquitination. De plus, les mécanismes animaux de sumoylation semblent conservés chez les plantes (Kurepa et al, 2003; Novatchkova et al, 2004) et touchent la signalisation, le développement et les réponses aux signaux hormonaux et environnementaux (Kurepa et al, 2003; Lois et al, 2003; Miura et al, 2007; Miura et al, 2005; Murtas et al, 2003; Saracco et al, 2007; Yoo et al, 2006).

#### II.4.2 LEC2 et la sumoylation

Dans le but de déterminer si des acides aminés de la protéine LEC2 peuvent être potentiellement la cible d'une sumoylation, la séquence protéique de LEC2 a été analysée avec le logiciel de prédiction sumosp (<http://sumosp.biocuckoo.org>). En réglant le seuil de sensibilité au plus bas, j'ai pu identifier deux sites potentiels (Figure 21A). J'ai modifié ces deux lysines (K) en arginines (R) afin d'empêcher une fixation du résidu SUMO potentiel, sans modifier les propriétés de la protéine dans cette zone. Le deuxième site potentiel est situé dans le domaine B3 essentiel à la fixation du facteur de transcription à l'ADN. Trois construits ont été réalisés : deux construits avec chaque lysine mutée individuellement et un construit contenant les deux lysines mutées. Ces construits ont été réalisés dans pBIB-Hyg-GTW, introduits dans le mutant *lec2-4* et les descendants ont été analysés. J'ai sélectionné des plantes 100% résistantes à l'hygromycine et homozygotes pour la mutation *lec2-4* puis confirmé l'absence de phénotypes *lec2* mutants dans ces lots de graines (Figure 21B). J'ai donc pu montrer que, si sumoylation il y a, elle n'a aucun impact sur la fonction de la protéine LEC2, les protéines modifiées permettant une restauration complète du phénotype. L'absence de sumoylation aurait également pu mener à une sur-accumulation de la protéine LEC2, dans ce cas, le phénotype de prolifération cellulaire observé chez des surexprimeurs de la protéine LEC2 aurait été détectés, ce qui n'a pas été le cas (Stone et al, 2001). L'influence



**B**

Fond génétique	Construit	
WT	Aucun	
<i>lec2</i>	Aucun	
	PLEC2:LEC2	
	PLEC2:LEC2K <sub>136</sub>	
	PLEC2:LEC2K <sub>190</sub>	
	PLEC2:LEC2K <sub>136+190</sub>	

**Figure 21: Effet de la mutation des sites potentiels de sumoylation sur la protéine LEC2.** (A) Les 2 lysine (K) prédites comme des sites potentiels de sumoylation dans la protéine LEC2 sont indiquées en gras et mutées en arginine (R). (B) Photographies des graines des plantes sauvages, *lec2.4* et *lec2.4* complémentées par des construits PLEC2:LEC2, PLEC2:LEC2K<sub>136</sub>, PLEC2:LEC2K<sub>190</sub> et PLEC2:LEC2K<sub>136+190</sub>.

de la sumoylation ne sera donc pas étudiée plus avant, bien que la présence d'un site de sumoylation dans le domaine B3 pouvait laisser envisager une modification possible du site de liaison à l'ADN.

## **II.5 Outils pour la détection de la protéine LEC2**

### **II.5.1 Anticorps anti-LEC2**

Un des objectifs de mon travail était d'étudier la régulation post-traductionnelle de LEC2, il était donc nécessaire de tenter d'obtenir des outils de détection et/ou de purification de la protéine LEC2 *in planta*.

Il n'existait pas d'anticorps anti-LEC2 disponibles au début de ce travail, j'ai donc tenté par plusieurs approches d'en obtenir. Tout d'abord, j'ai produit dans *Escherichia coli* la protéine LEC2 entière, fusionnée à un tag 6xHIS, qui apparaissait comme soluble. J'ai envoyé ces échantillons purifiés à l'entreprise BIOGENE qui m'a fourni des anticorps qui reconnaissaient LEC2 produit dans *E.coli* mais pas extrait de végétaux. Nous avons tenté également une collaboration avec l'animalerie de INRA de Jouy-en-Josas grâce à l'aide d'Erwana Harscoët, technicienne au laboratoire, mais, cette fois encore, nos tentatives ont échoué et aucun de ces anticorps produits dans des lapins n'était utilisable. Après tous ces essais infructueux, nous nous sommes dirigés vers l'utilisation d'une protéine étiquetée fusionnée à LEC2.

### **II.5.2 Obtention d'une protéine étiquetée fonctionnelle**

J'ai réalisé un certain nombre de construits permettant d'exprimer *LEC2* sous le contrôle de différents promoteurs et fusionné à différentes étiquettes (Tableau 5). Le seul construit à avoir permis une réversion phénotypique du mutant *lec2-4* est le construit PLEC2 :Flag ::LEC2. Ce résultat récent est très encourageant car il va nous permettre, en premier lieu, de tenter de localiser la protéine LEC2 par immunocytochimie sur des coupes de graines, ce qui n'a jamais été réalisé. De plus, les modifications post-traductionnelles vont pouvoir être explorées plus avant. En dehors du cadre de cette thèse et de l'étude la régulation de l'expression de *LEC2*, cet outil va pouvoir être utilisé pour étudier les partenaires potentiels de LEC2 par des manipulations de pull-down ainsi que pour prouver, *in vivo*, par des manipulations de ChIP, la liaison à des cibles identifiées par des données génétiques. Il convient cependant de préciser

promo \ tag	GFP	STREP	HA	Flag
35S	N-ter,C-ter	N-ter,C-ter	N-ter,C-ter	N-ter,C-ter
PLEC2	N-ter,C-ter	N-ter,C-ter	N-ter,C-ter	<b>N-ter,C-ter</b>

**Tableau 5: Construits réalisés pour la production d'une protéine LEC2 étiquetée.** Les construits ont été réalisés sous le contrôle du promoteur 35S et du promoteur endogène (PLEC2) avec les différentes étiquettes en position N et C-terminal (N-ter, C-ter). Le construit en gras est le seul à compléter le mutant *lec2-4*.

que les anticorps anti-flag détectent cinq bandes dans un extrait protéique de feuilles d'*Arabidopsis*, et nous ne savons pas encore si ces bandes parasites seront présentes ou non dans des extraits de graines.



### III DISCUSSION GENERALE/CONCLUSION

Les objectifs de cette thèse s'étendaient sur 3 domaines : 1) la régulation transcriptionnelle de *LEC2*, 2) l'identification de nouveaux régulateurs et 3) l'obtention d'outils d'analyse biochimique de *LEC2*. Le point 1 a permis d'obtenir de nombreux résultats originaux, les deux autres dans une moindre mesure, et nous allons discuter dans ce chapitre des perspectives de ces travaux.

#### **III.1 Avancées dans la régulation de *LEC2* :**

Comme nous l'avons vu dans le chapitre II.1, l'expression cordonnée de *LEC2* dans la graine et son absence dans les parties végétative semble reposer sur trois séquences, ou « boîtes », de régulation. La première boîte identifiée essentielle pour l'activité du promoteur de *LEC2* correspondrait à une boîte MADS. Les différents cribles simple-hybride réalisés n'ayant donné aucun résultat, nous nous sommes donc concentrés sur les données bibliographiques qui désignaient AGL15 et AGL18 comme les seuls facteurs MADS connus pour intervenir dans la régulation de l'expression de *LEC2*. Comme nous l'avons vu au chapitre II.2.1, les données préliminaires, que nous avons obtenues avec des expériences de simple hybride, ne permettent d'émettre aucune hypothèse quant à l'implication de ces deux protéines au locus *LEC2*. Des études plus poussées devraient être menées en crible triple hybride pour tenter de trouver un 3<sup>ème</sup> partenaire potentiellement impliqué. Etudier l'interaction en contexte hétérologue n'est peut-être pas la méthode la plus adaptée dans ce cas, sachant qu'aucune manipulation de ce type n'a jamais été publiée avec AGL15. L'étude de l'activation de *PLEC2* dans des protoplastes de mousse (Thevenin et al, 2011), ou des expériences d'immunoprécipitation de chromatine (ChIP-PCR) *in planta* avec des anticorps polyclonaux spécifiques d'AGL15 obtenus par une autre équipe (Wang et al, 2002) nous permettraient d'avoir des renseignements plus fiables sur la zone de fixation d'AGL15 sur le promoteur de *LEC2*.

La 2<sup>ème</sup> boîte de régulation positive mise en évidence dans ce travail a permis de souligner l'impact possible des protéines BPC sur l'expression de *LEC2*, mais, hormis l'absence d'activité du promoteur *PLEC2* dans certains organes dans le triple mutant *bpc1bpc2bpc3*, aucune donnée ne semble corroborer une action réelle des protéines BPC sur la régulation de *LEC2* dans la graine. L'existence de six membres dans la famille BPC ne simplifie pas les choses pour étudier ces mécanismes. L'obtention d'un sextuple mutant sera nécessaire pour aller plus loin dans cette étude. Même si l'implication des BPCs n'a pas été





prouvée *in planta*, celle des MADS, des PcG et des HDAC a été démontrée par nous-même ou d'autres équipes (cf, introduction). L'introduction des BPC dans le modèle de régulation permet d'insérer une autre composante chromatinienne dans la régulation de *LEC2*, car les BPCs sont de plus en plus considérées comme des homologues fonctionnels du facteur GAGA de drosophile (chapitre II.2.2.2). En particulier, nous sommes en droit de nous demander si les BPCs interviendraient également dans le recrutement des MADS au locus.

La boîte de régulation négative nous a permis d'appréhender les mécanismes moléculaires liés à la structure chromatinienne. Nous savons désormais (cf chapitre II.1) qu'elle est nécessaire au dépôt de la marque H3K27me3 au locus et qu'elle est suffisante pour permettre le dépôt de cette marque faiblement, mais significativement, dans un autre environnement, ce qui correspondrait alors à un PRE (Polycomb Response Element). Pour être sûr d'avoir bien identifié un PRE végétal, il faudrait détecter la disparition d'un membre du complexe PRC2 au locus, en l'absence de la séquence RLE. Pour cela, nous sommes en train d'obtenir, par croisement, des plantes contenant les construits PLEC2rand :GUS (ou PLEC2 :GUS ainsi que PF3H+/-RLE) avec le construit 35S :CLF :GFP fourni par Justin Goodrich. Ces construits, placés dans un même fond génétique, devraient nous permettre de réaliser des expériences de CHIP-PCR sur la protéine CLF en utilisant les anticorps anti-GFP et de vérifier que la séquence RLE est capable de recruter des éléments du complexe PRC2. Enfin, il a été montré, chez les animaux, qu'aucun élément de PRC2 n'est capable de se fixer par lui-même à l'ADN. Il a été proposé que la protéine PHO fasse l'intermédiaire entre les deux éléments. Qu'en est-il chez les plantes? Est-ce un des répresseurs déjà identifié qui pourrait jouer le rôle de « recruteur »? Sont-ils plusieurs à avoir un rôle redondant, vu l'importance du mécanisme pour la répression d'un grand nombre de cibles? Le promoteur de *LEC2*, et la séquence RLE en particulier, semble donc être un excellent matériel pour tenter de répondre à ces questions.

Le premier objectif de cette thèse, qui était d'étudier précisément les mécanismes moléculaires de la régulation du gène *LEC2*, a donc pu être atteint en ce qui concerne la répression de l'expression de *LEC2* par PRC2. Le 2<sup>ème</sup> objectif, qui était de déterminer de nouveaux acteurs de la régulation de l'expression du gène *LEC2*, n'a été atteint que partiellement. J'ai en effet pu identifier les BPCs comme étant capables de se fixer au promoteur de *LEC2* et montrer que leur séquence de reconnaissance est essentielle à l'activité du promoteur bien que le construit PLEC2 :GUS ou l'ADNc endogène de *LEC2* ne montre pas de différence d'expression en fond génétique mutants *bpc*.



Une autre stratégie a été entreprise pour identifier de nouveaux régulateurs de *LEC2*, en particulier en exploitant la variabilité naturelle. Des résultats d'analyse d'eQTLs (expression des gènes à 10 jours après pollinisation) ont montré qu'un locus en haut du chromosome 5 contrôle en partie le niveau d'accumulation des ARNm de *LEC2*. Cette variation du niveau de transcrite a pu être confirmée en PCR quantitative puis dans les lignées isogéniques (HIF) mais le niveau de la variation et le type de caractère (mesure du transcrite à 10 jours) n'ont pas permis d'aller plus loin dans l'analyse de ce caractère dans le cadre de ce travail. Il est probable que ce type d'analyse, couplé aux nouvelles technologies de séquençage profond (RNAseq) permette d'identifier à l'avenir de nouveaux régulateurs de l'expression des AFLs en général et de *LEC2* en particulier.

### **III.2 *LEC2* est vraiment un très bon modèle pour l'étude de systèmes complexes de régulation**

Si l'on considère la complexité des systèmes de régulation d'un gène, *LEC2* représente un très bon modèle. En effet, son expression combine l'effet de facteurs de transcription (régulation transcriptionnelle pure) à des facteurs de modification de la chromatine (régulation épigénétique), ce qui permet d'avoir une vue globale des mécanismes mis en jeu. Dans ce sens, il apparaît comme plus « simple » que le gène *FLC* qui, lui, est régulé par à peu près tous les systèmes connus de modification de la chromatine ou de l'ADN mais aussi par des facteurs de transcriptions aussi bien activateurs que répresseurs. *LEC2* ne semble, lui, concerné que par des systèmes répresseurs, comme si son pic d'expression dans l'embryon n'était dû qu'à une machinerie de base de la transcription, avec *AGL15* et *AGL18* en appui, couplée à une absence ou une inactivité de ses répresseurs. L'étude du promoteur a déjà permis d'étudier les mécanismes moléculaires impliqués dans la modification H3K27me3, il pourrait permettre d'autres études.

Le fait que l'absence de la séquence RLE permette une dérpression totale de *PLEC2* dans les plantules suggère que les signaux de déacétylation sont présents également dans cette séquence RLE, en faisant un outil de choix pour étudier les mécanismes de la déacétylation au locus *LEC2*. Les expériences à suivre devront donc, en plus de conduire au profil de la marque d'acétylation au locus *LEC2* dans les graines, déterminer si cette acétylation est modifiée dans des construits génétiques *PLEC2* :*GUS* avec ou sans RLE. Un point essentiel à ne pas oublier est que l'on ne sait toujours pas si l'action des inhibiteurs des HDAC n'est pas



indirecte sur l'expression de *LEC2*, tant qu'une corrélation n'aura pas été établie entre la dépression de *LEC2* en présence de TSA et le taux d'acétylation au locus. Il est également possible que les signaux de déacétylation soient présents ailleurs dans le promoteur de *LEC2* mais que la déacétylation ne suffise pas à inhiber l'activité du promoteur, son effet étant inférieure à celui de la déméthylation des H3K27. Si on suppose que les HDAC sont recrutées à l'ADN par des partenaires de type facteurs de transcription, les HSI ou ASIL1 pourrait jouer ce rôle dans le cas de *LEC2*.

Dans cette étude, je n'ai pas eu le temps d'aborder la répression par les facteurs de transcription de type HSI ou ASIL1. Il serait très intéressant de déterminer si ces protéines interviennent directement en se fixant sur *PLEC2*, comme il a déjà été montré pour leurs autres cibles (Gao et al, 2009; Suzuki et al, 2007), ou si leur effet est indirect. Comme les auteurs ont pu réaliser ce type d'étude en contexte hétérologue (simple hybride), ces facteurs ne doivent pas avoir besoin de partenaires, ou ils ont des homologues dans la levure, et il devrait être relativement simple de tester en simple hybride si ces facteurs peuvent activer *LEC2*.

Les autres types de modifications les plus observées comme les H3K36 ou H3K9 n'ont jamais été étudiées au niveau de *LEC2*, cela constitue encore un point à éclaircir. *LEC2* n'a jamais été détecté comme une cible de petits ARNs non plus. J'ai testé si des ARNs non cappés de longue taille peuvent exister sur *LEC2* et je n'ai pu en détecter aucun, contrairement à *FLC*, mais ces aspects nécessiteraient des analyses approfondies. Enfin, la régulation par méthylation de l'ADN pourrait être une voie de régulation supplémentaire. Je n'ai pas du tout abordé cet aspect dans ce manuscrit car il semble, d'après les données existantes, que le locus *LEC2* n'est pas méthylé. J'ai, par ailleurs, vérifié par moi-même le profil de méthylation, et les données publiques ont pu être confirmées.

### ***III.3 Régulation de LEC2 et des AFLs pendant le développement de la graine et projet CERES***

Si *LEC2* est bien un facteur clé et déclencheur du processus de maturation chez *Arabidopsis*, les autres AFLs sont cependant également très intéressants à étudier. De nombreuses questions sont soulevées par cette étude, essentiellement liées aux mécanismes de répression de l'expression de *LEC2* et des AFLs dans les tissus végétatifs. Cependant, la régulation de l'expression spatiale et temporelle des AFLs durant le développement de la



graine n'a pas été réellement abordée. En effet, on ne sait pas, à l'heure actuelle, si les profils d'expression des *AFLs* dans la graine sont dûs à des facteurs de transcription seuls ou si les modifications de la chromatine, telles que les marques H3K27me3, interviennent également à ce stade de développement. En d'autres termes, la régulation que nous avons mise en lumière dans les parties végétatives reflète-t-elle la régulation ayant lieu dans la graine en développement ? Des analyses de ChIP sur la graine en développement seraient des expériences de premier choix pour répondre à cette question. L'absence d'information sur le sujet réside certainement dans le fait que les manipulations de ChIP nécessitent beaucoup de matériel, matériel devant être fixé rapidement avant congélation, ce qui ne laisse par conséquent pas beaucoup de temps pour la dissection des embryons. Des mises au point techniques sont donc nécessaires à la poursuite de ce travail, mises au point qui seront effectuées dans le cadre du projet ANR blanc CERES en cours au laboratoire (L. Lepiniec coordinateur). Ce projet consiste, en particulier, à étudier les profils chromatinien de tous les *AFLs* dans l'embryon pendant le développement de la graine. Si aucune différence dans l'accumulation de la marque H3K27me3, par exemple, n'est observée durant le développement de la graine, nous aurons montré que les facteurs chromatinien ne jouent aucun rôle dans l'expression des *AFL* dans l'embryon. Si, par contre, nous observons une corrélation entre l'expression et les profils réalisés, nous pourrions tenter de déterminer quelle modification est importante pour les profils d'expression des *AFLs* et quel(s) acteur(s) protéique(s) est (sont) impliqué(s) dans le contrôle de ces gènes pendant la formation de la graine. Les premiers profils seront réalisés sur les marques chromatinienne déjà impliquées dans les parties végétatives mais les autres modifications pourront être étudiées par la suite.

#### **III.4 Critique de la démarche de délétion de promoteurs et alternatives**

La démarche de délétion de promoteurs que j'ai menée est essentielle pour l'étude de la régulation d'un gène. En effet, les éléments cis-régulateurs sont les facteurs les plus importants dans cette régulation dans la majeure partie des cas. De plus, même si une expression différente entre le patron d'accumulation de l'ADNc endogène et le patron d'expression du gène rapporteur apparaît, cela peut permettre de mettre en évidence l'influence d'autres acteurs (zones cis-régulatrices dans les parties codantes ou modifications chromatinienne). Les limitations de cette approche ont largement été illustrées dans ce travail, car identifier des séquences nucléotidiques essentielles à une activité promotrice ne signifie pas en identifier les acteurs. En effet, nous n'avons pas pu déterminer, de façon non équivoque, *in planta*, de nouveaux acteurs de la régulation de l'expression de *LEC2* pour





aucune de nos séquences importantes. Enfin, si nous avons réalisé les délétions uniquement en 5'-3' comme c'est en général la règle dans ce genre d'études, et non en 3'-5', nous n'aurions pas pu détecter la présence de la boîte de répression et la régulation chromatinienne de *LEC2* n'aurait pas pu être abordée.

La puissance de l'analyse génétique, associée aux moyens récents de biologie moléculaire et séquençage permet d'envisager des cribles génétiques de plantes issues de mutagenèse chimique (EMS). Il serait maintenant tout à fait envisageable de cribler pour des caractères de modification de patron d'expression de *LEC2* en associant plusieurs marqueurs de gènes cibles pour éviter les mutations intragéniques ou encore des gains de fonction à l'aide de marqueurs de résistance à des antibiotiques. Le seul point négatif, quant à cette approche dans notre cas est que les facteurs de régulation de l'expression de *LEC2*, hormis FIE et ASIL1, agissent de façon redondante (les HSI, les HDAC, les MADS, PcG à domaine SET (CLF/SWN)). La probabilité d'identifier des acteurs redondants en réalisant une mutagenèse est quasi nulle, il était donc peu probable que cette technique nous permette de détecter une différence d'activité du promoteur. Cependant, si le laboratoire veut continuer à essayer de trouver de nouveaux acteurs pour la régulation de *LEC2* ou d'autres *AFLs*, cette démarche est probablement une des seules envisageables à ce stade.

La troisième démarche restant à envisager pour trouver des acteurs impliqués dans la régulation des *AFLs* pendant le développement de la graine est de refaire des cribles double-hybride, mais avec des banques d'ADNc spécifiques de la graine, à différents stades de développement. Il est en effet probable que les banques d'expression que nous avons utilisées ne contiennent pas tous les facteurs de transcription, en particulier ceux de la graine en développement qui peuvent être exprimés très faiblement. Il existe maintenant des kits permettant de réaliser directement et simplement, sans quantités très importantes d'ARNs, des banques facilement utilisables qui permettraient peut-être de détecter des interactants du complexe régulateur au niveau du promoteur de *LEC2*, sous-représentés dans des banques globales d'organes.

Bien que ce travail ait permis une meilleure compréhension des phénomènes mis en jeu dans le contrôle de l'expression de *LEC2*, il reste encore beaucoup de limitations techniques et conceptuelles pour avoir une vision globale et synthétique des mécanismes du contrôle de la formation de la graine et de l'accumulation des réserves.



## IV MATERIEL ET METHODES

### IV.1 Matériel

#### IV.1.1 Mutants d'*Arabidopsis thaliana*

Les oligonucléotides utilisés (Annexe 4) et la position des insertions dans le génome des mutants (Annexe 5) sont décrits en annexes.

Le mutant *bpc1* (SALK\_072966.43.30.x) (Kooiker et al, 2005) utilisé présente une insertion à -300pb par rapport à l'ATG qui doit être incomplète car les bordures gauche ou droite ne sont pas amplifiables dans nos conditions. La présence de l'insertion est vérifiée uniquement par l'absence de bande correspondant à la copie sauvage avec le couple d'oligonucléotides PBPC1(-600)f/BPC1rev donnant un fragment de 1173pb dans une plante sauvage.

Le mutant *bpc2* : on amplifie la bordure gauche du T-DNA du mutant *bpc2* utilisé dans ces manipulations (SALK\_090810C) avec le couple d'oligos sigLB1/BPC2(410)rev donnant une bande de 300bp et la copie sauvage avec les oligos BPC2f/BPC2(410)rev donnant une bande de 598pb.

Le mutant *bpc3* est un mutant EMS présentant une mutation de C vers T en position +691 par rapport à l'ATG qui insert un codon STOP ainsi qu'un site *DdeI* qui permet le génotypage. Une amplification avec les oligonucléotides BPC3(490)/BPC3AttB2 suivie d'une digestion avec l'enzyme de restriction *DdeI* permettent de détecter la présence de la mutation. Les trois mutants sont en fond Col0.

Les mutants multiples *bpc1bpc2bpc3bpc4bpc6* et *bpc1bpc2bpc4bpc6* ont été aimablement fournis par le Pr Gasser (Monfared et al, 2011).

Mutant *pkl.1* (Ogas et al, 1997)

Mutant *lec2.4* (Stone et al, 2001)

#### IV.1.2 Souches de levures et de bactéries

Les souches utilisées dans ce travail sont :



*Escherichia coli* souche DH10B : génotype : F<sup>-</sup> endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ<sup>-</sup>

*Escherichia coli* souche DB3.1 pour amplifier les vecteurs gateway contenant le gène de sélection négative ccdB: F- gyrA462 endA1 glnV44 Δ(sr1-recA) mcrB mrr hsdS20(r<sub>B</sub><sup>-</sup>, m<sub>B</sub><sup>-</sup>) ara14 galK2 lacY1 proA2 rpsL20(Sm<sup>r</sup>) xyl5 Δleu mtl1

*Agrobacterium tumefaciens* souche C58C1Rif. Cette souche désarmée est utilisée comme vecteur de transformation génétique. Le gène de résistance à la rifampicine est porté par le chromosome bactérien.

*Saccharomyces cerevisiae* souche YM4271. Génotype : MATa, *ura3-52*, *his3-200*, *ade2-101*, *ade5*, *lys2-801*, *leu2-3, 112*, *trp1-901*, *tyr1-501*, *gal4D*, *gal8D*, *ade5::hisG* (Clontech)

## **IV.2 Méthodes**

### **IV.2.1 Stérilisation des graines d'*Arabidopsis* et conditions de culture**

Les graines sont stérilisées durant 6 min dans une solution de Bayrochlore (1/4 de comprimé dissous dans 10 mL d'eau ultra-pure, la solution est diluée 10 fois dans l'alcool à 95%) puis rincées abondamment à l'éthanol 95%. Les graines sont semées *in vitro* sur un milieu Gamborg's B5 Basal Medium. Après 48 heures à 4°C, elles sont mises à germer 10 jours en chambre de culture (16 h de photopériode ; hygrométrie 65% ; température 20°C le jour, 18°C la nuit). Les plantules sont ensuite repiquées en terreau et cultivées en serre (16 h de photopériode ; hygrométrie 60% le jour, 55% la nuit ; température 20°C le jour, 18°C la nuit).

### **IV.2.2 Transformation génétique d'*Arabidopsis thaliana***

Une colonie isolée d'une souche d'*Agrobacterium tumefaciens* transformée avec le construit d'intérêt est cultivée dans 5 ml de LB (Annexe 7) contenant de la rifampicine et de la gentamicine pour sélectionner les agrobactéries et un antibiotique supplémentaire correspondant à la résistance du clone d'intérêt.

La culture est mise sous agitation à 28°C pendant 24h puis 1ml de cette préculture est utilisé pour ensemencer 500ml de LB contenant les mêmes antibiotiques. Au bout de 15h la culture est centrifugée à 6000g pendant 15 minutes à 4°C et le culot repris dans 250ml d'eau contenant 5% de saccharose. Du silwet (Witco, Tarrytown, NY) à 0,05% v/v jouant le rôle de



« mouillant » (réduisant la tension superficielle du liquide) est ajouté juste avant la transformation. Le protocole de transformation est celui de Clough (Clough & Bent, 1998).

### **IV.2.3 Constructions génétiques**

Les ADNc de *LEC2*, *AGL15*, *AGL18*, *BPC1*, *BPC2*, *BPC3*, *BPC4*, *BPC6* et *BPC7* ont été amplifiés avec des oligonucléotides contenant les sites recombinaison Gateway AttB1 et AttB2 avec ajout de 2 nucléotides en 5' pour pouvoir être en phase avec les vecteurs du laboratoire en cas de fusion traductionnelle ultérieure. (Oligonucléotides en annexe 4). Les ADNc amplifiés avec la polymérase à activité de relecture Phusion® (New England Biolabs, Ipswich, UK) sont ensuite transférés dans les vecteurs d'entrée par réaction BP (Gateway™, Invitrogen, Villebon sur Yvette, France), puis dans les vecteurs finaux par réaction LR en divisant par quatre les quantités préconisées par le fournisseur. Toutes les constructions ont été validées par séquençage.

### **IV.2.4 Colorations histochimiques GUS**

L'enzyme  $\beta$ -glucuronidase (GUS) hydrolyse le 5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronide (X-Gluc) en un indoxyl incolore. Ce composé intermédiaire forme alors, par dimérisation, des cristaux bleus insolubles. Les cellules ou tissus possédant de l'activité GUS présenteront donc une coloration bleue. Les explants à colorer sont placés dans la solution de coloration (Tampon phosphate de sodium pH7,2 100mM ; Na<sub>2</sub>-EDTA 10mM ; Tritonx100 0,1% ; X-Gluc 1mM) puis maintenus 1 heure sous-vide. Les échantillons sont ensuite placés une nuit à 37°C. Le lendemain, les échantillons sont lavés à l'eau puis à l'éthanol à 70% et laissés à décolorer sous agitation dans l'éthanol 70% pendant 24 à 48h. La décoloration dans l'alcool à 70% est renouvelée si beaucoup de pigments sont présents dans les échantillons (feuilles par exemple).

### **IV.2.5 Extraction d'acides nucléiques**

#### ***IV.2.5.1 Extraction d'ADN génomique rapide pour génotypage (Edwards et al, 1991):***





Prélever un fragment de feuilles de la taille d'un bouchon de tube eppendorf de 1,5 ml dans la glace. Ajouter un peu de sable de Fontainebleau et broyer à l'aide d'une pointe de pipetman de 1ml. Ajouter 300 µl de tampon d'extraction Edwards (Tris-HCl pH 7,5 200 mM ; NaCl 250 mM ; EDTA 25 mM ; SDS 0,5%) et centrifuger 10 min à vitesse maximale pour culotter les débris. Prélever le surnageant dans un nouveau tube, ajouter 250 µl d'isopropanol et centrifuger 10 min à vitesse maximale. Eliminer le maximum de liquide et reprendre le surnageant dans 30 µl d'eau. Utiliser 1 µl par PCR.

#### ***IV.2.5.2 Extraction d'ADN plasmidique***

Pour les extractions d'ADN plasmidique, le kit NucleoSpin® Plasmid (Macherey–nagel, Düren, Allemagne) a été utilisé selon les recommandations du fournisseur.

#### ***IV.2.5.3 Extraction d'ARNs et Réverse Transcription (RT)***

Les extractions d'ARNs sont réalisées à l'aide du kit Genelute™ Mammalian Total RNA miniprep Kit (Sigma-Aldrich, Saint-Quentin Fallavier, France), selon les recommandations du fournisseur. Les échantillons sont élués dans 50µl d'eau. On utilise 10,25 µl d'ARN totaux dans tous les cas sans doser car les analyses contiennent toujours un témoin interne. La rétrotranscription des ARNs en ADNc à l'aide du kit Superscript II RNase H-Reverse Transcriptase (Invitrogen) se décompose en quatre étapes. La première consiste à dénaturer les structures secondaires formées par les ARNs. Un mélange composé des matrices d'ARN traitées à la DNase I (1 [g) et de 1µl d'oligo dT (500 ng/ml), qui sert d'amorce à la synthèse du premier brin, est incubé 5 min à 65°C puis immédiatement placé dans la glace. Le premier brin peut alors être synthétisé : au mélange précédent est ajouté 1 µl de RNase Out (40U/ml) pour inhiber les RNases, 4 µl de tampon 5 X (Tris-HCl 250 mM pH 8,3, KCl 375 mM, MgCl2 15 mM), 1 µl de dNTPs (10 mM) et 2 µl de DTT (0,1 M). Le milieu réactionnel est préchauffé pendant 2 min à 42°C puis la synthèse du premier brin débute lorsque 1 µl de la reverse transcriptase SuperScript II (50U/ml) est ajouté. Cette réaction dure 50 min à 42°C. Elle est ensuite arrêtée en plaçant les échantillons 15 min à 70°C.



## IV.2.6 PCR en temps réel ou PCR quantitative

### IV.2.6.1 Principe

Au cours de la PCR en temps réel, un suivi du nombre de copies synthétisées au cours du temps va permettre d'estimer la richesse relative d'une séquence d'ADN d'intérêt comparée à une séquence de référence au sein d'un mélange d'ADNc servant de matrice. Ainsi, il sera possible d'évaluer le niveau d'expression relatif d'un gène par rapport à un autre gène utilisé comme référence. Le suivi du nombre de copies synthétisées est permis par la présence d'un agent intercalant, le Sybr Green I, dans le milieu réactionnel. Lorsqu'il est lié à de l'ADN double brin, celui-ci devient fluorescent (maximum d'émission à 530nm) après excitation, et la fluorescence détectée est directement proportionnelle à la quantité d'ADN présente. Au cours de la phase de dénaturation de l'ADN, le Sybr Green est libéré dans le milieu réactionnel et cesse d'émettre jusqu'à la séquence d'élongation suivante. Le suivi de la fluorescence à la fin de chaque cycle d'amplification permet de construire la cinétique d'accumulation des produits de la PCR qui a la forme d'une sigmoïde. Dans la portion linéaire de la courbe, le nombre de copies synthétisées  $X$  peut être calculé de la manière suivante :

$$X = X_0 \cdot (1 + \text{Eff})^n$$

$X_0$  représente le nombre de copies de la matrice initialement introduites dans le milieu réactionnel, Eff l'efficacité de la réaction, et  $n$  le nombre de cycles d'amplification réalisés. En comparant le nombre de cycles nécessaires à l'obtention d'une même quantité  $X$  de copies dans deux réactions conduites en parallèle sur un même mélange d'ADNc et avec deux couples d'oligonucléotides distincts, il est possible d'évaluer la richesse relative d'une séquence d'ADNc par rapport à une autre au sein du mélange. Le rapport des concentrations entre les deux séquences considérées est calculé de la manière suivante :

$$X_{0a}/X_{0b} = (1 + \text{Eff})^{(n_b - n_a)}$$

Avec  $n_a$  et  $n_b$  tels que  $X_a = X_b$  et  $\text{Eff}_a = \text{Eff}_b$

### IV.2.6.2 Protocole

Les réactions de PCR sont réalisées dans un thermocycleur à air pulsé couplé à un microfluorimètre (Light Cycler ; Roche Diagnostics). Le volume réactionnel est de 10  $\mu\text{l}$  et



comprend 2,5 µl de mélange de cDNA, 0,5 µl de chaque oligonuclotide. Chaque cycle de PCR comprend une phase de dénaturation de l'ADN à 95°C (15s), une étape d'hybridation des amorces à 55°C (5s) et une phase d'élongation à 72°C (6s). Au préalable, les ADNc ont été dénaturés pendant 8 min à 95°C, ce qui permet également d'activer la Taq polymérase utilisée (« hot start »). Afin de vérifier la spécificité du produit amplifié, la PCR à proprement parler est suivie d'une lente montée en température (0,1°C par seconde, jusqu'à 95°C). Au cours de cette rampe de température, la fluorescence va diminuant ; elle chute brusquement lorsque la T<sub>m</sub> du produit de PCR est atteinte. Il est ainsi possible de déterminer la nature et la pureté du produit PCR chaque fragment amplifié ayant une T<sub>m</sub> spécifique.

#### **IV.2.7 Immuno-précipitation de la chromatine (Le Page et al)**

Les manipulations de CHIP ont été réalisées à partir de plantules de 10 jours après semis, semées sur des boîtes de 120cm de diamètre sur milieu solide *Arabidopsis* Gamborg's B5 Basal Medium. Le protocole utilisé est celui mis au point dans le laboratoire de Vincent Colot, IBENS, Paris.

##### ***IV.2.7.1. Fixation des échantillons (jour 1)***

Prélever un volume de plantules équivalent à 20ml sans tasser dans un tube falcon de 50 ml rempli d'eau milli-Q dans la glace. On ne peut pas congeler les échantillons avant le cross-link, le passage dans l'eau glacée avant la fixation est donc impératif. Retirer un maximum d'eau et ajouter 40ml de formaldéhyde 1% (39ml d'eau milli-Q + 1 ml de formaldéhyde 40%) fraîchement préparé, sous une sorbonne. Immerger les plantules par retournement dans le tube falcon et faire le vide (toujours sous une sorbonne) 7 minutes, casser le vide, retourner le tube pour immerger les plantules qui remontent à la surface et refaire le vide 7 minutes. A ce stade, les plantules deviennent légèrement translucides. Ajouter 2,5 ml de glycine 2M dans le falcon et refaire le vide 5 minutes pour stopper la réaction de fixation. Laver 3 fois avec de l'eau milli-Q en prenant garde de jeter les déchets de formaldéhyde et les 2 premiers rinçages dans la poubelle adéquate sous la sorbonne. Après les rinçages, sécher la pelote de plantules dans du papier sopalin propre pour éviter les cristaux de glace lors de la congélation des échantillons. Placer les tubes dans l'azote liquide, la manipulation peut être arrêtée à cette étape en stockant les échantillons à -80°C. Il est possible de réaliser des manipulations de



ChIP avec 300 à 400 siliques à différents stades de développement prélevées sur des barquettes de plantes poussées en serre, toutes les étapes restent les mêmes.

#### ***IV.2.7.2. Extraction de la chromatine (jour 2)***

A partir de cette étape, les compositions de tous les tampons se trouvent en annexe 8. Les tampons doivent être réalisés extemporanément et les manipulations réalisées dans la glace. Broyer très finement au mortier les plantules congelées. Les peser et prélever de 0,5 g à 1 g dans un tube falcon de 50 ml propre. Reprendre la poudre dans 30ml du tampon d'extraction 1 glacé, laisser 5 minutes dans la glace et filtrer 2 fois à l'aide d'un morceau de miracloth en transférant dans un tube falcon propre à chaque fois pour retirer tous les fragments de tissus végétaux. Centrifuger 20 min à 4000rpm (à 4°C), retirer doucement le surnageant et resuspendre le culot dans 1ml de tampon d'extraction 2. Transférer dans un tube eppendorf de 1,5 ml. Centrifuger 10 min à 12000g (à 4°C), retirer le surnageant et reprendre le culot dans 300 µl de tampon d'extraction 3. Déposer ces 300 µl sur 300 µl de tampon d'extraction 3 propre, pipeté au préalable dans un tube de 1,5ml et centrifuger 1h à 16000g (à 4°C) pour purifier les noyaux. Pendant ce temps, faire la solution d'extraction de noyaux. Retirer le surnageant et reprendre le culot par aspiration-refoulement (sans faire de bulles) dans 300 µl de solution d'extraction de noyaux, ne laisser aucun fragment vert dans le tube. Prélever 5 µl de chromatine comme témoin d'avant sonication. Soniquer la chromatine dans un bain à soniquer (diagenode, bioruptor) à 4°C à une fréquence de 30sec/30sec pulsation/repos pendant 10 minutes à intensité maximum (Energie H). Prélever 5 µl de chromatine comme témoin de chromatine soniquée. Placer la totalité de la chromatine soniquée à -20°C en attendant d'avoir testé la sonication. Si vous utilisez toujours la même quantité de poudre au départ et la même machine à soniquer dans les mêmes conditions vous pouvez ne réaliser l'étape qui va suivre que lors de la 1<sup>ère</sup> manipulation.

#### ***IV.2.7.3. Vérification de l'efficacité de sonication (jour 3)***

Il est nécessaire de défixer (reverse crosslinker) les témoins avant et après sonication prélevés à l'étape précédente avant de les déposer sur gel d'agarose. Ajuster le volume d'échantillon à 10µl avec de l'eau, ajouter 0,5µl de NaCl 5M, parafilmer le tube et laisser à 65°C toute la

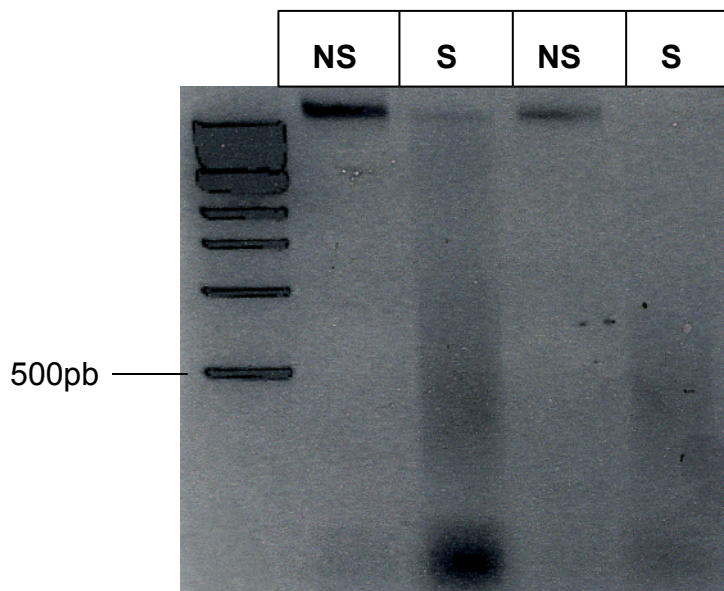




nuit. Le lendemain, procéder à un traitement à la RNase ((2,5µl EDTA ; 5 µl TrisHCl pH6,5 1M ; 0,5 µl RNase 20mg/ml ; 82 µl eau) en laissant 30 min à 42°C. Ajouter ensuite 0,5 µl de protéinase K (20mg/ml) et mettre à incuber 90min à 42°C. Nettoyer toutes les protéines grâce à une étape de phénol/chloroforme (biophenol/chloroform/alcool isoamylique 25 :24 :1, stabilisé ; Biosolve, cat :16972343) puis de lavage au chloroforme-alcool isoamylique. Prélever la phase supérieure et précipiter l'ADN (20µl NaAc pH5,2 ; 1µl glycogène ; 500µl éthanol 96%) en laissant 20 min à -80°C ou 1h à -20°C (plus efficace). Laver à l'éthanol 70%, sécher le culot et reprendre dans 7 µl d'eau. Déposer les 7 µl sur gel d'agarose, TAE 0,5X (Figure 22). Des fragments d'ADN de taille moyenne de 500pb sont souhaités, on s'attend donc à voir un « smear » autour de 500pb sur gel dans le puits correspondant à l'ADN soniqué. Si le « smear » est trop haut ou qu'il reste trop d'ADN génomique non fragmenté, on sonique à nouveau les échantillons de 2 min en 2 minutes jusqu'à obtenir la bonne taille de fragments. Une fois que les fragments obtenus sont de la taille attendue, on peut procéder à la purification de la chromatine.

#### ***IV.2.7.4. Purification de la chromatine (jour 4 et 5)***

Toutes les étapes doivent être réalisées à 4°C. Centrifuger la chromatine à 4°C à vitesse maximum pour culoter les débris et récupérer le surnageant. Prélever 50 µl de cette solution comme témoin avant purification appelé INPUT. CE TEMOIN INPUT EST ESSENTIEL AUX FUTURS CALCULS UTILISES POUR ANALYSER LES RESULTATS DU CHIP, CETTE ETAPE EST DONC CRUCIALE. Ajuster le volume à 300 µl avec le tampon ChIP dilution buffer et répartir 2 fois 150 µl dans des tubes eppendorf qui serviront pour faire les échantillons avec et sans anticorps (Ac). A partir de cette étape, les 2 tubes seront traités de la même façon. Ajouter 850 µl de ChIP dilution buffer et 40 µl de billes de protéineA-sepharose (dynabeads protein A ; Invitrogen, Villebon sur Yvette, France ; cat :1100-02D), préalablement lavées 3 fois avec du ChIP dilution buffer et laisser 1h à 4°C sur une roue. Cette étape permet d'éliminer du mélange de chromatine, les protéines qui se fixeraient non spécifiquement à la protéine A. Retirer les billes et ajouter 5 µl d'anticorps anti-H3K27me3 (Millipore ; cat :07.449) dans le tube + anticorps et laisser à 4°C agiter sur la roue pendant 4h minimum. On peut réaliser toutes ces étapes en une journée en prévoyant de rester 12h au laboratoire. Les étapes de lavage peuvent alors commencer. Placer les tubes sur un portoir



**Figure 22: Dépôt d'un test de sonication sur gel d'agarose lors d'une manipulation de ChIP.** On dépose l'ADN non-soniqué (NS) et soniqué (S) dans le but de vérifier la présence d'un « smear » dans le puits S autour de 500pb signifiant que la sonication est suffisante.

aimanté qui va retenir les billes et retirer le surnageant. Ajouter 1ml du tampon de lavage, retirer le surnageant puis ajouter 1 ml du même tampon en laissant cette fois le tube 5 minutes en agitation sur la roue à 4°C. Procéder de la même façon avec les 4 tampons de lavage (tampon à faible salinité, tampon à forte salinité, tampon LiCl et tampon TE dans cet ordre) différents préparés extemporanément. Eluer ensuite le complexe en ajoutant 2 fois 250 µl de tampon d'éluion avec une incubation 15 minute à 65°C à chaque fois. Combiner les 2 éluats, ajouter 20 µl de NaCl 5M et laisser à 65°C de 6h à toute la nuit. TRES IMPORTANT : FAIRE DE MEME AVEC LES 50 µL D'INPUT PRELEVES AVANT PURIFICATION APRES AVOIR REAJUSTE LE VOLUME A 500 µL AVEC DU TAMPON D'ELUTION. Ajouter 10 µl d'EDTA 0,5M et 20 µl de Tris-HCl 1M pH6,5 et 2 µl de protéinase K à 10 mg/ml, laisser une heure à 45°C. Laver les débris protéiques avec une étape de phénol/chloroforme et de chloroforme-alcool isoamylique, puis précipiter l'ADN avec 1/10 de volume de NaAc (pH 5,2), 1 volume d'éthanol à 96% et 2 µl de glycogène (20mg/ml). Laver le culot à l'éthanol à 70% et reprendre le culot d'ADN dans 50 µl d'eau milliQ.

#### ***IV.2.7.5. Méthode d'analyse du ChIP***

Les séquences des oligonucléotides utilisés sont décrites dans les annexes. Les primers ACT7 (témoin négatif, représente le bruit de fond de la manipulation car le locus amplifié n'est pas H3K27 triméthylé), FWA-5 (témoin positif qui est H3K27 triméthylé et qui n'est pas sensé varier dans les conditions que nous comparons dans notre étude) et les oligonucléotides permettant d'amplifier les cibles de l'étude. Pour chaque couple d'oligonucléotides, la manipulation de PCR quantitative sera réalisée sur les tubes INPUT, +Ac et -Ac. L'utilisation du tube -Ac permet de savoir si le ChIP a fonctionné, c'est-à-dire si nous sommes parvenus à purifier de la chromatine. Pour cela, il suffit de comparer les échantillons +/-Anticorps. Sur les échantillons sans anticorps, les valeurs de cycles de sortie en Q-PCR devront être très faibles par rapport aux valeurs avec anticorps et une amplification nulle peut être observée dans certains cas.

L'INPUT servira de référence pour tous les calculs.

$100 * 1 / (2^{(nbc+Ac-nbcINPUT)}) = \text{Enrichissement exprimé en \% d'INPUT. (Nbc : nombre de cycles)}$



## **IV.2.8 Manipulations de simple hybride**

### ***IV.2.8.1 Transformation intégrative des levures ou protocole haute efficacité:***

#### ***IV.2.8.1.1 Préparation du vecteur :***

Digérer 1 µg du vecteurs pHIS-i dans lequel le fragment de promoteur à tester a été cloné avec 1 µl de *ApaI* dans un volume de 20 µl final, selon les recommandations du fournisseur, 2h à 37°C. Ajouter 1 µl d'*ApaI* et laisser 2 à 3h à 37°C de nouveau. Vérifier le rendement de digestion par dépôt sur gel (le vecteur peut être conservé digéré à -20°C ou transformé directement).

#### ***IV.2.8.1.2 Préculture de 10ml:***

Ensemencer 10 ml de milieu YPDA liquide avec la souche de levure YM4271 fraîchement repiquée sur milieu YPDA (la quantité de cellules équivaut à un volume d'environ 20 µl dans un tube eppendorf de 1,5 ml). Laisser toute la nuit à 28°C sous agitation.

#### ***IV.2.8.1.3 Préculture de 50 ml (H-5) :***

Prélever 50 µl de la culture la veille dans 950 µl d'eau et mesurer la DO à 600nm.

Quantité de préculture à ensemer (ml) :

$(5 \times \text{vol culture final (50)}) / (\text{DO mesurée} \times \text{facteur de dilution (20)} \times 10)$

Ensemencer 50 ml de YPDA avec le volume calculé ci-dessus et laisser pousser entre 4 et 5h

#### ***IV.2.8.1.4 Transformation :***

Mesurer la DO à 600nm qui doit être comprise entre 1 et 2.

Centrifuger à 3000g pendant 5 min dans un tube flacon de 50ml. Retirer le surnageant et reprendre dans 25 ml d'eau stérile. Centrifuger à 3000g pendant 5min et reprendre le culot dans 1ml de LiAc 100 mM. Transférer dans un tube de 1,5ml. Culotter à vitesse maximale 1 min et retirer le LiAc à la pipette. Reprendre en fonction de la taille du



culot (environ 100  $\mu$ l en volume) dans 400-500  $\mu$ l de LiAc 100 mM. Vortexer et répartir dans des tubes selon le nombre de transformations (environ 50  $\mu$ l par transformation), un trop grand nombre de cellules peut être néfaste à l'efficacité de transformation. Préparer le mix (17  $\mu$ l plasmide + 10  $\mu$ l ADN carrier + eau qsp 84  $\mu$ l)

Ajouter aux cellules : 240  $\mu$ l de PEG (3350) 50% (p/v) + 36  $\mu$ l LiAc 1M + 84  $\mu$ l mix  
Vortexer 1min chaque tube jusqu'à complète dissolution du culot et placer à 28°C pendant 30 minutes. Vortexer pour homogénéiser et placer à 42°C pendant 25min. Culotter les cellules 1min à vitesse maximale et reprendre le culot dans 200  $\mu$ l d'eau stérile et tout étaler car l'efficacité est faible pour une transformation intégrative. Le milieu de sélection est le milieu SD-URA.

#### ***IV.2.8.2 Test de l'auto-activation du construit***

Une fois les clones transformés obtenus, il faut tester leur capacité à se développer seuls sur milieu sélectif. Il faut donc repiquer les clones (au moins 3 clones par construit) sur milieu SD-URA-His+3AT (3 aminotriazol) à différentes concentrations (0,5 mM, 10 mM, 15 mM, 25 mM, 50 mM). Prélever chaque clone obtenu et le diluer avec 50  $\mu$ l d'eau dans une plaque pour PCR stérile. On déposera ensuite de 5 à 10  $\mu$ l de cette solution de levures sur les différents milieux sélectifs pour être sûr de repiquer toujours la même quantité de levures sur chaque type de milieu et de pouvoir ainsi comparer les vitesses de croissance.

La croissance des clones doit être inhibée sur 3AT. Au delà d'une croissance sur 15 ou 25 mM de 3AT, les clones sont considérés comme s'autoactivant et ne sont pas utilisables pour une expérience de simple hybride en l'état. Dans le cas contraire, la manipulation peut continuer et on peut procéder à la transformation des vecteurs contenant les protéines à tester clonées dans de vecteur pDEST22.

#### ***IV.2.8.3 Protocole de transformation de levures « quick and easy »***

Suspendre des levures précédemment transformées et repiquées depuis moins d'une semaine dans 1 ml d'eau stérile (un volume de cellules d'environ 50  $\mu$ l). Culotter à vitesse maximale pendant 30sec, reprendre dans 1 ml de LiAc 100 mM et incubé 5 min à 30°C. Placer le volume représentant une transformation (environ 25  $\mu$ l) dans un tube eppendorf de 1,5 ml,



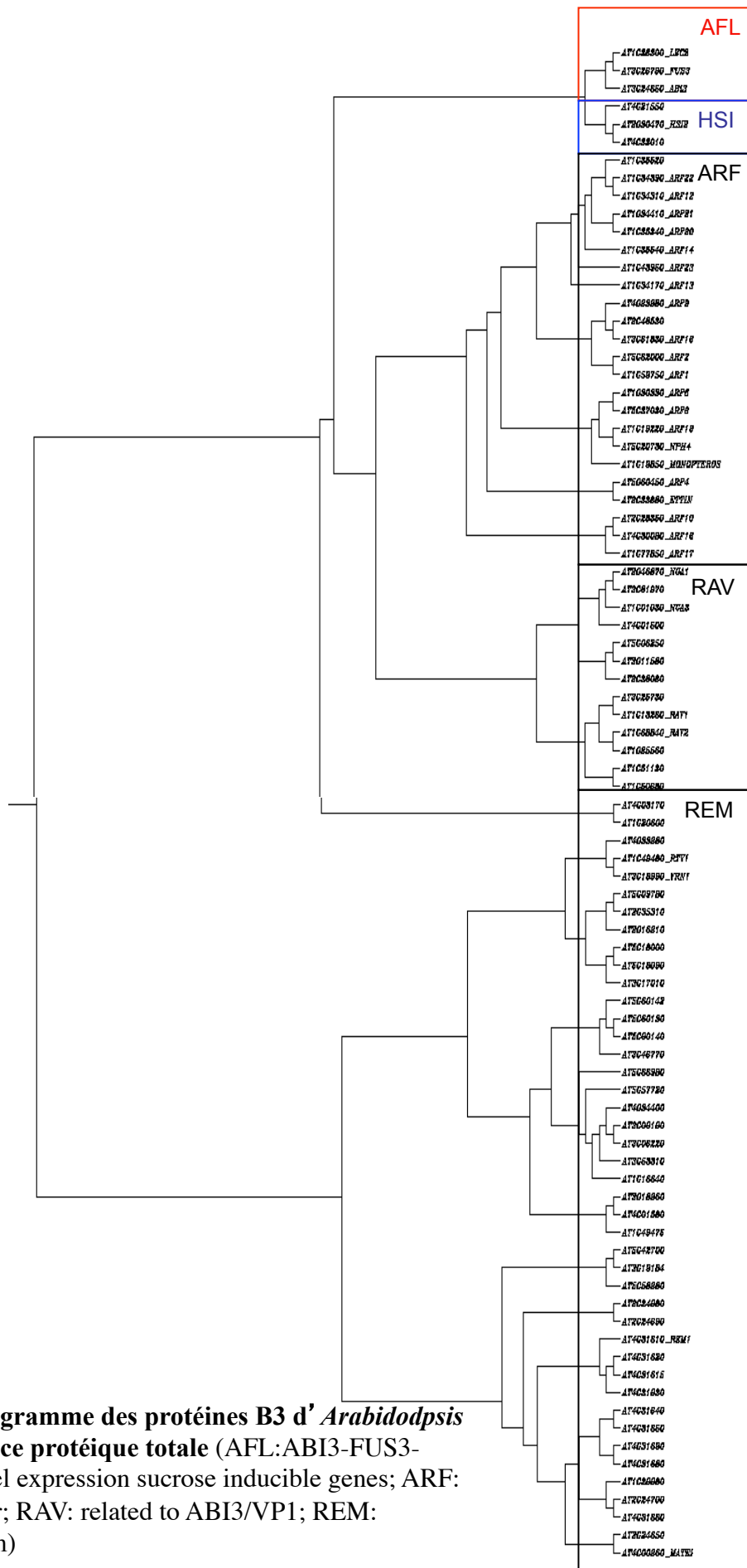


culotter à vitesse maximale 20 secondes et retirer le surnageant. Ajouter aux cellules dans cet ordre précis : 240 µl de PEG (50% p/v) + 36 µl de LiAc 10mM + 10 µl d'ADN carrier + 5 µl de plasmide (environ 500 ng) + 60 µl d'eau. Vortexer les tubes au moins 1 min pour reprendre le culot totalement (si le culot résiste, ne pas hésiter à reprendre avec des aspirations-refoulement avec un pipetman) et incuber à 42°C pendant 20 min. Culotter les cellules à vitesse maximale pendant 20 secondes, reprendre dans 400 µl d'eau et étaler 50 µl et 350 µl sur 2 boîtes différentes pour être sûr d'avoir des clones isolés. Le milieu de sélection avec le vecteur pDEST22 que nous utilisons est le SD-TRP

#### ***IV.2.8.4 Test simple hybride ou d'interaction***

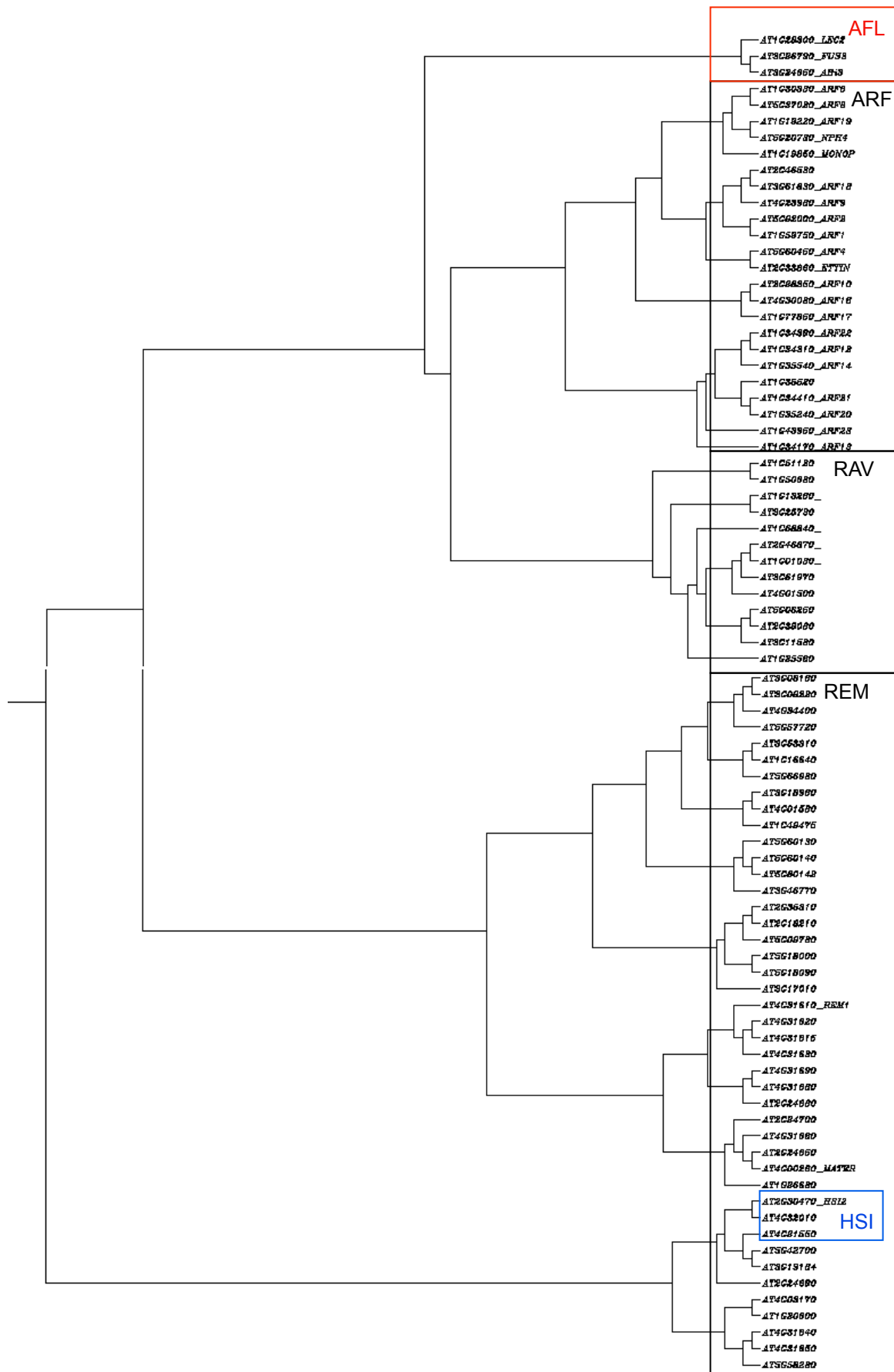
Chaque clone contenant le vecteurs pHIS-i et le vecteur pDEST22 contenant les construits à tester est repiquer de la même façon que pour tester l'auto-activation. Une interaction sera obtenue si les clones poussent sur un milieu contenant des concentrations de 3AT supérieures à celles notées lors du test d'autoactivation.





**ANNEXE 1: Dendrogramme des protéines B3 d' *Arabidopsis* à partir de la séquence protéique totale (AFL:ABI3-FUS3-LECs; HSI: High level expression sucrose inducible genes; ARF: Auxin response factor; RAV: related to ABI3/VP1; REM: reproductive meristem)**





**ANNEXE 2: Dendrogramme représentant les protéines B3 d' *Arabidopsis* en considérant uniquement leur domaine B3 (AFL:ABI3-FUS3-LECs; HSI: High level expression sucrose inducible genes; ARF: Auxin response factor; RAV: related to ABI3/VP1; REM: reproductive meristem)**



ANNEXE 3 : Description des fonctions liées aux HDAC chez *Arabidopsis*, adapté de Hollender et al., 2008

Nomenclature	AGI	TYPE	EXPRESSION	FONCTION CONNUE
HDA1 (HDA19, RPD3A)	At4G38130	RPD3 classe I	Ubiquitaire et plus élevée dans les tissus reproducteurs (1)	Réprime des gènes impliqués dans le développement de la graine (2), la germination (3,4,5), la floraison (6,7,8), la réponse à la lumière (9) et aux pathogènes.
HDA2	At5g26040	RPD3 classe III	Embryons en développement et graines sèches (1)	Aucune
HDA3 (HDT1, HD2A)	At3g44750	HD-tuin	Fleurs, jeunes siliques (Wobus & Weber), ovules, embryon, SAM et premières feuilles (11)	Polarité adaxiale/abaxiale de la feuille (13), dominance nucléaire et méthylation des histones et de l'ADN (12), développement de l'appareil reproducteur (14)
HDA4 (HDT2)	At5g22650	HD-tuin	Ubiquitaire (11)	Polarité adaxiale/abaxiale de la feuille (13), dominance nucléaire et méthylation des histones et de l'ADN (12), développement de l'appareil reproducteur (14)
HDA5	At5g61060	RPD3 classe II	Expression ubiquitaire faible et plus élevée dans le méristème inflorescentiel et dans les jeunes fleurs (1)	Réprime la densité du chevelu racinaire (15)





HDA6 AXE1, RPD3B, Si11	At5g63110	RPD3 classe I		Floraison, senescence, régulation par l'acide jasmonique (17,20,21), repression des caractéristiques embryonnaires (5), dominance nucléolaire ainsi que silencing de transgènes et de transposons, méthylation de l'ADN (16, 18,19)
HDA7	At5g35600	RPD3 classe I	Expression très faible ubiquitaire un peu plus élevée dans la fleur (1)	Inconnue
HDA8	At1g08460	RPD3 non classé	Expression ubiquitaire plus élevée dans la graine en fin de développement et le pollen (1).	Inconnue
HDA9	At3g44680	RPD3 classe I	S'exprime le plus fortement dans le méristème apical, les boutons floraux et les graines sèches	Inconnue
HDA10	At3g44660	RPD3 non classé	Expression presque exclusive dans le méristème inflorescentiel (1).	Inconnue
HDA14	At4g33470	RPD3 non classé	Majoritairement exprimé dans les feuilles et les	Densité racinaire augmentée dans le mutant hda14 (15)



			pédicelles (1).	
HDA15	At3g18520	RPD3 classe II	Expression ubiquitaire faible plus élevée dans le méristème inflorescentiel et les jeunes tissus floraux	Inconnue
HDA17	At3g44490	RPD3 non classé	Expression dans le méristème inflorescentiel (1)	Inconnue
HDA18	At5g61070	RPD3 classe II	Expression ubiquitaire faible plus élevée dans le méristème inflorescentiel et les jeunes tissus floraux et le pollen	Densité du chevelu racinaire augmentée et divisions cellulaires altérées dans l'épiderme du mutant <i>hda18</i> (15)
HDT1	At3g44750	HD-tuin	Expression dans les fleurs, les jeunes siliques (Wobus & Weber), les ovules, les embryons, le méristème apical et les premières feuilles (11)	Polarité adaxiale/abaxiale de la feuille (13), dominance nucléolaire (22) et développement du système reproducteur (12).
HDT2	At5g22650	HD-tuin	Expression ubiquitaire (11)	Polarité adaxiale/abaxiale de la feuille (13), dominance nucléolaire (22) et développement du système reproducteur (12).



HDT3	At5g03740	HD-tuin	Expression ubiquitaire (11)	Développement du système reproducteur et réponse à l'ABA et aux stress abiotiques (12,13, 11, 14)
HDT4	At2g27840	HD-tuin	Expression dans les tiges, fleurs et jeunes siliques (11)	Densité du chevelu racinaire augmente dans le mutant hdt4 (15)
SRT1	At2g27840	HD-tuin	Expression dans les tiges, fleurs et jeunes siliques (11)	Le développement de l'hypocotyle, de la racine et la vascularisation sont affectés en présence de sirtinol (23). Chez le riz, elle joue un rôle majeur dans la réponse aux pathogène (24)
SRT2	At5g55760	Sirtuin	Majoritairement exprimé dans les tissus reproducteurs (11)	Le développement de l'hypocotyle, de la racine et la vascularisation sont affectés en présence de sirtinol (23). Chez le riz, elle joue un rôle majeur dans la réponse aux pathogène (24)

(1) Schmid et al., 2005; (2) Long et al., 2006; (3) Tai et al., 2005; (4) Chen et al., 2010; (5) Tanaka et al., 2008; (6) Wu et al., 2008; (7) Yu et al., 2011; (8) He et al., 2003; (9) Benhamed et al., 2006; (Wobus & Weber) Wu et al., 2000a; (11) Zhou et al., 2004 (12) Lawrence et al., 2004; (13) Ueno et al., 2007; (14) Wu et al., 2000b; (15) Xu et al., 2005; (16) Aufsatz et al., 2002; (17) Devoto et al., 2002; (18) Lippman et al., 2003; (Brownlee) Probst et al., 2004; (20) Wu et al., 2008; (21) Zhou et al., 2005; (22) Pontes et al., 2007; (23) Grozinger et al., 2001; (24) Huang et al., 2007



ANNEXE 4 : liste et séquence des oligonucléotides utilisés.

Amplification des BPC

PBPC1(-600)f

BPC1rev	AGGTAAACCAGAAATGTCCA
BPC2f	ACCATTGAYCGGAACACTAA
BPC2(410)rev	CTTTTGCATTGGTCTTTCTC
BPC3(490)f	GAGATGGTGATAAATGGTGT
BPC3rev	GAAATCAAAACCATCAGAAG
BPC1AttB1	ggggacaagtttgtaaaaaagcaggctcgATGGACGATGATGGATTTC
BPC1AttB2	ggggaccactttgtacaagaaagctgggtTTATCTGATCGTGACAAAC
BPC2AttB1	ggggacaagtttgtaaaaaagcaggctcgATGGATGACGATGGGTTTC
BPC2AttB2	ggggaccactttgtacaagaaagctgggtTCATCTGATTGTGACGAAC
BPC3AttB1	ggggacaagtttgtaaaaaagcaggctcgATGATGGAAGAAGATGG
BPC3AttB2	ggggaccactttgtacaagaaagctgggtTTATCTGATGGTGACGAAC
BPC4AttB1	ggggacaagtttgtaaaaaagcaggctcgATGGAGAATGGTGGTCAG
BPC4AttB2	ggggaccactttgtacaagaaagctgggtCTACTTGATAGTGATGTA
BPC6AttB1	ggggacaagtttgtaaaaaagcaggctcgATGGATGATGGTGGGCATCG
BPC6AttB2	ggggaccactttgtacaagaaagctgggtTCATTTAATCGTAATGTAGC
BPC7AttB1	ggggacaagtttgtaaaaaagcaggctcgATGAATTCATTTCTGCAC
BPC7AttB2	ggggaccactttgtacaagaaagctgggtCTACTTGATGGTCACAAAC
SigLB1	cggaaccaccatcaaacag

Amplification des cDNA divers:

AttB1AGL1

5 ggggacaagtttgtaaaaaagcaggctcgATGGGTCGTGGAAAAATCG

AttB2AGL1

5 ggggaccactttgtacaagaaagctgggtCTAAACAGAGAACCTTTGTC

AttB1AGL1 ggggacaagtttgtaaaaaagcaggctcgATGGGGAGAGGAAGGATT

8 G

AttB2AGL1

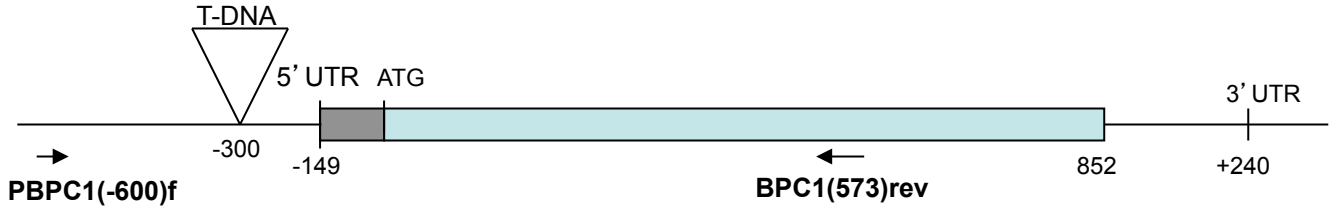
8 ggggaccactttgtacaagaaagctgggtTCAATCAGAAGCCACTTGAC



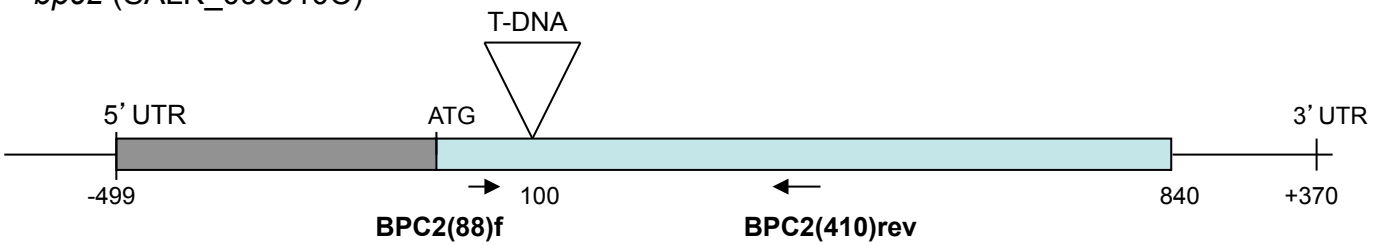


ANNEXE-5: génotypage des mutants *bpc*

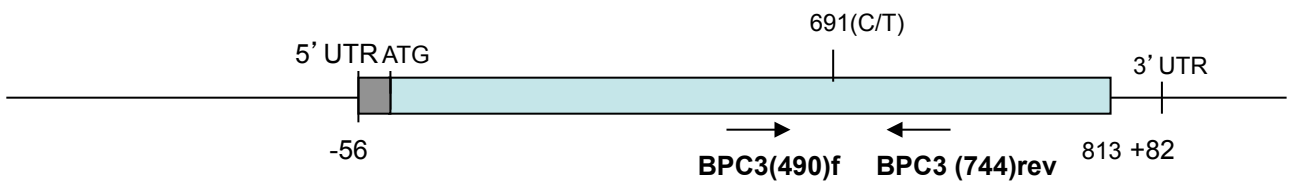
- *bpc1* (SALK\_072966.43.30.x)



- *bpc2* (SALK\_090810C)

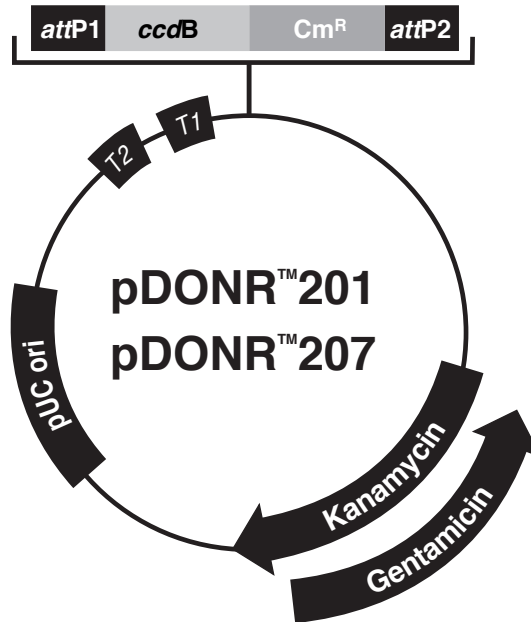


- *bpc3.2* (CS89719)





ANNEXE-6: carte des vecteurs utilisés



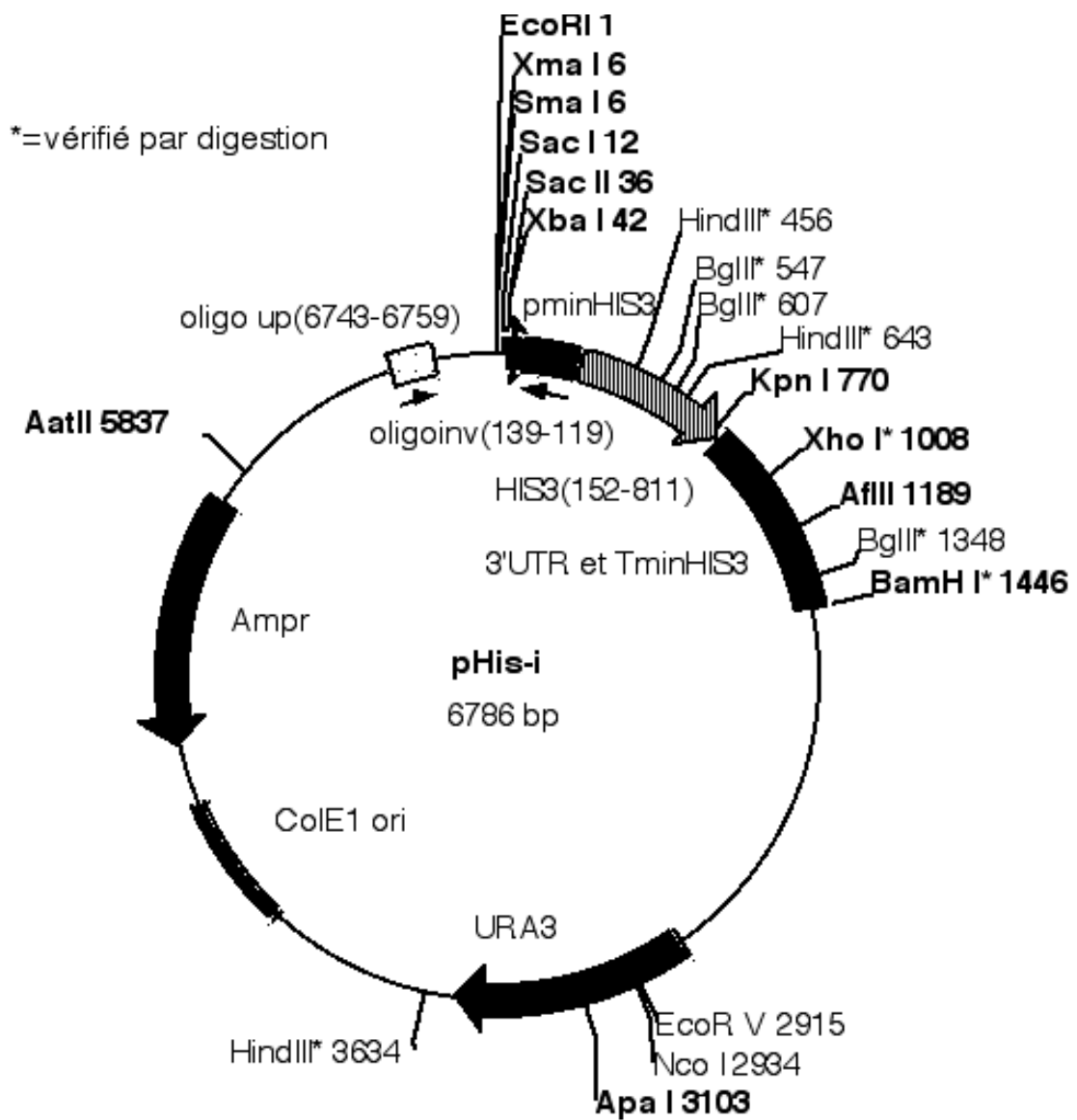
Comments for:

	pDONR™201 4470 nucleotides	pDONR™207 5585 nucleotides
<i>rrnB</i> T2 transcription termination sequence (c):	73-100	73-100
<i>rrnB</i> T1 transcription termination sequence (c):	232-275	232-275
Recommended forward priming site:	300-324	300-324
<i>attP1</i> :	332-563	332-563
<i>ccdB</i> gene (c):	959-1264	959-1264
Chloramphenicol resistance gene (c):	1606-2265	1606-2265
<i>attP2</i> (c):	2513-2744	2513-2744
Recommended reverse priming site:	2769-2792	2769-2792
Kanamycin resistance gene:	2868-3677	---
Gentamicin resistance gene (c):	---	3528-4061
pUC origin:	3794-4467	4909-5582

(c) = complementary strand

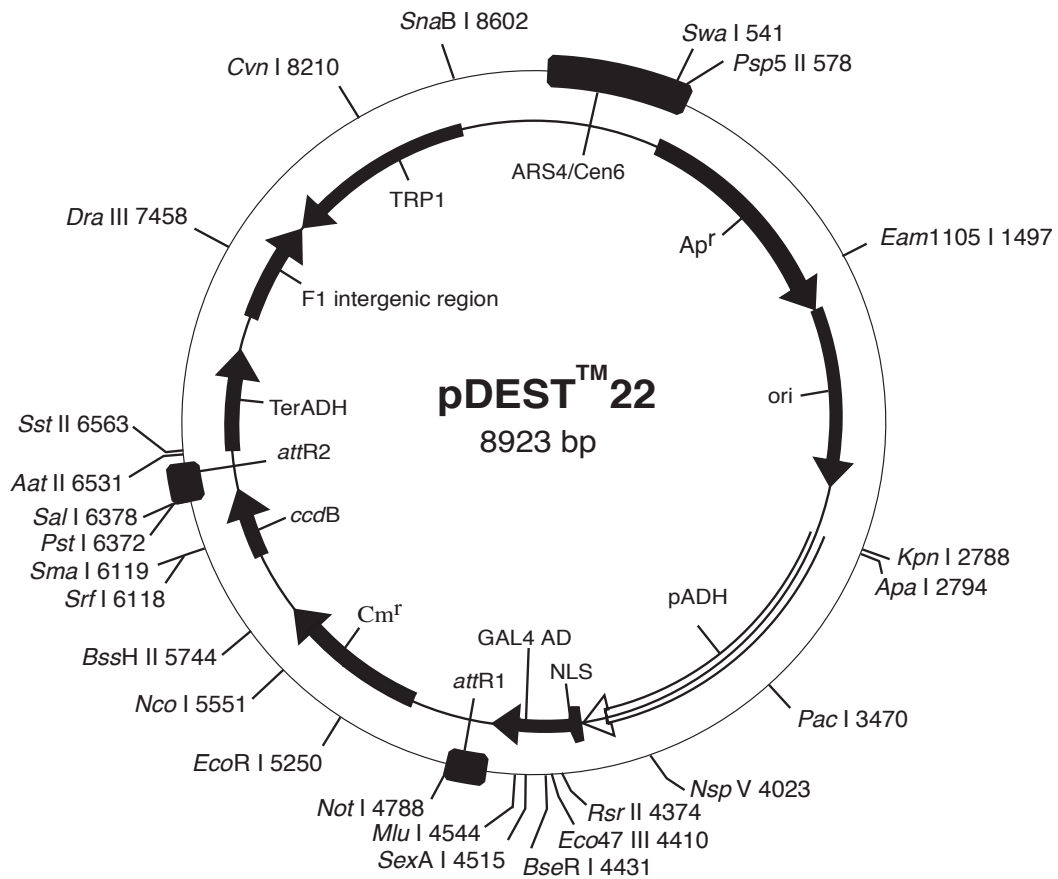


Annexe 6: carte des vecteurs utilisés





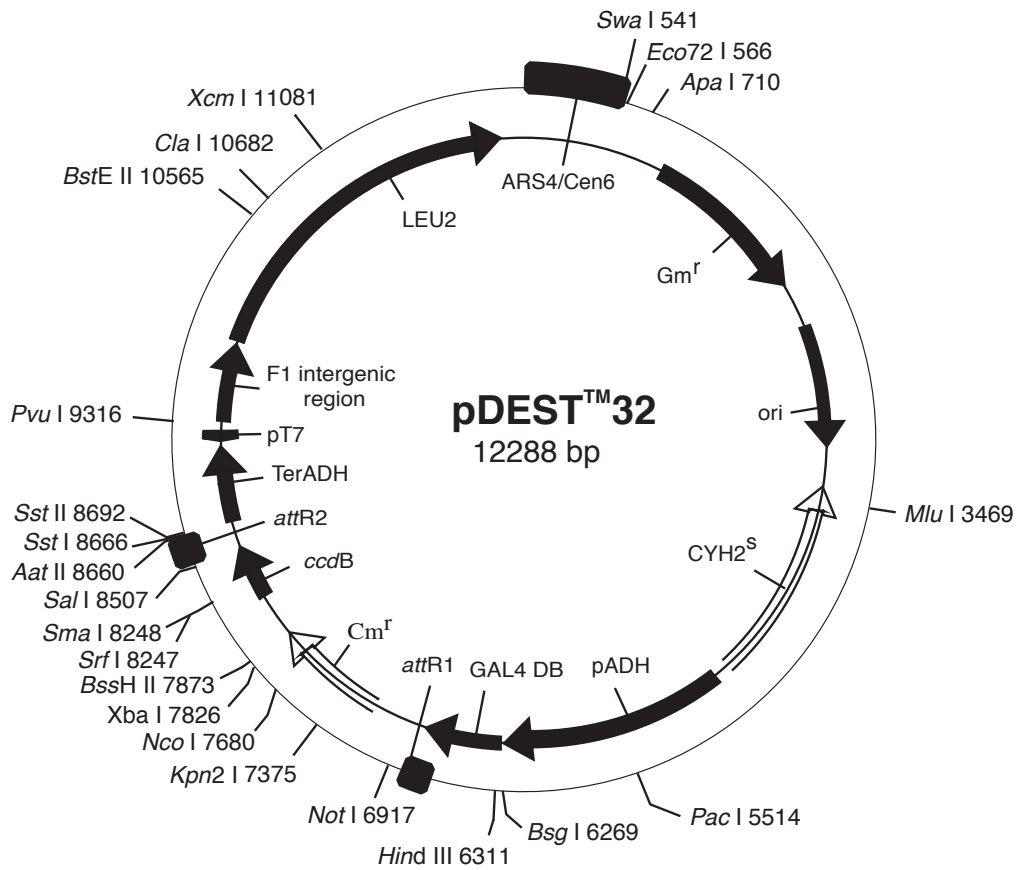
ANNEXE-6: carte des vecteurs utilisés







ANNEXE-6: carte des vecteurs utilisés





## ANNEXE 7 : milieux de culture

### Milieu de culture de bactéries LB:

10g/l tryptone

5g/l d'extrait de levures

10g/l de NaCl

### Milieu de culture d'*Arabidopsis* Gamborg's B5 Basal Medium :

0.5 g/l MES, pH 5.7

0.05 mg/l Kinetin

0.5 mg/l 2,4-dichlorophenoxyacetic Acid (2,4-D)

0.8% (p/v) Agar

2% (p/v) Glucose

### Milieu de culture de levures SD :

26,7g/l milieu SD

700 mg/l aa en mélange

100 mg/l aa individuel selon le milieu désiré

100 mg/l Ade

Ajuster le pH à 5,8 puis ajouter

20 mg/l Agar

### Milieu de culture de levure YPDA :

50g/l de YPD en poudre

100 mg/l Ade



## ANNEXE 8 : Tampons pour les manipulations de ChIP

Tous les tampons pour la manipulation de ChIP sont à réaliser le jour même.

PMSF 0,2 M : diluer 35mg de PMSF dans 1 ml d'isopropanol (long à dissoudre)

Inhibiteurs de protéases : Complete, Mini Protease Inhibitor (Roche)

1 minitablette diluée dans 1ml d'eau pour les tampons d'extraction 2 et 3

### Tampon d'extraction 1 :

	Pour 100 ml	Pour 150 ml ou 4 extractions
0,4M sucrose (MM :342,3)	20 ml de 2M	30 ml de 2M
10 mM Tris-HCl pH8	1 ml de 1M	1,5 ml de 1M
10 mM MgCl <sub>2</sub> (MM : 203,3)	1 ml de 1M	1,5 ml de 1M
A ajouter au dernier moment		
5mM βmercaptoethanol	35 µl de 14,3M	52,5 µl de 14,3M
0,1mM PMSF	50 µl de 0,2M	75 µl de 0,2M
Inhibiteurs de protéases	2 mini-tablettes	3 mini-tablettes
Eau	Qsp 100 ml	Qsp 150 ml

### Tampon d'extraction 2 :

	Pour 1 0ml	Pour 5 ml ou 4 extractions
0,25M sucrose (MM :342,3)	1,25 ml de 2M	625 µl de 2M
10 mM Tris-HCl pH8	100 µl de 1M	50 µl de 1M
10 mM MgCl <sub>2</sub> (MM : 203,3)	100 µl de 1M	50 µl de 1M
1% Triton X-100	100 µl de 100%	50 µl de 100%
A ajouter au dernier moment		
5 mM βmercaptoethanol	3,5 µl de 14,3M	1,75 µl de 14,3M
0,1 mM PMSF	5 µl de 0,2M	2,5 µl de 0,2M
Inhibiteurs de protéases	1 ml	0,5 ml
Eau	Qsp 10 ml	Qsp 5 ml

### Tampon d'extraction 3 :

	Pour 10 ml	Pour 5 ml ou 4 extractions
1,7M sucrose (MM :342,3)	8,5 ml de 2M	4,25 ml de 2M



10 mM Tris-HCl pH8	100 µl de 1M	50 µl de 1M
2 mM MgCl <sub>2</sub> (MM : 203,3)	20 µl de 1M	10 µl de 1M
0,15% Triton X-100	15 µl de 100%	7,5 µl de 100%
A ajouter au dernier moment		
5 mM βmercaptoethanol	3,5 µl de 14,3M	1,75 µl de 14,3M
0,1 mM PMSF	5 µl de 0,2M	2,5 µl de 0,2M
Inhibiteurs de protéases	1 ml	0,5 ml
Eau	Qsp 10 ml	Qsp 5 ml

Solution d'extraction des noyaux :

	Pour 5 ml	Pour 2,5 ml
50mM Tris-HCl pH8	250 µl de 1M	125 µl de 1M
10mM EDTA	100 µl de 0,5M	50 µl de 0,5M
1% SDS	250 µl de 20%	125 µl de 20%
A ajouter au dernier moment		
Inhibiteurs de protéases	250 µl	125 µl
Eau	Qsp 5 ml	Qsp 2,5 ml

Tampon de ChIP dilution :

	Pour 10 ml	Pour 30 ml
1 ,1% Triton X-100	110 µl de 100%	330 µl de 100%
1,2mM EDTA	24 µl de 0,5M	72 µl de 0,5M
16,7mM Tris-HCl pH 8	167 µl de 1M	500 µl de 1M
167mM NaCl	334 µl de 5M	1ml de 5M
Eau	Qsp 10 ml	Qsp 30 ml

Tampons de lavage :

1-Tampon à faible salinité :

150 mM NaCl ; 0,1% SDS ; 1% TritonX-100 ; 2 mM EDTA ; 20 mM Tris-HCl (pH 8,1)

2-Tampon à forte salinité :

500 mM NaCl ; 0,1% SDS ; 1% TritonX-100 ; 2 mM EDTA ; 20 mM Tris-HCl (pH 8,1)





3-Tampon LiCl :

0,25M LiCl, 1%NP40 ; 1% sodium deoxycholate ; 1 mM EDTA ; 10 mM Tris-HCl (pH8,1)

4-Tampon TE :

10 mM Tris-HCl pH 8,1 ; 1 mM EDTA

## V. BIBLIOGRAPHIE

- Aasland R, Stewart AF (1995) The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. *Nucleic Acids Res* **23**: 3168-3173
- Adamczyk BJ, Lehti-Shiu MD, Fernandez DE (2007) The MADS domain factors AGL15 and AGL18 act redundantly as repressors of the floral transition in Arabidopsis. *Plant J* **50**: 1007-1019
- Ahringer J (2000) NuRD and SIN3 histone deacetylase complexes in development. *Trends Genet* **16**: 351-356
- Aichinger E, Villar CB, Di Mambro R, Sabatini S, Kohler C (2011) The CHD3 chromatin remodeler PICKLE and polycomb group proteins antagonistically regulate meristem activity in the Arabidopsis root. *Plant Cell* **23**: 1047-1060
- Aichinger E, Villar CB, Farrona S, Reyes JC, Hennig L, Kohler C (2009) CHD3 proteins and polycomb group proteins antagonistically determine cell identity in Arabidopsis. *PLoS Genet* **5**: e1000605
- Albani D, Robert LS (1995) Cloning and characterization of a Brassica napus gene encoding a homologue of the B subunit of a heteromeric CCAAT-binding factor. *Gene* **167**: 209-213
- Alonso R, Onate-Sanchez L, Weltmeier F, Ehlert A, Diaz I, Dietrich K, Vicente-Carbajosa J, Droge-Laser W (2009) A pivotal role of the basic leucine zipper transcription factor bZIP53 in the regulation of Arabidopsis seed maturation gene expression based on heterodimerization and protein complex formation. *Plant Cell* **21**: 1747-1761
- Angelovici R, Galili G, Fernie AR, Fait A (2010) Seed desiccation: a bridge between maturation and germination. *Trends Plant Sci* **15**: 211-218
- Arteaga-Vazquez MA, Chandler VL (2010) Paramutation in maize: RNA mediated trans-generational gene silencing. *Current opinion in genetics & development* **20**: 156-163
- Aufsatz W, Mette MF, van der Winden J, Matzke M, Matzke AJ (2002) HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. *Embo J* **21**: 6832-6841



Ausin I, Alonso-Blanco C, Jarillo JA, Ruiz-Garcia L, Martinez-Zapater JM (2004) Regulation of flowering time by FVE, a retinoblastoma-associated protein. *Nat Genet* **36**: 162-166

Ayadi M, Delaporte V, Li YF, Zhou DX (2004) Analysis of GT-3a identifies a distinct subgroup of trihelix DNA-binding transcription factors in Arabidopsis. *FEBS Lett* **562**: 147-154

Bao X, Shorrosh BS, Ohlrogge JB (1997) Isolation and characterization of an Arabidopsis biotin carboxylase gene and its promoter. *Plant Mol Biol* **35**: 539-550

Baud S, Boutin J-P, Miquel M, Lepiniec L, Rochat C (2002) An integrated overview of seed development in *Arabidopsis thaliana* ecotype Ws. *Plant Physiology and Biochemistry* **40**: 151-160

Baud S, Dichow NR, Kelemen Z, d'Andrea S, To A, Berger N, Canonge M, Kronenberger J, Viterbo D, Dubreucq B, Lepiniec L, Chardot T, Miquel M (2009) Regulation of HSD1 in seeds of *Arabidopsis thaliana*. *Plant Cell Physiol* **50**: 1463-1478

Baud S, Mendoza MS, To A, Harscoet E, Lepiniec L, Dubreucq B (2007) WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis. *Plant J* **50**: 825-838

Baumbusch LO, Hughes DW, Galau GA, Jakobsen KS (2004) LEC1, FUS3, ABI3 and Em expression reveals no correlation with dormancy in Arabidopsis. *J Exp Bot* **55**: 77-87

Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumann K, Assalkhou R, Schulz I, Reuter G, Aalen RB (2001) The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res* **29**: 4319-4333

Baumlein H, Misera S, Luerssen H, Kollé K, Horstmann C, Wobus U, Müller AJ (1994) The FUS3 gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. *Plant J* **6**: 379-387

Belozero V, Majumder P, Shen P, Cai HN (2003) A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of *Drosophila*. *Embo J* **22**: 3113-3121

Benhamed M, Bertrand C, Servet C, Zhou DX (2006) Arabidopsis GCN5, HD1, and TAF1/HAF2 interact to regulate histone acetylation required for light-responsive gene expression. *Plant Cell* **18**: 2893-2903

Benhamed M, Martin-Magniette ML, Taconnat L, Bitton F, Servet C, De Clercq R, De Meyer B, Buyschaert C, Rombauts S, Villarreal R, Aubourg S, Beynon J, Bhalerao RP, Coupland G, Gruissem W, Menke FL, Weisshaar B, Renou JP, Zhou DX, Hilson P (2008) Genome-scale Arabidopsis promoter array identifies targets of the histone acetyltransferase GCN5. *Plant J* **56**: 493-504



Benyajati C, Mueller L, Xu N, Pappano M, Gao J, Mosammaparast M, Conklin D, Granok H, Craig C, Elgin S (1997) Multiple isoforms of GAGA factor, a critical component of chromatin structure. *Nucleic Acids Res* **25**: 3345-3353

Berger N, Dubreucq B, Roudier F, Dubos C, Lepiniec L (2011) Transcriptional Regulation of Arabidopsis LEAFY COTYLEDON2 Involves RLE, a cis-Element That Regulates Trimethylation of Histone H3 at Lysine-27. *The Plant cell* **23**: 4065-4078

Bhat KM, Farkas G, Karch F, Gyurkovics H, Gausz J, Schedl P (1996) The GAGA factor is required in the early Drosophila embryo not only for transcriptional regulation but also for nuclear division. *Development* **122**: 1113-1124

Bolle C, Herrmann RG, Oelmuller R (1996) Different sequences for 5'-untranslated leaders of nuclear genes for plastid proteins affect the expression of the beta-glucuronidase gene. *Plant Mol Biol* **32**: 861-868

Bonet C, Fernandez I, Aran X, Bernues J, Giralt E, Azorin F (2005) The GAGA protein of Drosophila is phosphorylated by CK2. *J Mol Biol* **351**: 562-572

Borlaug N (2007) Feeding a hungry world. *Science* **318**: 359

Bouazoune K, Brehm A (2006) ATP-dependent chromatin remodeling complexes in Drosophila. *Chromosome Res* **14**: 433-449

Bouyer D, Roudier F, Heese M, Andersen ED, Gey D, Nowack MK, Goodrich J, Renou JP, Grini PE, Colot V, Schnittger A (2011) Polycomb Repressive Complex 2 Controls the Embryo-to-Seedling Phase Transition. *PLoS Genet* **7**: e1002014

Brady SM, Sarkar SF, Bonetta D, McCourt P (2003) The ABSCISIC ACID INSENSITIVE 3 (ABI3) gene is modulated by farnesylation and is involved in auxin signaling and lateral root development in Arabidopsis. *Plant J* **34**: 67-75

Bratzel F, Lopez-Torrejon G, Koch M, Del Pozo JC, Calonje M (2010) Keeping cell identity in Arabidopsis requires PRC1 RING-finger homologs that catalyze H2A monoubiquitination. *Curr Biol* **20**: 1853-1859

Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL, Goldberg RB, Harada JJ (2006) Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc Natl Acad Sci USA* **103**: 3468-3473

Brocard-Gifford IM, Lynch TJ, Finkelstein RR (2003) Regulatory Networks in Seeds Integrating Developmental, Abscisic Acid, Sugar, and Light Signaling. *Plant Physiol* **131**: 78-92

Brown JL, Fritsch C, Mueller J, Kassis JA (2003) The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. *Development* **130**: 285-294



Brown JL, Mucci D, Whiteley M, Dirksen ML, Kassis JA (1998) The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol Cell* **1**: 1057-1064

Brownlee CU-hwscsaBV-JF-adfeff (2000) Plant development: Keeping your distance. *Current Biology* **10**: R555-R557

Burke TW, Kadonaga JT (1997) The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. *Genes Dev* **11**: 3020-3031

Busk PK, Pujal J, Jessop A, Lumbreras V, Pages M (1999) Constitutive protein-DNA interactions on the abscisic acid-responsive element before and after developmental activation of the rab28 gene. *Plant Mol Biol* **41**: 529-536

Busturia A, Lloyd A, Bejarano F, Zavortink M, Xin H, Sakonju S (2001) The MCP silencer of the Drosophila Abd-B gene requires both Pleiohomeotic and GAGA factor for the maintenance of repression. *Development* **128**: 2163-2173

Butenko Y, Ohad N (2011) Polycomb-group mediated epigenetic mechanisms through plant evolution. *Biochim Biophys Acta*

Cai H, Levine M (1995) Modulation of enhancer-promoter interactions by insulators in the Drosophila embryo. *Nature* **376**: 533-536

Carson CB, Hattori T, Rosenkrans L, Vasil V, Vasil IK, Peterson PA, McCarty DR (1997) The quiescent/colorless alleles of viviparous1 show that the conserved B3 domain of VP1 is not essential for ABA-regulated gene expression in the seed. *Plant J* **12**: 1231-1240

Cartwright IL, Elgin SC (1986) Nucleosomal instability and induction of new upstream protein-DNA associations accompany activation of four small heat shock protein genes in Drosophila melanogaster. *Mol Cell Biol* **6**: 779-791

Casson SA, Lindsey K (2006) The turnip mutant of Arabidopsis reveals that LEAFY COTYLEDON1 expression mediates the effects of auxin and sugars to promote embryonic cell identity. *Plant Physiol* **142**: 526-541

Causier B, Cook H, Davies B (2003) An Antirrhinum ternary complex factor specifically interacts with C-function and SEPALLATA-like MADS-box factors. *Plant Mol Biol* **52**: 1051-1062

Cernac A, Benning C (2004) *WRINKLED1* encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. *Plant J* **40**: 575-585

Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon Y-H, Sung ZR, Goodrich J (2004a) Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**: 5263-5276





Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, Goodrich J (2004b) Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**: 5263-5276

Chaudhury AM, Ming L, Miller C, Craig S, Dennis ES, Peacock WJ (1997) Fertilization-independent seed development in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **94**: 4223-4228

Che N, Yang Y, Li Y, Wang L, Huang P, Gao Y, An C (2009) Efficient LEC2 activation of OLEOSIN expression requires two neighboring RY elements on its promoter. *Sci China C Life Sci* **52**: 854-863

Chen D, Molitor A, Liu C, Shen WH (2010a) The Arabidopsis PRC1-like ring-finger proteins are necessary for repression of embryonic traits during vegetative growth. *Cell Res* **20**: 1332-1344

Chen LT, Luo M, Wang YY, Wu K (2010b) Involvement of Arabidopsis histone deacetylase HDA6 in ABA and salt stress response. *J Exp Bot* **61**: 3345-3353

Chen WJ, Chang SH, Hudson ME, Kwan WK, Li J, Estes B, Knoll D, Shi L, Zhu T (2005) Contribution of transcriptional regulation to natural variations in Arabidopsis. *Genome Biol* **6**: R32

Chen X, Hu Y, Zhou DX (2011) Epigenetic gene regulation by Plant Jumonji group of histone demethylase. *Biochim Biophys Acta*

Chopra VS, Srinivasan A, Kumar RP, Mishra K, Basquin D, Docquier M, Seum C, Pauli D, Mishra RK (2008) Transcriptional activation by GAGA factor is through its direct interaction with dmTAF3. *Dev Biol* **317**: 660-670

Chua YL, Brown AP, Gray JC (2001) Targeted histone acetylation and altered nuclease accessibility over short regions of the pea plastocyanin gene. *Plant Cell* **13**: 599-612

Chua YL, Mott E, Brown AP, MacLean D, Gray JC (2004) Microarray analysis of chromatin-immunoprecipitated DNA identifies specific regions of tobacco genes associated with acetylated histones. *The Plant journal : for cell and molecular biology* **37**: 789-800

Chua YL, Watson LA, Gray JC (2003) The transcriptional enhancer of the pea plastocyanin gene associates with the nuclear matrix and regulates gene expression through histone acetylation. *Plant Cell* **15**: 1468-1479

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**: 735-743

Combiér JP, Frugier F, de Billy F, Boualem A, El-Yahyaoui F, Moreau S, Vernie T, Ott T, Gamas P, Crespi M, Niebel A (2006) MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. *Genes & development* **20**: 3084-3088



- Conaway RC, Conaway JW (2009) The INO80 chromatin remodeling complex in transcription, replication and repair. *Trends Biochem Sci* **34**: 71-77
- Crowe AJ, Abenes M, Plant A, Moloney MM (2000) The seed-specific transactivator, ABI3, induces oleosin gene expression. *Plant Sci* **151**: 171-181.
- Curaba J, Moritz T, Blervaque R, Parcy F, Raz V, Herzog M, Vachon G (2004) AtGA3ox2, a Key Gene Responsible for Bioactive Gibberellin Biosynthesis, Is Regulated during Embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis. *Plant Physiol*: pp.104.047266
- Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**: 185-196
- De Amicis F, Patti T, Marchetti S (2007) Improvement of the pBI121 plant expression vector by leader replacement with a sequence combining a poly(CAA) and a CT motif. *Transgenic Res* **16**: 731-738
- de Boer GJ, Testerink C, Pielage G, Nijkamp HJ, Stuitje AR (1999) Sequences surrounding the transcription initiation site of the Arabidopsis enoyl-acyl carrier protein reductase gene control seed expression in transgenic tobacco. *Plant Mol Biol* **39**: 1197-1207
- de Folter S, Immink RGH, Kieffer M, Parenicova L, Henz SR, Weigel D, Busscher M, Kooiker M, Colombo L, Kater MM, Davies B, Angenent GC (2005) Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. *Plant Cell*: tpc.105.031831
- de la Paz Sanchez M, Gutierrez C (2009) Arabidopsis ORC1 is a PHD-containing H3K4me3 effector that regulates transcription. *Proc Natl Acad Sci U S A* **106**: 2065-2070
- De Lucia F, Crevillen P, Jones AM, Greb T, Dean C (2008) A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proc Natl Acad Sci U S A* **105**: 16831-16836
- Devoto A, Nieto-Rostro M, Xie DX, Ellis C, Harmston R, Patrick E, Davis J, Sherratt L, Coleman M, Turner JG (2002) COI1 links jasmonate signalling and fertility to the SCF ubiquitin-ligase complex in Arabidopsis. *Plant J* **32**: 457-466
- Doebley JF, Gaut BS, Smith BD (2006) The Molecular Genetics of Crop Domestication. *Cell* **127**: 1309-1321
- Dong G, Ma DP, Li J (2008) The histone methyltransferase SDG8 regulates shoot branching in Arabidopsis. *Biochem Biophys Res Commun* **373**: 659-664
- Dubos C, Le Gourrierec J, Baudry A, Huet G, Lanet E, Debeaujon I, Routaboul JM, Alboresi A, Weisshaar B, Lepiniec L (2008) MYBL2 is a new regulator of flavonoid biosynthesis in Arabidopsis thaliana. *Plant J* **55**: 940-953



- Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. *Trends in plant science* **15**: 573-581
- Earley KW, Shook MS, Brower-Toland B, Hicks L, Pikaard CS (2007) In vitro specificities of Arabidopsis co-activator histone acetyltransferases: implications for histone hyperacetylation in gene activation. *Plant J* **52**: 615-626
- Ebbs ML, Bartee L, Bender J (2005) H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. *Mol Cell Biol* **25**: 10507-10515
- Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic acids research* **19**: 1349
- Enjuto M, Lumberras V, Marin C, Boronat A (1995) Expression of the Arabidopsis HMG2 gene, encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase, is restricted to meristematic and floral tissues. *Plant Cell* **7**: 517-527
- Erhard KF, Jr., Hollick JB (2011) Paramutation: a process for acquiring trans-generational regulatory states. *Current opinion in plant biology* **14**: 210-216
- Espinas ML, Canudas S, Fanti L, Pimpinelli S, Casanova J, Azorin F (2000) The GAGA factor of Drosophila interacts with SAP18, a Sin3-associated polypeptide. *EMBO Rep* **1**: 253-259
- Espinas ML, Jimenez-Garcia E, Vaquero A, Canudas S, Bernues J, Azorin F (1999) The N-terminal POZ domain of GAGA mediates the formation of oligomers that bind DNA with high affinity and specificity. *J Biol Chem* **274**: 16461-16469
- Exner V, Aichinger E, Shu H, Wildhaber T, Alfarano P, Caflisch A, Grussem W, Kohler C, Hennig L (2009) The chromodomain of LIKE HETEROCHROMATIN PROTEIN 1 is essential for H3K27me3 binding and function during Arabidopsis development. *PLoS One* **4**: e5335
- Ezcurra I, Wycliffe P, Nehlin L, Ellerstrom M, 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* **24**: 57-66
- Ezcurra I, Wycliffe P, Nehlin L, Ellerstrom M, 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* **24**: 57-66.
- Farkas G, Gausz J, Galloni M, Reuter G, Gyurkovics H, Karch F (1994) The Trithorax-like gene encodes the Drosophila GAGA factor. *Nature* **371**: 806-808
- Faucheux M, Roignant JY, Netter S, Charollais J, Antoniewski C, Theodore L (2003) batman Interacts with polycomb and trithorax group genes and encodes a BTB/POZ protein that is included in a complex containing GAGA factor. *Mol Cell Biol* **23**: 1181-1195



Filion GJ, van Bommel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, van Steensel B (2010) Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **143**: 212-224

Finkelstein R, Reeves W, Ariizumi T, Steber C (2008) Molecular aspects of seed dormancy. *Annu Rev Plant Biol* **59**: 387-415

Finkelstein RR, Lynch TJ (2000) The Arabidopsis abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. *Plant Cell* **12**: 599-609.

Finkelstein RR, Somerville CR (1990) Three Classes of Abscisic Acid (ABA)-Insensitive Mutations of Arabidopsis Define Genes that Control Overlapping Subsets of ABA Responses. *Plant Physiol* **94**: 1172-1179

Flaus A, Martin DM, Barton GJ, Owen-Hughes T (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* **34**: 2887-2905

Franco-Zorrilla JM, Cubas P, Jarillo JA, Fernandez-Calvin B, Salinas J, Martinez-Zapater JM (2002) AtREM1, a member of a new family of B3 domain-containing genes, is preferentially expressed in reproductive meristems. *Plant Physiol* **128**: 418-427

Fritsch C, Beuchle D, Muller J (2003) Molecular and genetic analysis of the Polycomb group gene *Sex combs extra/Ring* in *Drosophila*. *Mech Dev* **120**: 949-954

Fritsch C, Brown JL, Kassis JA, Muller J (1999) The DNA-binding polycomb group protein pleiohomeotic mediates silencing of a *Drosophila* homeotic gene. *Development* **126**: 3905-3913

Fujimori S, Washio T, Higo K, Ohtomo Y, Murakami K, Matsubara K, Kawai J, Carninci P, Hayashizaki Y, Kikuchi S, Tomita M (2003) A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription. *FEBS letters* **554**: 17-22

Fujita H, Takemura M, Tani E, Nemoto K, Yokota A, Kohchi T (2003) An Arabidopsis MADS-Box Protein, AGL24, is Specifically Bound to and Phosphorylated by Meristematic Receptor-Like Kinase (MRLK). *Plant Cell Physiol* **44**: 735-742

Gamboa A, Paez-Valencia J, Acevedo GF, Vazquez-Moreno L, Alvarez-Buylla RE (2001) Floral transcription factor AGAMOUS interacts in vitro with a leucine-rich repeat and an acid phosphatase protein complex. *Biochem Biophys Res Commun* **288**: 1018-1026.

Gao MJ, Lydiate DJ, Li X, Lui H, Gjetvaj B, Hegedus DD, Rozwadowski K (2009) Repression of seed maturation genes by a trihelix transcriptional repressor in Arabidopsis seedlings. *Plant Cell* **21**: 54-71

Gazzarrini S, Tsuchiya Y, Lumba S, Okamoto M, McCourt P (2004) The Transcription Factor FUSCA3 Controls Developmental Timing in Arabidopsis through the Hormones Gibberellin and Abscisic Acid. *Dev Cell* **7**: 373-385





Geiss-Friedlander R, Melchior F (2007) Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol* **8**: 947-956

Georgel PT (2005) Chromatin potentiation of the hsp70 promoter is linked to GAGA-factor recruitment. *Biochem Cell Biol* **83**: 555-565

Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, Li L, Gilmour DS, Adelman K (2008) NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev* **22**: 1921-1933

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the Arabidopsis ABI3 gene by positional cloning. *Plant Cell* **4**: 1251-1261.

Gonzalez D, Bowen AJ, Carroll TS, Conlan RS (2007) The transcription corepressor LEUNIG interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (HEN3) to repress transcription. *Mol Cell Biol* **27**: 5306-5315

Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz EM, Coupland G (1997) A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. *Nature* **386**: 44-51

Gorman MJ, Kaufman TC (1995) Genetic analysis of embryonic cis-acting regulatory elements of the Drosophila homeotic gene *sex combs reduced*. *Genetics* **140**: 557-572

Grafi G, Zemach A, Pitto L (2007) Methyl-CpG-binding domain (MBD) proteins in plants. *Biochim Biophys Acta* **1769**: 287-294

Grossniklaus U, Vielle-Calzada JP, Hoepfner MA, Gagliano WB (1998) Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. *Science* **280**: 446-450

Grozinger CM, Chao ED, Blackwell HE, Moazed D, Schreiber SL (2001) Identification of a class of small molecule inhibitors of the sirtuin family of NAD-dependent deacetylases by phenotypic screening. *J Biol Chem* **276**: 38837-38843

Gusmaroli G, Tonelli C, Mantovani R (2001) Regulation of the CCAAT-Binding NF-Y subunits in Arabidopsis thaliana. *Gene* **264**: 173-185

Gutzat R, Borghi L, Futterer J, Bischof S, Laizet Y, Hennig L, Feil R, Lunn J, Grissem W (2011) RETINOBLASTOMA-RELATED PROTEIN controls the transition to autotrophic plant development. *Development* **138**: 2977-2986

Hagstrom K, Muller M, Schedl P (1997) A Polycomb and GAGA dependent silencer adjoins the Fab-7 boundary in the Drosophila bithorax complex. *Genetics* **146**: 1365-1380

Hall JA, Georgel PT (2007) CHD proteins: a diverse family with strong ties. *Biochem Cell Biol* **85**: 463-476

Harding EW, Tang W, Nichols KW, Fernandez DE, Perry SE (2003) Expression and Maintenance of Embryonic Potential Is Enhanced through Constitutive Expression of AGAMOUS-Like 15. *Plant Physiol* **133**: 653-663



Harikrishnan KN, Chow MZ, Baker EK, Pal S, Bassal S, Brasacchio D, Wang L, Craig JM, Jones PL, Sif S, El-Osta A (2005) Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nat Genet* **37**: 254-264

Hartmann U, Sagasser M, Mehrtens F, Stracke R, Weisshaar B (2005) Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. *Plant Mol Biol* **57**: 155-171

Haslekas C, Grini PE, Nordgard SH, Thorstensen T, Viken MK, Nygaard V, Aalen RB (2003) ABI3 mediates expression of the peroxiredoxin antioxidant AtPER1 gene and induction by oxidative stress. *Plant Mol Biol* **53**: 313-326

Haslekas C, Stacy RA, Nygaard V, Culianez-Macia FA, Aalen RB (1998) The expression of a peroxiredoxin antioxidant gene, AtPer1, in *Arabidopsis thaliana* is seed-specific and related to dormancy. *Plant Mol Biol* **36**: 833-845.

Hay RT (2005) SUMO: a history of modification. *Molecular cell* **18**: 1-12

He Y, Michaels SD, Amasino RM (2003) Regulation of Flowering Time by Histone Acetylation in *Arabidopsis*. *Science*: 1091109

Hebbes TR, Thorne AW, Crane-Robinson C (1988) A direct link between core histone acetylation and transcriptionally active chromatin. *Embo J* **7**: 1395-1402

Hill K, Wang H, Perry SE (2008) A transcriptional repression motif in the MADS factor AGL15 is involved in recruitment of histone deacetylase complex components. *Plant J* **53**: 172-185

Hiratsu K, Matsui K, Koyama T, Ohme-Takagi M (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. *Plant J* **34**: 733-739

Hiratsu K, Ohta M, Matsui K, Ohme-Takagi M (2002) The SUPERMAN protein is an active repressor whose carboxy-terminal repression domain is required for the development of normal flowers. *Febs Letters* **514**: 351-354

Hodgson JW, Argiropoulos B, Brock HW (2001) Site-specific recognition of a 70-base-pair element containing d(GA)(n) repeats mediates bithoraxoid polycomb group response element-dependent silencing. *Mol Cell Biol* **21**: 4528-4543

Hollender C, Liu Z (2008) Histone deacetylase genes in *Arabidopsis* development. *J Integr Plant Biol* **50**: 875-885

Horard B, Tatout C, Poux S, Pirrotta V (2000) Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. *Mol Cell Biol* **20**: 3187-3197

Hublitz P, Albert M, Peters AH (2009) Mechanisms of transcriptional repression by histone lysine methylation. *Int J Dev Biol* **53**: 335-354



- Hur MW, Laney JD, Jeon SH, Ali J, Biggin MD (2002) Zeste maintains repression of Ubx transgenes: support for a new model of Polycomb repression. *Development* **129**: 1339-1343
- Jackson JP, Lindroth AM, Cao X, Jacobsen SE (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**: 556-560
- Jacob Y, Feng S, LeBlanc CA, Bernatavichute YV, Stroud H, Cokus S, Johnson LM, Pellegrini M, Jacobsen SE, Michaels SD (2009) ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nature structural & molecular biology* **16**: 763-768
- Jenik PD, Barton MK (2005) Surge and destroy: the role of auxin in plant embryogenesis. *Development* **132**: 3577-3585
- Jenik PD, Gillmor CS, Lukowitz W (2007) Embryonic patterning in Arabidopsis thaliana. *Annu Rev Cell Dev Biol* **23**: 207-236
- Jiang D, Gu X, He Y (2009) Establishment of the winter-annual growth habit via FRIGIDA-mediated histone methylation at FLOWERING LOCUS C in Arabidopsis. *Plant Cell* **21**: 1733-1746
- Jiang D, Wang Y, Wang Y, He Y (2008) Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb repressive complex 2 components. *PLoS One* **3**: e3404
- Jiang D, Yang W, He Y, Amasino RM (2007) Arabidopsis relatives of the human lysine-specific Demethylase1 repress the expression of FWA and FLOWERING LOCUS C and thus promote the floral transition. *Plant Cell* **19**: 2975-2987
- Johnson CN, Adkins NL, Georgel P (2005) Chromatin remodeling complexes: ATP-dependent machines in action. *Biochem Cell Biol* **83**: 405-417
- Jullien PE, Katz A, Oliva M, Ohad N, Berger F (2006) Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in Arabidopsis. *Curr Biol* **16**: 486-492
- Jung M, Brosch G, Kolle D, Scherf H, Gerhauser C, Loidl P (1999) Amide analogues of trichostatin A as inhibitors of histone deacetylase and inducers of terminal cell differentiation. *J Med Chem* **42**: 4669-4679
- Kagale S, Rozwadowski K (2011) EAR motif-mediated transcriptional repression in plants: an underlying mechanism for epigenetic regulation of gene expression. *Epigenetics* **6**: 141-146
- Kagaya Y, Ohmiya K, Hattori T (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucl Acids Res* **27**: 470-478



Kagaya Y, Toyoshima R, Okuda R, Usui H, Yamamoto A, Hattori T (2005a) LEAFY COTYLEDON1 Controls Seed Storage Protein Genes through Its Regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3. *Plant Cell Physiol*: pci048

Kagaya Y, Toyoshima R, Okuda R, Usui H, Yamamoto A, Hattori T (2005b) LEAFY COTYLEDON1 Controls Seed Storage Protein Genes through Its Regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3. *Plant Cell Physiol* **46**: 399-406

Katsani KR, Hajibagheri MA, Verrijzer CP (1999) Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology. *Embo J* **18**: 698-708

Kaufmann K, Melzer R, Theissen G (2005) MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene* **347**: 183-198

Keith K, Kraml M, Dengler NG, McCourt P (1994) fusca3: A Heterochronic Mutation Affecting Late Embryo Development in Arabidopsis. *Plant Cell* **6**: 589-600

Kellum R, Schedl P (1991) A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**: 941-950

Kellum R, Schedl P (1992) A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol Cell Biol* **12**: 2424-2431

Kerrigan LA, Croston GE, Lira LM, Kadonaga JT (1991) Sequence-specific transcriptional antirepression of the Drosophila Kruppel gene by the GAGA factor. *J Biol Chem* **266**: 574-582

Kim IS, Sinha S, de Crombrughe B, Maity SN (1996) Determination of functional domains in the C subunit of the CCAAT-binding factor (CBF) necessary for formation of a CBF-DNA complex: CBF-B interacts simultaneously with both the CBF-A and CBF-C subunits to form a heterotrimeric CBF molecule. *Mol Cell Biol* **16**: 4003-4013

Kirik V, Kollé K, Balzer HJ, Baumlein H (1996) Two new oleosin isoforms with altered expression patterns in seeds of the Arabidopsis mutant fus3. *Plant Mol Biol* **31**: 413-417.

Köhler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Grissem W (2003) Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *EMBO J* **22**: 4804-4814

Köhler C, Villar CB (2008) Programming of gene expression by Polycomb group proteins. *Trends Cell Biol* **18**: 236-243

Kooiker M, Airoidi CA, Losa A, Manzotti PS, Finzi L, Kater MM, Colombo L (2005) BASIC PENTACYSTEINE1, a GA Binding Protein That Induces Conformational Changes in the Regulatory Region of the Homeotic Arabidopsis Gene SEEDSTICK. *Plant Cell*: tpc.104.030130

Koornneef M, Bentsink L, Hilhorst H (2002) Seed dormancy and germination. *Curr Opin Plant Biol* **5**: 33-36.





Koornneef M, Hanhart CJ, Hilhorst HW, Karssen CM (1989) In Vivo Inhibition of Seed Development and Reserve Protein Accumulation in Recombinants of Abscisic Acid Biosynthesis and Responsiveness Mutants in *Arabidopsis thaliana*. *Plant Physiol* **90**: 463-469

Koornneef M, Reuling G, Karssen CM (1984) The Isolation and Characterization of Abscisic-Acid Insensitive Mutants of *Arabidopsis-Thaliana*. *Physiol Plant* **61**: 377-383

Kotak S, Vierling E, Baumlein H, von Koskull-Doring P (2007) A novel transcriptional cascade regulating expression of heat stress proteins during seed development of *Arabidopsis*. *Plant Cell* **19**: 182-195

Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128**: 693-705

Kroj T, Savino G, Valon C, Giraudat J, Parcy F (2003) Regulation of storage protein gene expression in *Arabidopsis*. *Development* **130**: 6065-6073

Kunert N, Brehm A (2009) Novel Mi-2 related ATP-dependent chromatin remodelers. *Epigenetics* **4**: 209-211

Kurepa J, Walker JM, Smalle J, Gosink MM, Davis SJ, Durham TL, Sung DY, Vierstra RD (2003) The small ubiquitin-like modifier (SUMO) protein modification system in *Arabidopsis*. Accumulation of SUMO1 and -2 conjugates is increased by stress. *The Journal of biological chemistry* **278**: 6862-6872

Kwong RW, Bui AQ, Lee H, Kwong LW, Fischer RL, Goldberg RB, Harada JJ (2003) LEAFY COTYLEDON1-LIKE Defines a Class of Regulators Essential for Embryo Development. *Plant Cell* **15**: 5-18

Lara P, Onate-Sanchez L, Abraham Z, Ferrandiz C, Diaz I, Carbonero P, Vicente-Carbajosa J (2003) Synergistic activation of seed-storage protein gene expression in *Arabidopsis* by ABI3 and two bZIPs related to OPAQUE2. *J Biol Chem*

Lauria M, Rossi V (2011a) Epigenetic control of gene regulation in plants. *Biochim Biophys Acta*

Lauria M, Rossi V (2011b) Epigenetic control of gene regulation in plants. *Biochimica et biophysica acta* **1809**: 369-378

Laux T, Wurschum T, Breuninger H (2004) Genetic regulation of embryonic pattern formation. *Plant Cell* **16 Suppl**: S190-202

Lawrence RJ, Earley K, Pontes O, Silva M, Chen ZJ, Neves N, Viegas W, Pikaard CS (2004) A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. *Mol Cell* **13**: 599-609

Le Page S, Kwiatkowski F, Paulin C, Mohamed F, Pezet D, Chipponi J, Benhamed M, Gilly FN, Glehen O (2006) In vitro thermochemotherapy of colon cancer cell lines with irinotecan alone and combined with mitomycin C. *Hepato-gastroenterology* **53**: 693-697



Lee C, Li X, Hechmer A, Eisen M, Biggin MD, Venters BJ, Jiang C, Li J, Pugh BF, Gilmour DS (2008) NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol Cell Biol* **28**: 3290-3300

Lee H, Fischer RL, Goldberg RB, Harada JJ (2003) Arabidopsis LEAFY COTYLEDON1 represents a functionally specialized subunit of the CCAAT binding transcription factor. *Proc Natl Acad Sci U S A* **100**: 2152-2156

Lee WY, Lee D, Chung WI, Kwon CS (2009) Arabidopsis ING and Alfin1-like protein families localize to the nucleus and bind to H3K4me3/2 via plant homeodomain fingers. *Plant J* **58**: 511-524

Leibovich BA (2002) [Chromatin structure, heterochromatin, and transposable genetic elements--are they from one team?]. *Mol Biol (Mosk)* **36**: 252-260

Li G, Bishop KJ, Chandrasekharan MB, Hall TC (1999) beta-Phaseolin gene activation is a two-step process: PvALF-facilitated chromatin modification followed by abscisic acid-mediated gene activation. *Proc Natl Acad Sci U S A* **96**: 7104-7109

Li G, Chandler SP, Wolffe AP, Hall TC (1998) Architectural specificity in chromatin structure at the TATA box in vivo: nucleosome displacement upon beta-phaseolin gene activation. *Proc Natl Acad Sci U S A* **95**: 4772-4777

Li H-C, Chuang K, Henderson JT, Rider SD, Bai Y, Zhang H, Fountain M, Gerber J, Ogas J (2005) PICKLE acts during germination to repress expression of embryonic traits. *Plant J* **44**: 1010-1022

Li XY, MacArthur S, Bourgon R, Nix D, Pollard DA, Iyer VN, Hechmer A, Simirenko L, Stapleton M, Luengo Hendriks CL, Chu HC, Ogawa N, Inwood W, Sementchenko V, Beaton A, Weizmann R, Celniker SE, Knowles DW, Gingeras T, Speed TP, Eisen MB, Biggin MD (2008) Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol* **6**: e27

Li XY, Mantovani R, Hooft van Huijsduijnen R, Andre I, Benoist C, Mathis D (1992) Evolutionary variation of the CCAAT-binding transcription factor NF-Y. *Nucleic Acids Res* **20**: 1087-1091

Li Y, Jin K, Zhu Z, Yang J (2010) Stepwise origin and functional diversification of the AFL subfamily B3 genes during land plant evolution. *J Bioinform Comput Biol* **8 Suppl 1**: 33-45

Lippman Z, May B, Yordan C, Singer T, Martienssen R (2003) Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol* **1**: E67

Lis J, Wu C (1993) Protein traffic on the heat shock promoter: parking, stalling, and trucking along. *Cell* **74**: 1-4

Liu C, Lu F, Cui X, Cao X (2010) Histone methylation in higher plants. *Annu Rev Plant Biol* **61**: 395-420



Liu F, Quesada V, Crevillen P, Baurle I, Swiezewski S, Dean C (2007) The Arabidopsis RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate FLC. *Mol Cell* **28**: 398-407

Lois LM, Lima CD, Chua NH (2003) Small ubiquitin-like modifier modulates abscisic acid signaling in Arabidopsis. *Plant Cell* **15**: 1347-1359

Long JA, Szemenyei H, Hannon M (2006) TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. *Science* **319**: 1384-1386

Lotan T, Ohto M, Yee KM, West MA, Lo R, Kwong RW, Yamagishi K, Fischer RL, Goldberg RB, Harada JJ (1998) Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. 93 Thesis, Cell,

Lu F, Cui X, Zhang S, Jenuwein T, Cao X (2011) Arabidopsis REF6 is a histone H3 lysine 27 demethylase. *Nature genetics* **43**: 715-719

Lu F, Cui X, Zhang S, Liu C, Cao X (2010) JMJ14 is an H3K4 demethylase regulating flowering time in Arabidopsis. *Cell Res* **20**: 387-390

Lu F, Li G, Cui X, Liu C, Wang XJ, Cao X (2008) Comparative analysis of JmjC domain-containing proteins reveals the potential histone demethylases in Arabidopsis and rice. *J Integr Plant Biol* **50**: 886-896

Lu Q, Wallrath LL, Granok H, Elgin SC (1993) (CT)<sub>n</sub> (GA)<sub>n</sub> repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the Drosophila hsp26 gene. *Mol Cell Biol* **13**: 2802-2814

Luerssen H, Kirik V, Herrmann P, 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* **15**: 755-764.

Luo M, Bilodeau P, Koltunow A, Dennis ES, Peacock WJ, Chaudhury AM (1999) Genes controlling fertilization-independent seed development in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **96**: 296-301

Mahmoudi T, Zuijderduijn LM, Mohd-Sarip A, Verrijzer CP (2003) GAGA facilitates binding of Pleiohomeotic to a chromatinized Polycomb response element. *Nucleic Acids Res* **31**: 4147-4156

Maier VK, Chioda M, Becker PB (2008) ATP-dependent chromosome remodeling. *Biol Chem* **389**: 345-352

Makarevich G, Leroy O, Akinci U, Schubert D, Clarenz O, Goodrich J, Grossniklaus U, Kohler C (2006) Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep* **7**: 947-952

Malagnac F, Bartee L, Bender J (2002) An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. *Embo J* **21**: 6842-6852



- Mantovani R (1999) The molecular biology of the CCAAT-binding factor NF-Y. *Gene* **239**: 15-27
- Martin C, Ellis N, Rook F (2010) Do transcription factors play special roles in adaptive variation? *Plant physiology* **154**: 506-511
- Martin CH, Mayeda CA, Davis CA, Ericsson CL, Knafels JD, Mathog DR, Celniker SE, Lewis EB, Palazzolo MJ (1995) Complete sequence of the bithorax complex of *Drosophila*. *Proc Natl Acad Sci U S A* **92**: 8398-8402
- Masiero S, Imbriano C, Ravasio F, Favaro R, Pelucchi N, Gorla MS, Mantovani R, Colombo L, Kater MM (2002) Ternary complex formation between MADS-box transcription factors and the histone fold protein NF-YB. *J Biol Chem* **277**: 26429-26435.
- Matharu NK, Hussain T, Sankaranarayanan R, Mishra RK (2010) Vertebrate homologue of *Drosophila* GAGA factor. *J Mol Biol* **400**: 434-447
- Mathieu O, Probst AV, Paszkowski J (2005) Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis*. *Embo J* **24**: 2783-2791
- McCarty DR, Hattori T, Carson CB, Vasil V, Lazar M, Vasil IK (1991) The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator. *Cell* **66**: 895-905
- McNabb DS, Xing Y, Guarente L (1995) Cloning of yeast HAP5: a novel subunit of a heterotrimeric complex required for CCAAT binding. *Genes Dev* **9**: 47-58
- Meinke DW (1992) A Homeotic Mutant of *Arabidopsis-Thaliana* with Leafy Cotyledons. *Science* **258**: 1647-1650
- Meinke DW, Franzmann LH, Nickle TC, Yeung EC (1994) Leafy Cotyledon Mutants of *Arabidopsis*. *Plant Cell* **6**: 1049-1064
- Meister RJ, Williams LA, Monfared MM, Gallagher TL, Kraft EA, Nelson CG, Gasser CS (2004) Definition and interactions of a positive regulatory element of the *Arabidopsis* INNER NO OUTER promoter. *Plant J* **37**: 426-438
- Meluh PB, Koshland D (1995) Evidence that the MIF2 gene of *Saccharomyces cerevisiae* encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol Biol Cell* **6**: 793-807
- Mihaly J, Hogga I, Barges S, Galloni M, Mishra RK, Hagstrom K, Muller M, Schedl P, Sipos L, Gausz J, Gyurkovics H, Karch F (1998) Chromatin domain boundaries in the Bithorax complex. *Cell Mol Life Sci* **54**: 60-70
- Mishra K, Chopra VS, Srinivasan A, Mishra RK (2003) Trl-GAGA directly interacts with lola like and both are part of the repressive complex of Polycomb group of genes. *Mech Dev* **120**: 681-689
- Mishra RK, Mihaly J, Barges S, Spierer A, Karch F, Hagstrom K, Schweinsberg SE, Schedl P (2001) The *iab-7* polycomb response element maps to a nucleosome-free region of chromatin





and requires both GAGA and pleiohomeotic for silencing activity. *Mol Cell Biol* **21**: 1311-1318

Miura K, Jin JB, Hasegawa PM (2007) Sumoylation, a post-translational regulatory process in plants. *Curr Opin Plant Biol* **10**: 495-502

Miura K, Rus A, Sharkhuu A, Yokoi S, Karthikeyan AS, Raghothama KG, Baek D, Koo YD, Jin JB, Bressan RA, Yun D-J, Hasegawa PM (2005) The Arabidopsis SUMO E3 ligase SIZ1 controls phosphate deficiency responses. *Proc Natl Acad Sci U S A*: 0500778102

Mohd-Sarip A, Venturini F, Chalkley GE, Verrijzer CP (2002) Pleiohomeotic can link polycomb to DNA and mediate transcriptional repression. *Mol Cell Biol* **22**: 7473-7483

Monfared MM, Simon MK, Meister RJ, Roig-Villanova I, Kooiker M, Colombo L, Fletcher JC, Gasser CS (2011) Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. *Plant J* **66**: 1020-1031

Monke G, Altschmied L, Tewes A, Reidt W, Mock HP, Baumlein H, Conrad U (2004) Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA. *Planta* **219**: 158-166

Moon YH, Chen L, Pan RL, Chang HS, Zhu T, Maffeo DM, Sung ZR (2003) EMF genes maintain vegetative development by repressing the flower program in Arabidopsis. *Plant Cell* **15**: 681-693

Mulholland NM, King IF, Kingston RE (2003) Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins Zeste and GAGA. *Genes Dev* **17**: 2741-2746

Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE, Simon JA (2002) Histone methyltransferase activity of a Drosophila polycomb group repressor complex. *Cell* **111**: 197-208

Muller J, Kassis JA (2006) Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Curr Opin Genet Dev* **16**: 476-484

Murtas G, Reeves PH, Fu YF, Bancroft I, Dean C, Coupland G (2003) A nuclear protease required for flowering-time regulation in Arabidopsis reduces the abundance of SMALL UBIQUITIN-RELATED MODIFIER conjugates. *The Plant cell* **15**: 2308-2319

Mutskov V, Gerber D, Angelov D, Ausio J, Workman J, Dimitrov S (1998) Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding. *Mol Cell Biol* **18**: 6293-6304

Mylne JS, Barrett L, Tessadori F, Mesnage S, Johnson L, Bernatavichute YV, Jacobsen SE, Fransz P, Dean C (2006) LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. *Proc Natl Acad Sci U S A* **103**: 5012-5017

Nakamura S, Lynch TJ, Finkelstein RR (2001) Physical interactions between ABA response loci of Arabidopsis. *Plant J* **26**: 627-635



- Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of Arabidopsis. *Plant Mol Biol* **60**: 51-68
- Nakayama T, Nishioka K, Dong YX, Shimojima T, Hirose S (2007) Drosophila GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading. *Genes Dev* **21**: 552-561
- Nambara E, Hayama R, Tsuchiya Y, Nishimura M, Kawaide H, Kamiya Y, 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.
- Naumann K, Fischer A, Hofmann I, Krauss V, Phalke S, Irmeler K, Hause G, Aurich AC, Dorn R, Jenuwein T, Reuter G (2005) Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in Arabidopsis. *Embo J* **24**: 1418-1429
- Newman T, de Bruijn FJ, Green P, Keegstra K, Kende H, McIntosh L, Ohlrogge J, Raikhel N, Somerville S, Thomashow M, et al. (1994) Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones. *Plant Physiol* **106**: 1241-1255
- Ng DW-K, Chandrasekharan MB, Hall TC (2006) Ordered Histone Modifications Are Associated with Transcriptional Poising and Activation of the phaseolin Promoter. *Plant Cell* **18**: 119-132
- Ng DW-K, Hall TC (2008) PvALF and FUS3 activate expression from the phaseolin promoter by different mechanisms. *Plant Mol Biol* **66**: 233-244
- North H, Baud S, Debeaujon I, Dubos C, Dubreucq B, Grappin P, Jullien M, Lepiniec L, Marion-Poll A, Miquel M, Rajjou L, Routaboul JM, Caboche M (2010) Arabidopsis seed secrets unravelled after a decade of genetic and omics-driven research. *Plant J* **61**: 971-981
- Novatchkova M, Budhiraja R, Coupland G, Eisenhaber F, Bachmair A (2004) SUMO conjugation in plants. *Planta* **220**: 1-8
- O'Brien T, Wilkins RC, Giardina C, Lis JT (1995) Distribution of GAGA protein on Drosophila genes in vivo. *Genes Dev* **9**: 1098-1110
- Ogas J, Cheng JC, Sung ZR, Somerville C (1997) Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. *Science* **277**: 91-94
- Ogas J, Kaufmann S, Henderson J, Somerville C (1999) PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in Arabidopsis. *Proc Natl Acad Sci U S A* **96**: 13839-13844.
- Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S (2003) Gibberellin Biosynthesis and Response during Arabidopsis Seed Germination. *Plant Cell* **15**: 1591-1604



Ohad N, Yadegari R, Margossian L, Hannon M, Michaeli D, Harada JJ, Goldberg RB, Fischer RL (1999) Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell* **11**: 407-416

Ohtsuki S, Levine M (1998) GAGA mediates the enhancer blocking activity of the eve promoter in the Drosophila embryo. *Genes Dev* **12**: 3325-3330

Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D (1998) FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92**: 105-116

Orphanides G, Wu WH, Lane WS, Hampsey M, Reinberg D (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* **400**: 284-288

Pagans S, Ortiz-Lombardia M, Espinas ML, Bernues J, Azorin F (2002) The Drosophila transcription factor tramtrack (TTK) interacts with Trithorax-like (GAGA) and represses GAGA-mediated activation. *Nucleic Acids Res* **30**: 4406-4413

Pandey R, Muller A, Napoli CA, Selinger DA, Pikaard CS, Richards EJ, Bender J, Mount DW, Jorgensen RA (2002) Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucl Acids Res* **30**: 5036-5055

Papp B, Muller J (2006) Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev* **20**: 2041-2054

Parcy F, Valon C, Kohara A, Misera S, Giraudat J (1997) The ABSCISIC ACID-INSENSITIVE3, FUSCA3, and LEAFY COTYLEDON1 loci act in concert to control multiple aspects of Arabidopsis seed development. *Plant Cell* **9**: 1265-1277.

Parcy F, Valon C, Raynal M, Gaubier-Comella P, Delseny M, Giraudat J (1994) Regulation of gene expression programs during Arabidopsis seed development. *roles of the ABI3 locus and or endogenous abscisic acid*: 1567-1582

Pauli S, Rothnie HM, Chen G, He X, Hohn T (2004) The cauliflower mosaic virus 35S promoter extends into the transcribed region. *J Virol* **78**: 12120-12128

Pedone PV, Ghirlando R, Clore GM, Gronenborn AM, Felsenfeld G, Omichinski JG (1996) The single Cys2-His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding. *Proc Natl Acad Sci U S A* **93**: 2822-2826

Pelaz S, Gustafson-Brown C, Kohalmi SE, Crosby WL, Yanofsky MF (2001) APETALA1 and SEPALLATA3 interact to promote flower development. *Plant J* **26**: 385-394

Perruc E, Kinoshita N, Lopez-Molina L (2007) The role of chromatin-remodeling factor PKL in balancing osmotic stress responses during Arabidopsis seed germination. *The Plant Journal* **52**: 927-936



- Perry J, Zhao Y (2003) The CW domain, a structural module shared amongst vertebrates, vertebrate-infecting parasites and higher plants. *Trends Biochem Sci* **28**: 576-580
- Perry SE, Lehti MD, Fernandez DE (1999) The MADS-domain protein AGAMOUS-like 15 accumulates in embryonic tissues with diverse origins. *Plant Physiol* **120**: 121-130.
- Perry SE, Nichols KW, Fernandez DE (1996) The MADS domain protein AGL15 localizes to the nucleus during early stages of seed development. *Plant Cell* **8**: 1977-1989.
- Peserico A, Simone C (2011) Physical and functional HAT/HDAC interplay regulates protein acetylation balance. *J Biomed Biotechnol* **2011**: 371832
- Pien S, Fleury D, Mylne JS, Crevillen P, Inze D, Avramova Z, Dean C, Grossniklaus U (2008) ARABIDOPSIS TRITHORAX1 dynamically regulates FLOWERING LOCUS C activation via histone 3 lysine 4 trimethylation. *Plant Cell* **20**: 580-588
- Pien S, Grossniklaus U (2007) Polycomb group and trithorax group proteins in Arabidopsis. *Biochim Biophys Acta* **1769**: 375-382
- Probst AV, Fagard M, Proux F, Mourrain P, Boutet S, Earley K, Lawrence RJ, Pikaard CS, Murfett J, Furner I, Vaucheret H, Mittelsten Scheid O (2004) Arabidopsis histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. *Plant Cell* **16**: 1021-1034
- Puig OM, Belles E, Lopez-Rodas G, Sendra R, Tordera V (1998) Interaction between N-terminal domain of H4 and DNA is regulated by the acetylation degree. *Biochim Biophys Acta* **1397**: 79-90
- Qin F, Sakuma Y, Tran LS, Maruyama K, Kidokoro S, Fujita Y, Fujita M, Umezawa T, Sawano Y, Miyazono K, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K (2008) Arabidopsis DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress-responsive gene expression. *Plant Cell* **20**: 1693-1707
- Racki LR, Narlikar GJ (2008) ATP-dependent chromatin remodeling enzymes: two heads are not better, just different. *Curr Opin Genet Dev* **18**: 137-144
- Reidt W, Wohlfarth T, Ellerstrom M, Czihal A, Tewes A, Ezcurra I, Rask L, Baumlein H (2000) Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. *Plant J* **21**: 401-408.
- Rider SD, Henderson JT, Jerome RE, Edenberg HJ, Romero-Severson J, Ogas J (2003) Coordinate repression of regulators of embryonic identity by PICKLE during germination in Arabidopsis. *Plant J* **35**: 33-43
- Riechmann JL, Meyerowitz EM (1997) MADS domain proteins in plant development. *Biol Chem* **378**: 1079-1101.
- Riechmann JL, Ratcliffe OJU-hwscsaBV-B-Hadafdcc (2000) A genomic perspective on plant transcription factors. *Curr opin Plant Biol* **3**: 423-434





Ringrose L, Paro R (2007) Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* **134**: 223-232

Roeder RG (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* **21**: 327-335

Roeder RG (2005) Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett* **579**: 909-915

Romanel EA, Schrago CG, Counago RM, Russo CA, Alves-Ferreira M (2009) Evolution of the B3 DNA binding superfamily: new insights into REM family gene diversification. *PLoS One* **4**: e5791

Roschztardt H, Fuentes I, Vasquez M, Corvalan C, Leon G, Gomez I, Araya A, Holuigue L, Vicente-Carbajosa J, Jordana X (2009) A nuclear gene encoding the iron-sulfur subunit of mitochondrial complex II is regulated by B3 domain transcription factors during seed development in Arabidopsis. *Plant Physiol* **150**: 84-95

Roudier F, Ahmed I, Berard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L, Giraut L, Despres B, Drevensek S, Barneche F, Derozier S, Brunaud V, Aubourg S, Schnittger A, Bowler C, Martin-Magniette ML, Robin S, Caboche M, Colot V (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *Embo J* **30**: 1928-1938

Ruthenburg AJ, Allis CD, Wysocka J (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* **25**: 15-30

Sakata Y, Nakamura I, Taji T, Tanaka S, Quatrano RS (2010) Regulation of the ABA-responsive Em promoter by ABI3 in the moss *Physcomitrella patens*: role of the ABA response element and the RY element. *Plant Signal Behav* **5**: 1061-1066

Sala A, Toto M, Pinello L, Gabriele A, Di Benedetto V, Ingrassia AM, Lo Bosco G, Di Gesu V, Giancarlo R, Corona DF (2011) Genome-wide characterization of chromatin binding and nucleosome spacing activity of the nucleosome remodelling ATPase ISWI. *Embo J* **30**: 1766-1777

Saleh A, Al-Abdallat A, Ndamukong I, Alvarez-Venegas R, Avramova Z (2007) The Arabidopsis homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish 'bivalent chromatin marks' at the silent AGAMOUS locus. *Nucleic Acids Res* **35**: 6290-6296

Saleh A, Alvarez-Venegas R, Yilmaz M, Le O, Hou G, Sadler M, Al-Abdallat A, Xia Y, Lu G, Ladunga I, Avramova Z (2008) The highly similar Arabidopsis homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. *Plant Cell* **20**: 568-579

Sangwan I, O'Brian MR (2002) Identification of a soybean protein that interacts with GAGA element dinucleotide repeat DNA. *Plant Physiol* **129**: 1788-1794



Santi L, Wang Y, Stile MR, Berendzen K, Wanke D, Roig C, Pozzi C, Muller K, Muller J, Rohde W, Salamini F (2003) The GA octodinucleotide repeat binding factor BBR participates in the transcriptional regulation of the homeobox gene *Bkn3*. *Plant J* **34**: 813-826

Santos Mendoza M, Dubreucq B, Miquel M, Caboche M, Lepiniec L (2005) LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves. *FEBS Lett* **579**: 4666-4670

Saracco SA, Miller MJ, Kurepa J, Vierstra RD (2007) Genetic analysis of SUMOylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. *Plant physiology* **145**: 119-134

Saurin AJ, Shao Z, Erdjument-Bromage H, Tempst P, Kingston RE (2001) A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. *Nature* **412**: 655-660

Saze H, Shiraishi A, Miura A, Kakutani T (2008) Control of genic DNA methylation by a *jmjC* domain-containing protein in Arabidopsis thaliana. *Science* **319**: 462-465

Schonrock N, Bouveret R, Leroy O, Borghi L, Kohler C, Grissem W, Hennig L (2006) Polycomb-group proteins repress the floral activator AGL19 in the FLC-independent vernalization pathway. *Genes Dev* **20**: 1667-1678

Schubert D, Primavesi L, Bishopp A, Roberts G, Doonan J, Jenuwein T, Goodrich J (2006) Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *Embo J* **25**: 4638-4649

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G (2007) Genome regulation by polycomb and trithorax proteins. *Cell* **128**: 735-745

Schultz TF, Quatrano RS (1997) Characterization and expression of a rice RAD23 gene. *Plant Mol Biol* **34**: 557-562

Schwartz YB, Pirrotta V (2007) Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet* **8**: 9-22

Schweinsberg SE, Schedl P (2004) Developmental modulation of Fab-7 boundary function. *Development* **131**: 4743-4749

Scott KS, Geyer PK (1995) Effects of the *su(Hw)* insulator protein on the expression of the divergently transcribed Drosophila *yolk* protein genes. *Embo J* **14**: 6258-6267

Sheldon CC, Conn AB, Dennis ES, Peacock WJ (2002) Different regulatory regions are required for the vernalization-induced repression of FLOWERING LOCUS C and for the epigenetic maintenance of repression. *Plant Cell* **14**: 2527-2537

Shimell MJ, Peterson AJ, Burr J, Simon JA, O'Connor MB (2000) Functional analysis of repressor binding sites in the *iab-2* regulatory region of the abdominal-A homeotic gene. *Dev Biol* **218**: 38-52



Shimajima T, Okada M, Nakayama T, Ueda H, Okawa K, Iwamatsu A, Handa H, Hirose S (2003) Drosophila FACT contributes to Hox gene expression through physical and functional interactions with GAGA factor. *Genes Dev* **17**: 1605-1616

Shiu S-H, Shih M-C, Li W-H (2005) Transcription Factor Families Have Much Higher Expansion Rates in Plants than in Animals. *Plant Physiol* **139**: 18-26

Sieburth LE, Meyerowitz EM (1997) Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. *Plant Cell* **9**: 355-365

Sinha S, Kim IS, Sohn KY, de Crombrughe B, Maity SN (1996) Three classes of mutations in the A subunit of the CCAAT-binding factor CBF delineate functional domains involved in the three-step assembly of the CBF-DNA complex. *Mol Cell Biol* **16**: 328-337

Sinha S, Maity SN, Lu J, de Crombrughe B (1995) Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3. *Proc Natl Acad Sci U S A* **92**: 1624-1628

Soppe WJ, Jasencakova Z, Houben A, Kakutani T, Meister A, Huang MS, Jacobsen SE, Schubert I, Fransz PF (2002) DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *Embo J* **21**: 6549-6559

Spedaletti V, Polticelli F, Capodaglio V, Schinina ME, Stano P, Federico R, Tavladoraki P (2008) Characterization of a lysine-specific histone demethylase from Arabidopsis thaliana. *Biochemistry* **47**: 4936-4947

Spillane C, MacDougall C, Stock C, Kohler C, Vielle-Calzada JP, Nunes SM, Grossniklaus U, Goodrich J (2000) Interaction of the Arabidopsis polycomb group proteins FIE and MEA mediates their common phenotypes. *Curr Biol* **10**: 1535-1538.

Springer NM, Napoli CA, Selinger DA, Pandey R, Cone KC, Chandler VL, Kaeppler HF, Kaeppler SM (2003) Comparative analysis of SET domain proteins in maize and Arabidopsis reveals multiple duplications preceding the divergence of monocots and dicots. *Plant Physiol* **132**: 907-925

Sridhar VV, Surendrarao A, Liu Z (2006) APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development. *Development* **133**: 3159-3166

Sterner DE, Berger SL (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev* **64**: 435-459

Stone S, Kwong L, Yee K, Pelletier J, Lepiniec L, Fischer R, Goldberg R, Harada J (2001) LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proc Natl Acad Sci U S A* **98**: 11806 - 11811

Stone SL, Braybrook SA, Paula SL, Kwong LW, Meuser J, Pelletier J, Hsieh TF, Fischer RL, Goldberg RB, Harada JJ (2008) Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. *Proc Natl Acad Sci U S A* **105**: 3151-3156



- Sun Q, Zhou DX (2008) Rice jmjC domain-containing gene JMJ706 encodes H3K9 demethylase required for floral organ development. *Proc Natl Acad Sci U S A* **105**: 13679-13684
- Suzuki M, Kao CY, McCarty DR (1997) The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. *Plant Cell* **9**: 799-807
- Suzuki M, Wang HHY, McCarty DR (2007) Repression of the LEAFY COTYLEDON 1/B3 regulatory network in plant embryo development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes. *Plant Physiol* **143**: 902-911
- Tai H, Tai G, Beardmore T (2005) Dynamic Histone Acetylation of Late Embryonic Genes during Seed Germination. *Plant Mol Biol* **59**: 909-925
- Takada S, Jurgens G (2007) Transcriptional regulation of epidermal cell fate in the Arabidopsis embryo. *Development* **134**: 1141-1150
- Tan H, Yang X, Zhang F, Zheng X, Qu C, Mu J, Fu F, Li J, Guan R, Zhang H, Wang G, Zuo J (2011) Enhanced Seed Oil Production in Canola by Conditional Expression of Brassica napus LEAFY COTYLEDON1 and LEC1-LIKE in Developing Seeds. *Plant Physiol* **156**: 1577-1588
- Tanaka M, Kikuchi A, Kamada H (2008) The Arabidopsis histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant Physiol* **146**: 149-161
- Tang W, Perry SE (2003) Binding site selection for the plant MADS-domain protein AGL15: an in vitro and in vivo study. *J Biol Chem*: M212976200
- Tang X, Hou A, Babu M, Nguyen V, Hurtado L, Lu Q, Reyes JC, Wang A, Keller WA, Harada JJ, Tsang EW, Cui Y (2008) The Arabidopsis BRAHMA chromatin-remodeling ATPase is involved in repression of seed maturation genes in leaves. *Plant Physiol* **147**: 1143-1157
- Tariq M, Saze H, Probst AV, Lichota J, Habu Y, Paszkowski J (2003) Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin. *Proc Natl Acad Sci U S A* **100**: 8823-8827
- Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**: 1025-1040
- Terui Y, Saad N, Jia S, McKeon F, Yuan J (2004) Dual role of sumoylation in the nuclear localization and transcriptional activation of NFAT1. *The Journal of biological chemistry* **279**: 28257-28265
- Thakare D, Tang W, Hill K, Perry SE (2008) The MADS-domain transcriptional regulator AGAMOUS-LIKE15 promotes somatic embryo development in Arabidopsis and soybean. *Plant Physiol* **146**: 1663-1672





- Thevenin J, Dubos C, Xu W, Le Gourrierec J, Kelemen Z, Charlot F, Nogue F, Lepiniec L, Dubreucq B (2011) A new system for fast and quantitative analysis of heterologous gene expression in plants. *New Phytol*
- Thorstensen T, Grini PE, Aalen RB (2011) SET domain proteins in plant development. *Biochim Biophys Acta*
- Tian L, Chen ZJ (2001) Blocking histone deacetylation in Arabidopsis induces pleiotropic effects on plant gene regulation and development. *Proc Natl Acad Sci U S A* **98**: 200-205
- Tian L, Fong MP, Wang JJ, Wei NE, Jiang H, Doerge RW, Chen ZJ (2005) Reversible Histone Acetylation and Deacetylation Mediate Genome-Wide, Promoter-Dependent and Locus-Specific Changes in Gene Expression During Plant Development. *Genetics* **169**: 337-345
- Tian L, Wang J, Fong MP, Chen M, Cao H, Gelvin SB, Chen ZJ (2003) Genetic control of developmental changes induced by disruption of Arabidopsis histone deacetylase 1 (AtHD1) expression. *Genetics* **165**: 399-409
- Tiwari SB, Hagen G, Guilfoyle TJ (2004) Aux/IAA proteins contain a potent transcriptional repression domain. *Plant Cell* **16**: 533-543
- To A, Valon C, Savino G, Guilleminot J, Devic M, Giraudat J, Parcy F (2006) A network of local and redundant gene regulation governs Arabidopsis seed maturation. *Plant Cell* **18**: 1642-1651
- Tolhuis B, de Wit E, Muijers I, Teunissen H, Talhout W, van Steensel B, van Lohuizen M (2006) Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. *Nat Genet* **38**: 694-699
- Tsuchiya Y, Nambara E, Naito S, McCourt P (2004) The FUS3 transcription factor functions through the epidermal regulator TTG1 during embryogenesis in Arabidopsis. *Plant J* **37**: 73-81
- Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439**: 811-816
- Tsukagoshi H, Morikami A, Nakamura K (2007) Two B3 domain transcriptional repressors prevent sugar-inducible expression of seed maturation genes in Arabidopsis seedlings. *Proc Natl Acad Sci USA* **104**: 2543-2547
- Tsukagoshi H, Saijo T, Shibata D, Morikami A, Nakamura K (2005) Analysis of a Sugar Response Mutant of Arabidopsis Identified a Novel B3 Domain Protein That Functions as an Active Transcriptional Repressor. *Plant Physiol* **138**: 675-685
- Tsukiyama T, Becker PB, Wu C (1994) ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* **367**: 525-532



- Turck F, Roudier F, Farrona S, Martin-Magniette ML, Guillaume E, Buisine N, Gagnot S, Martienssen RA, Coupland G, Colot V (2007) Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet* **3**: e86
- Uddenberg D, Valladares S, Abrahamsson M, Sundstrom JF, Sundas-Larsson A, von Arnold S (2011) Embryogenic potential and expression of embryogenesis-related genes in conifers are affected by treatment with a histone deacetylase inhibitor. *Planta*
- Ulmasov T, Murfett J, Hagen G, Guilfoyle TJ (1997) Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9**: 1963-1971
- Ulrich HD (2005) Mutual interactions between the SUMO and ubiquitin systems: a plea of no contest. *Trends in cell biology* **15**: 525-532
- Utley RT, Ikeda K, Grant PA, Cote J, Steger DJ, Eberharter A, John S, Workman JL (1998) Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* **394**: 498-502
- van Steensel B (2011) Chromatin: constructing the big picture. *Embo J* **30**: 1885-1895
- van Steensel B, Delrow J, Bussemaker HJ (2003) Genomewide analysis of Drosophila GAGA factor target genes reveals context-dependent DNA binding. *Proc Natl Acad Sci U S A* **100**: 2580-2585
- Vaquero A, Blanch M, Espinas ML, Bernues J (2008) Activation properties of GAGA transcription factor. *Biochim Biophys Acta* **1779**: 312-317
- Vaquero A, Espinas ML, Azorin F, Bernues J (2000) Functional mapping of the GAGA factor assigns its transcriptional activity to the C-terminal glutamine-rich domain. *J Biol Chem* **275**: 19461-19468
- Vasanthi D, Anant M, Srivastava S, Mishra RK (2010) A functionally conserved boundary element from the mouse HoxD locus requires GAGA factor in Drosophila. *Development* **137**: 4239-4247
- Vasil V, Marcotte WR, Jr., Rosenkrans L, Cocciolone SM, Vasil IK, Quatrano RS, McCarty DR (1995) Overlap of Viviparous1 (VP1) and abscisic acid response elements in the Em promoter: G-box elements are sufficient but not necessary for VP1 transactivation. *Plant Cell* **7**: 1511-1518
- Vicente-Carbajosa J, Carbonero P (2005) Seed maturation: developing an intrusive phase to accomplish a quiescent state. *Int J Dev Biol* **49**: 645-651
- Wang D, Tyson MD, Jackson SS, Yadegari R (2006) Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis. *Proc Natl Acad Sci U S A* **103**: 13244-13249



Wang H, Caruso LV, Downie AB, Perry SE (2004a) The Embryo MADS Domain Protein AGAMOUS-Like 15 Directly Regulates Expression of a Gene Encoding an Enzyme Involved in Gibberellin Metabolism. *Plant Cell* **16**: 1206-1219

Wang H, Tang W, Zhu C, Perry SE (2002) A chromatin immunoprecipitation (ChIP) approach to isolate genes regulated by AGL15, a MADS domain protein that preferentially accumulates in embryos. *Plant J* **32**: 831-843

Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS (2004b) Hierarchical recruitment of polycomb group silencing complexes. *Mol Cell* **14**: 637-646

Wang Z, Zang C, Cui K, Schones DE, Barski A, Peng W, Zhao K (2009) Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**: 1019-1031

Wanke D, Hohenstatt ML, Dynowski M, Bloss U, Hecker A, Elgass K, Hummel S, Hahn A, Caesar K, Schleifenbaum F, Harter K, Berendzen KW (2011) Alanine zipper-like coiled-coil domains are necessary for homotypic dimerization of plant GAGA-factors in the nucleus and nucleolus. *PLoS One* **6**: e16070

Wehmeyer N, Vierling E (2000) The expression of small heat shock proteins in seeds responds to discrete developmental signals and suggests a general protective role in desiccation tolerance. *Plant Physiol* **122**: 1099-1108.

Weigel RR, Pfitzner UM, Gatz C (2005) Interaction of NIMIN1 with NPR1 modulates PR gene expression in Arabidopsis. *Plant Cell* **17**: 1279-1291

Weijers D, Jurgens G (2005) Auxin and embryo axis formation: the ends in sight? *Curr opin Plant Biol* **8**: 32-37

West M, Yee KM, Danao J, Zimmerman JL, Fischer RL, Goldberg RB, Harada JJ (1994) LEAFY COTYLEDON1 Is an Essential Regulator of Late Embryogenesis and Cotyledon Identity in Arabidopsis. *Plant Cell* **6**: 1731-1745

Wobus U, Weber HU-hwscsaBV-WSN-dfcfece (1999) Seed maturation: genetic programmes and control signals. *Curr opin Plant Biol* **2**: 33-38

Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA (2006) The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 14631-14636

Woodcock CL, Skoultchi AI, Fan Y (2006) Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res* **14**: 17-25

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano L (2003) The Evolution of Transcriptional Regulation in Eukaryotes. *Mol Biol Evol*: msg140

Wu K, Malik K, Tian L, Brown D, Miki B (2000) Functional analysis of a RPD3 histone deacetylase homologue in Arabidopsis thaliana. *Plant Mol Biol* **44**: 167-176



- Wu K, Zhang L, Zhou C, Yu CW, Chaikam V (2008) HDA6 is required for jasmonate response, senescence and flowering in Arabidopsis. *J Exp Bot* **59**: 225-234
- Xiao H, Sandaltzopoulos R, Wang HM, Hamiche A, Ranallo R, Lee KM, Fu D, Wu C (2001) Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. *Mol Cell* **8**: 531-543
- Xu C-R, Liu C, Wang Y-L, Li L-C, Chen W-Q, Xu Z-H, Bai S-N (2005) Histone acetylation affects expression of cellular patterning genes in the Arabidopsis root epidermis. *PNAS* **102**: 14469-14474
- Xu G, Goodridge AG (1998) A CT repeat in the promoter of the chicken malic enzyme gene is essential for function at an alternative transcription start site. *Arch Biochem Biophys* **358**: 83-91
- Xu L, Shen WH (2008) Polycomb silencing of KNOX genes confines shoot stem cell niches in Arabidopsis. *Curr Biol* **18**: 1966-1971
- Xu L, Zhao Z, Dong A, Soubigou-Taconnat L, Renou JP, Steinmetz A, Shen WH (2008) Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in Arabidopsis thaliana. *Mol Cell Biol* **28**: 1348-1360
- Yamagishi K, Nagata N, Yee KM, Braybrook SA, Pelletier J, Fujioka S, Yoshida S, Fischer RL, Goldberg RB, Harada JJ (2005) TANMEI/EMB2757 encodes a WD repeat protein required for embryo development in Arabidopsis. *Plant Physiol* **139**: 163-173
- Yamamoto A, Kagaya Y, Toyoshima R, Kagaya M, Takeda S, Hattori T (2009) Arabidopsis NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. *Plant J*
- Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Tomo Y, Hayami N, Terada T, Shirouzu M, Osanai T, Tanaka A, Seki M, Shinozaki K, Yokoyama S (2004) Solution Structure of the B3 DNA Binding Domain of the Arabidopsis Cold-Responsive Transcription Factor RAV1. *Plant Cell* **16**: 3448-3459
- Yoo CY, Miura K, Jin JB, Lee J, Park HC, Salt DE, Yun DJ, Bressan RA, Hasegawa PM (2006) SIZ1 small ubiquitin-like modifier E3 ligase facilitates basal thermotolerance in Arabidopsis independent of salicylic acid. *Plant physiology* **142**: 1548-1558
- Yoshida M, Horinouchi S, Beppu T (1995) Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays* **17**: 423-430
- Yu CW, Liu X, Luo M, Chen C, Lin X, Tian G, Lu Q, Cui Y, Wu K (2011) HISTONE DEACETYLASE6 interacts with FLOWERING LOCUS D and regulates flowering in Arabidopsis. *Plant Physiol* **156**: 173-184





Yu X, Li L, Li L, Guo M, Chory J, Yin Y (2008) Modulation of brassinosteroid-regulated gene expression by Jumonji domain-containing proteins ELF6 and REF6 in Arabidopsis. *Proc Natl Acad Sci U S A* **105**: 7618-7623

Zhang H, Rider SD, Jr., Henderson JT, Fountain M, Chuang K, Kandachar V, Simons A, Edenberg HJ, Romero-Severson J, Muir WM, Ogas J (2008) The CHD3 remodeler PICKLE promotes trimethylation of histone H3 lysine 27. *J Biol Chem* **283**: 22637-22648

Zhang K, Sridhar VV, Zhu J, Kapoor A, Zhu JK (2007a) Distinctive core histone post-translational modification patterns in Arabidopsis thaliana. *PLoS ONE* **2**: e1210

Zhang X, Germann S, Blus BJ, Khorasanizadeh S, Gaudin V, Jacobsen SE (2007b) The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation. *Nat Struct Mol Biol* **14**: 869-871

Zhao Z, Shen WH (2004) Plants contain a high number of proteins showing sequence similarity to the animal SUV39H family of histone methyltransferases. *Ann N Y Acad Sci* **1030**: 661-669

Zheng B, Chen X (2011) Dynamics of histone H3 lysine 27 trimethylation in plant development. *Curr Opin Plant Biol* **14**: 123-129

Zheng Y, Ren N, Wang H, Stromberg AJ, Perry SE (2009) Global identification of targets of the Arabidopsis MADS domain protein AGAMOUS-Like15. *Plant Cell* **21**: 2563-2577

Zhou C, Labbe H, Sridha S, Wang L, Tian L, Latoszek-Green M, Yang Z, Brown D, Miki B, Wu K (2004) Expression and function of HD2-type histone deacetylases in Arabidopsis development. *Plant J* **38**: 715-724

Zhou C, Zhang L, Duan J, Miki B, Wu K (2005) HISTONE DEACETYLASE19 is involved in jasmonic acid and ethylene signaling of pathogen response in Arabidopsis. *Plant Cell* **17**: 1196-1204

Zhou J, Wang X, He K, Charron JB, Elling AA, Deng XW (2010) Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in Arabidopsis reveals correlation between multiple histone marks and gene expression. *Plant Mol Biol* **72**: 585-595



## Résumé :

Le sujet de cette thèse est l'étude de la régulation de l'expression du facteur de transcription *LEAFY COTYLEDON 2 (LEC2)*, qui est un régulateur clé du développement de la graine d'*Arabidopsis thaliana*. La graine offre un mode de propagation et de protection des espèces végétales à graines, indispensable à leur survie. La graine est l'élément de base de l'agriculture, en tant que semence, et de l'alimentation humaine, aussi bien sous forme brute que sous forme transformée (farines, huiles...). Elle a aussi de très nombreuses applications industrielles dans l'industrie et les biocarburants. Le développement de la graine, comme beaucoup d'étapes nécessaires à la vie de la plante, est contrôlé par des phénomènes complexes, incluant des facteurs de transcription. *LEC2*, ainsi que 2 autres facteurs de transcription de type B3 (*FUSCA 3* et *ABI3*) et un facteur à domaine de fixation CAAT (*LEAFY COTYLEDON 1*), sont les quatre facteurs clés du développement de la graine d'*Arabidopsis thaliana*, appelés AFLs (*ABI3*, *FUS3*, *LEC1*, 2). Les gènes *AFLs* s'expriment spécifiquement dans l'embryon et sont fortement réprimés dans les parties végétatives, par de multiples mécanismes impliquant, entre autre, des facteurs de transcription et des acteurs plus généraux modifiant la structure de la chromatine. Bien que la régulation de *LEC2* ait déjà été largement étudiée, les mécanismes précis de répression ainsi que l'activation de ce gène, sont encore méconnus.

Le travail présenté est basé sur un travail de délétion de promoteurs, qui a révélé l'existence de 3 boîtes de régulation, essentielles à une activité correcte du promoteur de *LEC2*. Deux boîtes, dont une correspondrait à une boîte de fixation de facteurs MADS et une à une boîte GAGA dans la séquence transcrite non traduite de *LEC2*, sont essentielles à l'activité du promoteur. La 3<sup>ème</sup> séquence, d'une longueur de 50pb, est nécessaire à la répression de *LEC2* dans les parties végétatives par des mécanismes épigénétiques. La corrélation de la présence de cette boîte avec un enrichissement dans la marque H3K27me3 au locus nous a permis de rapprocher cette séquence de répression d'une PRE (Polycomb response element) de type végétal.

## Abstract :

The aim of this work is to study the regulation of the transcription factor *LEAFY COTYLEDON 2*, which is a key regulator of seed development in *Arabidopsis thaliana*.

Seeds have essential functions in the environment for plant propagation and as embryo protective tools. Furthermore, numerous products issued from the agriculture or involved in human food (e.g. cereals, oil, flour) are based on seeds. Several industrial applications depend, as well, on this organ such as oil production for human consumption, additives for some industrial processes, or biofuel synthesis.

Seed developmental phases are dependant of a complex regulatory network composed in major part with transcription factors, that were found to be central components of plant evolution and domestication.

*LEC2*, *FUS3*, *ABI3* (three B3 type factors) and *LEC1* (a CATT binding factor) are named AFL (*ABI3*, *FUS3*, *LEC1*, 2) genes, and are key regulators of *Arabidopsis* seed development. AFL genes are specifically expressed in embryo and repressed in vegetatives tissues. This repression has been principally studied in germinating seedlings and was shown to be caused by a set of transcription factors and chromatin structure modifiers.

Beside the fact that *LEC2* regulation has been extensively studied within the past few years, it was known that other mechanisms of repression and activation were still to be discovered. The work carried out on *LEC2* presented here, mainly based on an extensive promoter deletion analysis, has allowed the discovery of three essential nucleotidic sequences necessary for a proper *LEC2* promoter activity. The two first regulatory sequences are similar to a MADS box binding element and a GAGA binding site, and were found to be essentials for *LEC2* promoter activity. The third sequence (named RLE for Repression of *LEC2* Element) is 50bp long and lead to the repression of *LEC2* promoter activity after onset of seed germination. A very strong correlation between RLE and the enrichment of H3K27me3 mark deposition at specific loci, suggests this sequence is the first PRE-like (Polycomb response element) element identified in plants.

