



**HAL**  
open science

# Rôles des voies régulées par LEAFY dans l'initiation et la régulation du méristème floral

Grégoire Denay

► **To cite this version:**

Grégoire Denay. Rôles des voies régulées par LEAFY dans l'initiation et la régulation du méristème floral. Sciences agricoles. Université Grenoble Alpes, 2016. Français. NNT : 2016GREAV045 . tel-01938828

**HAL Id: tel-01938828**

**<https://theses.hal.science/tel-01938828>**

Submitted on 29 Nov 2018

**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.

## THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ  
GRENOBLE ALPES**

Spécialité : **Biologie végétale**

Arrêté ministériel : 7 août 2006

Présentée par

**Grégoire DENAY**

Thèse dirigée par **François PARCY** et **Gabrielle TICHTINSKY**

préparée au sein du **Laboratoire de Physiologie Cellulaire et Végétale**  
dans l'**École Doctorale de Chimie et Sciences du Vivant**

## **Rôles des voies régulées par LEAFY dans l'initiation et la régulation du méristème floral**

Thèse soutenue publiquement le **28 Novembre 2016**,  
devant le jury composé de :

**Pr Marianne DELARUE**

Professeure, Université Paris-Sud, Rapporteuse

**Pr Michel HERNOULD**

Professeur, Université de Bordeaux, Rapporteur

**Pr Jean-Gabriel VALAY**

Professeur, Université Grenoble Alpes, Président

**Dr Soazig GUYOMARC'H**

Maître de Conférences, Université Montpellier 2, Examinatrice

**Dr Téva VERNOUX**

Directeur de Recherche, CNRS - ENS Lyon, Examineur

**Dr Gabrielle TICHTINSKY**

Maître de Conférences, Université Grenoble Alpes, Co-Directrice de thèse

**Dr François PARCY**

Directeur de Recherche, CNRS - LPCV Grenoble, Directeur de thèse









*A Claude,  
A Oscar,*









## RESUME

---

Les plantes conservent la capacité à former de nouveaux organes tout au long de leur vie grâce au maintien de structures contenant des cellules souches, les méristèmes. La formation des fleurs, structures reproductives de la plante, est une étape essentielle de son cycle de vie. Afin d'assurer un développement floral complet, un méristème doit être formé *de novo* au sein du jeune bouton floral. Des données éparses de la littérature indiquent que le facteur de transcription LEAFY, en plus d'être un régulateur clé de l'identité florale, est aussi impliqué dans la mise en place du méristème floral.

Dans la première partie de ce travail nous explorons le rôle de LEAFY dans l'initiation du méristème floral. Cette étude est concentrée sur un gène cible de LEAFY, le facteur de transcription REGULATOR OF AXILLARY MERISTEMS1 (RAX1). Nous montrons notamment que la voie régulée par LEAFY/RAX1 agit en parallèle du facteur de transcription REVOLUTA pour permettre la mise en place du méristème floral.

Dans la deuxième partie de ce travail nous étudions les propriétés du domaine N-terminal de LEAFY. Ce domaine permet l'oligomérisation de LEAFY ainsi que potentiellement sa liaison aux régions fermées de la chromatine. Nous étudions également de manière plus exploratoire le rôle de ce domaine dans la régulation de l'expression du gène *AGAMOUS*, un important régulateur du développement floral.



## ABSTRACT

---

Plants have the capacity to continuously produce organs throughout their life because they maintain stem cells containing structures called meristems. The formation of flowers is an essential step of the plant's life-cycle. In order to ensure flower development, a new meristem must be formed within the young flower bud. Various data across the literature indicate that the transcription factor LEAFY is involved flower meristem formation in addition to its role as a master regulator of flower identity.

In the first part of this work we explore the role of LEAFY in the initiation of flower meristem. This study focuses on a LEAFY target gene, the transcription factor REGULATOR OF AXILLARY MERISTEMS1 (RAX1). We show that the LEAFY/RAX1 pathway acts in parallel of the transcription factor REVOLUTA to allow flower meristem formation.

In the second part of this work we study the properties of the N-terminal domain of LEAFY. This domain mediates LEAFY oligomerization and potentially its binding to closed chromatin regions. We also study in a more prospective manner the role of this domain in the transcriptional regulation of *AGAMOUS*, an important regulator of flower development.



## REMERCIEMENTS

---

Ce travail est le fruit de la participation de nombreuses personnes qui ont été impliquées, de manière fortuite ou non, dans cette aventure. Sans eux ce travail n'aurait pas été possible et je tiens donc à les remercier ici.

Je souhaite remercier Norbert Rolland qui m'a accueilli à PCV et Jérôme Garin pour son accueil à l'iRTSV/BIG. Je souhaite également remercier les membres du jury qui ont pris le temps d'évaluer ce travail. Je souhaite également remercier Jan Lohmann et Pradeep Das qui ont acceptés de faire partie de mon comité de suivi et m'ont apporté une aide précieuse.

Je remercie chaleureusement François qui m'a guidé tout au long de ce travail. Merci de m'avoir accordé ta confiance et de m'avoir fait participer à de nombreux projets annexes. Merci aussi pour ta franchise et tes conversations passionnées sur LEAFY, la science, l'escalade... J'espère pouvoir continuer à travailler avec toi dans le futur.

Un grand merci également à 'maman Gaby'. Merci d'avoir toujours été prête à m'accorder une oreille attentive, à m'aider, même (surtout ?) sur les manip pourries. Tu as su m'encourager quand tout semblait bancal. Merci aussi pour ta formation écharpe de portage sur peluche qui m'a été bien utile.

Je remercie aussi Marie de m'avoir apporté tant d'aide au cours de la thèse. Ça a été un plaisir de travailler avec toi. Cristel pour m'avoir accompagné à travers les *in situ* et de m'avoir donné l'opportunité de faire un peu d'enseignement. Merci aussi Renaud de m'avoir (presque) fait aimer la biochimie. Manu qui m'a initié aux EMSAs et à la purification de protéine et qui répondu à mes incessantes questions. Un grand merci à Sophie, Danielle, Nassima et Sébastien, les supers gestionnaires. Une mention spéciale pour Sylvie. Tu as réussi à transformer une cave glauque en véritable zone de culture. Les mouches et les thrips réfléchissent à deux fois depuis que tu es là ! (Je te promets de ramasser mes graines avant de partir définitivement). Un merci aussi aux stagiaires que j'ai eu la chance d'encadrer, le grand Greg et Salimata. J'espère que vous avez appris autant que vous m'avez appris.

De nombreuses personnes ont apportées à ce travail par leurs conseils et leurs expertises : Christine Lancelon-Pin pour la microscopie électronique, Ludivine Soubignou-Taconnat, Marie Martin-Magniette et Etienne Delannoy pour le RNA-seq, Didier Grunwald pour la microscopie confocale et Antoine Larrieu qui gentiment partagé de nombreux protocoles.

Un grand merci à tous ceux qui ont fait de ces trois ans un moment agréable, collègues et voisins de bureaux, compagnons d'afterworks et de pause-café et à tous ceux qui sont des amis avant d'être des collègues.

Merci à Oscar qui n'a pas bénéficié de toute l'attention qu'il méritait ces derniers mois et à Lina, pour son aide, sa patience et son support inconditionnel.





## ABREVIATIONS

---

<b>ABA</b>	Acide abscissique
<b>AFB</b>	AUXIN SIGNALING F6BOX
<b>AG</b>	AGAMOUS
<b>AGO</b>	ARGONAUTE
<b>AHK</b>	ARABIDOPSIS HISTIDINE KINASE
<b>AHP</b>	HISTIDINE PHOSPHOTRANSFER PROTEIN
<b>AM</b>	méristème axillaire
<b>AP</b>	APETALLA
<b>Arabidopsis</b>	<i>Arabidopsis thaliana</i> (L.) Heynh., 1842
<b>ARF</b>	AUXINE RESPONSE FACTOR
<b>ARR</b>	ARABIDOPSIS RESPONSE REGULATOR
<b>AS</b>	ASYMMETRIC LEAF
<b>AUX</b>	AUXINE RESISTANT
<b>BDL</b>	BODENLOS
<b>BRM</b>	BRAHMA
<b>BSA</b>	Albumine sérique bovine
<b>CHAPS</b>	Acide 3-[(3-cholamidopropyl)diméthylammonio]-1-propane sulfonique
<b>ChIP</b>	Immuno-précipitation de chromatine
<b>CK</b>	Cytokinine
<b>CLV</b>	CLAVATA
<b>CNA</b>	CORONA
<b>CRN</b>	CORYNE
<b>CUC</b>	CUP-SHAPED COTYLEDONS
<b>CZ</b>	Zone centrale
<b>DAP</b>	Purification d'ADN par affinité
<b>DBD</b>	Domaine de liaison à l'ADN
<b>DEG</b>	Gènes différentiellement exprimés
<b>DEX</b>	Dexaméthasone
<b>DMSO</b>	Diméthylsulfoxyde
<b>DTT</b>	Dithiothréitol
<b>EDTA</b>	Éthylène diamine tétra-acétique
<b>ER</b>	ERECTA
<b>EXB</b>	EXCESSIVE BRANCHING
<b>FIL</b>	FILAMENTOUS FLOWER
<b>FM</b>	méristème floral
<b>FT</b>	Facteur de transcription
<b>GA</b>	Gibbérélline
<b>GR</b>	Domaine récepteur aux glucocorticoïdes du rat
<b>HEC</b>	HECATE
<b>HEPES</b>	Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
<b>IAA</b>	INDOLE-3-ACETIC INDUCIBLE
<b>IPT</b>	ISOPENTENYLTRANSFERASE
<b>JBA</b>	JABBA

<b>JC</b>	Jour court
<b>JL</b>	Jour long
<b>KAN</b>	KANADI
<b>LAS</b>	LATERAL SUPPRESSOR
<b>LAX</b>	LIKE AUXIN RESISTANT
<b>LFY</b>	LEAFY
<b>LSM</b>	méristème de tige latérale
<b>MAB</b>	MACCHI-BOU
<b>MBP</b>	Protéine de liaison au maltose
<b>MES</b>	Acide 2-( <i>N</i> -morpholino)ethane sulfonique
<b>M-MLV</b>	Virus de la leucémie murine Moloney
<b>MP</b>	MONOPTEROS
<b>MS</b>	Murashige et Skoog
<b>NPA</b>	Naptalame
<b>OC</b>	Centre organisateur
<b>PAGE</b>	Electrophorese en gel de poly-acrylamide
<b>PBM</b>	Protein-binding miro-array
<b>PBM</b>	Puce de liaison de protéines
<b>PBS</b>	Tampon phosphate salin
<b>PHB</b>	PHABULOSA
<b>PHV</b>	PHAVOLUTA
<b>PI</b>	PISTALATA
<b>PID</b>	PINOID
<b>PIN</b>	PIN-FORMED
<b>PNH</b>	PINHEAD
<b>PSSM</b>	Matrice de score de position spécifique
<b>PVDF</b>	Poly(fluorure de vinylidène)
<b>PWM</b>	Matrice de poids-position
<b>PZ</b>	Zone périphérique
<b>RAM</b>	méristème apical racinaire
<b>RAX</b>	REGULATOR OF AXILLARY MERISTEMS
<b>REV</b>	REVOLUTA
<b>RM</b>	Méristème médullaire
<b>ROX</b>	REGULATOR OF AXILLARY MERISTEM FORMATION
<b>SAM</b>	Meristème apical caulinaire
<b>SAM</b>	Méristème apical caulinaire
<b>SDS</b>	Dodécylsulfate de sodium
<b>SELEX</b>	Evolution systématique de ligands par enrichissement exponentiel
<b>SEP</b>	SEPALLATA
<b>STM</b>	SHOOTMERISTEMLESS
<b>SYD</b>	SPLAYED
<b>TAA</b>	TRYPTOPHANE AMINOTRANSFERASE OF ARABIDOPSIS
<b>TBS</b>	Tampon tris salin
<b>TCEP</b>	Tris(2-carboxyethyl)phosphine HCl
<b>TPL</b>	TOPLESS
<b>UFO</b>	UNUSUAL FLORAL ORGAN

**WUS**  
**YAB**  
**YUC**  
**ZLL**

WUSCHEL  
YABBI  
YUCCA  
ZWILLE



# TABLE DES MATIERES

RÉSUMÉ .....	- 9 -
ABSTRACT .....	- 11 -
REMERCIEMENTS .....	- 13 -
ABRÉVIATIONS .....	- 17 -
TABLE DES MATIÈRES .....	- 21 -
<b>INTRODUCTION .....</b>	<b>- 25 -</b>
I. LE MÉRISTÈME APICAL CAULINAIRE .....	- 27 -
1. <i>L'homéostasie des cellules souches</i> .....	- 28 -
2. <i>La phyllotaxie</i> .....	- 34 -
II. LE DÉVELOPPEMENT FLORAL .....	- 40 -
1. <i>Article 1: 'A Flower is Born: Recent Progress on Arabidopsis Flower Meristem Formation'</i> .....	- 41 -
2. <i>Que peut-on apprendre de la formation des méristèmes végétatifs ?</i> .....	- 50 -
3. <i>Le développement floral</i> .....	- 55 -
III. IDENTIFICATION DES GÈNES LIÉS PAR DES FACTEURS DE TRANSCRIPTION .....	- 62 -
1. <i>Détermination expérimentale des séquences liées par un TF</i> .....	- 62 -
2. <i>Modélisation et prédiction des sites de liaison d'un FT</i> .....	- 65 -
IV. LEAFY, REGULATOR OF AXILLARY MERISTEM 1 ET REVOLUTA, TROIS FACTEURS DE TRANSCRIPTION IMPLIQUÉS DANS LA FORMATION DES FLEURS ? .....	- 68 -
1. <i>LEAFY, régulateur central du développement floral</i> .....	- 68 -
2. <i>Rôles de RAX1 dans la mise en places des LSM</i> .....	- 72 -
3. <i>REVOLUTA, un gène de développement aux fonctions complexes</i> .....	- 73 -
OBJECTIFS .....	- 77 -
<b>CHAPITRE PREMIER .....</b>	<b>- 79 -</b>
I. INTRODUCTION .....	- 79 -
II. ARTICLE 2 : "CONTROL OF STEM-CELL NICHE ESTABLISHMENT IN ARABIDOPSIS FLOWERS BY LEAFY, ITS TARGET RAX1, AND REVOLUTA" .....	- 80 -
III. RÉSULTATS COMPLÉMENTAIRES .....	- 103 -
1. <i>LFY pourrait réguler l'expression des HD-ZIPIII via AGO10</i> .....	- 103 -
2. <i>Mutation d'un élément cis du promoteur de RAX1</i> .....	- 104 -
3. <i>Quelques commentaires sur les lignées CRISPR/Cas9 ciblant RAX1 et REV</i> .....	- 106 -
IV. CONCLUSION ET PERSPECTIVES .....	- 108 -
<b>CHAPITRE SECOND .....</b>	<b>- 113 -</b>
I. RÔLES DU DOMAINE N-TERMINAL DE LFY .....	- 114 -
1. <i>Introduction</i> .....	- 114 -
2. <i>Résultats</i> .....	- 116 -
3. <i>Discussion</i> .....	- 121 -
II. ÉTUDE DE L'INTERACTION DE LFY ET WUS AU LOCUS AGAMOUS .....	- 124 -
1. <i>Introduction</i> .....	- 124 -
2. <i>Résultats</i> .....	- 129 -

3.	<i>Discussion</i> .....	- 133 -
<b>MATÉRIEL &amp; MÉTHODES</b> .....		<b>- 137 -</b>
I.	MATÉRIEL VÉGÉTAL ET MÉTHODES DE TRAVAIL SUR LES PLANTES.....	- 137 -
1.	<i>Matériel végétal</i> .....	- 137 -
2.	<i>Conditions de cultures</i> .....	- 138 -
3.	<i>Extraction d'ADN génomique et génotypage</i> .....	- 139 -
4.	<i>Génération de lignées transgéniques stables d'Arabidopsis thaliana</i> .....	- 141 -
5.	<i>Phénotypage</i> .....	- 145 -
6.	<i>Analyse des quantités de protéines par Western Blot</i> .....	- 145 -
7.	<i>Analyse de niveau d'expression de gènes par qRT-PCR</i> .....	- 146 -
8.	<i>Microscopie confocale</i> .....	- 149 -
9.	<i>Analyse de l'expression des gènes par hybridation in situ</i> .....	- 150 -
10.	<i>Microscopie électronique</i> .....	- 153 -
II.	PRODUCTION ET CARACTÉRISATION DES PROTÉINES <i>IN VITRO</i> .....	- 153 -
1.	<i>Production de protéines recombinantes</i> .....	- 153 -
2.	<i>Protein-binding micro-array</i> .....	- 154 -
3.	<i>Gels retards</i> .....	- 155 -
III.	MÉTHODES BIO-INFORMATIQUES.....	- 156 -
1.	<i>Analyse des données de transcriptomiques</i> .....	- 156 -
2.	<i>Détermination in silico des sites de liaison à l'ADN des facteurs de transcription</i> .....	- 157 -
3.	<i>Analyse des données de ChIP-seq et d'état chromatinien</i> .....	- 159 -
<b>CONCLUSION &amp; DISCUSSION</b> .....		<b>- 163 -</b>
<b>RÉFÉRENCES</b> .....		<b>- 169 -</b>
<b>ANNEXES</b> .....		<b>- 187 -</b>
	ANNEXE 1 : JASPAR 2016 .....	- 189 -
	ANNEXE 2 : LFY SAM DOMAIN .....	- 195 -
	ANNEXE 3 : EDGER SCRIPT .....	- 209 -
	ANNEXE 4 : BIOPYTHON SCRIPT .....	- 211 -



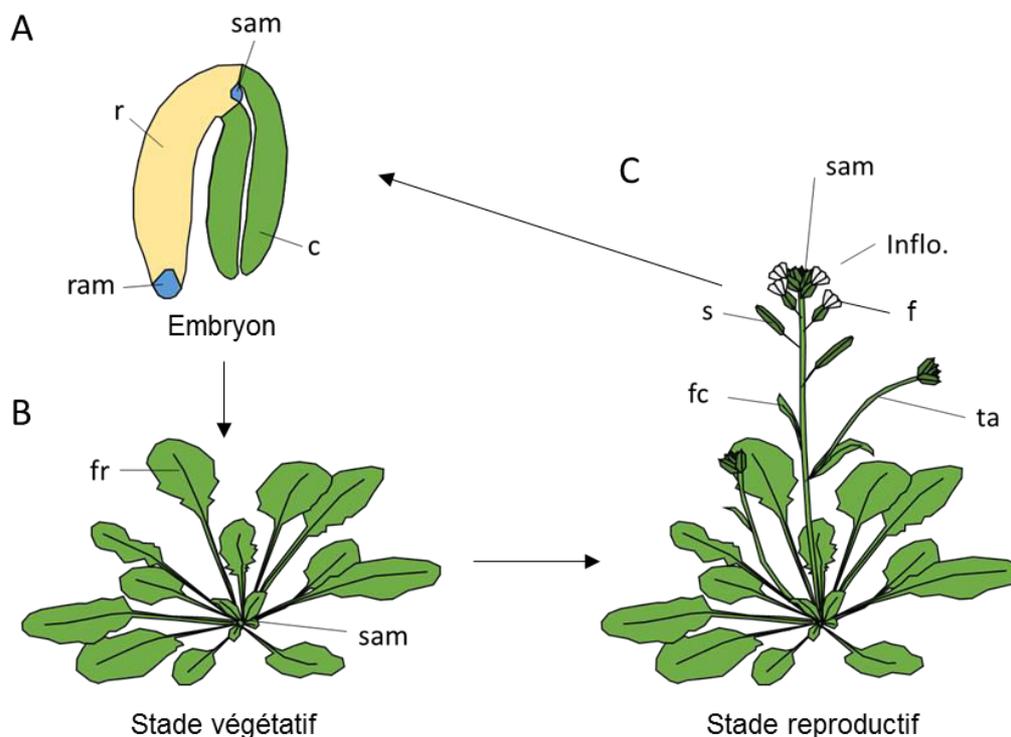


# INTRODUCTION

Les plantes, au contraire de la plupart des animaux, conservent la capacité à former de nouveaux organes tout au long de leur vie. Cela est dû au fait qu'elles maintiennent des niches de cellules souches actives dans des structures spécialisées, les méristèmes, au cours de leur développement. L'embryon contient deux méristèmes : le méristème racinaire, situé à l'extrémité de la radicule donnera naissance à la racine primaire, et le méristème apical caulinaire, niché à la base du ou des cotylédons donnera naissance à tous les organes aériens de la plante (Figure Intro. 1). Ces structures génèrent en permanence de nouvelles cellules qui, sous l'action d'éléments internes ou externes, se différencieront pour donner les différents organes spécialisés de la plante.

Une étape centrale du développement de la plante et essentielle à la survie et la dissémination de l'espèce, est la formation de structures reproductives. Chez les angiospermes ces structures sont organisées en fleurs. Lorsque les conditions sont réunies, la plante effectue une transition, de la production de tissus végétatifs (*i.e.* feuilles et boutons de tiges) vers la production d'organes reproductifs, les fleurs. Chez la plante modèle *Arabidopsis thaliana* (nommée *Arabidopsis* par la suite) cela se traduit par une profonde modification des entrecœuds, l'unité développementale de la plante. En effet, pendant la phase

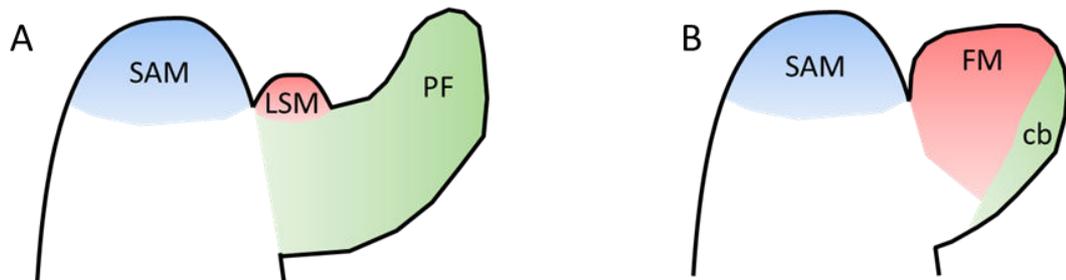
végétative, la plante produit des entrenœuds consistant en une feuille à l'aisselle de laquelle se développe, en général, un méristème caulinaire qui donnera par la suite une tige secondaire. Après la transition florale, le méristème caulinaire est converti en un méristème floral et le développement de la feuille (qu'on nomme bractée lorsqu'elle est à la base de la fleur) est inhibé (Figure Intro. 2). Alors que le méristème caulinaire perdure dans le temps grâce au maintien des cellules souches qu'il contient, le méristème floral est transitoire et les cellules souches qu'il contient ne sont pas maintenues plus de quelques jours.



**Figure Intro. 1. Architecture d'*Arabidopsis*.** (A) Schéma d'embryon d'*Arabidopsis*. Le méristème apical caulinaire (*shoot apical meristem*; sam) est situé entre les deux cotylédons (c). Le méristème apical racinaire (*root apical meristem*; ram) est situé à la base de la radicule (r). (B) Au stade végétatif le sam produit des feuilles de rosette (fr). (C) Après la transition florale, le sam produit des feuilles caulines (fc) portant des tiges axillaires (ta), puis des fleurs (f). Après fécondation, les fleurs forment des siliques (s) contenant les graines.

Au niveau moléculaire la différenciation et la croissance des organes est orchestrée par la transcription de sous-ensembles spécifiques de gènes. L'organisation dans le temps et l'espace de cette régulation transcriptionnelle est le fait des facteurs de transcription, une classe particulière de protéines permettant l'activation ou l'inhibition de la transcription de gènes *via* des

interactions avec des séquences spécifiques d'ADN. L'un des acteurs les plus importants du développement floral est le facteur de transcription LEAFY (LFY), qui régule à la fois la transition florale et l'initiation des méristèmes floraux, et orchestre l'identité de la fleur et des organes floraux.



**Figure Intro. 2. Organisation des entre-nœuds avant et après la transition florale.** Les organes latéraux sont initiés sur les flancs du SAM. (A) Au stade végétatif le méristème caulinaire latéral (*lateral shoot meristem*; LSM) est formé à l'aisselle du primordia floral (PF). Le LSM est quiescent et ne se développe en inflorescence secondaire qu'après la transition florale. (B) Après la transition florale la bractée cryptique (*cryptic bract*; cb) qui remplace la feuille est consommée par le développement du méristème floral (*floral meristem*; FM).

Dans ce travail nous discuterons de comment LFY contrôle l'initiation et la terminaison des méristèmes floraux au point de vue génétique et biochimique et comment certaines de ces propriétés peuvent lui permettre de contrôler les importantes transitions développementales nécessaires à la production de fleurs.

## I. Le méristème apical caulinaire

Chez *Arabidopsis*, tous les organes aériens à l'exception des cotylédons sont formés par les cellules dérivant des cellules souches du méristème apical caulinaire (*shoot apical meristem* SAM). Le SAM est établi lors de l'embryogénèse et perdure au cours de la vie de la plante. Comment les cellules souches sont établies et maintenues dans le SAM ? Comment des cellules sont recrutés pour l'organogénèse ? Comment est déterminé le site d'initiation d'un nouvel organe ? Ce sont les questions auxquelles j'essaierai de répondre dans cette partie.

## 1. L'homéostasie des cellules souches

Les méristèmes sont des structures très organisées dont l'architecture est maintenue au fil du temps, malgré la prolifération des cellules souches, grâce au maintien d'un équilibre entre la prolifération et la différenciation. Le maintien de cet équilibre est rendu possible par la communication entre les différentes parties fonctionnelles du méristème.

### *Organisation du méristème*

On peut distinguer quatre zones dans les méristèmes aériens d'*Arabidopsis* (Figure Intro. 3A) : la zone centrale (*central zone* CZ), qui contient les cellules souches *per se*, est confinée au centre du méristème et s'étend sur trois couches cellulaires distinctes (L1-L3). Le centre organisateur (*organizing centre* OC) est situé juste sous la CZ et contient des cellules qui se divisent très peu mais sont importantes pour le maintien du méristème. La zone périphérique (PZ) qui est située autour de la CZ et de l'OC contient les cellules quittant latéralement la CZ pour former les organes latéraux. C'est dans la PZ que se déroulent les premières étapes de l'organogénèse. Le méristème médullaire (*rib meristem* RM), situé sous l'OC, contient les cellules quittant vers le bas la CZ et l'OC et qui participeront à l'extension de la tige et la vascularisation. Au cours du développement et notamment au cours de la transition florale, la morphologie du SAM change de manière subtile. Le SAM s'élargit et s'allonge en hauteur en quelques jours après le début de la transition florale (Jacqmard et al., 2003). On peut aussi remarquer que lorsque les plantes sont artificiellement conservées dans des conditions empêchant la floraison, le SAM s'élargit drastiquement au fil du temps (Gruel et al., 2016). Malgré ces changements morphologiques les différents domaines du SAM maintiennent leur organisation respective et leurs proportions relatives par rapport au SAM sont maintenues (Gruel et al., 2016).

### *Maintien du méristème par une boucle de signalisation*

L'expression du marqueur de l'OC, le facteur de transcription (FT) WUSCHEL (WUS) est le premier signe visible de la mise en place du SAM au stade 16-cellules de l'embryon (Mayer et al., 1998), précédant l'expression du peptide

marqueur des cellules souches CLAVATA 3 (CLV3) dans la CZ (Fletcher et al., 1999) (Figure Intro. 3B). De plus l'expression de WUS dans la racine convertit celle-ci en SAM (Gallois et al., 2004), indiquant que WUS induit la mise en place du méristème aérien. Lors du développement post-embryonnaire WUS est crucial pour la maintenance du méristème et la prolifération des cellules souches, et donc l'organogénèse. Les mutants *wus* n'ont pas de SAM fonctionnels et initient rarement des feuilles après germination. Quand éventuellement des fleurs sont produites, celles-ci sont incomplètes, indiquant une réduction de la population de cellules souches dans le méristème floral (Laux et al., 1996). De plus l'augmentation de la transcription de *WUS* entraîne la prolifération des cellules souches (Lenhard et al., 2002). L'expression de CLV3 par les cellules souches est induite directement par WUS, protéine qui déplace de l'OC vers la CZ *via* les plasmodesmes (Yadav and Reddy, 2012; Daum et al., 2014). De manière opposée aux mutants *wus*, les mutants *clv3* ont un SAM excessivement large et des fleurs avec de nombreux organes additionnels (Fletcher et al., 1999). CLV3 diffuse également dans le SAM et limite la prolifération des cellules souches en réprimant WUS, formant un rétrocontrôle négatif (Brand et al., 2000).

CLV3 s'oppose à l'expression de *WUS* *via* deux récepteurs, CLV1 et CLV2 (Brand et al., 2000; Jeong et al., 1999). En accord avec cela les phénotypes des mutants *clv1*, *clv2* et *clv3* sont très similaires (Clark et al., 1993; Kayes and Clark, 1998). L'expression des récepteurs CLV dans les cellules latéralement et apicalement adjacentes à l'OC y piège CLV3 et permet de restreindre l'expression de WUS dans l'OC (Schoof et al., 2000; Lenhard and Laux, 2003). Un récepteur additionnel, CORYNE (CRN) lie CLV3 dans l'OC et limite l'expression de WUS (Müller et al., 2008). Il a longtemps été pensé que CLV1/CLV2 et CRN agissait dans des voies parallèles (Müller et al., 2008), cependant des données plus récentes suggèrent que la liaison de CLV3 déclenche la formation de complexes CLV1/CLV2/CRN rassemblés dans des sous-domaines membranaires (Somssich et al., 2015). Il n'est cependant pas clair si ces complexes ont un rôle dans la transduction du signal ou dans le recyclage des récepteurs.

WUS maintient les cellules de la CZ indifférenciées en réprimant les programmes de différenciation (Figure Intro. 3B), contrôlés notamment par les FT KANADI (KAN) ou ASYMETRIC LEAVES (AS) et en induisant la transcription du répresseur de transcription TOPLESS (TPL) (Yadav et al., 2013; Busch et al., 2010). L'analyse des transcrits de la CZ révèle un enrichissement en gènes liés à la réplication de l'ADN et la division cellulaire, reflétant la haute activité mitotique de ce tissu (Yadav et al., 2009).

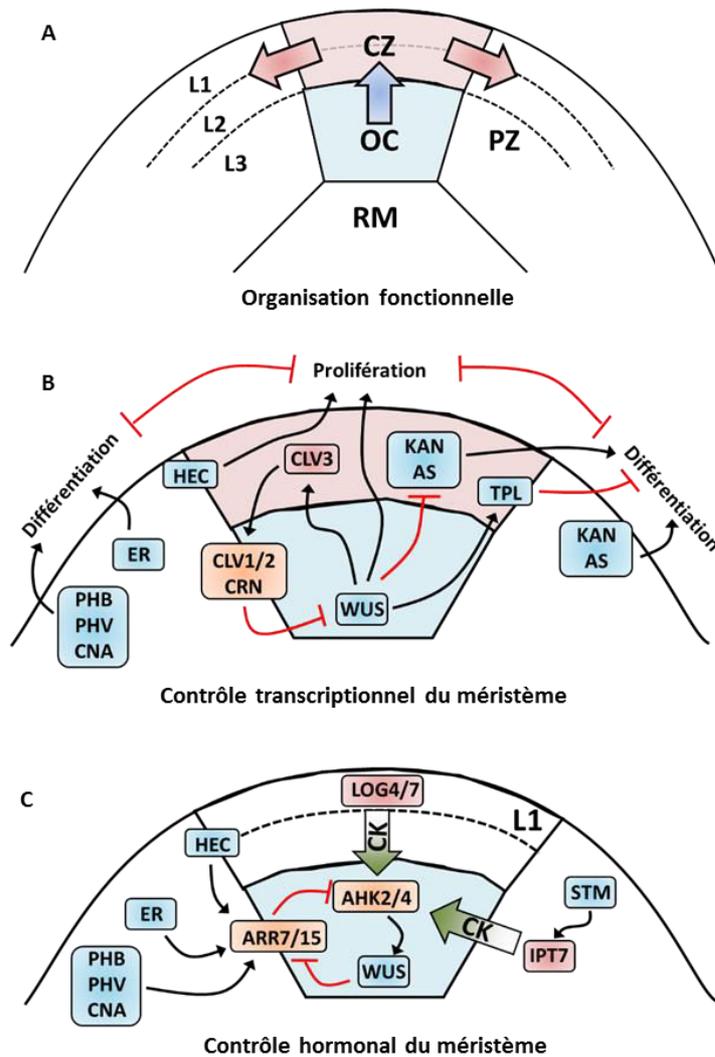
#### *Le maintien du méristème par les cytokinines*

Une fois le méristème en place, l'expression de WUS est finement régulée pour s'adapter aux changements de morphologie du SAM au cours du développement et en réponse aux conditions extérieurs.

En plus de WUS, le facteur de transcription SHOTMERISTEMLESS (STM) est nécessaire à la formation et au maintien du SAM. Les mutants forts *stm* n'initient pas de SAM et les mutants plus faibles échouent à le maintenir après la germination (Long and Barton, 1998; Long et al., 1996; Endrizzi et al., 1996). STM est exprimé dans l'ensemble du SAM. Alors que WUS régule la prolifération des cellules souches dans la CZ, STM inhibe la différenciation des cellules du SAM. Ces fonctions sont contrôlées indépendamment : STM ne requiert pas la présence de WUS pour inhiber la différenciation cellulaire, de même que WUS ne requiert pas la présence de STM pour stimuler la prolifération des cellules méristématique. En revanche, leur expression simultanée stimule drastiquement la prolifération des cellules souches, indiquant qu'ils agissent de manière synergique sur l'activité du méristème (Lenhard et al., 2002). STM agit en activant la synthèse des cytokinines (CK) *via* l'activation de la transcription de l'enzyme de biosynthèse ISOPENTENYLTRANSFERASE7 (IPT7) et *via* la répression de la synthèse de gibbérellines (GA) en inhibant la transcription de l'enzyme de biosynthèse GA20ox (Jasinski et al., 2005; Yanai et al., 2005) (Figure Intro. 3C). Alors que les GA induisent la différenciation des tissus (Ogas et al., 1997), les CK induisent la prolifération des cellules en stimulant la progression du cycle cellulaire *via* la régulation des Cyclines D (Riou-Khamlichi et al., 1999; Lenhard et al., 2002). Dans le SAM, les CK sont perçues par les récepteurs

ARABIDOPSIS HISTIDINE KINASE2/4 (AHK2/4) qui activent et répriment dans l'OC l'expression de WUS et CLV1, respectivement (Gordon et al., 2009; Chickarmane et al., 2012). De plus, WUS sensibilise les cellules de l'OC aux CK en réprimant les inhibiteurs de réponse aux CK ARABIDOPSIS RESPONSE REGULATORS de type A (A-ARR) (Leibfried et al., 2005; Zhao et al., 2010), formant une boucle d'auto-activation (Figure Intro. 3C).

Il est également prédit par des modélisations que le maintien d'un rapport de taille constant entre le SAM et l'OC est le résultat d'un gradient de CK entre les couches externes et l'intérieure du SAM (Adibi et al., 2016; Gruel et al., 2016; Chickarmane et al., 2012). Ce gradient résulte probablement de l'activité des enzymes de synthèse des CK LONELYGUY4/7 (LOG4/7) dans la L1 et de la diffusion des CK dans le tissu (Kuroha et al., 2009; Chickarmane et al., 2012; Gruel et al., 2016). Ce modèle repose en partie sur l'existence d'un signal encore non-identifié qui diffuserait de la L1 à moins grande distance que les CK et restreindrait l'expression des d'AHK2/4 au centre du SAM (Adibi et al., 2016; Gruel et al., 2016).



**Figure Intro. 3. Organisation et maintenance du méristème.** (A) Organisation fonctionnelle du méristème. Le centre organisateur (*organizing centre*; OC; bleu) organise l'architecture du méristème en induisant la prolifération des cellules-souches dans la zone centrale (*central zone*; CZ; rouge). Les cellules quittant la CZ vers la zone périphérique (*peripheral zone*; PZ) peuvent se différencier pour former les organes axillaires. Les cellules quittant l'OC par le bas vers le méristème médullaire (*rib meristem*; RM) formeront la vasculature. (B) Contrôle transcriptionnel du méristème. WUS est exprimé dans l'OC et diffuse vers la CZ où il induit la prolifération des cellules souches en partie en induisant TPL et en réprimant KAN et AS. De plus WUS induit l'expression du peptide CLV3 qui diffuse dans le SAM. CLV3 est perçu par les récepteurs CLV1/2/CRN en bordure de l'OC pour réprimer l'expression de WUS. De plus HEC induit la prolifération des cellules souches et PHB/PHV/CNA/ER induisent la différenciation. (C) Contrôle hormonal du SAM. La synthèse des CK par LOG4/7 dans la L1 et par IPT7, sous contrôle de STM, crée un gradient de CK dans le SAM. La forte concentration en CK dans l'OC induit l'activité des AHK2/4 stimulant l'expression de WUS. WUS inhibe les régulateurs négatifs de la transduction du signal CK ARR7/15. HEC active la transcription d'ARR7/15 ce qui limite l'étendue de l'OC et PHB/PHV/CNA/ER induisent ARR7/15 facilitant ainsi la différenciation

### *Des voies indépendantes de WUS régulent l'activité du méristème*

En plus de la boucle de régulation canonique CLV/WUS, l'activité du méristème est contrôlée par d'autres facteurs plus récemment identifiés (Figure Intro. 3B). Le facteur de transcription HECATE (HEC) est directement réprimé par WUS dans l'OC. L'expression de *HEC* dans l'OC entraîne l'arrêt du méristème. En revanche son expression dans la CZ entraîne la prolifération des cellules souches, indépendamment de WUS et CLV3 (Schuster et al., 2014). HEC et WUS partagent un même sous-ensemble de cibles, qu'ils régulent différemment. HEC induit notamment la progression du cycle cellulaire et la transcription des *ARR7/15*. HEC jouerait donc à la fois sur la prolifération des cellules souches et sur la restriction de la taille du méristème (Schuster et al., 2014).

Au moins deux autres voies de régulation de la morphologie du méristème existent en parallèle de la boucle CLV/WUS. Les mutants *clv3*, *erecta* (*er*) et *jabba* (*jba1D* ; allèle dominant) ont des interactions synergiques sur la taille du méristème (Mandel et al., 2016; Lee and Clark, 2015). JBA contrôle la transcription d'un microARN responsable de la dégradation des ARNm des facteurs de transcription de type HD-ZIPIII PHABULOSA (PHB), PHAVOLUTA (PHV) et CORONA (CNA) (Williams et al., 2005). PHB, PHV et CNA sont des inhibiteurs de l'activité du méristème dans différents tissus, incluant l'ovule et la fleur (Yamada et al., 2015; Landau et al., 2015). Comme les quadruples mutants *wus phb phv cna* montrent des signes d'activité méristématique, il est possible que PHB, PHV et CNA répriment des voies indépendantes de WUS et CLV3 dans le SAM (Lee and Clark, 2015). Si CLV3, ER et les HD-ZIPIII agissent en parallèles pour contrôler la morphologie du SAM, ils en contrôlent des aspects distincts. Les mutants *clv3* présentent un agrandissement du SAM en hauteur alors que le mutants *jba1D* présente un SAM extrêmement allongé latéralement. L'addition de la mutation *er* à *jba1D* entraîne l'épaississement du SAM dans la direction perpendiculaire à l'allongement des *jba1D* (Mandel et al., 2016). Cela suggère que les HD-ZIPIII, contrôlés par JBA d'un côté, et ER de l'autre, régulent l'expansion du SAM latéralement dans des directions particulières. Une analyse des transcrits de ces mutants révèle que ER et les HD-ZIPIII répriment l'activité mitotique et la transduction du signal CK (*via*

l'induction des ARR7/15) dans la PZ (Mandel et al., 2016). Les HD-ZIPIII et ER pourraient donc jouer un rôle dans la 'dé-méristématisation' des cellules de la PZ, permettant leur différenciation.

- L'organisation du méristème est contrôlée par la boucle régulatrice CLV/WUS
- Les cytokinines jouent un rôle essentiel dans le maintien du méristème
- D'autres facteurs régulent la balance prolifération/différenciation indépendamment de WUS

## 2. La phyllotaxie

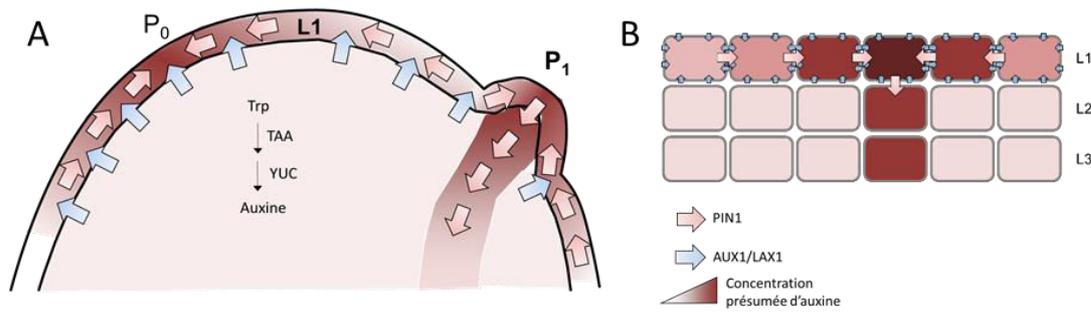
Au fur et à mesure de la division des cellules souches, leur progéniture est repoussée latéralement dans la PZ. Alors que la différenciation est inhibée dans les cellules souches de la CZ, c'est dans PZ que va s'initier l'organogénèse. Cependant toutes les cellules ne sont pas recrutées pour l'organogénèse et les nouveaux organes ne sont pas formés continuellement par la PZ mais s'initient un par un, à partir d'un petit groupe de cellules fondatrices, selon un motif régulier qui conditionne la phyllotaxie, c'est à dire la disposition des feuilles et par extension des fleurs sur une tige. Plusieurs mutant produisant une inflorescence nue après transition florale ont été isolés. Ces mutants sont majoritairement impliqués dans la synthèse, le transport ou la signalisation de l'hormone auxine (Okada et al., 1991; Christensen et al., 2000; Cheng et al., 2006; Przemeck et al., 1996), indiquant son rôle central dans l'initiation de nouveaux organes. La question du rôle de l'auxine dans la phyllotaxie et l'initiation de nouveaux organes a été abondamment discutée dans les dernières années (Armengot et al., 2016; Taylor-Teeples et al., 2016; Sassi and Vernoux, 2013; Vernoux et al., 2010). Je me concentrerai donc ici sur certains aspects importants de cette question dans le contexte de l'initiation des fleurs.

### *Le transport de l'auxine contrôle l'organogénèse*

L'initiation des fleurs repose fortement sur le transport polaire de l'auxine. L'utilisation de l'inhibiteur du transport de l'auxine naptalame (NPA) ou la

mutation des transporteurs d'efflux PIN-FORMED1 (PIN1) abolit presque totalement la production de fleurs (Okada et al., 1991). La perte de PIN1 peut être compensée par l'application locale d'auxine, entraînant la formation d'un nouvel organe au site d'application (Reinhardt et al., 2000, 2003). Cela suggère que la concentration locale d'auxine, plutôt que sa simple présence, est un facteur déterminant de l'organogénèse. L'étude de la répartition du transporteur PIN1, à l'échelle cellulaire montre que PIN1 a une localisation polarisée dans la cellule (Reinhardt et al., 2003; Heisler et al., 2005). De plus, les transporteurs, qui sont localisés principalement dans la L1, s'organisent à l'échelle tissulaire afin de faire converger l'auxine à travers la L1 vers le site d'initiation d'un nouvel organe (Figure Intro. 4A) (Heisler et al., 2005; Kierzkowski et al., 2013). S'il est encore impossible de visualiser les flux d'auxine dans le SAM, cette observation est confirmée par l'expression du marqueur de réponse à l'auxine DR5 aux sites de convergence des PIN1 (Heisler et al., 2005).

Bien que le transport de l'auxine à travers la L1 soit primordial pour l'organogénèse, sa synthèse a lieu de manière moins localisée dans l'ensemble du SAM (Cheng et al., 2006). La synthèse de l'auxine s'effectue à partir du tryptophane par les enzymes TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS (TAA) et YUCCA (YUC) (Won et al., 2011; Mashiguchi et al., 2011; Zhao et al., 2001). La synthèse de l'auxine semble être aussi importante que son transport pour l'initiation des fleurs car la combinaison de mutants dans les gènes *YUC* exprimés dans le SAM entraîne une réduction dramatique de la production d'organes et parfois une inflorescence nue (Cheng et al., 2006, 2007a). Il a été proposé que l'auxine soit concentrée dans la L1 par les transporteurs d'influx AUXIN RESISTANT1 (AUX1) et Like-AUX1 (LAX1). Cette hypothèse est étayée par le fait que les AUX1/LAX1 sont principalement exprimés dans la L1, sont distribuées de manière homogène à la membrane des cellules et sont aussi impliqués dans la régulation de la phyllotaxie (Stieger et al., 2002; Reinhardt et al., 2003; Kierzkowski et al., 2013). Les AUX1/LAX1 pourraient alors activement pomper l'auxine vers la L1 et l'y piéger (Figure Intro. 4A-B).



**Figure Intro. 4. La création de maxima d'auxine contrôle la phyllotaxie.** (A) Représentation schématique des flux d'auxines dans l'inflorescence. La synthèse de l'auxine à partir du tryptophane (Trp) a lieu dans le SAM et nécessite l'activité des enzymes des familles TAA et YUC. L'auxine est pompée vers la L1 par les transporteurs d'influx AUX1/LAX1. Les maxima d'auxine sont créés par la convergence des flux induits par les transporteurs d'efflux PIN1 dans la L1 en direction du site d'initiation de l'organe (P<sub>0</sub>). Une fois l'organe initié (P<sub>1</sub>) les transporteurs PIN1 des cellules adaxiales ("au-dessus") à l'organe changent de polarisation pour créer un flux d'auxine vers le site d'initiation du prochain organe. Le maximum d'auxine entraîne également la relocalisation basale d'une partie des transporteurs PIN1, permettant la redirection du flux d'auxine à l'intérieur du tissu. (B) La polarisation des transporteurs PIN1 dans la L1 est orientée par la concentration relative en auxine des cellules adjacentes, de manière à transporter l'auxine à l'encontre du gradient. Lorsqu'une quantité critique d'auxine est atteinte PIN1 se relocalise à la membrane basale de la cellule, ce qui permet l'export de l'auxine vers les couches internes du méristème.

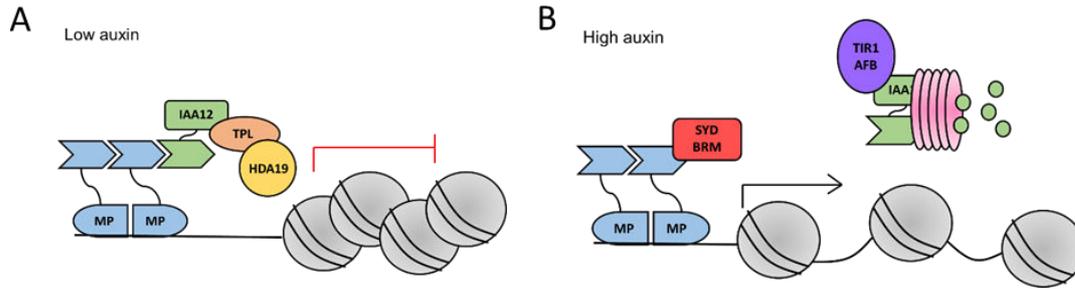
Ces observations ont été intégrées dans plusieurs modèles du transport de l'auxine dans le SAM (Jönsson et al., 2006; de Reuille et al., 2006; Smith et al., 2006). D'après ces modèles, la polarisation des transporteurs PIN1 serait contrôlée par la concentration relative des cellules adjacentes, les PIN se localisant à la membrane adjacente à la cellule avec la plus grande concentration (Figure Intro. 4B). L'implémentation de cette règle combinée à la synthèse d'auxine et sa rétention dans la L1 permet de reproduire la localisation de PIN1 dans un SAM *in silico* (Jönsson et al., 2006; de Reuille et al., 2006; Smith et al., 2006). Ce système implique néanmoins une grande dynamique de la localisation des PIN1 dans la cellule. En effet après la création d'un maximum d'auxine, la polarité des PIN s'inverse très rapidement dans les cellules adaxiales ("au-dessus") au primordia afin de se réorienter vers le nouveau maximum (le futur primordium) (Figure Intro. 4A) (Heisler et al., 2005; Jönsson et al., 2006; Smith et al., 2006).

La polarisation de PIN1 est contrôlée par des modifications post-traductionnelles, régulées notamment par la kinase PINOID (PID) (Michniewicz et al., 2007), exprimée dans l'inflorescence (Christensen et al., 2000). En accord avec cela, les mutants *pid* sont extrêmement similaires aux

mutants *pin* ou aux plantes traités au NPA (Bennett et al., 1995). De la même manière, les plantes exprimant *PID* constitutivement présentent des phénotypes suggérant une perturbation du transport polarisé de l'auxine (Christensen et al., 2000; Benjamins et al., 2001). Il a été montré que *PID* induit la localisation apicale de *PIN1* dans les cellules de l'épiderme de l'inflorescence, alors que chez les mutants *pid*, *PIN1* est localisé au pôle basal des cellules (Friml et al., 2004). La polarisation de *PIN1* est également régulée par la kinase *PP2A*, qui agit de manière antagoniste à *PID* (Michniewicz et al., 2007). La position de *PIN1* dans la cellule est extrêmement dynamique. *PIN1* est constamment internalisé puis recyclé vers la membrane (Geldner et al., 2001; Dhonukshe et al., 2007). La perturbation de l'endocytose dépendante des clathrines provoque l'accumulation de *PIN1* dans les membranes et la dépolarisation de sa localisation dans l'embryon (Kitakura et al., 2011). Ces résultats suggèrent que les kinases et phosphatases contrôlant la polarisation de *PIN* régulent son recyclage à une membrane particulière plutôt que son internalisation, permettant une relocalisation rapide de *PIN* en fonction des flux d'auxines.

#### *La transduction du signal auxinique*

La perception nucléaire de l'auxine requiert une machinerie complexe et implique un mécanisme de 'dégradation de répresseur' (Figure Intro. 5). En présence de niveaux d'auxine bas, les protéines de la famille Auxin/INDOL-3-ACETIC ACID (Aux/IAA) sont liées aux facteurs de transcription AUXIN RESPONSE FACTOR (ARF) et recrutent le corépresseur TPL pour inhiber la transcription des gènes de réponse à l'auxine (Tiwari et al., 2001; Overvoorde et al., 2005; Okushima et al., 2005; Szemenyei et al., 2008; Long et al., 2006; Weijers et al., 2005; Ulmasov et al., 1997). L'ajout d'auxine entraîne l'adressage des Aux/IAA au protéasome par un complexe *ubiquitin protein ligase* SCF<sup>TIR1/AFB</sup> (Gray et al., 2001, 1999). Les ARFs ainsi libérés de la répression par TPL entraînent la transcription des gènes de réponse à l'auxine.



**Figure Intro. 5. Activation des gènes de réponse à l'auxine.** (A) MP (bleu) lie l'ADN en dimère. En présence de peu d'auxine, IAA12/BDL (vert) lie MP et recrute le corépresseur TPL (orange) qui permet le maintien à l'état réprimé des gènes de réponses à l'auxine. TPL agit en partie en recrutant indirectement HDA19 (jaune) qui modifie localement la chromatine et empêche ainsi la transcription. (B) La présence d'auxine déstabilise l'interaction MP/IAA12. IAA12 est alors recruté par le complexe TIR1/AFB qui l'adresse pour dégradation au protéasome. MP recrute les remodelleurs chromatinien SYD et BRM, qui augmentent l'accessibilité de la chromatine, permettant la transcription des gènes de réponses à l'auxine.

Il existe chez *Arabidopsis* 23 ARFs (Okushima et al., 2005), 29 Aux/IAA (Overvoorde et al., 2005) et 6 AFBs (AUXIN SIGNALING F-BOX) (Dharmasiri et al., 2005). Dans le cadre de l'initiation du méristème, les principaux acteurs sont ARF5/MONOPTEROS (MP) et IAA12/BODENLOS (BDL). MP est important au stade embryonnaire pour la mise en place des méristèmes apical et racinaire, puis après la transition florale pour l'initiation des méristèmes floraux, les mutants hypomorphes *mp-S319* ayant un phénotype semblable à celui des mutants *pin* après la transition florale (Cole et al., 2009). MP interagit directement avec BDL et les mutants insensibles à l'auxine *bdl* ont un phénotype embryonnaire semblable à celui des mutants *mp* (Hamann et al., 1999, 2002).

L'importance de la répression par TPL des cibles de MP a été démontrée de manière spectaculaire par la transformation de la tige en racine chez les mutants dominants *tpl-1* (Long et al., 2006). Le recrutement de TPL par les Aux/IAAs aux sites de liaisons de MP sur l'ADN permet le recrutement de l'HISTONE DEACETYLASE19 (HDA19) qui entraîne la stabilisation de la répression des gènes cibles par des modifications chromatinien (Wu et al., 2015; Krogan et al., 2012). L'auxine permet de lever la répression par TPL ce qui permet le recrutement des remodelleurs chromatinien SPLAYED (SYD) et BRAHMA (BRM) par MP (Wu et al., 2015). SYD and BRM augmentent l'accessibilité des loci liés par MP à la transcription, et donc l'activation des

gènes de réponse à l'auxine (Wu et al., 2015). De plus, il a été montré que MP et les complexes MP-BDL sont capables de former des oligomères sur l'ADN peut-être pour lier de multiples sites, et ainsi former des boucles d'ADN pour l'activation et/ou la répression des gènes de réponse à l'auxine (Nanao et al., 2014).

TPL est induit par WUS (Busch et al., 2010; Yadav et al., 2013) dans le méristème, permettant probablement de maintenir réprimés les programmes de différenciation dans les cellules quittant la CZ. Le maximum d'auxine créé par convergence des PIN1 permettrait alors l'activation ponctuelle par MP de ces programmes dans les cellules fondatrices. Cependant, il est intéressant de noter que la fusion directe du domaine de liaison à l'ADN de MP au domaine recrutant SYD et BRM, permettant donc de court-circuiter la répression par les Aux/IAA, ne semble pas perturber profondément la phyllotaxie (Wu et al., 2015). Il existe donc probablement des couches régulatrices supplémentaires, indépendantes de MP, régulant l'organisation de l'organogénèse.

#### *La création d'un champ d'inhibition permet de stabiliser de la phyllotaxie*

Après la création d'un maximum d'auxine, déclenchant la réponse transcriptionnelle contrôlée par MP, le flux d'auxine est canalisé vers l'intérieur des tissus depuis la pointe du primordium, permettant d'initier la vascularisation. Cela est caractérisé par la localisation basale du transporteur PIN1 dans la L1 et sa présence dans les tissus proto-vasculaires (Reinhardt et al., 2003; Heisler et al., 2005; Furutani et al., 2014; Bayer et al., 2009). Cette relocalisation de PIN1 est contrôlée par la famille de protéines MACCHI-BOU4 (MAB4), induites directement par MP (Furutani et al., 2014). Les mutants dans les gènes de la famille *MAB4* ou une combinaison de mutants *mab4 yuc1 yuc4* présentent des défauts d'initiation de primordia floraux (Cheng et al., 2007b; Furutani et al., 2014).

La création d'un puit d'auxine permet d'une part la vascularisation du primordium, et d'autre part, par la canalisation de l'auxine vers l'intérieur du tissu, la création d'une zone pauvre en auxine autour du primordium, prévenant la formation d'un nouveau maximum (Smith et al., 2006; de Reuille et al., 2006; Stoma et al., 2008).

Les CK apportent un niveau de régulation supplémentaire de la phyllotaxie. Le régulateur négatif de transduction du signal CK, ARABIDOPSIS HISTIDINE PHOSPHOTRANSFER PROTEIN6 (AHP6), est nécessaire pour empêcher l'initiation simultanée d'organes au niveau du SAM (Besnard et al., 2014a). AHP6 est directement induit dans les primordia par MP et diffuse dans le SAM, inhibant la signalisation des CK (Besnard et al., 2014b, 2014a). Or il a été montré que les CK permettent l'activation ectopique des réponses auxiniques en complément d'un traitement à l'auxine (Besnard et al., 2014b), indiquant une synergie dans la signalisation de l'auxine et des CK. Les organes seraient donc initiés en réponse à l'auxine aux sites où AHP6 est minimal (*i.e.* où la signalisation CK est maximale). Cela révèle un double rôle de MP dans la régulation des CK pour le contrôle de la phyllotaxie : d'un côté MP active la formation des organes en inhibant localement ARR7/15, de l'autre, MP inhibe la formation de primordia dans le voisinage en induisant l'expression d'AHP6 qui diffuse.

- La phyllotaxie dépend de la création de maxima locaux d'auxine
- L'auxine est localement concentrée par la polarisation des transporteurs PIN1
- Le signal auxinique est intégré par MP au niveau transcriptionnel
- L'interaction auxine/CK permet de stabiliser la phyllotaxie

## II. Le développement floral

Après la spécification d'un groupe de cellules comme cellules fondatrices, le premier signe visible de l'organogénèse est la formation d'une excroissance sphérique (stade 1) qui se détache progressivement du SAM (stade 2). A ce stade cette structure se met à exprimer des marqueurs du méristème, indiquant qu'elle acquiert *de novo* une identité méristématique. La formation d'un méristème floral est essentielle pour maintenir la prolifération cellulaire nécessaire à l'organogénèse florale. Cette identité méristématique perdure jusqu'au stade 6, le bouton floral acquiert alors une croissance déterminée et le méristème est consommé dans l'organogénèse. Comment le méristème est-il établi au sein de la jeune structure florale ? Comment sont spécifiées

l'identité de la fleur puis celle des différentes couronnes d'organes ? Comment l'identité méristématique s'éteint-elle ? Ce sont les questions auxquelles j'essaierai de répondre dans cette partie.

### 1. Article 1 : 'A Flower is Born : Recent Progress on Arabidopsis Flower Meristem Formation'

Une fois le site d'initiation d'un organe déterminé, MP entraîne la transcription de nombreux gènes permettant la formation d'une fleur. De nombreuses avancées récentes dans l'étude des gènes régulés par MP et leurs effets sur le développement floral nous permettent de formuler quelques hypothèses sur le contrôle de l'initiation du méristème floral. Ces avancées et les hypothèses qu'elles ont engendrées sont présentées dans l'article 1 (Denay et al., 2017).

# A flower is born: an update on Arabidopsis floral meristem formation

Grégoire Denay<sup>1</sup>, Hicham Chahtane<sup>2</sup>, Gabrielle Tichtinsky<sup>1</sup> and François Parcy<sup>1\*</sup>

<sup>1</sup>Laboratory of Plant & Cell Physiology, CNRS, CEA, Univ. Grenoble Alpes, INRA, 38000 Grenoble, France, <sup>2</sup>Department of Plant Biology and Institute for Genetics and Genomics in Geneva (iGE3), University of Geneva, Geneva, Switzerland

\* Author for correspondence. email: [francois.parcy@cea.fr](mailto:francois.parcy@cea.fr)

Accepted in the 2017 Growth and Development issue, Current Opinion in Plant Biology

## Highlights

- Flower development requires *de novo* formation of a stem cell niche
- Auxin and cytokinin signaling interact to trigger emergence of floral meristems
- Meristem growth and floral fate determination are intimately coupled
- Several key floral regulators share unexpected structural features

## Abstract

In Arabidopsis, floral meristems appear on the flanks of the inflorescence meristem. Their stereotypic development, ultimately producing the four whorls of floral organs, is essentially controlled by a network coordinating growth and cell-fate determination. This network integrates hormonal signals, transcriptional regulators, and mechanical constraints. Mechanisms regulating floral meristem formation have been studied at many different scales, from protein structure to tissue modeling. In this paper, we review recent findings related to the emergence of the floral meristem and floral fate determination and examine how this field has been impacted by recent technological developments.

## Introduction

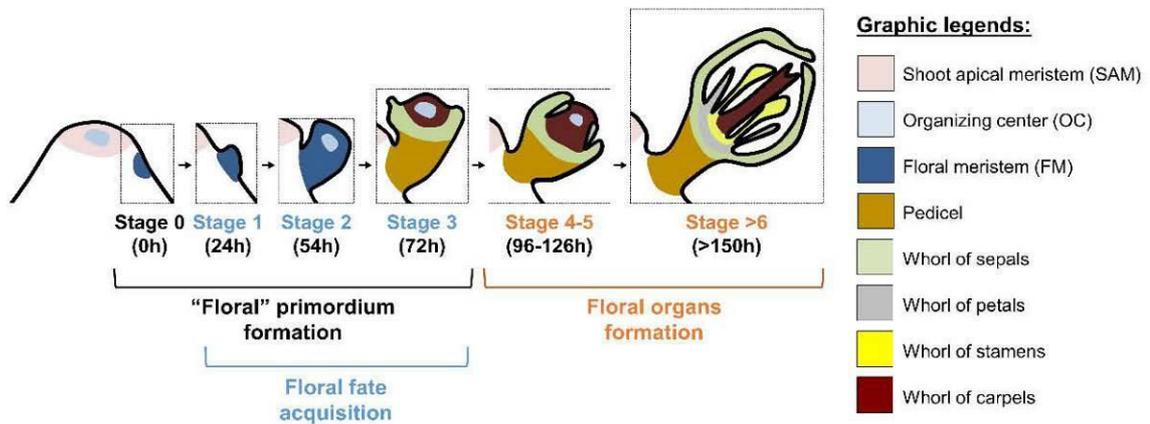
Since the first genes controlling flower development in Arabidopsis and snapdragon were cloned, enormous progress has been accomplished and regularly reviewed [1,2]. For this review, we chose to focus on the early stages of floral meristem (FM) development in Arabidopsis, highlighting recent advances in the field. We also examine how new technologies and methods have improved our understanding of floral development.

### Emergence of Floral Primordia

In Arabidopsis, the emergence of flowers on the flanks of the shoot apical meristem (SAM) follows a precise pattern (Figure 1). The reason for this positioning of flower primordia (phyllotaxis) can largely be explained by a localized accumulation of auxin due to its polarized transport in the L1 layer through the PIN-FORMED1 (PIN1) efflux carrier [3]. This transient auxin peak releases the AUXIN RESPONSE FACTOR5/MONOPTEROS (ARF5/MP) from Aux/IAA repression [4], which is essential for FM emergence as attested by the pin-shaped inflorescence of the *mp-S319* mutant [5] (Figure 2). The MP protein acts at several levels: it first upregulates the *MACCHI-BOU 4 (MAB4)* family genes that control PIN basipetal relocalization [6]; it also induces *ARABIDOPSIS HISTIDINE PHOSPHOTRANSFER PROTEIN6 (AHP6)*, a negative regulator of cytokinin (CK) signaling. AHP6

diffuses to neighboring initiation positions where it inhibits meristem initiation, thus reinforcing the auxin phyllotactic pattern [7,8]. Finally, MP directly induces the expression of the FM identity gene, *LEAFY (LFY)* [5], which acts in concert with *AINTEGUMENTA (ANT)*, *AINTEGUMENTA-LIKE6 (AIL6)* and *FILAMENTOUS FLOWER (FIL)* in this process [5,9]. Feedback and feed-forward loops come into action as ANT and AIL6 both induce LFY, while LFY in turn reinforces auxin transport [10–12]. LFY also activates *REGULATOR OF AXILLARY MERISTEMS1 (RAX1* - a Myb-like transcription factor (TF)), which, as its name suggests, regulates meristem formation ([13] and references therein) but also cytokinin signaling through inhibition of *ARABIDOPSIS RESPONSE REGULATOR7 (ARR7)* [13].

At the molecular level, MP was shown to recruit the SWI/SNF ATPases SPLAYED (SYD) and BRAHMA (BRM) to target loci where they increase chromatin accessibility [9]. Interestingly, constitutive expression of a fusion protein composed of the MP DNA-binding domain and BUSHI (a member of the SWI/SNF ATPase complex that should allow the recruitment of SYD and BRM) is sufficient to restore a wild-type inflorescence architecture in *mp* mutants. The fusion protein is constitutively expressed and lacks the PB1 domain mediating Aux/IAA repression, therefore this result suggests the existence of an additional pathway which can pattern the SAM to determine the site of floral emergence [9]. This pathway



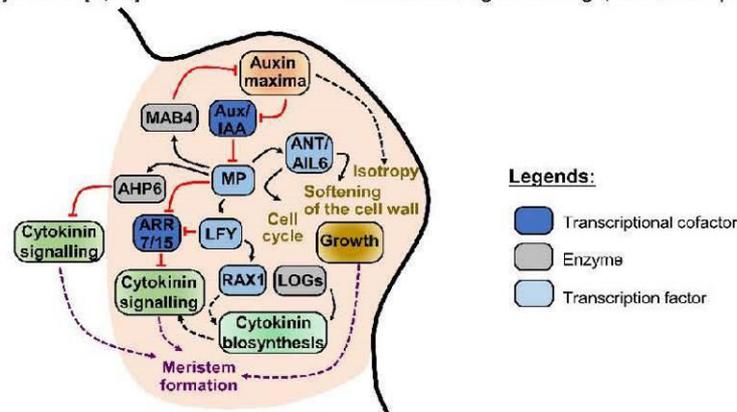
**Figure 1. Schematic representation of flower development in *Arabidopsis*.** The floral bud emerges on the flanks of the shoot apical meristem (stage 0). Bud outgrowth during stages 1 to 3 allows a meristem-organizing center to be established at stage 2. During these early stages floral identity is acquired, leading to definition of the floral organ whorls, starting with sepal outgrowth at stage 3. The meristem-organizing center finally disappears at stage 6, once all floral organs have been initiated and the pedicel has begun to expand. Hours indicate approximately the age of the flower at the end of a given stage as defined in ref. [63].

has, as yet, remained elusive, and no other information on the components involved is available. After reaching its peak level, auxin is channeled away from the primordium. This is achieved by PIN1 relocation to the internal pole of the cell combined with the local action of auxin-influx carriers. This combination effectively creates an auxin sink in the inner tissues that contributes to vascularization and phyllotaxis [6,14].

outgrowth, as indicated by their localized expression patterns in the inflorescence [18]. Moreover, it has also been shown that auxin can alter tissue plasticity through its effect on microtubule reorganization and dynamics, controlling both the stiffness and isotropy of cell walls [19].

#### CK Establishes FM Meristematic Competence

After an initial growth stage, the FM acquires meristematic



**Figure 2. Genetic network involved in floral meristem formation.** In the nascent floral bud, the size of the primordium and the presence of a cytokinin-induced signal combine to control the emergence of a stem-cell niche. This process is largely initiated by the auxin response factor, MP, which is activated in response to a local auxin maximum after the release of Aux/IAA inhibition. Induction of MAB4 triggers relocation of the PIN1 auxin efflux carrier to channel auxin away from the primordium. By inducing ANT and AIL6, MP stimulates the cell cycle and cell wall softening to allow meristem growth. Plasticity is reinforced by the effect of auxin on cell wall stiffness and isotropy. Cytokinin biosynthesis is triggered by LOG genes, and probably by the action of MP through LFY and RAX1. MP and LFY also stimulate cytokinin signalling by inhibiting ARR7/15. Finally, MP induces AHP6 which diffuses from the primordium to inhibit cytokinin signalling in surrounding areas, and thus control the emergence of new primordia.

#### Floral Bud Outgrowth

Once its position has been established, the floral meristem enters its growth phase. This involves the ANT and AIL6 [15,16] proteins, cell cycle regulators which also control pectin dimethyl-esterification of SAM cell walls [16], a process which is thought to facilitate organ outgrowth [17]. Cell-wall-forming enzymes play a specific role in floral

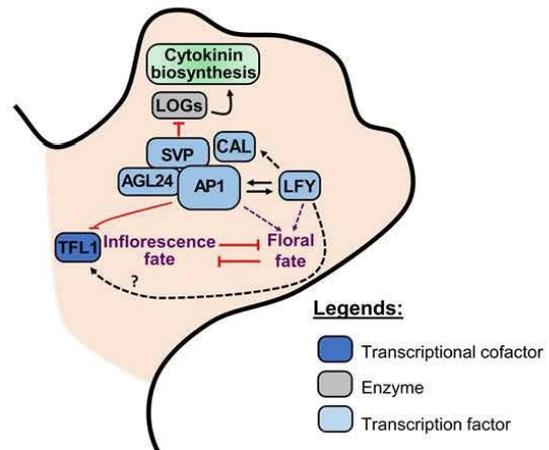
features (Figure 2). Around stage 2, activation of the CK signaling reporter pTCS [20] is closely followed by the expression of *WUSCHEL* (*WUS*) and *CLAVATA3* (*CLV3*) which mark the establishment of an organizing center and a stem cell niche, respectively. Inter-regulations of *WUS* and *CLV3*, CK and its receptors, and a hypothetical diffusible signal from the L1, have been integrated in a model recapitulating gene expression and stem cell

organization in the SAM in various conditions [21,22]. Remarkably, this model can also explain how *WUS* and *CLV3* expression are triggered in the early flower, as soon as the primordium has attained a size threshold [22]. The central role played by CK in this process is illustrated by a flurry of recent results: (i) MP represses the negative regulators of the CK response, *ARR7* and *15* [23], while *LFY* represses *ARR7* [13], thereby locally increasing the CK response in FMs. In support of this, *ARR7/15* inhibition by amiRNA was also shown to compensate for the loss of primordia initiation in *mp* mutants [23]. (ii) *RAX1* may also act on CK levels in the SAM as overexpression of a CK biosynthesis enzyme can compensate for loss of axillary meristems in *rax1* mutants [24]. (iii) As described in the SAM network model mentioned above, cytokinin biosynthesis is thought to occur mainly in the L1, as suggested by the expression of the *LONELY GUY4* and *7 (LOG4/7)* reporters [22]. Although few data are available regarding their involvement in flower formation and their regulation, *LOG7* is specifically expressed in FM, while *LOG4* is found throughout the inflorescence, strongly hinting at a role in primordia development [22]. In addition to the CK receptor *AHK4* [21,22,25], the *HECATE* family (bHLH TFs) [26] and *REVOLUTA* (a HD-ZIP III TF) [27,28] were also shown to stimulate meristem activity and are expressed in young flower primordia.

#### Floral Fate Acquisition

Acquisition of the floral fate is determined by a gene regulatory network (Figure 3) based on the central players *LEAFY*, the MADS-domain TF *APETALA1 (AP1)*, *CAULIFLOWER (CAL)*, *AGAMOUS-LIKE 24 (AGL24)*, and *SHORT VEGETATIVE PHASE (SVP)*. Other MADS-domain TFs - *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)*, *FRUITFULL (FUL)*, and *XAANTAL2/AGL14 (XAL2)* - might be involved or at least can compensate for *AGL24* and *SVP* function in various mutant backgrounds [29,30]. *AP1* and *LFY* positively regulate each other and share partially overlapping functions. Genome-wide analyses revealed that these TFs act as activators and repressors [31–33]. *AP1* for example activates *LFY* but represses a myriad of other genes expressed in the SAM and in early floral stages [32]. The analysis of common *AP1* and *LFY* targets identified multiple genes (including *EUI-LIKE P450 A1 (ELA1)*), genes involved in peptide signaling, nitrate transport, UV-B signaling and carbohydrate homeostasis) paving the way for future functional analysis [31,34]. Several gene network models describe the regulatory relationships and predict steady states corresponding to organ-identity gene expression in all four floral whorls [30,35,36]. The mechanistic basis for whorl-specific activation of ABC floral organ homeotic genes through the combined actions of TF and their cofactors remains unclear and this field has progressed relatively slowly. Recent data suggest that ubiquitination could be involved as *SVP* upregulates the *WDR55 DCAP* gene, which is involved in restricting the expression pattern of *AG* and potentially targets proteins

or ubiquitination [37].



**Figure 3. Genetic network involved in floral identity.**

Inflorescence and floral identity are mutually exclusive statuses in the nascent meristem. In the young floral meristem, expression of *LFY* induces the *AP1* and indirectly *CAL* MADS-box genes. These three genes work together with other MADS-box genes, *SVP* and *AGL24*, to specify the floral identity. A feed-forward regulatory loop between *LFY* and *AP1* stabilizes the floral identity. At this stage, the inflorescence identity gene, *TFL1*, is repressed - mostly by *AP1* and other MADS-domain TFs. As discussed in the text, repression of *TFL1* by *LFY* might be only indirect through activation of *AP1* and *LFY* might be a direct activator of *TFL1* in the inflorescence stem. *AP1* also inhibits cytokinin biosynthesis through its effect on the *LOG* genes, thus preventing meristem formation at the sepal axil.

Repression of the inflorescence identity gene *TERMINAL FLOWER1 (TFL1)* is also essential for proper floral development [38] (Figure 3). This repression involves *LFY*, *AP1*, *CAL*, *SVP*, a variety of remote regulatory elements in the 3' region of *TFL1* and DNA looping [38,39]. *LFY* was initially thought to directly repress *TFL1*, a hypothesis supported by ChIP-Seq assays showing *LFY* binding to the 3' region of *TFL1*, mediated by the *LFY* oligomerization domain [31,40] (Figure 4). However, when analyzing raw data from ref [31], we realized that *LFY-GR* appears to induce *TFL1* expression in seedlings and it was recently found that deletion of a region encompassing the *LFY* binding sites resulted in reduced *TFL1* expression in the stem [39]. Thus, *LFY* could actually be a direct activator of *TFL1* in the stem, and its repressive capacity in the flower might be indirect, through activation of *AP1* and *CAL*. *AP1* also suppresses another inflorescence trait, the production of axillary buds which grow in the axil of first-whorl organs in *ap1* mutants. To achieve this inhibition, *AP1* represses *LOG1*, a gene encoding a CK biosynthetic enzyme, and induces the expression of the *CYTOKININ OXIDASE/DEHYDROGENASE3* CK catabolic enzyme [41]. The resultant reduction in CK levels prevents the sepal axil from acquiring meristematic competence [41].

It is interesting to note that meristem emergence and floral fate acquisition may be coupled by the continuous action of *LFY*, acting in these two processes throughout flower development.

### Impact of Novel Technologies on our Understanding of Flower Development

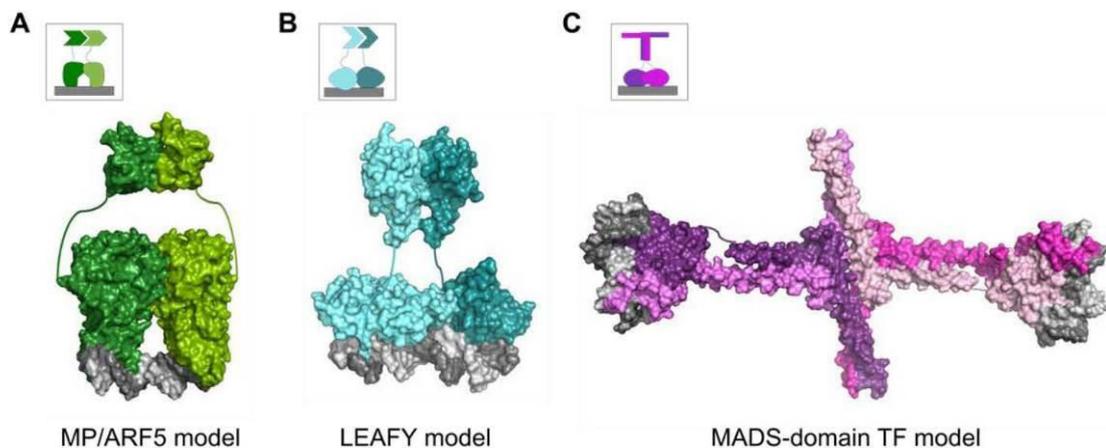
Floral development will only be understood once it can be fully modeled. The formation of flowers is a robust process, indicating that founder-cells interpret information they possess (genome and other molecules inherited from cell division) as well as external cues (such as hormonal cues or mechanical constraints imposed by their neighbors) to determine their development. However, for us it is extremely difficult to understand the rules that these cells so easily follow every time a flower is formed.

Over the last 30 years, genetics has extremely efficiently identified a number of regulators affecting floral development in a fairly specific way [1,2]. Here, we will analyze how the advent of new technological developments has generated new information which should help us to understand and model this process at multiple levels: atomic, genomic, cellular, and organ.

them common functional properties (such as the capacity to loop DNA, to bind closed chromatin, or to interact with chromatin remodelers SYD and BRM). These properties could never have been inferred from their primary amino-acid sequences [9,40,42,44,45].

### Impact of Genome-Wide Techniques

When applied to floral development, Next-Generation Sequencing (NGS) based genome-wide studies (describing gene expression, TF binding, chromatin compaction, histone marks, nucleosome positioning or 3D genome architecture) hold great promise as they can be used to infer general rules, thus extending beyond the bounds of individual case-studies. These types of study also represent important data collections that can easily be reused by the broader community. Genome browsers can be utilized as 'NGS microscopes' to help generate hypotheses for genome-wide testing. The diversity of the



**Figure 4. Structural models of complexes between transcription factors and DNA.** A) MONOPTEROS (MP) model based on the complex between ARF1 and DNA (pdb 4LDX) and the MP PB1 oligomerization domain (4CHK) B) LFY model based on the complex between LFY, DNA (2VY1) and the GbLFY SAM oligomerization domain (4UDE) C) MADS-domain TF model involving the MEF2 (1TQE) DBD and the SEP3 tetramerization domain (4OX0). Arbitrarily-shaped linkers (narrow lines) have been added between domains for which the structures are known. The C-terminal domain of MADS-domain TFs is not represented. The insets aim at showing the similarity between the three proteins that share a face-to-face dimeric DNA-binding domain and another domain that allows higher-order complex formation (oligomer for LFY and MP and tetramer for MADS-domain TF) with the potential to form DNA loops.

### Structural Advances

Although the technique is far from new, protein crystallography has recently been applied to several key floral regulators such as MP, LFY, and SEPALLATA3 (SEP3) MADS-domain TF [40,42–44] (Figure 4). Beyond the satisfaction of obtaining a picture of these important proteins, the results obtained offer the possibility to understand how the atomic features of major proteins (for example elements mediating DNA binding, di-, tetra- or oligo-merization) contribute to their functions. In contrast to genetics, structural data make targeted modifications of protein properties possible in the search to unravel new functions, as in the case of LFY [13,40]. Structural advances also revealed unexpected similarities in the overall organization of LFY, ARF and MADS-domain TFs (Figure 4): they all harbor a dimeric face-to-face DNA-binding domain associated with a higher-order oligomerization domain. The shared organization of these 3 TFs, which form an activation cascade, might confer

floral cell types present in the inflorescence is a clear hindrance to genome-wide studies, but several technical advances have made it possible to sort tissues or cells [46,47], even if more specific promoters are needed to dissect floral meristems over both space and time.

In some cases, large-scale studies identified new actors (such as with AHP6 or ELA1) which could then be investigated by classical genetics techniques [7,34]. In others, they highlighted novel or understudied features: genes involved in pathogen defense are overrepresented in the genes expressed in the epidermis [46]; this group of genes is shut off by LFY and ANT/AIL6 [16,31]; expression of cell cycle genes displays a burst during the floral transition [48]. One of the most fascinating questions that can be addressed through genome-wide analysis is the target-specificity of MADS-domain TF complexes. This specificity was elegantly analyzed for SVP and FUL, and their roles in regulating common or specific target genes was identified [49]. However, for organ-identity MADS-domain TF tetramers, ChIP-Seq data suggest that they

bind to similar genomic regions ([1] and references therein), indicating that their specific capacity to trigger separate developmental programs lies in elusive - sequence specific - co-regulators that have yet to be identified. Exhaustive proteomics-based studies of protein interactions might help to identify these co-regulators [45,50].

A fertile area that will certainly be more exploited in the future is the comparison of various types of genome-wide data. Two such analyses, from very different angles, proposed that LFY and the MADS-domain TFs AP1 and SEP3 could act as pioneer TFs with the capacity to bind closed chromatin regions allowing the subsequent recruitment of other TFs [40,51,52]. This attractive hypothesis remains to be confirmed, but is consistent with earlier reports showing that some floral genes are controlled by Polycomb Group proteins repression and that their activation require chromatin remodelers such as SYD and BRM [53]. The cross-talk between transcriptional and chromatin regulation appears particularly prominent during the early stages of flower development [45,54,55].

### Modeling Regulatory Networks

Built with information relating to key regulators and their relationships, models of gene regulatory networks (GRN) can be used to understand whether known relationships explain gene expression patterns or morphogenesis, or if additional actors or connections need to be identified. Models are becoming increasingly complex [22,30,36], but there is still a gap between what can be modeled and the wealth of information available from genome-wide studies. Gene expression cannot yet be modeled based only on genome sequences plus some information about chromatin, but predictive tools are improving, in particular with regard to the location of TF binding sites [56,57]. A major challenge is to understand when TF binding leads to gene regulation. New CRISPR-Cas9 tools which can directly modify regulatory elements at their genomic locations [58] will provide a novel means to test GRN models and better understand complex gene regulation such as the control of *TFL1*'s expression [39].

Ultimately, GRN will have to be integrated into cellular models such as those modeling growth and integrating mechanical constraints [59–61]. Novel techniques such as Atomic Force Microscopy or Fluorescence emission-Brillouin imaging are also being developed or adapted to provide access to the mechanical properties of living cells within developing tissues [62].

### Conclusion

Over the last 30 years, genetic approaches have identified the main regulators of FM development. We must now build multi-scale models that integrate other types of information such as tissue mechanics, morphology, hormone transport and signaling. Although they may remain incomplete, these models will play an essential role in the search to identify missing elements, inconsistencies or network behavior that cannot be simply apprehended. While data on gene expression, cell growth and mechanics will soon coexist in individual models, in the future we hope they can be included in multi-scale models bridging atomic and genomic resolutions.

### Acknowledgments

Work in our laboratory is supported by the French National Agency for Research (ANR Blanc – SVSE2 – 2011-Charmful; GRAL, ANR-10-LABX-49-01 and ANR-12-BSV6-0005-Auxiflo). We thank G. Vachon for comments on the manuscript. We apologize to our colleagues whose work could not be cited due to space limitations.

### Bibliography

1. Stewart D, Graciet E, Wellmer F: **Molecular and regulatory mechanisms controlling floral organ development.** *FEBS J* 2016, **283**:1823–1830.
  2. Pajoro A, Biewers S, Dougali E, Leal Valentim F, Mendes MA, Porri A, Coupland G, Van de Peer Y, van Dijk AD, Colombo L, et al.: **The (r)evolution of gene regulatory networks controlling Arabidopsis plant reproduction: a two-decade history.** *J Exp Bot* 2014, **65**:4731–4745.
  3. Sassi M, Vernoux T: **Auxin and self-organization at the shoot apical meristem.** *J Exp Bot* 2013, **64**:2579–2592.
  4. Pierre-Jerome E, Moss BL, Nemhauser JL: **Tuning the auxin transcriptional response.** *J Exp Bot* 2013, **64**:2557–2563.
  5. Yamaguchi N, Wu MF, Winter CM, Berns MC, Nole-Wilson S, Yamaguchi A, Coupland G, Krizek BA, Wagner D: **A molecular framework for auxin-mediated initiation of flower primordia.** *Dev Cell* 2013, **24**:271–282.
  6. Furutani M, Nakano Y, Tasaka M: **MAB4-induced auxin sink generates local auxin gradients in Arabidopsis organ formation.** *Proc Natl Acad Sci U S A* 2014, **111**:1198–1203.
  - \*7. Besnard F, Refahi Y, Morin V, Marteaux B, Brunoud G, Chambrier P, Rozier F, Mirabet V, Legrand J, Lainé S, et al.: **Cytokinin signalling inhibitory fields provide robustness to phyllotaxis.** *Nature* 2014, **505**:417–421.
- AHP6 is a cytokinin signaling inhibitor activated by MONOPTEROS downstream of auxin in floral meristems. AHP6 migrates to surround newly formed primordia, creating inhibitory fields where CK signaling is decreased to prevent the initiation of new floral primordia. Together with data showing that MP activates CK signaling in developing primordia, this study reveals a dual role for MP in regulating phyllotaxis.
8. Besnard F, Rozier F, Vernoux T: **The AHP6 cytokinin signaling inhibitor mediates an auxin-cytokinin crosstalk that regulates the timing of organ initiation at the shoot apical meristem.** *Plant Signal Behav* 2014, **9**.
  - \*\*9. Wu MF, Yamaguchi N, Xiao J, Bargmann B, Estelle M, Sang Y, Wagner D: **Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate.** *elife* 2015, **4**:e09269.

This paper presents how MONOPTEROS can interpret auxin levels to influence the chromatin status of its target genes. Thus, when auxin is present, MP recruits the BRAHMA and SPLAYED SWI/SNF ATPases to increase chromatin accessibility. In the absence of auxin, IAA corepressors promote chromatin closure by recruiting TOPLESS. Through these effects, auxin triggers a switch in the floral meristem that links transcriptional and chromatin regulation.

10. Yamaguchi N, Wu MF, Winter CM, Wagner D: **LEAFY and Polar Auxin Transport Coordinately Regulate Arabidopsis Flower Development.** *Plants* 2014, **3**:251–265.
11. Yamaguchi N, Jeong CW, Nole-Wilson S, Krizek BA, Wagner D: **AINTEGUMENTA and AINTEGUMENTA-LIKE6/PLETHORA3 Induce LEAFY Expression in Response to Auxin to Promote the Onset of Flower Formation in Arabidopsis.** *Plant Physiol* 2016, **170**:283–293.
12. Li W, Zhou Y, Liu X, Yu P, Cohen JD, Meyerowitz EM: **LEAFY controls auxin response pathways in floral primordium formation.** *Sci Signal* 2013, **6**:ra23.
13. Chahtane H, Vachon G, Le Masson M, Thévenon E, Périgon S, Mihajlovic N, Kalinina A, Michard R, Moyroud E, Monniaux M, et al.: **A variant of LEAFY reveals its capacity to stimulate meristem development by inducing RAX1.** *Plant J* 2013, **74**:678–689.
14. Kierzkowski D, Lenhard M, Smith R, Kuhlemeier C: **Interaction between meristem tissue layers controls phyllotaxis.** *Dev Cell* 2013, **26**:616–628.
15. Mudunkothge JS, Krizek BA: **Three Arabidopsis AIL/PLT genes act in combination to regulate shoot apical meristem function.** *Plant J* 2012, **71**:108–121.
16. Krizek BA, Bequette CJ, Xu K, Blakley IC, Fu ZQ, Stratmann J, Loraine AE: **RNA-Seq links AINTEGUMENTA and AINTEGUMENTA-LIKE6 to cell wall remodeling and plant defense pathways in Arabidopsis.** *Plant Physiol* 2016, doi:10.1104/pp.15.01625.
17. Peaucelle A, Braybrook SA, Le Guillou L, Bron E, Kuhlemeier C, Höfte H: **Pectin-induced changes in cell wall mechanics underlie organ initiation in Arabidopsis.** *Curr Biol* 2011, **21**:1720–1726.
18. Yang W, Schuster C, Beahan CT, Charoensawan V, Peaucelle A, Bacic A, Doblin MS, Wightman R, Meyerowitz EM: **Regulation of Meristem Morphogenesis by Cell Wall Synthases in Arabidopsis.** *Curr Biol* 2016, **26**:1404–1415.
- \*19. Sassi M, Ali O, Boudon F, Cloarec G, Abad U, Cellier C, Chen X, Gilles B, Milani P, Friml J, et al.: **An auxin-mediated shift toward growth isotropy promotes organ formation at the shoot meristem in Arabidopsis.** *Curr Biol* 2014, **24**:2335–2342.
20. Yoshida S, Mandel T, Kuhlemeier C: **Stem cell activation by light guides plant organogenesis.** *Genes Dev* 2011, **25**:1439–1450.
21. Adibi M, Yoshida S, Weijers D, Fleck C: **Centering the Organizing Center in the Arabidopsis thaliana Shoot Apical Meristem by a Combination of Cytokinin Signaling and Self-Organization.** *PLoS ONE* 2016, **11**:e0147830.
- \*\*22. Gruel J, Landrein B, Tarr P, Schuster C, Refahi Y, Sampathkumar A, Hamant O, Meyerowitz EM, Jönsson H: **An epidermis-driven mechanism positions and scales stem cell niches in plants.** *Science advances* 2016, **2**:e1500989.
23. Zhao Z, Andersen SU, Ljung K, Dolezal K, Miotk A, Schultheiss SJ, Lohmann JU: **Hormonal control of the shoot stem-cell niche.** *Nature* 2010, **465**:1089–1092.
24. Wang Y, Wang J, Shi B, Yu T, Qi J, Meyerowitz EM, Jiao Y: **The Stem Cell Niche in Leaf Axils Is Established by Auxin and Cytokinin in Arabidopsis.** *Plant Cell* 2014, **26**:2055–2067.
25. Chickarmane VS, Gordon SP, Tarr PT, Heisler MG, Meyerowitz EM: **Cytokinin signaling as a positional cue for patterning the apical-basal axis of the growing Arabidopsis shoot meristem.** *Proc Natl Acad Sci U S A* 2012, **109**:4002–4007.
26. Schuster C, Gaillochet C, Lohmann JU: **Arabidopsis HECATE genes function in phytohormone control during gynoecium development.** *Development* 2015, **142**:3343–3350.
27. Shi B, Zhang C, Tian C, Wang J, Wang Q, Xu T, Xu Y, Ohno C, Sablowski R, Heisler MG, et al.: **Recent data on the regulation of stem cell niche homeostasis are integrated into a model that is robust across various conditions. This model predicts that CK diffusion from the L1 and diffusible signals restricting CK signaling to the center of the meristem are sufficient to control gene expression. The model also predicts that meristem initiation in growing flowers is unlocked by organ growth.**

- Two-Step Regulation of a Meristematic Cell Population Acting in Shoot Branching in Arabidopsis.** *PLoS Genet* 2016, **12**:e1006168.
28. Lee C, Clark SE: **A WUSCHEL-Independent Stem Cell Specification Pathway Is Repressed by PHB, PHV and CNA in Arabidopsis.** *PLoS ONE* 2015, **10**:e0126006.
  29. Lee J, Lee I: **Regulation and function of SOC1, a flowering pathway integrator.** *J Exp Bot* 2010, **61**:2247–2254.
  30. Pérez-Ruiz RV, García-Ponce B, Marsch-Martínez N, Ugartechea-Chirino Y, Villajuana-Bonequi M, de Folter S, Azpeitia E, Dávila-Velderrain J, Cruz-Sánchez D, Garay-Arroyo A, et al.: **XAANTAL2 (AGL14) Is an Important Component of the Complex Gene Regulatory Network that Underlies Arabidopsis Shoot Apical Meristem Transitions.** *Mol Plant* 2015, **8**:796–813.
  31. Winter CM, Austin RS, Blanvillain-Baufumé S, Reback MA, Monniaux M, Wu MF, Sang Y, Yamaguchi A, Yamaguchi N, Parker JE, et al.: **LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response.** *Dev Cell* 2011, **20**:430–443.
  32. Kaufmann K, Wellmer F, Muiño JM, Ferrier T, Wuest SE, Kumar V, Serrano-Mislata A, Madueño F, Krajewski P, Meyerowitz EM, et al.: **Orchestration of floral initiation by APETALA1.** *Science* 2010, **328**:85–89.
  33. Moyroud E, Minguet EG, Ott F, Yant L, Posé D, Monniaux M, Blanchet S, Bastien O, Thévenon E, Weigel D, et al.: **Prediction of regulatory interactions from genome sequences using a biophysical model for the Arabidopsis LEAFY transcription factor.** *Plant Cell* 2011, **23**:1293–1306.
  34. Winter CM, Yamaguchi N, Wu MF, Wagner D: **Transcriptional programs regulated by both LEAFY and APETALA1 at the time of flower formation.** *Physiol Plant* 2015, **155**:55–73.
  35. Van Mourik S, van Dijk AD, de Gee M, Immink RG, Kaufmann K, Angenent GC, van Ham RC, Molenaar J: **Continuous-time modeling of cell fate determination in Arabidopsis flowers.** *BMC Syst Biol* 2010, **4**:101.
  36. Azpeitia E, Davila-Velderrain J, Villarreal C, Alvarez-Buylla ER: **Gene regulatory network models for floral organ determination.** *Methods Mol Biol* 2014, **1110**:441–469.
  37. Gregis V, Andrés F, Sessa A, Guerra RF, Simonini S, Mateos JL, Torti S, Zambelli F, Prazzoli GM, Bjerkan KN, et al.: **Identification of pathways directly regulated by SHORT VEGETATIVE PHASE during vegetative and reproductive development in Arabidopsis.** *Genome Biol* 2013, **14**:R56.
  38. Baumann K, Venail J, Berbel A, Domenech MJ, Money T, Conti L, Hanzawa Y, Madueno F, Bradley D: **Changing the spatial pattern of TFL1 expression reveals its key role in the shoot meristem in controlling Arabidopsis flowering architecture.** *J Exp Bot* 2015, **66**:4769–4780.
  39. Serrano-Mislata A, Fernández-Nohales P, Doménech MJ, Hanzawa Y, Bradley D, Madueño F: **Separate elements of the TERMINAL FLOWER 1 cis-regulatory region integrate pathways to control flowering time and shoot meristem identity.** *Development* 2016, doi:10.1242/dev.135269.
  - \*40. Sayou C, Nanao MH, Jamin M, Posé D, Thévenon E, Grégoire L, Tichtinsky G, Denay G, Ott F, Peirats Llobet M, et al.: **A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor.** *Nat Commun* 2016, **7**:11222.
- Structural analysis revealed that LFY's N-terminal conserved region encodes a SAM oligomerization domain which is essential for LFY protein function. Using biochemical and genome-wide approaches on an oligomerization-defective LFY variant showed that the SAM domain helps LFY bind to genomic regions lacking high-affinity binding sites or closed chromatin regions. This later result suggests the SAM oligomerization domain may confer LFY TF pioneer properties.
41. Han Y, Zhang C, Yang H, Jiao Y: **Cytokinin pathway mediates APETALA1 function in the establishment of determinate floral meristems in Arabidopsis.** *Proc Natl Acad Sci U S A* 2014, **111**:6840–6845.
  42. Boer DR, Freire-Rios A, van den Berg WA, Saaki T, Manfield IW, Kepinski S, López-Vidriero I, Franco-Zorrilla JM, de Vries SC, Solano R, et al.: **Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors.** *Cell* 2014, **156**:577–589.
  43. Nanao MH, Vinos-Poyo T, Brunoud G, Thévenon E, Mazzoleni M, Mast D, Lainé S, Wang S, Hagen G, Li H, et al.: **Structural basis for oligomerization of auxin transcriptional regulators.** *Nat Commun* 2014, **5**:3617.
  44. Puranik S, Acajjaoui S, Conn S, Costa L, Conn V, Vial A, Marcellin R, Melzer R, Brown E, Hart D, et al.: **Structural basis for the oligomerization of the MADS domain transcription factor SEPALLATA3 in Arabidopsis.** *Plant Cell* 2014, **26**:3603–3615.
  45. Smaczniak C, Immink RG, Muiño JM, Blanvillain R, Busscher M, Busscher-Lange J, Dinh QD, Liu S, Westphal AH, Boeren S, et al.:

- Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development.** *Proc Natl Acad Sci U S A* 2012, **109**:1560–1565.
46. Yadav RK, Tavakkoli M, Xie M, Girke T, Reddy GV: **A high-resolution gene expression map of the Arabidopsis shoot meristem stem cell niche.** *Development* 2014, **141**:2735–2744.
47. Deal RB, Henikoff S: **The INTACT method for cell type-specific gene expression and chromatin profiling in Arabidopsis thaliana.** *Nat Protoc* 2011, **6**:56–68.
48. Klepikova AV, Logacheva MD, Dmitriev SE, Penin AA: **RNA-seq analysis of an apical meristem time series reveals a critical point in Arabidopsis thaliana flower initiation.** *BMC Genomics* 2015, **16**:466.
49. Mateos JL, Madrigal P, Tsuda K, Rawat V, Richter R, Romera-Branchat M, Fornara F, Schneeberger K, Krajewski P, Coupland G: **Combinatorial activities of SHORT VEGETATIVE PHASE and FLOWERING LOCUS C define distinct modes of flowering regulation in Arabidopsis.** *Genome Biol* 2015, **16**:31.
50. Smaczniak C, Li N, Boeren S, America T, van Dongen W, Goerdayal SS, de Vries S, Angenent GC, Kaufmann K: **Proteomics-based identification of low-abundance signaling and regulatory protein complexes in native plant tissues.** *Nat Protoc* 2012, **7**:2144–2158.
51. Iwafuchi-Doi M, Zaret KS: **Cell fate control by pioneer transcription factors.** *Development* 2016, **143**:1833–1837.
- \*52. Pajoro A, Madrigal P, Muiño JM, Matus JT, Jin J, Mecchia MA, Debernardi JM, Palatnik JF, Balazadeh S, Arif M, et al.: **Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development.** *Genome Biol* 2014, **15**:R41.
- Transcriptome, chromatin accessibility, and DNA-binding data were combined for AP1 and SEP3 at several steps of flower development. The results indicate that AP1 and SEP3 binding appear to precede chromatin accessibility and gene expression, suggesting that these proteins could act as pioneer transcription factors during flower formation.
53. Wu MF, Sang Y, Bezhani S, Yamaguchi N, Han SK, Li Z, Su Y, Slewinski TL, Wagner D: **SWI2/SNF2 chromatin remodeling ATPases overcome polycomb repression and control floral organ identity with the LEAFY and SEPALLATA3 transcription factors.** *Proc Natl Acad Sci U S A* 2012, **109**:3576–3581.
54. Moreau F, Thévenon E, Blanvillain R, Lopez-Vidriero I, Franco-Zorrilla JM, Dumas R, Parcy F, Morel P, Trehin C, Carles CC: **The Myb-domain protein ULTRAPETALA1 INTERACTING FACTOR 1 controls floral meristem activities in Arabidopsis.** *Development* 2016, **143**:1108–1119.
55. Sacharowski SP, Gratkowska DM, Sarnowska EA, Kondrak P, Jancewicz I, Porri A, Bucior E, Rolicka AT, Franzen R, Kowalczyk J, et al.: **SWP73 Subunits of Arabidopsis SWI/SNF Chromatin Remodeling Complexes Play Distinct Roles in Leaf and Flower Development.** *Plant Cell* 2015, **27**:1889–1906.
56. Mathelier A, Fornes O, Arenillas DJ, Chen CY, Denay G, Lee J, Shi W, Shyr C, Tan G, Worsley-Hunt R, et al.: **JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles.** *Nucleic Acids Res* 2016, **44**:D110–5.
57. O'Malley RC, Huang SS, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR: **Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape.** *Cell* 2016, **165**:1280–1292.
58. Yan W, Chen D, Kaufmann K: **Efficient multiplex mutagenesis by RNA-guided Cas9 and its use in the characterization of regulatory elements in the AGAMOUS gene.** *Plant Methods* 2016, **12**:23.
59. Boudon F, Chopard J, Ali O, Gilles B, Hamant O, Boudaoud A, Traas J, Godin C: **A computational framework for 3D mechanical modeling of plant morphogenesis with cellular resolution.** *PLoS Comput Biol* 2015, **11**:e1003950.
60. Barbier de Reuille P, Routier-Kierzkowska AL, Kierzkowski D, Bassel GW, Schüpbach T, Tauriello G, Bajpai N, Strauss S, Weber A, Kiss A, et al.: **MorphoGraphX: A platform for quantifying morphogenesis in 4D.** *elife* 2015, **4**:05864.
61. Milani P, Mirabet V, Cellier C, Rozier F, Hamant O, Das P, Boudaoud A: **Matching Patterns of Gene Expression to Mechanical Stiffness at Cell Resolution through Quantitative Tandem Epifluorescence and Nanoindentation.** *Plant Physiol* 2014, **165**:1399–1408.
62. Elsayad K, Werner S, Gallemí M, Kong J, Sánchez Guajardo ER, Zhang L, Jaillais Y, Greb T, Belkhadir Y: **Mapping the subcellular mechanical properties of live cells in tissues with fluorescence emission-Brillouin imaging.** *Sci Signal* 2016, **9**:rs5.
63. Smyth DR, Bowman JL, Meyerowitz EM: **Early flower development in Arabidopsis.** *Plant Cell* 1990, **2**:755–767.

## 2. Que peut-on apprendre de la formation des méristèmes végétatifs ?

Bien que les organes axillaires initiés avant et après la transition florale soient très différents, la structure des entrenœuds est fondamentalement similaire (Figure Intro. 2). L'entrenœud est en général constitué d'un méristème axillaire (*axillary meristem* AM) sous-tendu par une feuille. Lors du développement végétatif le méristème axillaire est un méristème caulinaire latéral (*lateral shoot meristem* LSM). Son développement n'est pas toujours simultané avec celui de la feuille et il reste quiescent à l'aisselle de la feuille. Après la transition florale, le méristème axillaire est un méristème floral (*flower meristem* FM) et est sous-tendu par une feuille modifiée, la bractée qui ne se développe pas chez *Arabidopsis* et est consommée par la croissance du FM. Malgré ces différences de développement, les réseaux de gènes contrôlant la mise en place des LSM et des FM partagent de nombreux acteurs. De récentes avancées dans la mise en place des LSM sont donc potentiellement transférables à l'étude des FM.

### *La conservation des acteurs régulant l'initiation des méristèmes axillaires et floraux*

Un des premiers signes moléculaires de l'initiation d'un organe axillaire, après la réponse auxinique, est l'extinction de STM. Cette réponse est conservée pour la mise en place des LSM et FM (Long and Barton, 2000). L'activation de la signalisation CK, précède ensuite l'établissement des marqueurs du méristème WUS et CLV3, et la réactivation de STM, lorsque le primordium est déjà visible comme un groupe de cellules distinct (Shi et al., 2016; Wang et al., 2014b, 2014a; Long and Barton, 2000; Yoshida et al., 2011; Denay et al., 2017).

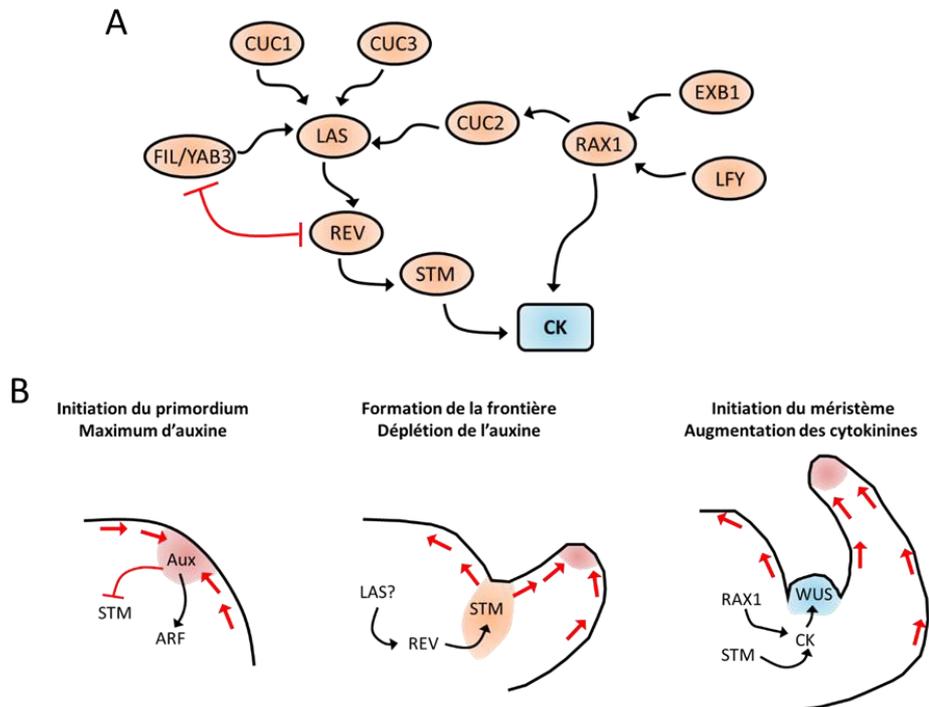
En plus des mécanismes discutés dans l'article 1, certains acteurs ont été beaucoup plus étudiés pour leur rôle dans la mise en place des LSM mais sont ou pourraient aussi être impliqués dans la mise en place des FM (Figure Intro. 6A). Le facteur de transcription de la famille HD-ZIP III REVOLUTA (REV) est nécessaire à la fois pour la formation des LSM et des FM (Otsuga et al., 2001;

Prigge et al., 2005). Les mutants *rev* initient moins d'AM que les plantes sauvages et les fleurs produites sont souvent incomplètes, reflétant probablement un problème de maintenance des FM. En accord avec cela, l'expression de REV précède l'émergence du FM et les mutants *rev* sont épistatiques à *stm*, indiquant une possible régulation de *STM* par REV (Otsuga et al., 2001). Plus récemment, il a été montré que les HD-ZIPIII PHB, PHV et CNA, qui agissent de manière antagoniste à REV dans la régulation des méristèmes (Prigge et al., 2005), sont impliqués dans la restriction de la taille du méristème floral (Landau et al., 2015). De manière intéressante, *REV* est aussi exprimé dans la PZ du SAM et l'expression de *REV* dans le domaine basal de l'embryon suffit à le convertir en SAM (Otsuga et al., 2001; Smith and Long, 2010). Malgré les efforts pour identifier les cibles génomiques de REV, ces approches ont fourni peu de résultats en ce qui concerne son rôle dans la régulation des cellules souches et ont plutôt étayé un autre rôle de *REV*, dans la régulation de l'identité adaxiale des organes (Brandt et al., 2012; Reinhart et al., 2013).

Les facteurs de transcription YABBY (YAB) et notamment YAB1/FILAMENTOUS FLOWER (FIL) agissent de manière antagoniste à REV pour la détermination de la polarité des organes axillaires (Eshed et al., 2001). Cependant les mutants *fil* ont des phénotypes similaires aux mutants *rev* pour l'initiation des méristèmes floraux, avec la formation de structures filamenteuses (Chen et al., 1999; Sawa et al., 1999). En revanche, les fleurs initiées par les mutants *fil* et *fil yab3* forment parfois cinq sépales ou plus (Goldshmidt et al., 2008; Chen et al., 1999). De plus FIL et YAB3 restreignent les domaines d'expression de WUS et CLV3 dans le SAM *via* l'induction du facteur de transcription présomptif LATERAL SUPPRESSOR (LAS) à la frontière organe-SAM (Goldshmidt et al., 2008). Les relations entre REV et FIL sont donc assez intrigantes : antagonistes dans le contrôle de la polarité des organes et exprimés dans des domaines différents ils contrôlent cependant de manière similaire la mise en place du FM.

LAS est un régulateur connu de la formation des LSM et les mutants *las* échouent presque complètement à former des tiges secondaires à l'aisselle des feuilles de rosettes (Greb et al., 2003). En revanche les plantes *las* forment

normalement des tiges secondaires et des fleurs après la transition florale. Cela indique que soit LAS n'est pas requis après la transition florale, malgré son expression dans le SAM, soit d'autres voies redondantes régulent l'initiation des AM après la transition florale (Greb et al., 2003). Il a été proposé que LAS agisse en régulant l'expression de REV (Greb et al., 2003).



**Figure Intro. Intro. 6 . Réseau de gènes dans la frontière et leur rôle dans la formation des LSM.** (A) Réseau présumptif de gènes exprimés contrôlant la mise en place des AM. *RAX1* est induit par *EXB1* dans les tissus végétatifs et par *LFY* dans les fleurs. *RAX1* induit l'expression de *CUC2*. Les gènes *CUCs* régulent *LAS* qui induit lui-même l'expression de *REV*. *LAS* est aussi sous le contrôle de *FIL* qui est exprimé de manière antagoniste à *REV*. *REV* induit l'expression de *STM*. *STM* et *RAX1* agissent probablement *via* la synthèse des cytokinines. (B) Mise en place des LSM. Le maximum d'auxine induit l'activité transcriptionnelle des ARFs et la répression de *STM*. L'expression de *STM* dans le primordium n'est pas totalement abolie mais réduite à un niveau basal. L'expression de *REV*, peut-être dépendante de *LAS*, permet la ré-induction de *STM* dans la zone frontière. *STM*, potentiellement aidé par *RAX1* permet la synthèse de CK dans la frontière. Les CK sont des activateurs de *WUS*.

En plus de sa régulation par *FIL*, l'expression de *LAS* est contrôlée par les gènes de frontière de la famille *CUP-SHAPED COTYLEDON* (*CUC*), *CUC1* et *CUC2* (Raman et al., 2008). Le troisième membre de cette famille, *CUC3*, régule également la formation des AM, avec un rôle plus important dans la formation des LSM à l'aisselle des feuilles de rosettes. Bien que l'activité des *CUC1/2/3* ne semble pas absolument requise pour l'initiation des FM,

l'expression de versions de *CUC1* ou *CUC2* insensibles à la dégradation de leur ARNm par des miRNA entraîne la formation de méristèmes accessoires, à l'aisselle d'un AM existant et à la formation de cluster de fleurs pour *CUC2* (Raman et al., 2008; Peaucelle et al., 2007).

L'expression de *CUC2* est en partie régulée par les facteurs de transcription Myb R2R3, REGULATOR OF AXILLARY MERISTEMS1/2/3 (*RAX1/2/3*) (Müller et al., 2006; Keller et al., 2006). Les *RAX* sont exprimés dans les frontières entre le SAM et les organes axillaires et régulent de manière redondante la formation des LSM, avec un rôle prépondérant pour *RAX1*. De plus les gènes *RAX* agissent en parallèle avec *LAS* pour la régulation des LSM (Müller et al., 2006; Keller et al., 2006). Comme les défauts d'initiation de LSM des mutants *rax1* peuvent être compensés l'expression locale du gène *IPT* de biosynthèse des CK, *RAX1* pourrait agir *via* la régulation de la synthèse des CK (Wang et al., 2014b). L'expression de *RAX1/2/3* à l'aisselle des feuilles est contrôlée par le facteur de transcription EXCESSIVE BRANCHING1 (*EXB1*) dont l'allèle dominant *exb-1D* induit la prolifération de tiges secondaires (Guo et al., 2015). Il a également été proposé que *LFY* puisse contrôler l'initiation des FM *via* l'induction directe de *RAX1* (Chahtane et al., 2013).

#### *Un mécanisme pour l'établissement des méristèmes caulinaires latéraux*

De profondes avancées dans la compréhension de la mise en place des LSM sont venues ces deux dernières années des travaux des équipes de Yuling Jiao et Klaus Theres. Ces travaux ont révélé qu'un minimum d'auxine dans la frontière "SAM – organe" est requis pour l'établissement des LSM (Wang et al., 2014a, 2014b). Ce minimum est obtenu par la canalisation de l'auxine vers la pointe de la feuille et vers les sites d'initiation des prochains primordia (Figure Intro. 6B). Le minimum d'auxine est un prérequis à l'expression de *STM* dans la frontière, et est suivi par une augmentation de la signalisation des CK permettant l'établissement du méristème (Wang et al., 2014b, 2014a). L'imagerie de l'expression de *STM* au cours de ce processus a révélé qu'il reste exprimé dans la frontière à très bas niveau (Shi et al., 2016). L'expression de *REV* dans la frontière permet de renforcer l'expression de *STM* par liaison directe de son promoteur lors de l'initiation du LSM (Shi et al.,

2016). STM serait alors probablement en partie responsable de l'augmentation des CK, *via* son action sur leur synthèse (Jasinski et al., 2005). Cette étude montre en revanche que REV n'est pas capable d'initier *de novo* l'expression de STM, et que le maintien de bas niveaux de STM est requis pour la mise en place des LSM (Shi et al., 2016).

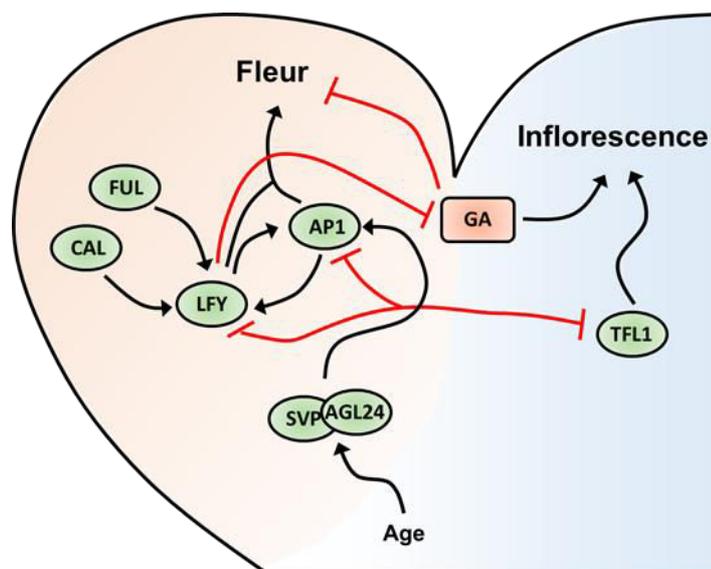
Dans les fleurs, STM est également exprimé à un niveau basal dans la frontière des primordia puis son expression augmente pendant les stades 1 et 2 du développement du FM (Landrein et al., 2015). Lors de ces premiers stades, l'expression de REV pourrait coïncider avec celle de STM (Heisler et al., 2005), indiquant que le mécanisme observé pour la mise en place des LSM pourrait être aussi valide pour la mise en place des FM. Il est également intéressant de noter que chez les mutants dans les gènes de la famille *MAB4*, l'absence de relocalisation basale de PIN1 provoque l'accumulation d'auxine dans la L1 et le marqueur de l'initiation florale *LFY* ne s'exprime pas, bien que le marqueur de réponse à l'auxine DR5 montre un comportement normal (Furutani et al., 2014). Le minimum d'auxine requis pour la mise en place des LSM est donc probablement aussi requis pour la mise en place des FM.

La large proportion de gènes exprimés dans les frontières et régulant la formation des AM illustre le rôle central de cette région pour l'initiation du méristème. Cela s'explique facilement pour les LSM car ils sont initiés directement dans cette zone, à l'aisselle des feuilles (Figure Intro. 6B). En revanche, les FM sont initiés au cœur du primordium, qui n'exprime pas les gènes marqueurs des frontières. Il est possible que l'expression très précoce des gènes de frontière prépare la mise en place du méristème. Une autre hypothèse serait que la frontière produise des signaux mobiles permettant l'établissement du méristème. Il est intéressant de noter que le SAM lui-même est initié dans une zone de frontière entre les deux cotylédons (Aida et al., 1999, 1997).

- Un même réseau contrôle la mise en place des LSM et des FM
- Les flux d'auxine et la synthèse de cytokinines sont requis pour l'initiation des méristèmes
- La zone de frontière est particulièrement importante pour l'initiation des méristèmes

### 3. Le développement floral

Le développement d'un méristème axillaire en une fleur ou en une tige latérale est la conséquence d'un réseau de gènes déterminant son identité. Une fois l'identité florale établie, de profonds changements morphologiques se mettent en place : changement de la phyllotaxie, répression de la croissance de la bractée, développement des organes floraux et arrêt de l'activité du méristème. Une abondante littérature décrit ces processus (Chandler, 2012; Stewart et al., 2016; Sablowski, 2015; Ó'Maoiléidigh et al., 2014; Pajoro et al., 2014b; Siriwardana and Lamb, 2012b; Bouché et al., 2016), je donnerai donc ici une vue d'ensemble des mécanismes gouvernant les différents aspects du développement floral, de la spécification de l'identité des organes floraux à la terminaison du méristème floral.



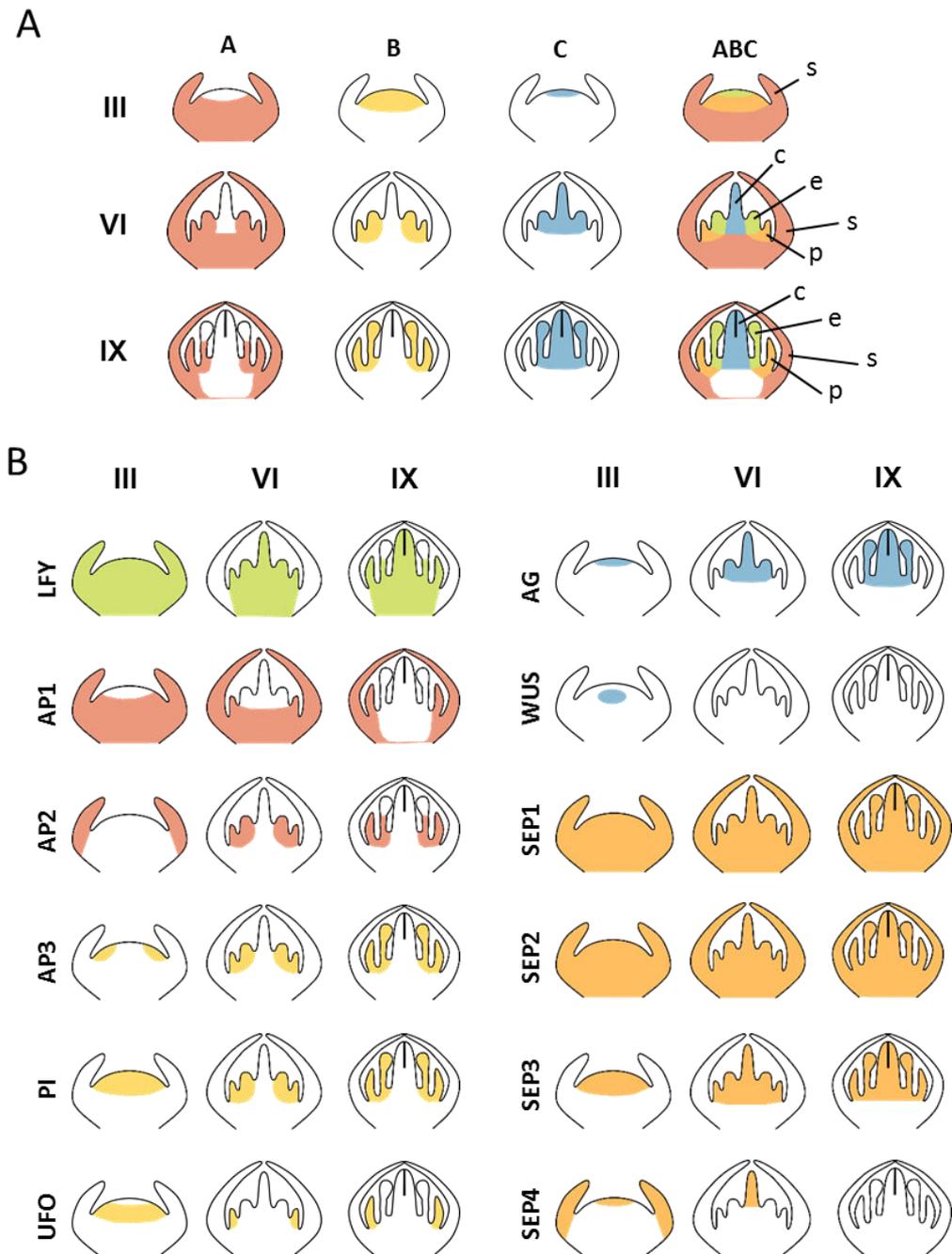
**Figure Intro. 7 . Détermination de l'identité florale.** L'identité florale est induite par LFY et AP1, alors que l'identité de l'inflorescence est établie par TFL1. LFY/AP1 et TFL1 sont exprimés de manière antagoniste dans le SAM. L'expression de LFY est maintenue en partie par CAL et FUL. LFY induit directement AP1 qui maintient l'identité florale via une boucle de régulation positive avec LFY. L'identité florale peut être induite de manière indépendante de LFY par SVP et AGL24, lorsque la plante atteint un certain âge.

#### *Détermination de l'identité florale*

L'identité fleur ou inflorescence est déterminée par l'expression de LFY ou TERMINAL FLOWER1 (TFL1) respectivement (Figure Intro. 7). Les mutants *lfy* produisent des structures feuillues à la place des fleurs (Weigel et al.,

1992), structure qui sont de plus sous-tendues par une bractée et ont une phyllotaxie spiralée rappelant celle de l'inflorescence (Weigel et al., 1992; Chandler and Werr, 2014). De plus l'expression ectopique de LFY entraîne la conversion des tiges secondaires en fleurs (Weigel and Nilsson, 1995). Au contraire, chez les mutants *tfl1* le méristème d'inflorescence se termine très tôt en une fleur terminale (Shannon and Meeks-Wagner, 1991). En plus de LFY, le facteur de transcription de type MADS APETALA1 (AP1) est important pour déterminer l'identité florale. Les mutants *ap1* forment des fleurs secondaires à l'aisselle des feuilles remplaçant les sépales, indiquant une indétermination du développement (Mandel et al., 1992). De plus, la combinaison des mutations *ap1* et *lfy* entraîne une conversion complète des fleurs en tiges secondaires (Huala and Sussex, 1992). D'autres combinaisons de mutants induisent divers degrés de conversion des fleurs en tiges. Les gènes *CAULIFLOWER* (*CAL*) et *FRUITFUL* (*FUL*), très proches d'*AP1* lui sont partiellement redondants pour la détermination de l'identité florale (Ferrándiz et al., 2000). Les mutants *ap1 cal* initient des méristèmes d'inflorescences en place des méristèmes floraux, résultant en une inflorescence ressemblant au chou-fleur, qui fleurit tardivement (Kempin et al., 1995; Ferrándiz et al., 2000). L'introduction de la mutation *ful* dans *ap1 cal* provoque l'abolition quasi complète de la floraison (Ferrándiz et al., 2000). L'expression forcée de LFY dans ce mutant permet de restaurer partiellement la floraison, indiquant un rôle de ces gènes dans le maintien de l'expression de LFY (Ferrándiz et al., 2000). LFY induit l'expression d'*AP1* dans les primordia floraux, et en retour AP1 induit l'expression de *LFY*, formant un boucle d'amplification renforçant l'identité florale (Liljegren et al., 1999; Wagner et al., 1999). Cependant AP1 est également induit indépendamment de LFY par les facteurs SHORT VEGETATIVE PHASE (SVP) et AGAMOUS-LIKE24 (AGL24) (Grandi et al., 2012). AP1 réprime directement l'expression de *TFL1*, empêchant le primordium d'acquérir une identité d'inflorescence (Liljegren et al., 1999) et *TFL1* réprime l'expression de *AP1* et *LFY* (Gustafson-Brown et al., 1994; Liljegren et al., 1999; Ratcliffe et al., 1999), permettant la définition de frontières nettes entre les identités fleur et inflorescence (Figure Intro 7).

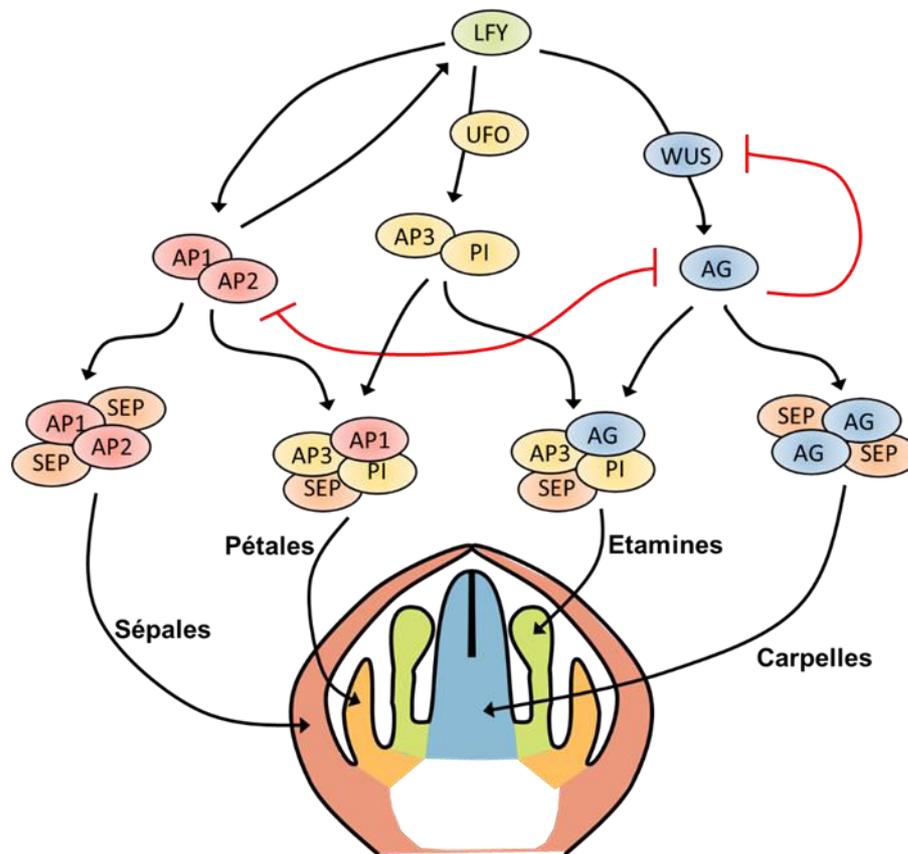
En plus de la voie génétique, il a été montré récemment que les GA jouaient également un rôle dans la détermination de l'identité florale (Winter et al., 2015; Yamaguchi et al., 2014a). Les GA sont des régulateurs connus de la transition florale (Blazquez et al., 1998). En revanche, l'application de GA après la transition florale retarde la formation de fleurs et les mutations dans divers gènes de synthèse des GA ou de signalisation par les GA l'accélère (Yamaguchi et al., 2014a). Dans les fleurs, LFY et AP1 induisent EUI-LIKE P450 A1 (ELA1), une enzyme potentiellement impliquée dans la dégradation des GA et STM inhibe la synthèse des GA par répression de l'enzyme GA20ox (Winter et al., 2015; Yamaguchi et al., 2014a; Jasinski et al., 2005).



**Figure Intro. 8 . Spécification des organes floraux par les gènes ABCE.** Profils d'expression des activités A (rouge), B (jaune), C (bleu) et des trois activités combinées (A) ou des gènes individuels (B) dans des fleurs de stade 3, 6 et 9. Les ébauches des différent organes sont indiqués par une lettre (s: sépales; p: pétales; e: étamines; c: carpelles). (A) L'activité combinée des gènes A et B est indiquée en orange et l'activité combinée des gènes B et C est indiquée en vert. (B) L'activité de *LFY* et des gènes *SEPs* est indiqués en vert et orange respectivement.

### Le modèle ABC

Une fois l'identité florale établie, l'identité de chacun des organes (sépales, pétales, étamines et carpelles) est établie par des combinaisons spécifiques de facteurs de transcription de la famille des FT MADS-box ou APETALA2 (AP2) dans chacune des couronnes d'organes (Figure Intro. 8A-B). Il a été mis en évidence que les activités de trois classes de gènes différentes sont nécessaires dans les couronnes d'organes consécutives pour parvenir à la structure florale. Ces trois classes de gènes ont été nommées A, B et C et ont données naissance au modèle ABC original (Coen and Meyerowitz, 1991; Bowman et al., 1991).



**Figure Intro. 9 . Action combinatoire des gènes ABC.** LFY induit l'expression des gènes A (rouge), des gènes B (jaune) avec son co-régulateur UFO et des gènes C (bleu) avec son co-régulateur WUS. L'expression des gènes A et des gènes C est mutuellement exclusive. L'identité des différents organes est déterminée par un quartet de protéines MADS comprenant les protéines d'activité A, A/B, B/C ou C et des protéines SEPs.

L'activité A est principalement supportée par AP1 et AP2 qui sont exprimés dans la totalité du primordium jusqu'au stade 3 mais dont l'expression se restreint ensuite aux ébauches de sépales et pétales (Gustafson-Brown et al., 1994). Les mutants *ap1* sont caractérisés par l'absence de pétales et par la

formation fleurs à l'aisselle des sépales convertis en feuilles (Gustafson-Brown et al., 1994). Chez les mutants *ap2* les sépales et pétales sont remplacés par les organes reproducteurs (Bowman et al., 1991; Coen and Meyerowitz, 1991). L'activité B est supportée par les gènes *PISTALATA (PI)* et *AP3* dont l'expression est restreinte aux pétales et étamines (Krizek and Meyerowitz, 1996). Les mutants *pi* et *ap3* sont caractérisés par la conversion des pétales et étamines en sépales et carpelles, respectivement (Bowman et al., 1991; Coen and Meyerowitz, 1991). La fonction C est supportée par *AGAMOUS (AG)* qui est exprimé au centre du FM et dans les ébauches d'étamines et de carpelles (Parcy et al., 1998). Les mutants *ag* sont caractérisés par un remplacement des étamines par des pétales et des carpelles par une nouvelle fleur, indiquant un défaut de spécification des couronnes internes accompagné d'une indétermination du méristème (Bowman et al., 1991; Coen and Meyerowitz, 1991).

Les profils d'expression des gènes ABC sont établis *via* des interactions entre gènes des classes A, B et C et orchestrés par LFY (Figure Intro. 9). LFY induit l'expression des gènes *AP1*, *AP3* et *AG* (Parcy et al., 1998). L'induction d'*AG* au centre du FM est dépendante de l'activité de LFY et de son corégulateur WUS (Lohmann et al., 2001; Busch et al., 1999; Lenhard et al., 2001). L'induction d'*AG* dépend également du remodelleur chromatinien *ULTRAPETALA1 (ULT1)*, qui agit en parallèle de LFY et WUS (Engelhorn et al., 2014). La définition de frontières nettes entre les profils d'expression des gènes A et C est produite par la répression mutuelle d'*AG* et *AP1/2* (Gustafson-Brown et al., 1994; Mizukami and Ma, 1992), restreignant donc *AP1/2* aux couronnes externes du FM. L'expression des gènes B dépend de la présence de LFY et du cofacteur *UNUSUAL FLOWER ORGAN (UFO)* exprimé dans les ébauches de pétales et d'étamines (Parcy et al., 1998; Honma and Goto, 2000). Cependant l'expression ectopique des gènes A, B ou C dans la feuille ne résulte pas en la formation d'organes floraux, indiquant la nécessité de facteurs floraux supplémentaires (Honma and Goto, 2001). Cette fonction (E) est remplie par les facteurs MADS-box de la famille *SEPALLATA (SEP)* 1-4 qui sont nécessaires à l'activité des gènes A, B et C (Ditta et al., 2004; Pelaz et al., 2000; Honma and Goto, 2001).

La base génétique de l'activité combinatoire des gènes ABCE s'explique par la capacité des FT MADS-box à agir en tétramères. Pour induire l'identité spécifique d'un organe le tétramère doit contenir certains MADS spécifiques (Figure Intro. 9) (Theissen, 2001; Krizek and Fletcher, 2005).

#### *Terminaison des cellules souches*

Contrairement aux tiges, les fleurs ont une croissance déterminée. Après la mise en place des organes floraux, l'activité du méristème s'arrête. La terminaison du méristème floral repose en grande partie sur l'activité d'AG. Chez les mutants *ag* une nouvelle fleur est formée au centre de la fleur de manière récursive, indiquant que le méristème maintient son activité (Coen and Meyerowitz, 1991; Bowman et al., 1991). AG induit l'expression de KNUCKLES (KNU) au centre du FM. Les mutants *knu* produisent des étamines et carpelles supplémentaires et chez les sur-expressseurs de KNU le FM est terminé précocement, indiquant que KNU est requis pour la terminaison du FM. KNU est normalement maintenu réprimé par la présence de répresseur de type Polycomb-group (PcG) sur son promoteur. L'activation de KNU ne semble pas dépendre de l'activité transcriptionnelle d'AG mais plutôt de sa compétition pour les sites de liaisons avec les PcG. Après l'éviction des PcG, les marques répressives de la chromatine déposées dans le promoteur de KNU sont retirées au fil des mitoses successives (Sun et al., 2009, 2014). En plus d'AG, les facteurs HD-ZIP III PHB, PHV et CNA et le facteur de transcription ER jouent un rôle dans la terminaison du FM. En effet dans un fond sensibilisé *clv3*, les mutations *er* et *jab1D* (régulateur négatif des HD-ZIP II, voir *1.1. Voies indépendantes de WUS*), entraînent la formation de carpelles surnuméraires et de couronnes d'organes supplémentaires, indiquant une activité prolongée du FM (Landau et al., 2015).

- LFY détermine l'identité florale et orchestre l'identité des organes floraux
- Continuité de l'identité florale et de l'identité des organes floraux *via* les réseaux contrôlés par LFY et AP1
- L'activité du méristème floral est arrêtée par AG

### III. Identification des gènes liés par des facteurs de transcription

Les processus développementaux comme la formation des fleurs requièrent de nombreux acteurs dont la plupart interagissent génétiquement, formant des réseaux de gènes complexes. La modélisation de ces réseaux est un moyen d'évaluer le niveau de compréhension que l'on a des processus développementaux qu'ils régulent, d'identifier des acteurs manquants, ou sous-estimés. Pour modéliser les réseaux, il faut pouvoir déterminer les relations hiérarchiques entre les acteurs en leur sein (expérimentalement ou par prédiction), être capable d'interconnecter les différents réseaux et enfin, pour une modélisation fine, connaître les propriétés de liaison des FT à l'ADN. La détermination du lexique des motifs d'ADN reconnus par les facteurs de transcription est une étape importante qui permet de répondre à certaines de ces questions. Les motifs de liaison des FT caractérisés jusque-là ont été rassemblés dans des bases de données, dont JASPAR (Mathelier et al., 2016) (Annexe 2), la base en accès libre de référence. Dans le cadre d'une collaboration j'ai participé à l'archivage d'une centaine de ces motifs pour des FT de plantes dans la version 2016 de JASPAR. Ce travail a constitué principalement en la vérification de données issue d'analyses haut-débit par comparaison avec les diverses données existantes dans la littérature.

Dans les paragraphes suivants j'expose d'une part les méthodes expérimentales permettant d'identifier le lexique partiel des motifs de liaison d'un FT à l'ADN et d'autre part les méthodes de prédiction de l'ensemble de ces motifs en contexte génomique.

#### 1. Détermination expérimentale des séquences liées par un TF

*Méthodes biochimiques à haut-débit pour déterminer in vitro les sites de liaison d'un FT*

Plusieurs méthodes biochimiques ou de biologie moléculaire permettent d'obtenir sans *a priori* les motifs d'ADN liés par un facteur de transcription *in vitro*. Deux méthodes principales existent : l'évolution systématique de ligands par enrichissement exponentiel (*Systematic Evolution of Ligands by*

*EXponential enrichment* ; SELEX) et les puces d'hybridation de protéines (*Protein Binding Micro-array* ; PBM).

Pour le SELEX, des oligonucléotides de séquences aléatoires sont sélectionnés par affinité pour le FT. Les oligonucléotides ainsi sélectionnés sont amplifiés par une étape de PCR. La répétition de ces deux étapes pendant plusieurs cycles permet d'enrichir considérablement la bibliothèque d'oligonucléotides en séquences ayant une haute affinité pour le FT. Cette bibliothèque d'oligonucléotides est ensuite séquencée pour identifier les séquences liées par le FT (Moyroud et al., 2011; Chai et al., 2011).

Le PBM consiste en l'hybridation d'un FT sur une puce ADN contenant la plupart des combinaisons de 10 à 12 nucléotides. L'hybridation d'un anticorps fluorescent spécifique du FT permet d'identifier précisément les séquences liées (Franco-Zorrilla and Solano, 2014).

Ces deux méthodes donnent des résultats similaires dans la plupart des cas et permettent de prédire de façon relativement solide les sites de liaisons *in vivo* (Orenstein and Shamir, 2014). La différence majeure entre ces deux techniques est leur débit. Le SELEX a été appliqué avec succès à plusieurs FT chez les plantes (pour des exemples voir Moyroud et al., 2011; Busch et al., 2010; Hosoda et al., 2002; Nole-Wilson and Krizek, 2000) mais nécessite un travail d'optimisation conséquent pour chaque facteur. En revanche le PBM est relativement simple à mettre en place et a permis l'analyse de nombreux FT chez *Arabidopsis* (Franco-Zorrilla et al., 2014; Weirauch et al., 2014). La principale limitation du PBM réside dans la taille des motifs pouvant être identifiés. La plupart des motifs déterminés par PBM correspondent à des séquences de 8 nucléotides environ alors que le SELEX a permis l'identification de motifs beaucoup plus longs, comme celui de LFY qui correspond à une séquence de 19 nucléotides (Moyroud et al., 2011; Franco-Zorrilla et al., 2014; Weirauch et al., 2014; Stormo and Zhao, 2010).

#### *La purification d'ADN génomique par affinité – CHIP in vitro ou DAPseq*

Une amélioration récente de ces techniques est la purification, non plus d'oligonucléotides mais de fragments d'ADN génomique isolés de noyaux, par affinité pour le FT d'intérêt produit *in vitro*. Cette purification est couplée au

séquençage haut-débit des fragments purifiés (*DNA Affinity Purification-sequencing* ; DAPseq) (O'Malley et al., 2016). Cette méthode présente de nombreux intérêts comparés au SELEX ou au PBM : les fragments d'ADN testés sont issus directement du génome, ils sont longs d'environ 200 nucléotides ce qui permet d'identifier des motifs de liaison à l'ADN bien plus longs qu'avec le PBM. Cette technique permet ainsi de caractériser les propriétés d'espacement et d'orientation des sites de liaison proches : ainsi si ARF2 lie préférentiellement des sites inversés distants d'une quinzaine de nucléotides, ARF5 qui lui est structurellement très proche et reconnaît les mêmes motifs d'ADN, lie des sites répétés distant d'une dizaine de nucléotides ou des sites palindromiques distant de 17 nucléotides (O'Malley et al., 2016). Cette technique, utilisable à haut débit, permet donc de tenir compte du contexte génomique et de préciser certaines propriétés de liaison d'un FT à l'ADN.

#### *L'immuno-précipitation de la chromatine*

En parallèle à ces techniques *in vitro*, la technique d'immuno-précipitation de la chromatine (ChIP) permet de déterminer les sites de liaison d'un facteur de transcription *in vivo*. Des organes, tissus ou cellules sont fixés de façon à rendre covalentes les liaisons protéine-ADN et protéine-protéine (Yamaguchi et al., 2014b; Saleh et al., 2008; Yant et al., 2010). Le FT est alors immuno-précipité permettant la purification des fragments d'ADN qu'il lie *in vivo* dans un contexte biologique précis (Yamaguchi et al., 2014b; Saleh et al., 2008; Yant et al., 2010). Cette méthode permet ainsi l'identification des régions du génome liées par une protéine, qu'elle lie l'ADN directement ou *via* des cofacteurs. Cependant, contrairement à ce qui est obtenu *in vitro*, les sites de liaison identifiés ne constituent pas tous les sites de liaison potentiels du FT mais uniquement un sous-ensemble des sites liés par la protéine, dépendant du contexte biologique (tissu, stade de développement, conditions de cultures...).

## 2. Modélisation et prédiction des sites de liaison d'un FT

### *Modéliser l'affinité un FT pour l'ADN*

Une fois qu'un ensemble de séquences liées par un FT est connu, il est possible de construire un modèle décrivant l'affinité de la protéine pour l'ADN. Ces modèles sont communément appelés matrices de poids-position (*position weight matrix* PWM) ou matrice de score position-spécifique (*position specific scoring matrix* PSSM). Ces modèles sont construits à partir d'alignements de séquences liées et permettent d'assigner à un motif d'ADN donné un score représentant l'affinité relative du FT pour le motif. Il existe deux méthodes de construction des PWM : une méthode probabiliste, et une méthode thermodynamique. Pour chacune de ces méthodes une matrice de fréquence est établie, représentant la fréquence de chaque nucléotide  $i$  à chaque position  $j$  de l'alignement de séquence liées (Wasserman and Sandelin, 2004; Stormo, 2000).

Dans la modélisation probabiliste la fréquence est considérée comme la probabilité du nucléotide à cette position. Le score attribué à chaque paire nucléotide-position est calculé comme le log de la fréquence  $f_{i,j}$  de la base à cette position par rapport à la fréquence attendue  $f_b$  dans le génome. Le score  $W$  de chaque motif d'ADN est donc obtenu comme la somme du score des paires nucléotide-position dans le motif :

$$W_{(i|j)} = \sum_{j=1}^n \log \left( \frac{f_{i,j}}{f_b} \right)$$

Le score représente ici la probabilité du motif d'être lié par la protéine, un score nul ayant la même probabilité que le bruit de fond (Staden, 1984).

Dans la modélisation thermodynamique, la fréquence est considérée comme proportionnelle à l'énergie libre de liaison. Le score attribué à chaque paire nucléotide-position est calculé comme l'opposée (une énergie devant être positive) du log de la fréquence  $f_{i,j}$  du nucléotide par rapport au nucléotide favorisé (*i.e.* minimisant l'énergie de liaison) à cette position  $f_{Max,j}$ . Le score pour le motif complet est donc :

$$W_{(i|j)} = \sum_{j=1}^n -\log\left(\frac{f_{i,j}}{f_{Max,j}}\right)$$

Le score obtenu est donc positif ou nul, avec un score nul pour le motif minimisant l'énergie de liaison (Berg and von Hippel, 1987b). Certaines méthodes de calcul (*e.g.* MORPHEUS) utilisent un log positif, produisant donc un score négatif, représentant une pénalité énergétique (Minguet et al., 2015).

#### *Détermination des sites de liés par un FT*

Si le score permet d'obtenir une valeur représentant l'affinité du motif, il ne permet pas de savoir si le site est lié ou non. L'approche la plus commune pour décider de la liaison d'un motif est l'établissement d'un seuil au-delà duquel une séquence sera considérée comme liée. Il est possible *via* l'estimation de la distribution des scores à partir des fréquences de la matrice et du bruit de fond de déterminer un seuil satisfaisant un taux d'erreur de type I (faux-positifs) ou II (faux-négatifs) ou un rapport des deux (Sayou et al., 2016; Cock et al., 2009). Certaines méthodes prennent également en compte l'information contenue dans la matrice pour établir ce seuil (Hertz and Stormo, 1999). Lorsque des données de ChIP sont disponibles, il est également possible de déterminer le seuil maximisant la discrimination des séquences liées et non liées (Sayou et al., 2016). Il existe également des méthodes permettant de prédire la probabilité pour une séquence d'être liée (Granek and Clarke, 2005) ou d'évaluer le nombre de molécules liées sur une séquence (Roeder et al., 2007) à partir de la PWM. De manière intéressante, ces méthodes permettent de prendre en compte la coopérativité ou la compétitivité de liaison d'un ou plusieurs facteurs dans le calcul de la probabilité de liaison. Ces méthodes reposent en revanche sur l'estimation de la concentration en protéine, donnée rarement disponible *in vivo* et qui doit donc être définie comme une constante.

La principale limitation à l'utilisation des matrices pour la prédiction de sites de liaison reste souvent la difficulté à réduire le taux de faux-positifs et faux-négatifs. Certaines approches tentent de réduire ce problème en modifiant la manière dont sont construites les PWM. Les approches classiques

considèrent la contribution de chaque nucléotide comme indépendante pour la détermination de l'affinité du score. Des approches récentes utilisent une variété de modélisation différente : énergétique, chaînes cachées de Markov ou caractéristiques physiques. Toutes prennent en compte la participation interdépendante des différents nucléotides du motif ou sa topologie (Mathelier and Wasserman, 2013; Sharon et al., 2008; Zhao et al., 2012; Yang et al., 2014). Ces améliorations augmentent sensiblement la robustesse de la prédiction des sites de liaison, mais sont toujours limitées dans la prédiction des sites de liaison au niveau génomique. En effet, tous les sites de liaison d'un FT ne sont pas occupés en permanence et cette liaison peut dépendre de plusieurs facteurs : statuts chromatinien, présence de co-facteurs, etc... (Won et al., 2010; Whittington et al., 2009). L'élaboration de modèles prédictifs au niveau génomique passera donc pas l'incorporation de ce type de données. L'ensemble de ces techniques donnent accès, par identification directe ou prédiction (voir ci-dessous) aux sites de fixation sur l'ADN d'un FT. C'est une information clef qui comparée à des données de transcriptomique peut notamment aider à identifier parmi les gènes régulés par un FT, ceux qui en sont des cibles directes. Ce type de technique prédictive est également particulièrement utile dans le cas d'espèces non-modèles ou pour des études 'evo-devo' sur des nombreuses espèces différentes (Moyroud et al., 2011). Ces approches permettent également d'étudier les propriétés de liaison à l'ADN des FT (Sayou et al., 2016), ou d'identifier les régulateurs potentiels d'un gène cible (McLeay and Bailey, 2010).

- Différentes techniques permettent l'obtention d'un modèle de liaison d'un FT à l'ADN
- L'utilisation de ces modèles permet la prédiction des sites de liaison à l'ADN

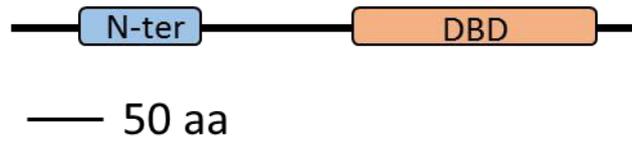
## IV. LEAFY, REGULATOR OF AXILLARY MERISTEM 1 et REVOLUTA, trois facteurs de transcription impliqués dans la formation des fleurs ?

### 1. LEAFY, régulateur central du développement floral

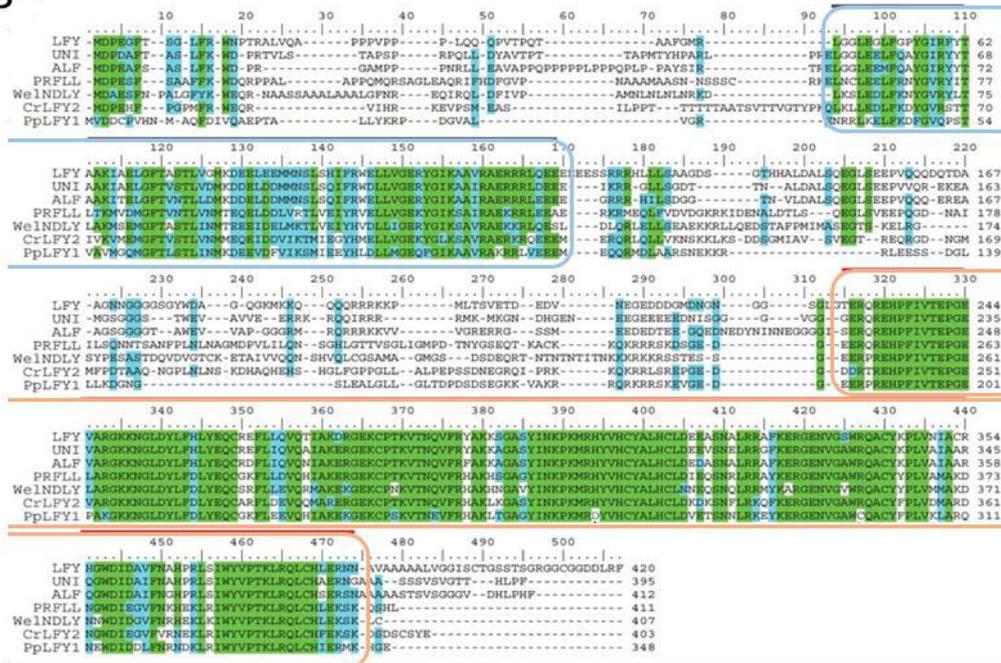
#### *LFY, un facteur de transcription conservé*

LFY est présent de manière ubiquitaire dans la lignée verte, des algues vertes charophytes aux angiospermes (Sayou et al., 2014). Si sa fonction dans les lignées les plus basales reste un mystère, sa fonction chez les angiospermes est très conservée. Chez les plantes modèles pour lesquels des mutants sont disponibles comme *Arabidopsis*, le maïs, le pétunia, la gueule-de-loup ou la tomate, les mutants *lfy* provoquent une conversion des fleurs en tiges, montrant un rôle conservé dans la détermination de l'identité florale (Molinero-Rosales et al., 1999; Souer et al., 1998; Bomblies et al., 2003; Coen et al., 1990; Weigel et al., 1992). Son rôle dans la formation des méristèmes semble également conservé chez le riz, le pois ou le maïs, la mutation des homologues de LFY dans ces espèces réduisant la complexité des panicules de riz et maïs et des feuilles de pois (Moyroud et al., 2010). Chez la mousse *Physcomitrella patens*, les deux homologues de LFY, PpLFY1 et 2, contrôlent la première division du zygote (Tanahashi et al., 2005). PpLFY1 complémente mal la formation des fleurs chez les mutants *lfy* d'*Arabidopsis*, indiquant une divergence fonctionnelle entre les plantes basales et les angiospermes (Maizel et al., 2005). Chez les gymnospermes LFY a un orthologue, NEEDLY (NLY). La fonction de NLY et LFY chez les gymnospermes est inconnue mais leur expression pourrait correspondre à un rôle dans le développement ou l'identité des cônes mâles et femelles (Vázquez-Lobo et al., 2007; Mellerowicz et al., 1998).

A



B

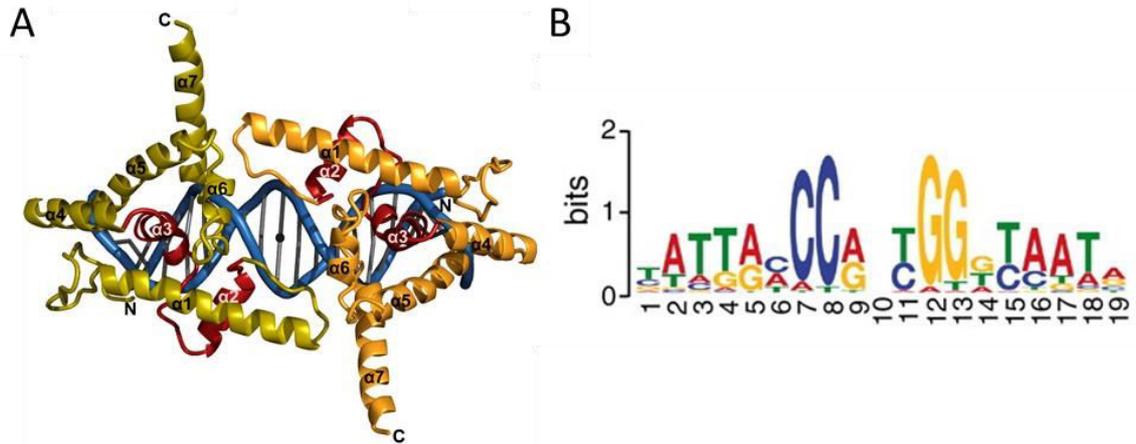


**Figure Intro. 10. Conservation de LFY chez les plantes terrestres.** (A) Représentation des domaines conservés de la protéine LFY. La fonction du domaine N-terminal (bleu) était inconnue au démarrage de ce travail de doctorat. Il s'agit d'un domaine d'oligomérisation de type SAM pour Sterile Alpha Motif (Sayou *et al.*, 2016). Le domaine C-terminal est un domaine de liaison à l'ADN de type HTH (orange). (B) Alignement des séquences protéiques de LFY chez les plantes terrestres (Adapté de Maizel *et al.*, 2005). Les domaines N- et C-terminaux sont encadrés en bleu et orange respectivement. LFY : *Arabidopsis thaliana*, UNI: *Pisum sativum*, ALF : Petunia, PRFLL : *Pinus radiata*, WeINDLY: *Welwitschia mirabilis*, CrLFY2 : *Ceratopteris richardii* , PpLFY1 : *Physcomitrella patens*.

LFY contient deux domaines très conservés en positions N- et C- terminales de la protéine (Figure Intro. 10). L'évolution de la liaison à l'ADN de LFY s'est faite principalement à travers des mutations discrètes du DBD alors que les multiples modifications de séquences du domaine peu conservé entre le DBD et le domaine N-terminal ont une influence minimale sur la fonction de la protéine (Sayou *et al.*, 2016; Maizel *et al.*, 2005).

### *Liaison de LFY à l'ADN*

Le domaine C-terminal est un domaine de liaison à l'ADN (*DNA binding domain*; DBD) fonctionnel, permettant la liaison au demi-site canonique (Figure Intro 11B). La résolution de sa structure a permis de révéler des similarités avec les domaines *helix-turn-helix* (HTH), typiquement impliqués dans les interactions protéine-ADN (Figure Intro. 11A) (Aravind et al., 2005; Hamès et al., 2008). Le DBD contient également une interface de dimérisation. Ainsi le DBD lie des sites pseudo-palindromiques composés de deux demi sites inversés. Les deux demi-sites sont séparés de trois bases, plaçant les deux monomères de DBD sur le même plan lors de la liaison à l'ADN (Hamès et al., 2008). La détermination du site de liaison par analyses des séquences liées en SELEX et en ChIP révèle un demi site de liaison de la forme NTGGNTAATN (Figure Intro. 11 B). On peut noter l'importance particulière des deux guanines contigus ainsi que la séquence non-aléatoire de l'espaceur. Bien que l'espaceur ne soit pas contacté directement par la protéine, sa séquence permet probablement de contraindre la topologie du site de liaison, en accord avec l'observation de la torsion particulière du brin d'ADN lié par LFY (Hamès et al., 2008; Moyroud et al., 2011). L'interface de dimérisation a une importance fonctionnelle sur la protéine. En effet un mutant de LFY touché spécifiquement dans cette interface présente une fonction florale réduite. La morphologie des fleurs est affectée et lorsque que la protéine est exprimée de manière constitutive, les fleurs ectopiques caractéristiques du 35S :LFY sont remplacées par des tiges (Chahtane et al., 2013). Les acides aminés impliqués dans la dimérisation ne sont pas conservés chez les homologues algaux de LFY (Sayou et al., 2014). En accord avec cette observation, le motif d'ADN reconnue par KsLFY (*Klebsormidium subtile*) ne comporte pas la séquence d'espaceur, les demi-sites étant directement accolés. Cette conformation résulte probablement en la liaison de deux monomères de LFY possédant un contact protéine-protéine différent (Sayou et al., 2014).



**Figure Intro. 11 . Propriétés du domaine de liaison à l'ADN de LFY.** (A) Modèle d'un dimère de liaison à l'ADN (DBD) de LFY lié à un site de liaison dans le promoteur d'*AP1* (bleu). Chaque monomère (jaune et orange) est composé de 7 hélices dont deux ( $\alpha 2$  et  $\alpha 3$ , rouge) contactent le grand sillon de l'ADN (Adapté de Hamès *et al.*, 2008). (B) Contenu d'information (en bits) de la matrice de liaison à l'ADN de LFY. La hauteur des lettres correspond à la probabilité du nucléotide à cette position. Chaque monomère reconnaît un site de 8 nucléotides, les deux sites sont séparés par un espaceur de 3 nucléotides (Adapté de Moyroud *et al.*, 2011).

#### *Stimulation de la formation des méristèmes par LFY*

Plusieurs combinaisons de mutants incluant des mutations dans *LFY* ont été rapportées comme étant affectées non pas dans l'identité florale, mais dans la mise en place des fleurs (Sawa *et al.*, 1999; Yamaguchi *et al.*, 2013; Wu *et al.*, 2015). Ces mutants produisent des filaments à la place des fleurs indiquant une implication de LFY et d'autres gènes tel que *ANT*, *AIL6* ou *FIL* dans l'établissement précoce de la morphogénèse florale et potentiellement dans la mise en place du FM. Il a de plus été montré récemment que l'expression ectopique d'une version de LFY avec une activité florale réduite est capable d'activer de manière précoce la formation de LSM à l'aisselle des feuilles de rosettes (Chahtane *et al.*, 2013). LFY régule plusieurs gènes potentiellement impliqués dans la mise en place des LSM tels que *RAX1*, *GRF5*, *ARR7* ou *TEL2*. Cependant la simple mutation de *RAX1* suffit à supprimer l'effet de LFY sur la stimulation de la formation des LSMs, indiquant que *RAX1* joue probablement un rôle majeur dans la fonction méristématique de LFY (Chahtane *et al.*, 2013). Si toutes ces données convergent pour indiquer un rôle de LFY dans la mise en place des FMs, ce rôle n'a jamais été décrit au niveau moléculaire.

## 2. Rôles de *RAX1* dans la mise en places des LSM

*RAX1* est un facteur de transcription de la famille des Myb R2R3, il contient un domaine de liaison à l'ADN dans sa partie N-terminal et sa partie C-terminale est prédite comme étant largement désordonnée (Müller et al., 2006; Keller et al., 2006). Son rôle potentiel dans la formation des fleurs n'a été évoqué que récemment (Chahtane et al., 2013). Initialement il a été identifié simultanément par deux approches différentes. *RAX1* est d'une part l'orthologue du gène *Blind* de la tomate, impliqué dans l'établissement des axes latéraux (Schmitz et al., 2002; Müller et al., 2006). Par ailleurs il a été identifié lors d'un criblage d'activateurs de la formation de méristèmes (Keller et al., 2006). La présence d'un T-DNA contenant plusieurs copies du promoteur constitutif *35S* dans le promoteur de *RAX1* induit une réduction de la formation de tiges axillaires et la formation SAM végétatifs ectopiques à l'aisselle des feuilles (Keller et al., 2006).

*RAX1* est exprimé principalement dans la racine mais aussi à l'aisselle des feuilles de rosette et dans les zones de frontières entre les jeunes méristèmes floraux et le SAM, cependant les mutants *rax1* ne sont pas affectés dans la formation des fleurs (Keller et al., 2006; Müller et al., 2006). *RAX1* et ses proches homologues *RAX2* et *RAX3* régulent la formation des LSM, *RAX1* jouant un rôle plus important que ses homologues dans ce processus *via* l'activation notamment de *CUC2* (Müller et al., 2006; Keller et al., 2006). *RAX1* régule également de manière négative l'accumulation des GA, les mutants *rax1* fleurissant tardivement (Keller et al., 2006). L'activité de *RAX1* dans la mise place des LSM est redondante avec celle d'autres régulateurs tels que REGULATOR OF AXILLARY MERISTEM FORMATION (ROX) et LAS (Yang et al., 2012). De plus, l'effet de la mutation *rax1* sur la mise en place des LSM peut être compensée par la synthèse locale de CK (Wang et al., 2014b). *RAX1* est également impliqué dans la signalisation de l'acide abscissique (ABA). Les sur-expresses *RAX1* sont hypersensibles à l'apport exogène d'ABA dans les processus de fermeture des stomates, délais de floraison, inhibition de la germination, croissance des jeunes plantules et tolérance à la dessiccation. De plus l'induction des gènes de réponse à l'ABA par l'apport exogène d'hormone est drastiquement accrue chez les sur-expresses de *RAX1* (Yu

et al., 2016). En revanche les mutants *rax1* ne montrent pas de réponse particulière à l'apport d'ABA, indiquant que d'autres facteurs sont redondants dans cette réponse (Yu et al., 2016).

Dans l'inflorescence, l'expression de *RAX1* est au moins en partie régulée par LFY (Chahtane et al., 2013). Bien que l'augmentation d'expression de *RAX1* déclenche la formation précoce de LSM chez les sur-expressseurs LFY (Chahtane et al., 2013) et qu'il soit exprimé à des stades compatibles avec la formation des FM dans les jeunes fleurs (Keller et al., 2006; Müller et al., 2006), son rôle dans la mise en place des FM reste inconnu. De même son mode d'action dans la formation des LSM reste très peu caractérisé.

### 3. *REVOLUTA*, un gène de développement aux fonctions complexes

REV est un FT de la famille des HD-ZIPIII, qui compte cinq membres : REV/IFL1, PHB/ATHB14, PHV/ATHB9, CNA/ATHB15 et ATHB8, dont les domaines d'expression se superposent largement (Prigge et al., 2005). Les relations au sein de cette famille de gènes sont complexes, les différents membres ayant des rôles complémentaires ou antagonistes selon les tissus et les processus observés (Prigge et al., 2005). Dans l'écotype *Nossen*, les mutants *rev* ont les feuilles très enroulées et anormalement larges, la tige est souvent plus épaisse que chez les plantes sauvages et de nombreux organes axillaires (tiges ou fleurs) sont remplacés par des filaments (Talbert et al., 1995; Otsuga et al., 2001; Zhong and Ye, 1999). En accord avec ces phénotypes, *REV* est exprimé dans le cambium inter-fasciculaire où il régule la différenciation des fibres interfasciculaire et réprime l'activité du méristème cambial (Zhong and Ye, 1999; Talbert et al., 1995; Liu et al., 2014). *REV* est également exprimé dans la partie adaxiale des primordia floraux où il limite la prolifération cellulaire et dans la partie adaxiale des AM où il semble nécessaire pour le développement de l'organe axillaire (Otsuga et al., 2001; Talbert et al., 1995). Le phénotype des mutants *rev* semble très écotype-dépendant : dans *Landsberg erecta* (Ler), les mutants *rev* produisent des fleurs normales et occasionnellement des fleurs sans carpelles et étamines, indiquant l'arrêt précoce de l'activité du méristème floral (Otsuga et al., 2001). De plus *REV* est requis pour la formation de méristèmes adventifs chez les

mutants *wus* et *stm*, indiquant là encore un rôle dans la régulation de l'activité méristématique (Otsuga et al., 2001). L'expression de *REV* dans les SAM reproductifs prédit la position des primordia floraux, indiquant que *REV* pourrait être impliqué dans le développement précoces des FM (Otsuga et al., 2001). L'effet des mutations dans *REV* est aggravé par des mutations dans différents composants du protéasome 26S, provoquant la formation de nombreux filaments en place des fleurs et l'arrêt de l'expression de *WUS* et *CLV3* dans les inflorescences et fleurs de plantes âgées (Zhang et al., 2011). Ces données indiquent que *REV* pourrait agir de manière redondante avec des voies de signalisation requérant la dégradation de protéines.

Au niveau moléculaire *REV* agit en partie *via* la régulation de la synthèse et du transport de l'auxine. Chez les mutants *rev* plusieurs transporteurs de l'auxine de la famille des PINs sont réprimés et les gènes de synthèses de l'auxine *YUC5* et *TAA1* sont activés (Zhong and Ye, 2001; Brandt et al., 2012). *REV* active également les gènes *HEC*, impliqués la régulation de l'activité du méristème et l'enzyme de synthèse des cytokinines *IPT7* ainsi que des gènes en lien avec la signalisation de l'ABA, des brassinostéroïdes et régulant la croissance et la différenciation cellulaire (Reinhart et al., 2013).

Au niveau protéique *REV* reste très peu caractérisé. En plus du domaine de liaison à l'ADN de type homéodomaine leucine-zipper (HD-ZIP) caractérisant la famille, les HD-ZIPIII contiennent un domaine *StAR-related lipid transfer* (START) (Prigge et al., 2005). Les domaines START peuvent lier les stéroïdes ou la phosphatidyl-choline, permettant de réguler la localisation et/ou l'activité de la protéine (Ponting and Aravind, 1999). Il a été proposé qu'un signal dérivé du centre (du SAM ou de la tige) active les HD-ZIPIII dans la partie adaxiale des organes axillaires et dans les tiges (Emery et al., 2003). Ce signal hypothétique pourrait donc être un lipide transporté de manière centripète vers les zones d'initiation d'organes.





## OBJECTIFS

---

Le facteur de transcription LEAFY est le maître régulateur de l'identité florale mais régule aussi d'autres processus liés à la formation des fleurs : transition florale, émergence du méristème floral, élongation du pédicelle, répression de la bractée, terminaison des cellules souches florales. Lors du développement de la fleur la régulation de l'activité du méristème est primordiale. Si le méristème est absent ou trop peu actif la fleur sera incomplète et stérile. Si la durée de vie du méristème est trop longue, la fleur produira trop d'organes, voire une autre fleur à l'intérieur de la précédente, au risque de diminuer l'efficacité de la reproduction.

Au cours de ce travail, j'ai exploré la fonction de LEAFY à deux niveaux dans la régulation du méristème floral : lors de sa mise en place après émergence du primordium, et lors d'une étape nécessaire à sa terminaison, le déclenchement de l'activité d'AG qui aboutit par rétrocontrôle à l'extinction de WUS.

Le chapitre premier présente ainsi la plus grosse partie de mon travail de thèse, sur le rôle de LEAFY et un de ses gènes cibles, *RAX1*, dans la mise en place du méristème floral, en parallèle de *REVOLUTA*.

Le second chapitre présente un travail plus exploratoire effectué d'une part sur les propriétés d'oligomérisation de LEAFY et d'autre part sur le rôle potentiel de LEAFY et son co-régulateur *WUSCHEL* dans la régulation d'*AGAMOUS*.



# CHAPITRE PREMIER

## ROLE DE LFY ET RAX1 DANS L'ETABLISSEMENT DE LA NICHE DE CELLULES SOUCHES FLORALES

### I. Introduction

De nombreuses données éparses de la littérature indiquent un rôle de LFY dans la mise en place des FM (Moyroud *et al.*, 2010). D'une part les combinaisons de mutants comprenant *lfy* sont affectées dans l'initiation des fleurs (Yamaguchi *et al.*, 2013; Wu *et al.*, 2015; Sawa *et al.*, 1999). D'autre part chez certaines espèces comme le pois ou le riz, les mutations des homologues de *LFY* provoquent une réduction de la complexité des organes latéraux, indiquant un rôle dans le contrôle de l'architecture de la plante (Moyroud *et al.*, 2010). Il a de plus été montré avant mon arrivé au laboratoire que LFY peut stimuler de manière ectopique la formation des LSM *via*

l'induction de *RAX1* (Chahtane *et al.*, 2013). Cependant, le rôle de LFY dans la mise en place des FM n'a jamais été formellement étudié.

Ni les mutants *lfy*, ni les mutants *rax1* ne sont affectés dans l'*initiation* des méristèmes, indiquant que d'autres voies parallèles peuvent compenser pour la perte de leur activité. Nous avons donc entrepris d'identifier dans la littérature des voies redondantes régulant la formation des FM. Nous avons concentré nos efforts sur le FT de la famille HD-ZIPIII, REV, qui régule d'une manière assez large la formation des méristèmes, aux stades embryonnaire, végétatif et reproductif (Otsuga *et al.*, 2001; Talbert *et al.*, 1995; Prigge *et al.*, 2005). En utilisant des mutants *rev* comme fond génétique sensibilisé, nous avons étudié le rôle de LFY et RAX1 dans la mise en place des FM. Les résultats de cette étude, qui a constitué mon principal sujet au cours de la thèse, sont exposés dans l'article 2 (*en préparation*), dans cette partie.

## II. Article 2 : "Control of stem-cell niche establishment in Arabidopsis flowers by LEAFY, its target RAX1, and REVOLUTA"

L'article présenté ici est en cours de rédaction pour être soumis prochainement.

# Control of stem-cell niche establishment in Arabidopsis flowers by LEAFY, its target RAX1, and REVOLUTA

Denay Grégoire<sup>1</sup>, Tichtinsky Gabrielle<sup>1</sup>, Le Masson Marie<sup>1</sup>, Chahtane Hicham<sup>1,§</sup>, Huguet Sylvie<sup>2</sup>, Wenzl Christian<sup>3</sup>, Carles Cristel<sup>1</sup>, Franco-Zorrilla José Manuel<sup>4</sup>, Lohmann Jan<sup>3</sup> and Parcy François<sup>1,\*</sup>

1 - Laboratoire de Physiologie Cellulaire et Végétale, LPCV, BIG, Univ. Grenoble Alpes, CNRS, CEA, INRA, LPCV, F38000 Grenoble, France. 2 - Institute of Plant Sciences Paris-Saclay IPS2, CNRS, INRA, Univ Paris Sud, Univ Evry, Univ Paris-Diderot, Sorbonne Paris-Cite, Université Paris-Saclay, Rue de Noetzlin, 91405 Orsay, France. 3 - Centre for Organismal Studies, Heidelberg University, Heidelberg, Baden-Württemberg, Germany. 4 - Genomics Unit, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, 28049 Madrid, Spain. § - Present address: Department of Plant Biology, University of Geneva, 30 Quai Ernest-Ansermet-Sciences III, 1211, Geneva 4, Switzerland. \* - Corresponding author: francois.parcy@cea.fr

Article in preparation

## Abstract

**Plants keep the ability of producing organs throughout their life by maintaining active pools of stem cells. These stem cells are organized in specialized structures called meristems. Flowers are formed from cells leaving the shoot apical meristem and initially do not contain stem cells nor express meristem markers. A meristem must be *de novo* established within the flower to enable subsequent growth. However, the molecular mechanisms by which flower meristem is established remain elusive. We show here that the transcription factors LEAFY and REVOLUTA control two redundant pathways leading to stemness acquisition in flowers. We found that LFY acts through the transcription factor REGULATOR OF AXILLARY MERISTEM1 that controls meristem domain formation together with REV and we provide mechanistic insights in how RAX1 allows meristem identity establishment in young flowers. Our work provides a molecular link between the process of flower initiation, meristem formation and identity determination.**

## Introduction

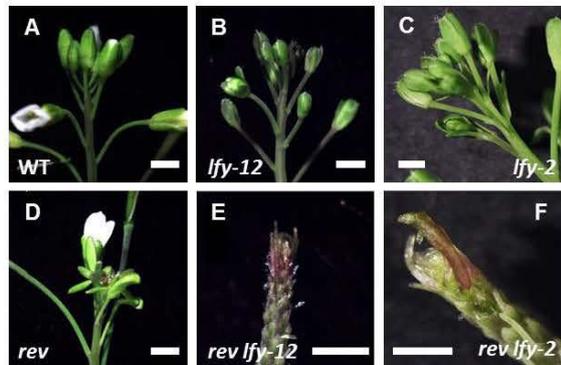
Plants retain the capacity to initiate new organs throughout their life. This is made possible by the maintenance of self-sustaining pools of stem cells organized in structures called meristems. The progeny of the stem cells contained in the shoot apical meristem (SAM) gives rise to all the post-embryonic aerial organs. Stem cell maintenance in the SAM involves a regulatory feedback loop between the transcription factor (TF) WUSCHEL (WUS) and the CLAVATA (CLV) signalling module (Schoof et al., 2000). WUS is expressed in the embryo from the 16-cells stage onward and becomes restricted to the organizing centre (OC) of the SAM, located a few cell layers below the epidermis (Mayer et al., 1998). WUS induces stem-cell fate (stemness) in the central zone (CZ), occupying the upper cell layers (L1 to L3) of the SAM, in a non-cell autonomous manner (Daum et al., 2014; Gallois et al., 2004; Yadav and Reddy, 2012). Additionally, WUS induces the expression of the signalling peptide CLAVATA3 (CLV3) in the CZ, which then diffuses and binds its receptor CLAVATA1 (CLV1) in the cells surrounding the OC (Fletcher et al., 1999; Lenhard and Laux, 2003; Schoof et al., 2000). CLV1 negatively regulates WUS expression (Schoof et al., 2000), thereby restricting the OC and forming a negative feedback loop controlling stem cell homeostasis. The hormone cytokinin (CK) is also important for meristem maintenance by regulating both WUS and CLV3 expression and promoting stemness (Gordon et al., 2009; Gruel et al., 2016; Leibfried et al., 2005; Zhao et al., 2010).

Flowers are initiated the peripheral zone (PZ) of the SAM, from the progeny of the stem cells as they leave the CZ.

The very young flowers (stages 0 and 1 from Smyth et al., 1990) initially do not express any of the meristem markers, however the expression of WUS, closely followed by that of CLV3, becomes visible from floral stage 2 (Yoshida et al., 2011). *De novo* meristem formation in flowers is especially important as failure to do so impedes reproduction (Laux et al., 1996; Lohmann et al., 2001). Although a lot of efforts were made to understand the transcriptional and spatial control of flower initiation (Besnard et al., 2014; Sassi and Vernoux, 2013; Wu et al., 2015), our understanding of FM formation remains partial.

Cell-fate transition to flower formation is determined by the activity of the MONOPTEROS (MP) TF as *mp* mutants largely fail to initiate flowers, resulting in a pin-shaped inflorescence (Yamaguchi et al., 2013). Several MP-controlled pathways are thus expected to play a role in FM initiation (Denay et al., 2017). Accordingly, some of its direct targets show flower initiation defects in specific mutant combinations such as *lfy ant ail6* or *lfy fil* (Sawa et al., 1999; Wu et al., 2015; Yamaguchi et al., 2013). However, neither of these TFs has been directly linked to meristem initiation. Current models predict the FM to be established as the result of CK synthesis in the L1 of the newly initiated organ once it reaches a threshold size (Gruel et al., 2016). Consistently, ANT and AIL6 were shown to stimulate cell-proliferation and cell-wall softening (Krzek et al., 2016), while both LFY and MP regulate CK signalling (Chahtane et al., 2013; Zhao et al., 2010). It was also shown that LFY ectopic expression can stimulate meristem formation at the axils of rosette leaves, through the upregulation of the Myb TF REGULATOR OF AXILLARY MERISTEM1 (RAX1) (Chahtane et al., 2013).

RAX1 is also expressed in young flower buds (Keller et al., 2006; Müller et al., 2006), however neither *lfy* nor *rax1* mutants show defects in the flower initiation per se - even though they lack floral identity in the case of *lfy* mutants - suggesting that several redundant pathways might control the initiation of FMs.



**Figure 1** **Figure 1. LFY and REV act in parallel pathways to control flower formation.** Representative inflorescences of WT (A), *lfy-12* (B), *lfy-2* (C), CRISPR-mutated line in REV (D) and combinations of *lfy-12 rev* (E) and *lfy-2 rev* (F). Scale bar: 1 mm.

The HD-ZIP/III family of TFs has been linked to *de novo* meristem formation in aerial tissues. Triple mutants *revoluta phabulosa phavoluta* (*rev phb phv*) fail to form an embryonic SAM (Prigge et al., 2005) and single *rev* mutants show pleiotropic defects (Otsuga et al., 2001; Talbert et al., 1995). These defects include failure to form axillary stems and occasionally flowers, resulting in the formation of filaments or a flowers looking like those of *wus* mutants (Laux et al., 1996; Otsuga et al., 2001). Additionally, REV was shown recently to be an essential component of axillary shoot meristem formation by stimulating the expression of *SHOOTMERISTEMLESS* (*STM*) in the leaf axils (Shi et al., 2016). Thereby preventing cell differentiation and locally increasing CK synthesis, thus allowing meristem formation (Endrizzi et al., 1996; Jasinski et al., 2005; Wang et al., 2014; Yanai et al., 2005).

Because REV appears as a possible candidate to act in parallel of the LFY/RAX1 pathway in flowers, we combined mutations in both pathways to study their potential role in flower meristem initiation. We show that LFY indeed regulates flower meristem formation *via* the regulation of *RAX1*. Additionally, we show that *RAX1* and REV control the establishment of FM domains and that *RAX1* may allow meristem establishment by repressing *CLV1* expression in the young flower.

## Results

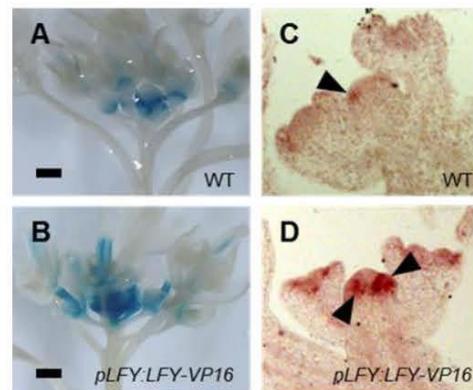
### *LFY and REV act together during flower formation*

In order to determine if REV could act redundantly with LFY in the process of flower initiation we combined mutations in *LFY* and *REV*. As *LFY* and *REV* are genetically linked we generated a CRISPR/Cas9 construct targeting the third exon of *REV* (Figure S1B). We recovered several mutants carrying insertions or deletions provoking the apparition of an early stop codon (Figures S1D, F). All lines recovered showed similar phenotypes to the *rev-6* mutant in the Landsberg *erecta* (*Ler*) accession (Otsuga et al., 2001) and a representative one was further

characterized (thereafter named *rev-ct1*): leaves were slightly curved downward, the number of axillary stems was reduced and a small proportion of flowers lacked internal whorls or were replaced by filaments (Figures S2A, C, D). When combined to the weak *lfy-2* or the strong *lfy-12* alleles, flowers were extensively replaced by filaments, whereas *lfy* mutants are only affected in flower identity but initiate “leafy flowers” normally (Figure 1). This indicates that REV and LFY act in parallel pathways that control flower initiation. Moreover, LFY activity seems critical in a *rev* background as even a weak allele of *LFY* almost completely prevented flower formation. Thus *rev* mutants appear as an excellent sensitized background to study a LFY-dependent pathway regulating FM formation.

### *LFY regulates RAX1 expression in flowers*

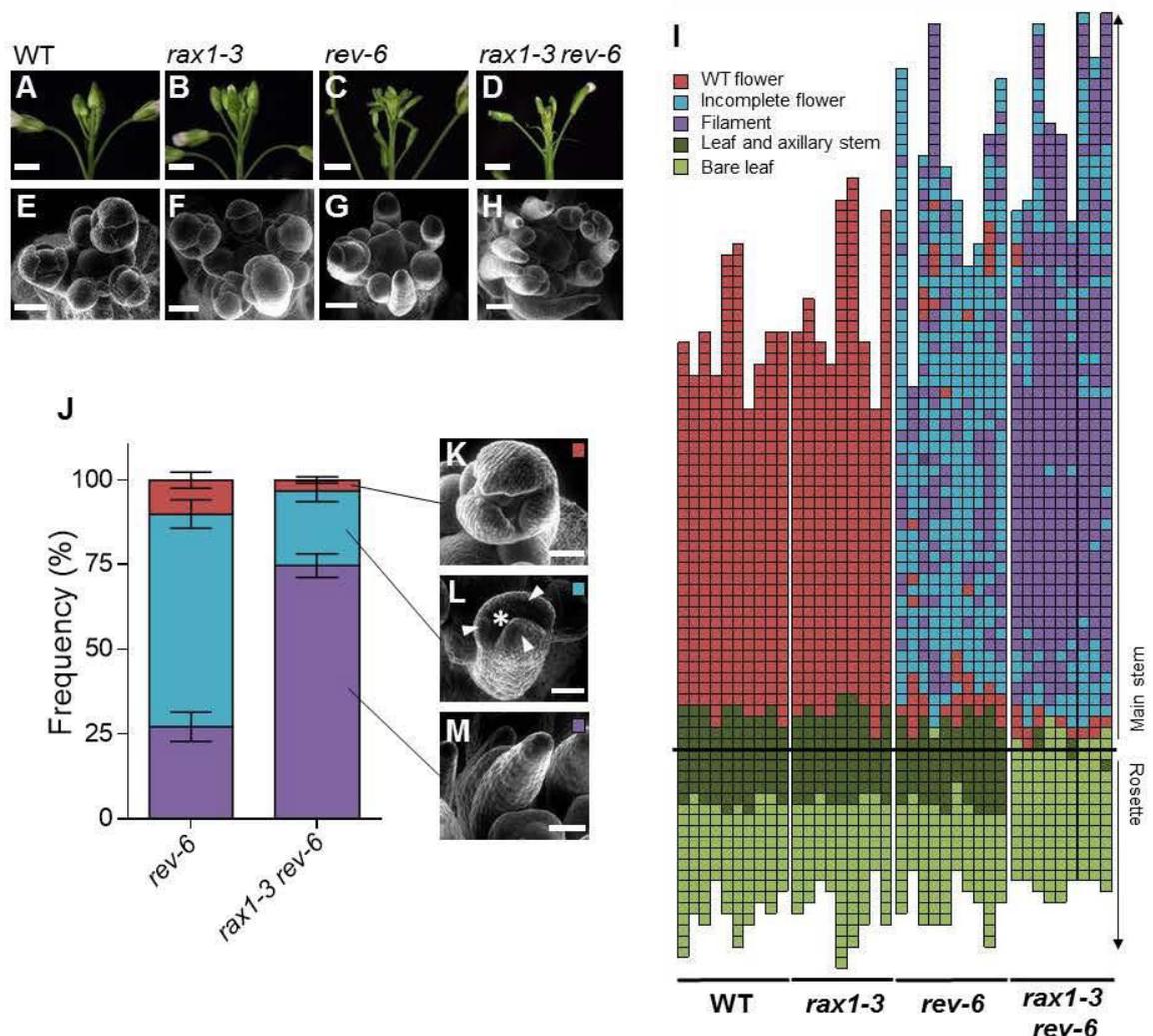
As it was previously shown that LFY could ectopically stimulate meristem formation by inducing *RAX1* (Chahtane et al., 2013), we checked if LFY could also stimulate *RAX1* expression in flower buds. *RAX1* is expressed in the boundaries between the SAM and very young flower buds (Keller et al., 2006; Müller et al., 2006). Using a GUS transcriptional reporter under the control of *RAX1* promoter, we showed that *RAX1* is also expressed at the abscission zones between the floral organs and the pedicel, and in the replum (Figure 2A). When expressed together with LFY fused to the trans-activation domain VP16 under the control of the endogenous *LFY* promoter, GUS reporter expression was broadly stronger and its expression domain was extended in the pedicel of young flowers (Figure 2B). In the SAM, *RAX1* expression was also accentuated by the expression of *pLFY:LFY-VP16* and *RAX1* expression could be detected within the flower buds (Figures 2C-D). This indicates that LFY is a potent regulator of *RAX1* in the inflorescence and can induce its expression in young flower buds.



**Figure 2** **LFY induces RAX1 expression in inflorescences.** (A, B) GUS staining of WT (A) or *pLFY:LFY-VP16* (B) plants expressing a *pRAX1:GUS* construct (scale bar: 1 mm). (C, D) *In situ* hybridization with a *RAX1* antisense probe in the WT (C) or *pLFY:LFY-VP16* (D) background. Black arrowheads indicate dig-labelled probe signal.

### *RAX1 and REV redundantly regulate flower formation*

*rax1* mutants do not display a phenotype in flowers (Keller et al., 2006; Müller et al., 2006), we therefore introduced the *rax1-3* mutation in the *rev-6* sensitized background in order to identify a putative role of *RAX1* in flower formation. As *rev-6* was isolated in the *Ler* background,



**Figure 3. RAX1 and REV act together to control flower initiation.** (A-H) Inflorescences of WT (A, E), *rax1-3* (B, F), *rev-6* (C, G) and *rax1-3 rev-6* (D, H) plants, observed by light (A-D; scale bar: 2 mm) or scanning electron microscopy (E-H; scale bar: 100  $\mu$ m). (I) Plant architecture diagram of *rax1-3* and *rev-6* single and double mutants three weeks after bolting. Each column represents an individual plant and each square represent a single internode. Squares below the thick black line represent internodes on the rosette and squares above the thick black line, internodes on the main stem. Structures initiated are color-coded. Dark green: leaf and axillary stem; light green: leaf lacking an axillary stem; red: wild-type flower; blue: flower lacking one or more internal whorls; purple: filamentous structure. (J) Proportion of floral structures initiated in *rev-6* and *rax1-3 rev-6* mutants. Color code is the same as for (I). (K-M) Typical structures scored in (J) observed by scanning electron microscopy: WT flower bud (K), flower bud lacking a meristematic dome (asterisk) between the sepals (white arrowheads) (L) and filament (M). Scale bar: 100  $\mu$ m.

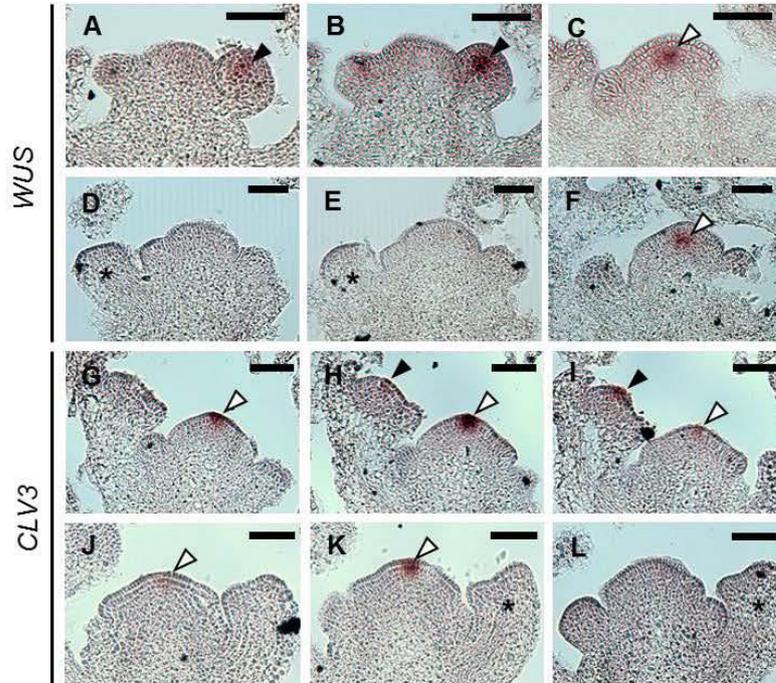
mutants were first backcrossed at least 3 times to the Columbia (Col) accession. The backcrossed plants were very similar to the phenotypes described earlier for *rev-6* (Otsuga et al., 2001) but lacked *Ler* hallmarks such as compact inflorescences, short stems and round leaves. However, the frequency of floral defects (*i.e.* lack of internal whorls and filaments) was increased compared to the *rev-c1* allele (Compare Figures 3J, I and S2C, E). The plants produced very few fertile flowers. When combined to the *rax1-3* mutation the *rev-6* phenotypes were considerably enhanced (Figures 3A-H, I). The proportion of fertile flowers was further decreased and most of the structures generated consisted of filaments. Additionally, the plants only occasionally formed axillary stems and growth was prolonged, resulting in the formation of an

abnormally long main stem (Figure S3). The plants also exhibited thickened and overgrown leaves as well as thicker stems that were reminiscent of the phenotypes described for the *rev-1* allele in the Nossen accession (Figure S3) (Talbert et al., 1995).

Observation of the young flower buds by scanning electron microscopy revealed that the lack of internal organs in some of the observed flowers most likely stemmed from the absence of a morphologically distinct meristematic dome between the sepal primordia (Figures 3E-H, K-M), preventing the formation of inner whorls. The filamentous structures did not exhibit any feature distinctive of a particular floral tissue, thus we hypothesize that they are formed by an aborting primordium that fails establishing a

meristem. Consistently, when treated with the auxin polar transport inhibitor naptalam (NPA), *rev-6* mutants formed less floral structures and this effect was accentuated in the double mutant, indicating that *RAX1* and *REV* act in an auxin-controlled process, like flower initiation (Figure S4).

flowers of the same stage that lacked a meristematic dome. These FMs lacking meristem marker expression are likely the FMs that will form empty flowers or filamentous structures.



**Figure 4. Aberrant flower meristem formation in *rax1 rev* mutants.** Expression patterns of *WUS* (A-F) and *CLV3* (G-L) observed by in situ hybridization in successive *rax1-3 rev-6* inflorescence sections. (A-C) Inflorescence showing WT *WUS* expression in a stage 2 flower bud (black arrowhead) and in the SAM (white arrowhead). (D-F) Inflorescence showing the lack of *WUS* expression in a stage 2 flower bud (star) whereas *WUS* is normally expressed in the SAM. (G-I) Inflorescence showing WT expression of *CLV3* in a stage 3 flower bud (black arrowhead) and in the SAM (white arrowhead). (J-L) Inflorescence showing the absence of *CLV3* expression in a stage 3 flower bud lacking a meristematic dome (star) but showing WT expression in the SAM. Scale bar: 50  $\mu$ m.

Observation of a segregating F2 population from the *rax1-3 rev-6* cross showed that the *rev-6* background was highly sensitized to variations in the number of *RAX1* active copies. *rev-6 rax1-3/RAX1* plants were more affected than *rev-6* mutants in both fertile flowers and axillary stems production (Table S1). Additionally, non-inductive short days conditions, where the expression of *LFY* is decreased (Müller-Xing et al., 2015), drastically enhanced the flower defects of both the single *rev-6* and double *rev-6 rax1-3* mutants (Figures S5A-D). Whereas *rev-6* mutants developed a filament at the axil of a morphologically distinct bract, the filaments were smaller in the double mutants and their bracts bore stipules (Figures S5 E-H). This suggests that *RAX1* and *LFY* levels are critical in the *rev-6* background. Altogether these results show that *REV* and *RAX1* share a role in the early development of the floral bud and are necessary for the formation of a morphologically distinct meristematic structure. Although *rax1* mutants have no flower phenotype, the number of active *RAX1* copies is critical in *rev* mutants.

#### *RAX1* and *REV* are required for FM organization

Because the absence of a distinctive meristematic dome in flowers could be a sign of the absence of a FM, we monitored the expression of the meristem markers *CLV3* and *WUS* in flowers using *in situ* hybridization. In the SAM, *CLV3* and *WUS* expression was normal in all mutants and double mutant (Figure 4). In the double mutant some flowers showed WT expression of *WUS*, in the centre of FM in stage 2 flowers, and of *CLV3*, in the upper layers of the FM in stage 3 flowers. However, some other flowers showed a complete lack of *CLV3* and *WUS* expression in

OC we generated CRISPR/Cas9 constructs targeting *RAX1* and *REV*, directly in lines expressing a transcriptional reporter for *WUS* expression. The construct targeting *RAX1* was directed against the C-terminal part of the protein, before the site of T-DNA insertion in the *rax1-3* allele (Figure S1A). We recovered several lines carrying insertions or deletions that resulted in the apparition of an early stop codon (Figures S1C, E). As *rax1* mutants show no obvious phenotype we counted the number of axillary buds produced in non-inductive short days conditions, which is a reported effect of *rax1* mutations (Keller et al., 2006; Müller et al., 2006). *rax1-c1* plants produced approximately 10% less axillary buds than the control, indicating that *RAX1* function was disturbed in *rax1-c1* (Figure S2F). We also generated double *rax1 rev* mutants by combining both CRISPR/Cas9 constructs. For these mutants we could recover double heteroallelic plants (*rax1-c2/3 rev-c2/3*) When *rax* and *rev* mutations were combined, a large proportion of flowers lacked internal organs and a small proportion was converted to filaments (Figures S2A-E). The weaker phenotype observed for CRISPR-mutated lines compared to the *rax1-3 rev-6* line is consistent with the weaker phenotype of the *rev-c1* allele compared to *rev-6*.

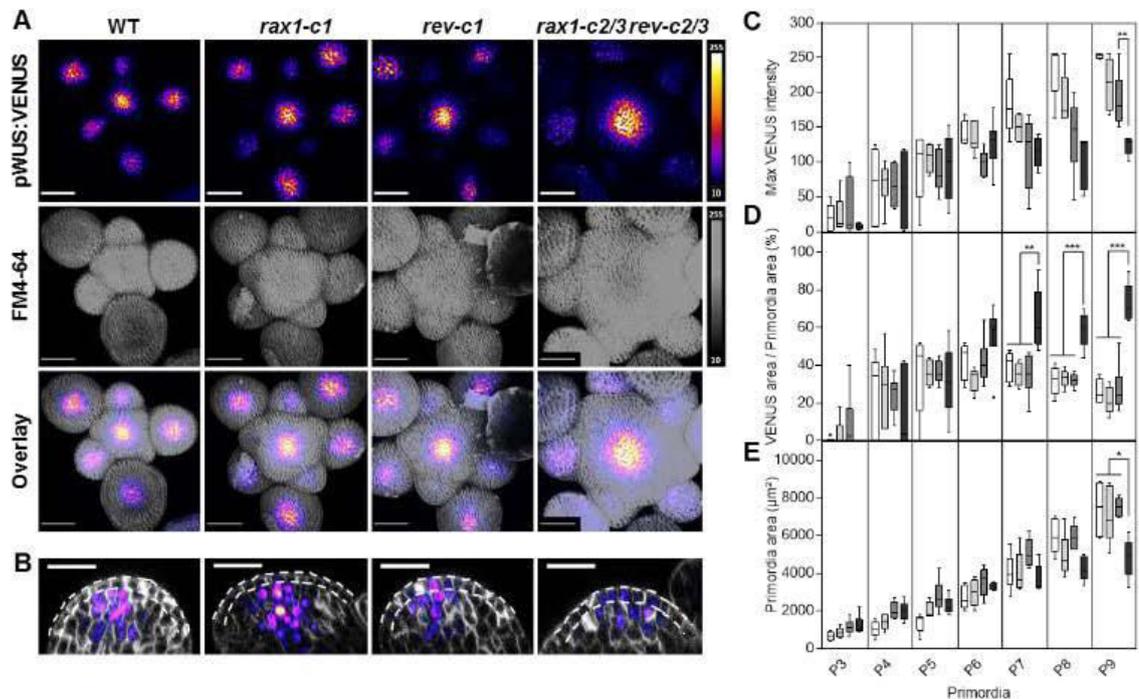
In *rax1-c1*, the expression of a *WUS* transcriptional reporter (*pWUS:2xVENUS*) was normal (Figure 5A). However, in *rev-c1* *WUS* expression was reduced in some flower buds and *WUS* was expressed in the L1 of 60% of the observed FMs (Figure 5A-C). In the double heteroallelic line *rax1-c2/3 rev-c2/3*, *WUS* expression was even further reduced in most of the FMs observed (Figures 5A, C). In all of these FM showing *WUS*

expression, WUS was expressed ectopically in the L1 and in the PZ, sometimes being expressed throughout the FM (Figures 5B, D). Additionally, FM growth was slightly delayed compared to either single mutant (Figure 5E). RAX1 and REV also showed an effect of the SAM morphology (Figure 5A). In the double mutants the SAM was drastically enlarged, so was the WUS expressing domain (Figures S6A-B). However, the size ratio of the SAM and the WUS-expression domain remained constant, resulting in a flatter SAM (Figures S6C-D). These observations are consistent with the existence of a scaling mechanism between the SAM and the OC size (Gruel et al., 2016).

Taken together these data indicate that RAX1 and REV are necessary for the proper initiation of the FM. Moreover, they are required for both restricting WUS expression in the OC and to stimulate its expression.

*RAX1 acts on hormonal, cell wall modification and defence response pathways*

constitutive *UBQ10* promoter (Geldner et al., 2009). The glucocorticoid receptor (GR) domain allows the retention of the fusion protein in the cytoplasm and its conditional translocation to the nucleus upon dexamethasone (DEX) treatment (Padidam, 2003). Protein expression was validated by western blot (Figure S7J) and protein translocation after DEX treatment was validated following mCHERRY fluorescence in root tips (data not shown). When RAX1 activity was induced by regular DEX treatments, plants appear stunted, retarded in growth, produced very small siliques with only a few seeds, and occasionally formed ectopic rosettes at the axils of cauline leaves, similarly to the *rax1-1D* activation mutant (Keller et al., 2006). Paired-end RNA-sequencing analysis was performed on 15 day-old seedlings 4h after DEX or mock treatment, yielding libraries of 26 to 37 million reads. Of these, 97% of read pairs mapped to unique regions annotated as exons or untranslated regions (TAIR10) with a mapping confidence score (MAPQ) of 30 or more. Principal component analysis of the top 500 variable genes across all datasets showed a clear segregation of



**Figure 5. Disturbed WUS expression pattern in *rax1 rev* mutants leads to primordia growth arrest.** (A,B) WUS expression pattern in representative inflorescences of CW039, *rax1-c1*, *rev-c1* and *rax1-c2/3 rev-c2/3* (left to right). (A) From top to bottom: z-maximal intensity projections of VENUS signal, membranes stained with FM4-64 and overlay. Scale bar: 50µm, relative color scales are indicated on the right side. (B) Orthogonal view of WUS expression pattern in representative floral primordia. The L1 is indicated between dashed lines. Scale bar: 25µm, color scales are the same as in (A). (C-E) Quantitative analysis in successive primordia (P3 to P9) of CW039 (white), *rax1-c1* (light-grey), *rev-c1* (medium grey) and *rax1-c2/3 rev-c2/3* (dark-grey) plants. Maximum VENUS intensity (C), relative WUS expression domain size (D), and primordia size (E). \* p < 0.05, \*\* p < 0.01, \*\*\*\* p < 0.001, determined by a Wilcoxon-Mann-Whitney test, N ≥ 5.

In order to gain insights into the process of FM establishment we examined the transcriptome of plants expressing a post-transcriptionally inducible version of RAX1 to compare it to existing similar data for REV (Reinhart et al., 2013). To do so we generated plants lines expressing either an mCHERRY-RAX1-GR or mCHERRY-GR fusion protein under the control of the

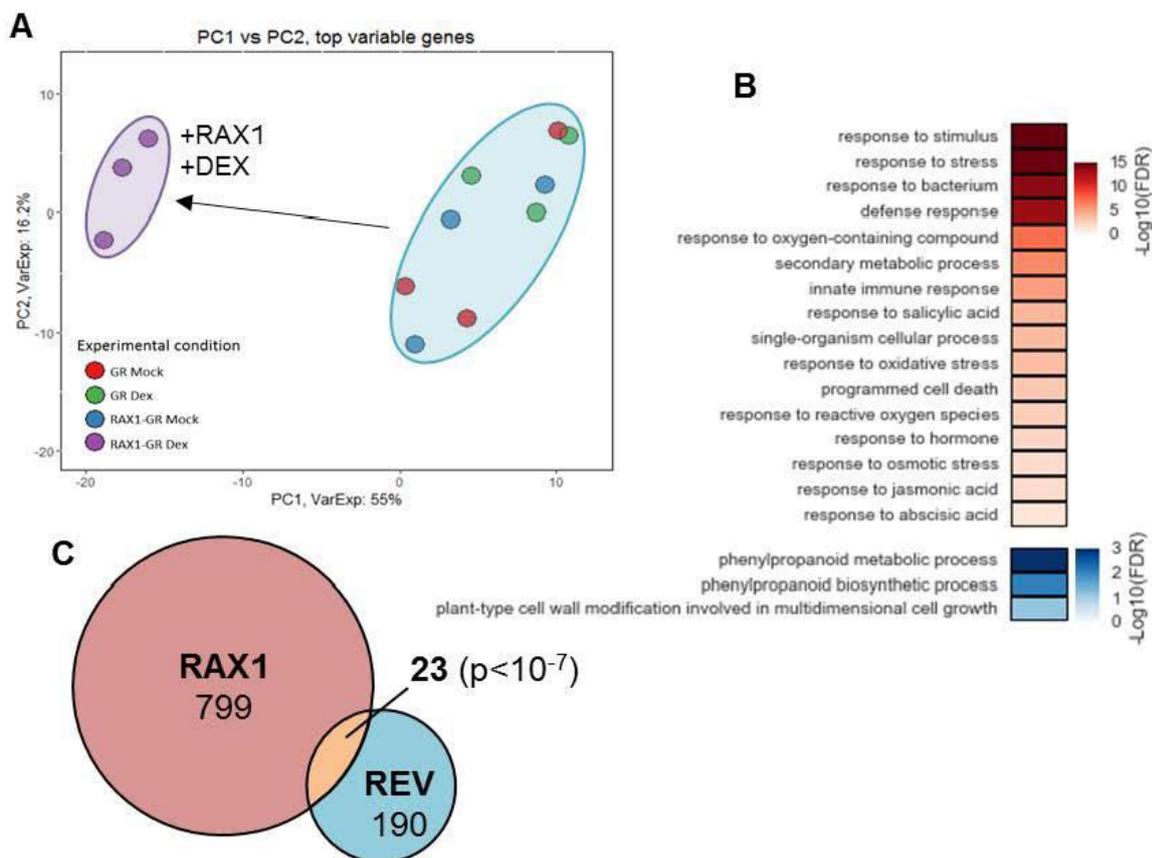
the DEX-treated RAX1 inducible lines along the first component axis (Figure 6A), indicating a specific transcriptome response in these samples. We identified 822 differentially expressed genes (DEG; FDR ≤ 0.01) in response to the treatment and the presence of RAX1 (Supplementary Dataset 1). Amongst them, 482 were down-regulated and 340 were up-regulated. Gene

ontology (GO) term enrichment of these genes showed an enrichment of genes involved in cell modifications and phenylpropanoid synthesis in genes up-regulated. The down-regulated genes were enriched in genes involved in the response to pathogens, oxidative and osmotic stresses as well as genes involved in the response to abscisic acid, jasmonate and ethylene (Figure 6B; Datasets 2-3).

We compared the DEG to a published dataset of genes regulated by REV (Reinhart et al., 2013). 23 genes were found to be coregulated by RAX1 and REV, which is 3.6 times more than what would be expected by chance (Figure 6C and Table S3). However, regarding the involvement of both TF in the same pathways this number is rather small. Additionally, none of the genes in this list appears as an important regulator of meristematic processes. This suggests that the phenotypes observed in the double mutants may be the results of synergistic pathways rather than transcriptional redundancy.

(PBM). A fusion protein of the full length RAX1 and the MBP-6His tags (RAX1<sub>full</sub>-MBP-6H) was produced in *E. coli* and used for PBM. The best motif obtained for this construct was expectedly a typical Myb R2R3 motif (Franco-Zorrilla et al., 2014) (Figure 7A). Using this DNA-binding model we determined both the best predicted binding site and the probability of occupancy (pOcc) for RAX1 in the genomic region of each differentially expressed gene. The scores obtained for RAX1 DNA-binding model ranged from -47.87 (worst possible binding site) to 14.37 (best possible binding site). We arbitrarily set a score threshold of 11.08, which corresponds to a type I error rate of 0.01% and a pOcc threshold of 0.2. Based on these predictions we identified 272 genes that could be direct targets of RAX1 (Dataset 4). Amongst these, some genes expressed in the SAM (Yadav et al., 2009, 2014) stood out as of special interest or potentially explaining GO term enrichment and are discussed below (Table S2).

The transcription factors ETHYLENE RESPONSE



**Figure 6. Transcriptomic response to the induction of RAX1 activity.** (A) PCA analysis of the 12 RNA-seq samples performed on the 500 most variable genes. Variance explained by the first and second components is indicated in the axis title. DEX-treated RAX1-GR samples segregate from all the other samples on the first component. (B) Selected GO term enrichment in the genes downregulated (red) or upregulated (blue) in DEX-treated RAX1-GR samples. (C) Venn-diagram showing the intersection of RAX1 transcriptomic data (this study) and REV micro-array data (Reinhart et al., 2013). The number of genes in each set is indicated. Co-enrichment p-value for the overlap between the two datasets is indicated between parentheses (hypergeometric test).

#### *In silico* prediction of putative direct RAX1 targets

To identify putative direct targets of RAX1 amongst the differentially expressed genes, we first determined its DNA-binding properties using protein-binding micro-array

FACTOR (ERF) 1 and 2 were strongly repressed and are involved in pathogen resistance and act upstream of ethylene and jasmonate signalling (Cheng et al., 2013). In addition, the copper transporter RESPONSE TO ANTAGONIST1 (RAN1) is involved in the synthesis of the

ethylene receptors (Binder et al., 2010). The regulation of these genes may thus explain the over-representation of defence response-related GO terms in the DEG set.

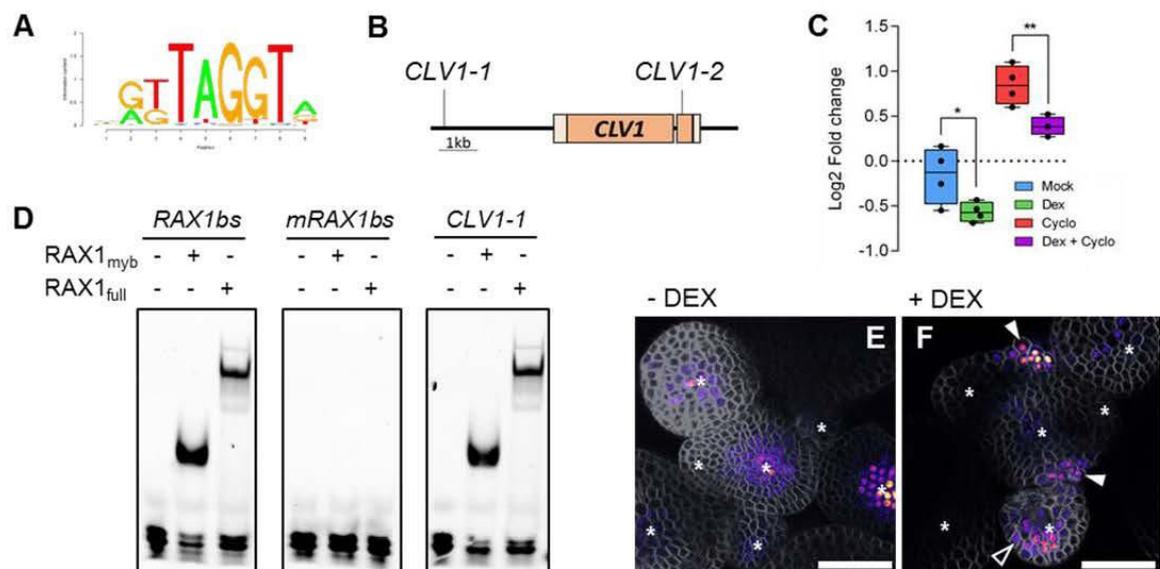
The presence of ABF2 in the putative direct targets may explain the enrichment in abscisic acid response in the GO term analysis as it was shown to play a role in ABA signal transduction, which is a reported role of RAX1 (Kim et al., 2004; Yu et al., 2016). Moreover, ABA was shown to be involved in osmotic stress response and involve ROS signalling, both of which are also enriched (Cruz de Carvalho, 2008).

RAX1 may also act on cell wall at several levels. It could directly regulate the cell wall modifying enzymes FUCOSYL TRANSFERASE4 (FUT4), PECTIN METHYLESTERASE3 (PME3) and EXPANSIN10 (EXPA10) that catalyse the modification of cell-wall components and act on cell growth (Goh et al., 2012; Hewezi et al., 2008; Liang et al., 2013; Wu et al., 2010). In addition, RAX1 may also directly regulate the pectin receptors WALL ASSOCIATED KINASE (WAK) 1 and 2. These two receptors were shown to be involved in cell expansion (Kohorn et al., 2006; Wagner and Kohorn, 2001).

roots (Tsukagoshi et al., 2010). The combined inhibition of UPB1 and CLV1 observed in response to RAX1 activity may thus allow the stimulation of cell proliferation.

#### RAX1 directly regulates CLV1 expression

The presence of CLV1 amongst the putative direct RAX1 targets was especially interesting as it is a negative regulator of WUS and it appeared more than two times less expressed in seedlings in response to RAX1. Moreover, *CLV1* presents two very good predicted binding sites (score > 12) for RAX1 in its promoter and towards the end of its coding sequence (Figure 7B). In order to determine if *CLV1* was a genuine target of RAX1 in the inflorescence meristem we measured transcript abundance in plants mock- or DEX-treated mCHERRY-RAX1-GR plants. This analysis showed a mild reduction in *CLV1* mRNA levels in response to DEX treatment (Figure 7C). When incubating the plants with the protein synthesis inhibitor cycloheximide (CHX) prior to RNA extraction the levels of *CLV1* were increased, which is a reported effect of CHX treatment (Gordon et al., 2009). When we combined DEX and CHX treatment, we could still detect a reduction of *CLV1* mRNA levels, indicating that *CLV1* regulation by RAX1 is likely direct (Figure 7C).



**Figure 7. CLV1 is a direct target of RAX1.** (A) DNA-binding model of RAX1 as determined by PBM. (B) Scheme of *CLV1* genomic locus with the two best RAX1 predicted binding site (CLV1-1 and CLV1-2). Non-coding regions and introns are indicated by a black line, exons (dark orange) and 5' and 3' UTR are indicated in light orange. (C) Quantification of *CLV1* transcripts in RAX1-GR inflorescences treated with mock (blue), DEX (green), cycloheximide (red) and cycloheximide + DEX (violet). \*  $p < 0.05$ , \*\*  $p < 0.01$ , determined by a t-test (N = 3). (D) *in vitro* binding assay of RAX1full and RAX1myb on sequences from *CLV1* genomic region (CLV1-1), the best possible binding site (RAX1bs) a mutated binding site (mRAX1bs). (E, F) pWUS:2xVENUS reporter expression in RAX1-GR plants after mock (E) or DEX (F) treatment. Ectopic pWUS activity between flower buds is indicated by filled arrowheads and ectopic pWUS expression in the flower bud is indicated by an empty arrowhead. Stars indicate meristem positions.

Consistently with its expression pattern in floral abscission zones (Figures 2A-B) RAX1 may directly inhibit the abscission activator HAESA (HAE), possibly playing a role in delaying abscission of floral organs (Jinn et al., 2000).

Very good RAX1 binding sites were detected in the *UPBEAT1* (*UPB1*) and *CLV1* loci. *UPB1* is involved in the control of ROS balance and was shown to control the balance between cell proliferation and differentiation in

Consistently with this result we found that RAX1full-MBP-6H as well as a tagged truncated version of RAX1 carrying only the Myb domain (RAX1myb-6H) can bind to the best predicted binding sites from *CLV1* genomic region (Figure 7D). As inhibition of *CLV1* is an important component of the *WUS* regulatory module (Schoof et al., 2000), we checked if the induction of RAX1 activity throughout the meristem was sufficient to perturb *WUS* expression pattern. When we combined mCHERRY-RAX1-GR to the

WUS transcriptional reporter line we could observe in some plants ectopic expression of WUS between FM and in the PZ of FM upon DEX treatment (Figures 7E-F). Additionally, the cells expressing WUS ectopically between the FM appeared swollen and the tissue organization in these regions resembled that of a callus, indicating that WUS expression was associated to cell growth and proliferation (Figure 7E-F). This effect of ectopic WUS expression on cell growth and proliferation is consistent with the reported effects of WUS expression in the root or the hypocotyl (Gallois et al., 2004; Zhang et al., 2013).

## Discussion

Several studies suggested that in addition to its well-studied role in flower identity determination, LFY is also involved in the initiation of FMs (Chahtane et al., 2013; Moyroud et al., 2010; Sawa et al., 1999; Wu et al., 2015; Yamaguchi et al., 2013). However, *lfy* mutants correctly initiate floral-shoot structures, indicating that other pathways could compensate for its meristem formation activity. Because *rev* mutants are affected in flower initiation and meristem formation (Otsuga et al., 2001; Prigge et al., 2005; Talbert et al., 1995) we used *rev* as a sensitized background to study the role of LFY in FM formation. We found that combinations of *lfy* and *rev* mutations almost completely prevented flower formation, which were replaced by small filamentous structures. These results confirm a role for LFY in flower initiation and indicate that LFY act in parallel of REV in this process.

It was proposed earlier that LFY ectopically induces meristem formation through the regulation of RAX1 (Chahtane et al., 2013). Although RAX1 is expressed in the first steps of flower development, *rax1* mutant do not have any flower phenotype (Keller et al., 2006; Müller et al., 2006). If RAX1 also acts downstream of LFY in flowers this might be because of the redundant action of REV. We show here that RAX1 is a potent target of LFY in flowers as the fusion of LFY to the VP16 transactivation domain enhances RAX1 transcription. However previous data showed that RAX1 is still expressed in *lfy* inflorescences (Chahtane et al., 2013). This might have two reasons: (1) *lfy* inflorescences can be considered as vegetative tissues and RAX1 is therefore regulated by the same factors as in vegetative tissues or (2) other factors regulate RAX1 in parallel of LFY.

We show by combining *rax1* and *rev* mutations that RAX1 and REV act in parallel pathways, consistently with a role of RAX1 downstream of LFY. The defects observed in *rax1 rev* mutants are milder than that of *lfy rev* mutants. This might be because the *rax1* alleles used here contain mutations in the C-terminal part of the proteins which can retain some activity. It has also been shown that LFY regulates additional genes involved in meristem formation such as A-type ARRs (Chahtane et al., 2013), and is thus likely to have a broader effect than RAX1. The absence of stem cell defects in lateral structures of *rax1* and *lfy* mutants (apart from flower identity) can be explained by the role of REV as a major regulator of meristem formation (Prigge et al., 2005). However, *rev* mutants are extremely sensitive to variations in the activity of RAX1 and LFY as illustrated by the effects of *rax1-3* heterozygous mutation in the *rev-6* background.

The lack of a morphologically distinct meristematic structures in most of the double mutant flowers indicates a role of RAX1 and REV in the initiation of the FM. Accordingly, these flowers lacked the expression of both WUS and CLV3 in the first flower development stages where they normally appear. In order to determine the role of RAX1 and REV on meristem initiation, we used CRISPR/Cas9 mutated lines for these two genes which showed a milder phenotype than the *rax1-3 rev-6* mutant. This showed that if both RAX1 and REV are involved in the stimulation of WUS expression and its inhibition in the L1, RAX1 has a specific effect in preventing WUS expression in the flower PZ. This indicates that RAX1 and REV are required for the proper definition of the FM domains, a step required for the maintenance of meristem activity throughout flower development.

Additionally, RAX1 and REV had a synergic effect on restricting the size of the OC in the SAM. A similar role was already described for FIL and YAB3 (Goldshmidt et al., 2008). It had been suggested that FIL could act non-cell autonomously via LAS in this process, moreover, it has been shown that LAS acts upstream of REV and that FIL antagonises REV expression (Greb et al., 2003; Nakata and Okada, 2013). Together with our data this suggest that the FIL and REV pathways are intertwined for the regulation of SAM size. The way REV would restrict SAM size remains elusive but is coherent with the observed enlargement of the WUS expression domain in FM of the double *rax1 rev* mutants.

To gain insights into the molecular mechanisms controlled by RAX1 we performed RNA sequencing analysis upon RAX1 activity induction in seedlings. This analysis showed a role of RAX1 in regulating genes involved in cell-wall modifications. Amongst those, we predicted several cell wall remodelling enzymes as well two pectin receptors, WAK1 and 2, to be direct targets of RAX1. WAKs are cell wall sensors required for cell elongation. Some members of the family were shown to have a role in cell differentiation (Kohorn et al., 2006; Lally et al., 2001; Wagner and Kohorn, 2001). This analysis also revealed a role of RAX1 in inhibiting plant defence response. This seems to be a common feature of floral regulators as it is also observed for the genes LFY and ANT/AIL6 (Krizek et al., 2016; Winter et al., 2011). Consistently we predicted the ethylene receptors ERF1 and 2 as being direct RAX1 targets. The ERFs can induce both abiotic and biotic defence pathways in response to a variety of stresses (Cheng et al., 2013). Interestingly we also detected a role of RAX1 in regulating ABA and ROS responses which are linked to cell differentiation in the SAM (Wilson et al., 2016). In particular, the ABA response regulator ABF2 and the ROS homeostasis regulator UPB1 are predicted direct RAX1 targets. UPB1 was shown to regulate the balance of cell proliferation and differentiation in the growing root via control of ROS homeostasis (Tsukagoshi et al., 2010). Interestingly, we did not detect any enrichment in CUC2 transcripts upon RAX1 induction. CUC2 was shown to be a direct target of RAX1 (Tian et al., 2014). CUC2 is a target for microRNA degradation and therefore its accumulation in the tissues used for this analysis may be prevented (Laufs et al., 2004).

Arguably one of the most interesting target of RAX1 in the FM initiation context is the receptor CLV1 which act as a negative regulator of WUS, restricting the size of the OC

(Lenhard and Laux, 2003). We confirmed that this regulation was direct in inflorescences, which provides a direct mechanistic link between *RAX1* and the regulation of meristem activity. In order to assess the capacity of *RAX1* to trigger FM formation, we followed *WUS* expression after ectopic induction of *RAX1* activity. We could observe ectopic expression of *WUS* in the PZ of FMs, which is consistent with an inhibition of *CLV1* in flowers. Interestingly, a similar phenomenon, although more localized, could be observed in the SAM. *RAX1* induction provoked the creation of foci of *WUS* expression in the basal part of the PZ of the SAM. This is consistent with the existence of an L1 signal preventing *WUS* expression at the tip of the SAM (Gruel et al., 2016). The cells in these foci appeared swollen and disorganized indicating a growth activity in those regions, which is compatible with organogenesis.

Interestingly, it was shown that *WUS*-mediated repression of *CLV1* fine tunes its expression, promoting the adaptation of the *CLV3/WUS* equilibrium (Busch et al., 2010). A similar mechanism might be leveraged by *RAX1*, prior to the establishment of the meristem. Fine-tuning of *CLV1* expression during FM establishment would be a requirement to establish the balance of *CLV3* and *WUS*, and thus define the different meristematic domains. Failure to do so would lead to a collapse of the meristem and developmental arrest, as observed in the *rax1 rev* double mutants.

The RNA-seq analysis yielded a surprisingly poorly interesting set of common targets with *REV*. Although the two datasets were statistically enriched in common targets, we did not find known stemness regulators amongst them. This reinforces the hypothesis that *RAX1* and *REV* act in different pathways that can compensate for each other. This is supported by two facts: (1) *rev* mutants are less sensitive to *clv1* mutation in the context of flower formation (Otsuga et al., 2001) and (2) HD-ZIP III factors were shown to regulate CK synthesis (Dello Iorio et al., 2012), and meristem initiation defects of *rax1* mutants can be compensated by CK synthesis (Wang et al., 2014), although we provide no support for a role of *RAX1* in CK synthesis or signalling.

## Experimental Procedures

### Plant material and treatments

The *rax1-3*, *lfy-12*, and *lfy-2* alleles have been previously described (Keller et al., 2006; Müller et al., 2006; Weigel et al., 1992). The *rev-6* allele (Otsuga et al., 2001) was backcrossed from *Ler* to *Col* in this study. Plants were cultivated in long-days conditions (16h of light) unless specified otherwise. For confocal and scanning electron microscopy, plants were grown in short-day conditions (8 h of light) for 6 weeks and transferred to long-days conditions for 2 weeks. Mutants phenotyping was performed three weeks after bolting. For NPA treatments plants were sprayed with either 10  $\mu$ M of NPA or 1/10 000 EtOH every other day from two weeks on. For DEX treatment the plants were sprayed with either 10  $\mu$ M DEX in 1/10 000 DMSO or 1/10 000 DMSO (mock) every other day from two weeks on.

### Transgenes

The *pRAX1:GUS* construct contains 2.1 kb of *RAX1* promoter. CRISPR spacers were designed using

CHOPCHOP (Montague et al., 2014). Spacers with no predicted off-targets were selected. CRISPR constructs were then cloned as described (Yan et al., 2016). *pUBQ10:mCHERRY-RAX-GR:tUBQ10* (pGD46) and *pUBQ10:mCHERRY-GR:tUBQ10* (pGD48) and *pWUS:2xVENUS:tWUS* (CW039) constructs were cloned using the GreenGate system (Lampropoulos et al., 2013). The GR sequence was amplified from the AP1-GR construct (Kaufmann et al., 2010) and *RAX1* cDNA sequence was provided by ABRC (stock # DQ446976). *WUS* promoter comprises 4.4 kb upstream the TSS (Lampropoulos et al., 2013). The *RAX1<sub>full</sub>-MBP-6H* (pGD14) and *RAX<sub>Myb</sub>-6H* (pGD19) constructs were cloned by inserting the full length cDNA or a version containing the 135 first amino-acids of *RAX1* into the pETM-41 and pETM-11 plasmids respectively (Dümmler et al., 2005). Detailed cloning procedures are available as supplementary methods.

### Identification of CRISPR mutants

CRISPR-induced mutations were identified in the T1 using poly-acrylamide gel electrophoresis (PAGE) separation of DNA heteroduplexes (Zhu et al., 2014). The region containing the mutation was PCR-amplified from T2 plants and subcloned in pCR-Blunt (ThermoFisher) for sequencing. CRISPR lines targeting *REV* were crossed to either *lfy-2* or *lfy-12* and T1 was screened for double mutant phenotype. As the plants were sterile, the mutation was only detected using the PAGE separation method.

### Western blotting

mCHERRY expression was checked in seedlings whole extract using the 6G6 anti-RFP antibody (ChromoTek). Western blotting was performed as described previously (Sayou et al., 2016).

### In situ hybridization

Plant samples were harvested shortly after bolting. Older flowers were swiftly removed, apices were collected in fixative, and *in situ* hybridization was performed as previously described (Carles et al., 2010). The *RAX1*, *WUS*, and *CLV3* probes were described in previous studies (Brand et al., 2000; Fletcher et al., 1999; Keller et al., 2006).

### Scanning electron microscopy

Older flowers were removed from the inflorescences. Apices were collected and swiftly fixed on a stub by carbon tape. A drop of water was added and the samples were placed in the FEI Quanta 250 chamber. Imaging was performed in ESEM mode with a pressure between 700 and 550 Pa and a temperature of 1°C to 2°C with a tension of 14 kV.

### Confocal microscopy and image treatment

Apices were dissected and placed on agarose 2% and cell walls were counter-stained with FM4-64. Cell wall and VENUS signals were recorded in two separate channels. A signal threshold was applied uniformly on all samples using FIDJI (Schindelin et al., 2012) to improve signal-to-noise ratio, as indicated by the minimal value on the figure scale. Fluorescence intensity and area was subsequently quantified using the ellipse tool of FDJI. For dexamethasone (DEX) induction, the apices were placed

on apex culture medium (Hamant et al., 2014) containing 1/10 000 DMSO with or without 10  $\mu$ M DEX. Samples were imaged after 8 h of incubation. SAM surface reconstruction was performed in the MorphoGraphX framework (Barbier de Reuille et al., 2015). We used the marching cube surface method with a cube size of 3  $\mu$ m. Then mesh was smoothed several times and the cell wall signal in the L1 was projected onto it. Individual cells were manually seeded and the seeds were propagating by watershed. The Gaussian curvature was then projected onto the mesh with a neighbouring of 20  $\mu$ m and represented as a cell heatmap.

#### RNA sequencing and analysis

14 day-old seedlings grown on MS plates were shortly immersed with a solution containing 0.03% L-77 Silwett, 1/10 000 DMSO, and 1/1 000 ethanol (mock), and alternatively 10  $\mu$ M DEX (DEX), 50  $\mu$ M cycloheximide (CHX) or both (DEX+CHX). Whole seedlings were harvested 4 h after treatment and RNA was extracted with the RNeasy kit (Qiagen). Libraries were synthesized with the TruSeq Stranded (Illumina) and paired-end sequencing was performed on an HiSeq2000 (Illumina). Adapter sequences were trimmed and duplicated and low-quality reads were discarded. Mapping was performed on TAIR10 assembly with HISAT2 (Kim et al., 2015) and mapped reads with a mapping quality score (MPAQ) below 30 or mapped at several locations were discarded. Reads mapped to exons or untranslated regions were counted with HTSeq (Anders et al., 2015). DEG discovery was performed with EdgeR (Robinson et al., 2010) using a multiparametric GLM model for the interaction genotype:treatment after TMM normalization (McCarthy et al., 2012; Robinson and Smyth, 2007; Zhou et al., 2014). Genes were considered differentially expressed if the likelihood-ratio test FDR was equal or below 0.01. GO-term enrichment analysis was performed in Araport (Krishnakumar et al., 2015).

#### RAX1 binding-site prediction

PBM was performed as previously described using an MBP tagged full-length RAX1 fusion protein (Franco-Zorrilla and Solano, 2014). Binding sites were predicted using the Biopython package for python2.7 (Cock et al., 2009). pOcc was calculated as described for the GOMER program (Granek and Clarke, 2005), using a custom python script.

#### qRT-PCR

Older flowers were dissected and 6 inflorescences per replicate were harvested. RNA was extracted with the RNeasy kit (Qiagen). Gene expression was quantified using AT2G28390 and AT4G34270 as internal reference as they were shown to be stable across a wide range of conditions (Czechowski et al., 2005). Statistical analysis was performed of the  $\Delta$ Cq values and fold-change was calculated solely for graphical representation purpose.

#### In vitro DNA-binding assay

Protein production was performed as previously described (Sayou et al., 2016). Proteins were purified on nickel-sepharose column and eluted with 150mM imidazole. Electrophoretic mobility-shift assay was performed as previously described (Sayou et al., 2016) in Promega1 buffer (Tris-HCl 10 mM; NaCl 50 mM; MgCl<sub>2</sub> 1 mM; 1%

glycerol; EDTA 0.5 mM; DTT 1 mM; pH7.5). Probe sequences are indicated in Table S4.

#### Author Contributions

FP, GT and GD designed the experiments. GD, GT, MLM, HC, SH and CC carried out experiments. GD, GT and FP analysed and interpreted results. CW and JL provided material. JMFZ and JL provided expertise. GD, FP, GT wrote the paper.

#### Acknowledgments

The authors wish to acknowledge the support of the Electron Microscopy facility of the ICMG Nanobio – Chemistry Platform and C. Lancelon-Pin in particular. The POPS transcriptomic platform and L. Soubignou-Taconnat. S. Figueat, D. Grunwald and L. Kurzawa for BIG's shared facilities. K. Kaufmann and W. Yan for sharing plasmids ahead of publication. E. Delannoy, M-L. Martin-Magniette, P. Das and A. Larrieu for their inputs on the project. pRAX1:GUS seeds were a generous gift from P. Doerner. This work was supported by the French National Agency for Research (ANR Blanc – SVSE2 – 2011-Charmaful).

#### References

- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169.
- Barbier de Reuille, P., Routier-Kierzkowska, A.L., Kierzkowski, D., Bassel, G.W., Schüpbach, T., Tauriello, G., Bajpai, N., Strauss, S., Weber, A., Kiss, A., et al. (2015). MorphoGraphX: A platform for quantifying morphogenesis in 4D. *Elife* 4, 05864.
- Besnard, F., Refahi, Y., Morin, V., Marteaux, B., Brunoud, G., Chambrier, P., Rozier, F., Mirabet, V., Legrand, J., Lainé, S., et al. (2014). Cytokinin signalling inhibitory fields provide robustness to phyllotaxis. *Nature* 505, 417–421.
- Binder, B.M., Rodríguez, F.I., and Bleecker, A.B. (2010). The copper transporter RAN1 is essential for biogenesis of ethylene receptors in Arabidopsis. *J Biol Chem* 285, 37263–37270.
- Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M., and Simon, R. (2000). Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by CLV3 activity. *Science* 289, 617–619.
- Busch, W., Miotk, A., Ariel, F.D., Zhao, Z., Forner, J., Daum, G., Suzuki, T., Schuster, C., Schultheiss, S.J., Leibfried, A., et al. (2010). Transcriptional control of a plant stem cell niche. *Dev Cell* 18, 849–861.
- Carles, C.C., Ha, C.M., Jun, J.H., Fiume, E., and Fletcher, J.C. (2010). Analyzing shoot apical meristem development. *Methods Mol Biol* 655, 105–129.
- Chahtane, H., Vachon, G., Le Masson, M., Thévenon, E., Périgon, S., Mihajlovic, N., Kalinina, A., Michard, R., Moyroud, E., Monniaux, M., et al. (2013). A variant of LEAFY reveals its capacity to stimulate meristem development by inducing RAX1. *Plant J* 74, 678–689.
- Cheng, M.C., Liao, P.M., Kuo, W.W., and Lin, T.P. (2013). The Arabidopsis ETHYLENE RESPONSE FACTOR1 regulates abiotic stress-responsive gene expression by binding to different cis-acting elements in response to different stress signals. *Plant Physiol* 162, 1566–1582.
- Cock, P.J., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for

- computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423.
- Cruz de Carvalho, M.H. (2008). Drought stress and reactive oxygen species: Production, scavenging and signaling. *Plant Signal Behav* 3, 156–165.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol* 139, 5–17.
- Daum, G., Medzihradsky, A., Suzuki, T., and Lohmann, J.U. (2014). A mechanistic framework for noncell autonomous stem cell induction in *Arabidopsis*. *Proc Natl Acad Sci U S A* 111, 14619–14624.
- Dello Iorio, R., Galinha, C., Fletcher, A.G., Grigg, S.P., Molnar, A., Willemsen, V., Scheres, B., Sabatini, S., Baulcombe, D., Maini, P.K., et al. (2012). A PHABULOSA/cytokinin feedback loop controls root growth in *Arabidopsis*. *Curr Biol* 22, 1699–1704.
- Denay, G., Chahtane, H., Tichtinsky, G., and Parcy, F. (2017). A flower is born: an update on *Arabidopsis* floral meristem formation. *Curr Opin Plant Biol*.
- Dümmler, A., Lawrence, A.M., and de Marco, A. (2005). Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb Cell Fact* 4, 34.
- Endrizzi, K., Moussian, B., Haecker, A., Levin, J.Z., and Laux, T. (1996). The SHOOT MERISTEMLESS gene is required for maintenance of undifferentiated cells in *Arabidopsis* shoot and floral meristems and acts at a different regulatory level than the meristem genes WUSCHEL and ZWILLE. *Plant J* 10, 967–979.
- Fletcher, J.C., Brand, U., Running, M.P., Simon, R., and Meyerowitz, E.M. (1999). Signaling of cell fate decisions by CLAVATA3 in *Arabidopsis* shoot meristems. *Science* 283, 1911–1914.
- Franco-Zorrilla, J.M., and Solano, R. (2014). High-throughput analysis of protein-DNA binding affinity. *Methods Mol Biol* 1062, 697–709.
- Franco-Zorrilla, J.M., López-Vidriero, I., Carrasco, J.L., Godoy, M., Vera, P., and Solano, R. (2014). DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc Natl Acad Sci U S A* 111, 2367–2372.
- Gallois, J.L., Nora, F.R., Mizukami, Y., and Sablowski, R. (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes Dev* 18, 375–380.
- Geldner, N., Dénervaud-Tendon, V., Hyman, D.L., Mayer, U., Stierhof, Y.D., and Chory, J. (2009). Rapid, combinatorial analysis of membrane compartments in intact plants with a multicolor marker set. *Plant J* 59, 169–178.
- Goh, H.H., Sloan, J., Dorca-Fornell, C., and Fleming, A. (2012). Inducible repression of multiple expansin genes leads to growth suppression during leaf development. *Plant Physiol* 159, 1759–1770.
- Goldshmidt, A., Alvarez, J.P., Bowman, J.L., and Eshed, Y. (2008). Signals derived from YABBY gene activities in organ primordia regulate growth and partitioning of *Arabidopsis* shoot apical meristems. *Plant Cell* 20, 1217–1230.
- Gordon, S.P., Chickarmane, V.S., Ohno, C., and Meyerowitz, E.M. (2009). Multiple feedback loops through cytokinin signaling control stem cell number within the *Arabidopsis* shoot meristem. *Proc Natl Acad Sci U S A* 106, 16529–16534.
- Granek, J.A., and Clarke, N.D. (2005). Explicit equilibrium modeling of transcription-factor binding and gene regulation. *Genome Biol* 6, R87.
- Greb, T., Clarenz, O., Schafer, E., Muller, D., Herrero, R., Schmitz, G., and Theres, K. (2003). Molecular analysis of the LATERAL SUPPRESSOR gene in *Arabidopsis* reveals a conserved control mechanism for axillary meristem formation. *Genes Dev* 17, 1175–1187.
- Gruel, J., Landrein, B., Tarr, P., Schuster, C., Refahi, Y., Sampathkumar, A., Hamant, O., Meyerowitz, E.M., and Jönsson, H. (2016). An epidermis-driven mechanism positions and scales stem cell niches in plants. *Science Advances* 2, e1500989.
- Hamant, O., Das, P., and Burian, A. (2014). Time-lapse imaging of developing meristems using confocal laser scanning microscope. *Methods Mol Biol* 1080, 111–119.
- Hewezi, T., Howe, P., Maier, T.R., Hussey, R.S., Mitchum, M.G., Davis, E.L., and Baum, T.J. (2008). Cellulose binding protein from the parasitic nematode *Heterodera schachtii* interacts with *Arabidopsis* pectin methylesterase: cooperative cell wall modification during parasitism. *Plant Cell* 20, 3080–3093.
- Jasinski, S., Piazza, P., Craft, J., Hay, A., Woolley, L., Rieu, I., Phillips, A., Hedden, P., and Tsiantis, M. (2005). KNOX action in *Arabidopsis* is mediated by coordinate regulation of cytokinin and gibberellin activities. *Curr Biol* 15, 1560–1565.
- Jinn, T.L., Stone, J.M., and Walker, J.C. (2000). HAESA, an *Arabidopsis* leucine-rich repeat receptor kinase, controls floral organ abscission. *Genes Dev* 14, 108–117.
- Kaufmann, K., Wellmer, F., Muiño, J.M., Ferrier, T., Wuest, S.E., Kumar, V., Serrano-Mislata, A., Madueño, F., Krajewski, P., Meyerowitz, E.M., et al. (2010). Orchestration of floral initiation by APETALA1. *Science* 328, 85–89.
- Keller, T., Abbott, J., Moritz, T., and Doerner, P. (2006). *Arabidopsis* REGULATOR OF AXILLARY MERISTEMS1 controls a leaf axil stem cell niche and modulates vegetative development. *Plant Cell* 18, 598–611.
- Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12, 357–360.
- Kim, S., Kang, J.Y., Cho, D.I., Park, J.H., and Kim, S.Y. (2004). ABF2, an ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. *Plant J* 40, 75–87.
- Kohorn, B.D., Kobayashi, M., Johansen, S., Riese, J., Huang, L.F., Koch, K., Fu, S., Dotson, A., and Byers, N. (2006). An *Arabidopsis* cell wall-associated kinase required for invertase activity and cell growth. *Plant J* 46, 307–316.
- Krishnakumar, V., Hanlon, M.R., Contrino, S., Ferlanti, E.S., Karamycheva, S., Kim, M., Rosen, B.D., Cheng, C.Y., Moreira, W., Mock, S.A., et al. (2015). Araport: the *Arabidopsis* information portal. *Nucleic Acids Res* 43, D1003–9.
- Krizek, B.A., Bequette, C.J., Xu, K., Blakley, I.C., Fu, Z.Q., Stratmann, J., and Loraine, A.E. (2016). RNA-Seq links AINTEGUMENTA and AINTEGUMENTA-LIKE6 to cell wall remodeling and plant defense pathways in *Arabidopsis*. *Plant Physiol*.
- Lally, D., Ingmire, P., Tong, H.Y., and He, Z.H. (2001). Antisense expression of a cell wall-associated protein kinase, WAK4, inhibits cell elongation and alters morphology. *Plant Cell* 13, 1317–1331.

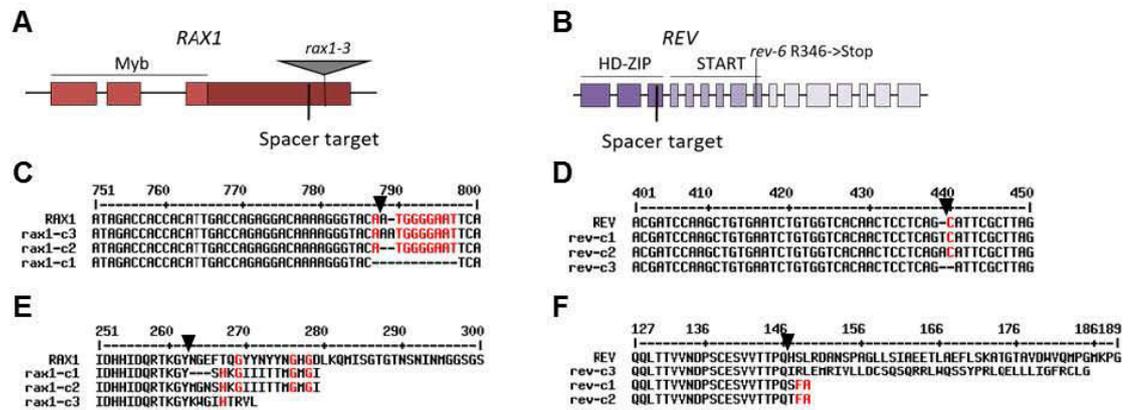
- Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J.U., and Forner, J. (2013). GreenGate—a novel, versatile, and efficient cloning system for plant transgenesis. *PLoS ONE* 8, e83043.
- Laufs, P., Peaucelle, A., Morin, H., and Traas, J. (2004). MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. *Development* 131, 4311–4322.
- Laux, T., Mayer, K.F., Berger, J., and Jürgens, G. (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development* 122, 87–96.
- Leibfried, A., To, J.P., Busch, W., Stehling, S., Kehle, A., Demar, M., Kieber, J.J., and Lohmann, J.U. (2005). WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* 438, 1172–1175.
- Lenhard, M., and Laux, T. (2003). Stem cell homeostasis in the Arabidopsis shoot meristem is regulated by intercellular movement of CLAVATA3 and its sequestration by CLAVATA1. *Development* 130, 3163–3173.
- Liang, Y., Basu, D., Pattathil, S., Xu, W.L., Venetos, A., Martin, S.L., Faik, A., Hahn, M.G., and Showalter, A.M. (2013). Biochemical and physiological characterization of fut4 and fut6 mutants defective in arabinogalactan-protein fucosylation in Arabidopsis. *J Exp Bot* 64, 5537–5551.
- Lohmann, J.U., Hong, R.L., Hobe, M., Busch, M.A., Parcy, F., Simon, R., and Weigel, D. (2001). A molecular link between stem cell regulation and floral patterning in Arabidopsis. *Cell* 105, 793–803.
- Mayer, K.F., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T. (1998). Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell* 95, 805–815.
- McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 40, 4288–4297.
- Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res* 42, W401–7.
- Moyroud, E., Kusters, E., Monniaux, M., Koes, R., and Parcy, F. (2010). LEAFY blossoms. *Trends Plant Sci* 15, 346–352.
- Müller, D., Schmitz, G., and Theres, K. (2006). Blind homologous R2R3 Myb genes control the pattern of lateral meristem initiation in Arabidopsis. *Plant Cell* 18, 586–597.
- Müller-Xing, R., Schubert, D., and Goodrich, J. (2015). Non-inductive conditions expose the cryptic bract of flower phytomer in Arabidopsis thaliana. *Plant Signal Behav* 10, e1010868.
- Nakata, M., and Okada, K. (2013). The Leaf Adaxial-Abaxial Boundary and Lamina Growth. *Plants* 2, 174–202.
- Otsuga, D., DeGuzman, B., Prigge, M.J., Drews, G.N., and Clark, S.E. (2001). REVOLUTA regulates meristem initiation at lateral positions. *Plant J* 25, 223–236.
- Padidam, M. (2003). Chemically regulated gene expression in plants. *Curr Opin Plant Biol* 6, 169–177.
- Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N., and Clark, S.E. (2005). Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. *Plant Cell* 17, 61–76.
- Reinhart, B.J., Liu, T., Newell, N.R., Magnani, E., Huang, T., Kerstetter, R., Michaels, S., and Barton, M.K. (2013). Establishing a framework for the Ad/abaxial regulatory network of Arabidopsis: ascertaining targets of class III homeodomain leucine zipper and KANADI regulation. *Plant Cell* 25, 3228–3249.
- Robinson, M.D., and Smyth, G.K. (2007). Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* 23, 2881–2887.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.
- Sassi, M., and Vernoux, T. (2013). Auxin and self-organization at the shoot apical meristem. *J Exp Bot* 64, 2579–2592.
- Sawa, S., Ito, T., Shimura, Y., and Okada, K. (1999). FILAMENTOUS FLOWER controls the formation and development of Arabidopsis inflorescences and floral meristems. *Plant Cell* 11, 69–86.
- Sayou, C., Nanao, M.H., Jamin, M., Posé, D., Thévenon, E., Grégoire, L., Tichtinsky, G., Denay, G., Ott, F., Peirats Llobet, M., et al. (2016). A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nat Commun* 7, 11222.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682.
- Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F., Jürgens, G., and Laux, T. (2000). The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* 100, 635–644.
- Shi, B., Zhang, C., Tian, C., Wang, J., Wang, Q., Xu, T., Xu, Y., Ohno, C., Sablowski, R., Heisler, M.G., et al. (2016). Two-Step Regulation of a Meristematic Cell Population Acting in Shoot Branching in Arabidopsis. *PLoS Genet* 12, e1006168.
- Smyth, D.R., Bowman, J.L., and Meyerowitz, E.M. (1990). Early flower development in Arabidopsis. *Plant Cell* 2, 755–767.
- Talbert, P.B., Adler, H.T., Parks, D.W., and Comai, L. (1995). The REVOLUTA gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of Arabidopsis thaliana. *Development* 121, 2723–2735.
- Tian, C., Zhang, X., He, J., Yu, H., Wang, Y., Shi, B., Han, Y., Wang, G., Feng, X., Zhang, C., et al. (2014). An organ boundary-enriched gene regulatory network uncovers regulatory hierarchies underlying axillary meristem initiation. *Mol Syst Biol* 10, 755.
- Tsakagoshi, H., Busch, W., and Benfey, P.N. (2010). Transcriptional regulation of ROS controls transition from proliferation to differentiation in the root. *Cell* 143, 606–616.
- Wagner, T.A., and Kohorn, B.D. (2001). Wall-associated kinases are expressed throughout plant development and are required for cell expansion. *Plant Cell* 13, 303–318.
- Wang, Y., Wang, J., Shi, B., Yu, T., Qi, J., Meyerowitz, E.M., and Jiao, Y. (2014). The Stem Cell Niche in Leaf Axils Is Established by Auxin and Cytokinin in Arabidopsis. *Plant Cell* 26, 2055–2067.
- Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F., and Meyerowitz, E.M. (1992). LEAFY controls floral meristem identity in Arabidopsis. *Cell* 69, 843–859.
- Wilson, M.E., Mixdorf, M., Berg, R.H., and Haswell, E.S. (2016). Plastid osmotic stress influences cell differentiation at the plant shoot apex. *Development*.

- Winter, C.M., Austin, R.S., Blanvillain-Baufumé, S., Reback, M.A., Monniaux, M., Wu, M.F., Sang, Y., Yamaguchi, A., Yamaguchi, N., Parker, J.E., et al. (2011). LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Dev Cell* 20, 430–443.
- Wu, M.F., Yamaguchi, N., Xiao, J., Bargmann, B., Estelle, M., Sang, Y., and Wagner, D. (2015). Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate. *Elife* 4, e09269.
- Wu, Y., Williams, M., Bernard, S., Driouch, A., Showalter, A.M., and Faik, A. (2010). Functional identification of two nonredundant Arabidopsis alpha(1,2)fucosyltransferases specific to arabinogalactan proteins. *J Biol Chem* 285, 13638–13645.
- Yadav, R.K., and Reddy, G.V. (2012). WUSCHEL protein movement and stem cell homeostasis. *Plant Signal Behav* 7, 592–594.
- Yadav, R.K., Girke, T., Pasala, S., Xie, M., and Reddy, G.V. (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. *Proc Natl Acad Sci U S A* 106, 4941–4946.
- Yadav, R.K., Tavakkoli, M., Xie, M., Girke, T., and Reddy, G.V. (2014). A high-resolution gene expression map of the Arabidopsis shoot meristem stem cell niche. *Development* 141, 2735–2744.
- Yamaguchi, N., Wu, M.F., Winter, C.M., Berns, M.C., Nole-Wilson, S., Yamaguchi, A., Coupland, G., Krizek, B.A., and Wagner, D. (2013). A molecular framework for auxin-mediated initiation of flower primordia. *Dev Cell* 24, 271–282.
- Yan, W., Chen, D., and Kaufmann, K. (2016). Efficient multiplex mutagenesis by RNA-guided Cas9 and its use in the characterization of regulatory elements in the AGAMOUS gene. *Plant Methods* 12, 23.
- Yanai, O., Shani, E., Dolezal, K., Tarkowski, P., Sablowski, R., Sandberg, G., Samach, A., and Ori, N. (2005). Arabidopsis KNOX1 proteins activate cytokinin biosynthesis. *Curr Biol* 15, 1566–1571.
- Yoshida, S., Mandel, T., and Kuhlemeier, C. (2011). Stem cell activation by light guides plant organogenesis. *Genes Dev* 25, 1439–1450.
- Yu, Y.T., Wu, Z., Lu, K., Bi, C., Liang, S., Wang, X.F., and Zhang, D.P. (2016). Overexpression of the MYB37 transcription factor enhances abscisic acid sensitivity, and improves both drought tolerance and seed productivity in Arabidopsis thaliana. *Plant Mol Biol* 90, 267–279.
- Zhang, D., Wang, X., Wang, M., Li, J., Guo, X., Chong, K., and Xu, Y. (2013). Ectopic expression of WUS in hypocotyl promotes cell division via GRP23 in Arabidopsis. *PLoS ONE* 8, e75773.
- Zhao, Z., Andersen, S.U., Ljung, K., Dolezal, K., Miotk, A., Schultheiss, S.J., and Lohmann, J.U. (2010). Hormonal control of the shoot stem-cell niche. *Nature* 465, 1089–1092.
- Zhou, X., Lindsay, H., and Robinson, M.D. (2014). Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic Acids Res* 42, e91.
- Zhu, X., Xu, Y., Yu, S., Lu, L., Ding, M., Cheng, J., Song, G., Gao, X., Yao, L., Fan, D., et al. (2014). An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. *Sci Rep* 4, 64

**Supplementary information**

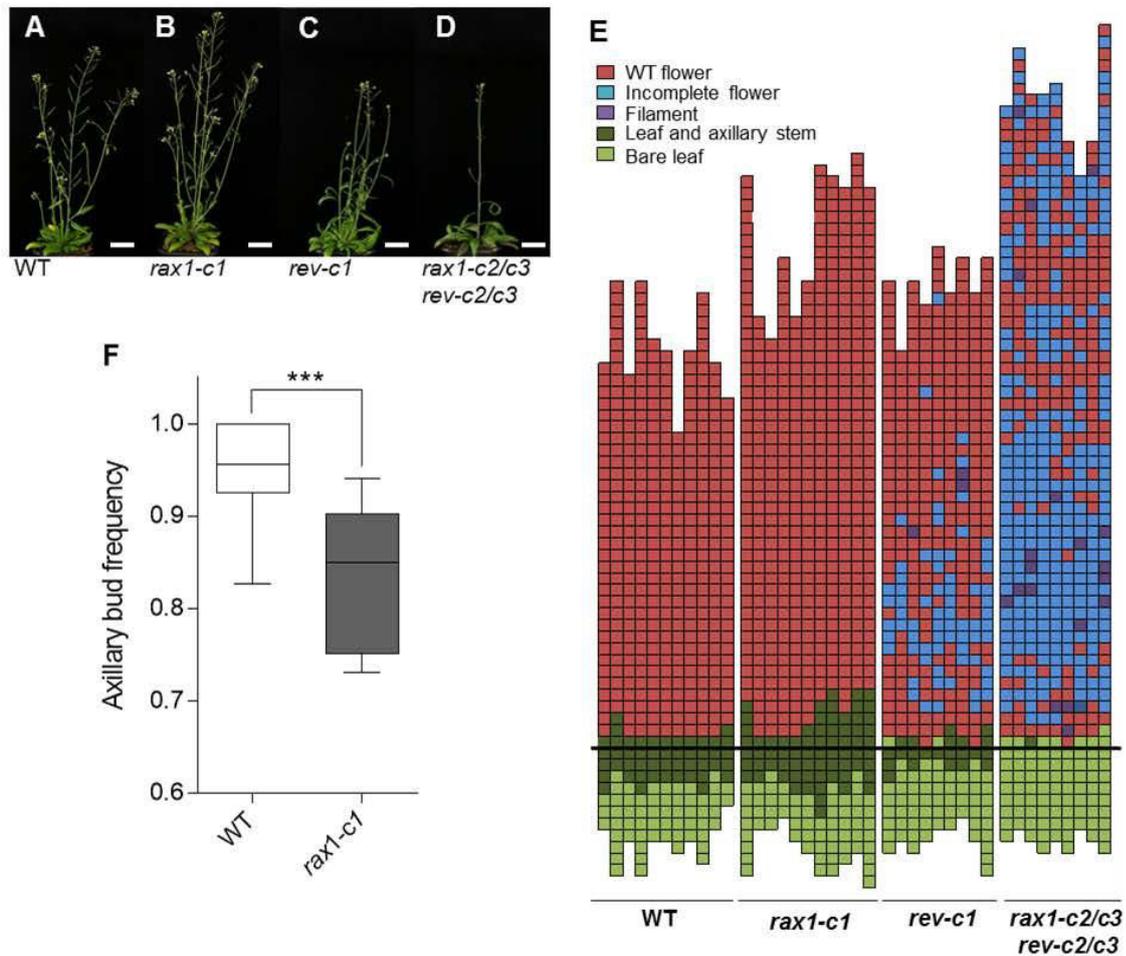
**Figures S1 to S7**

**Tables S1 to S4**



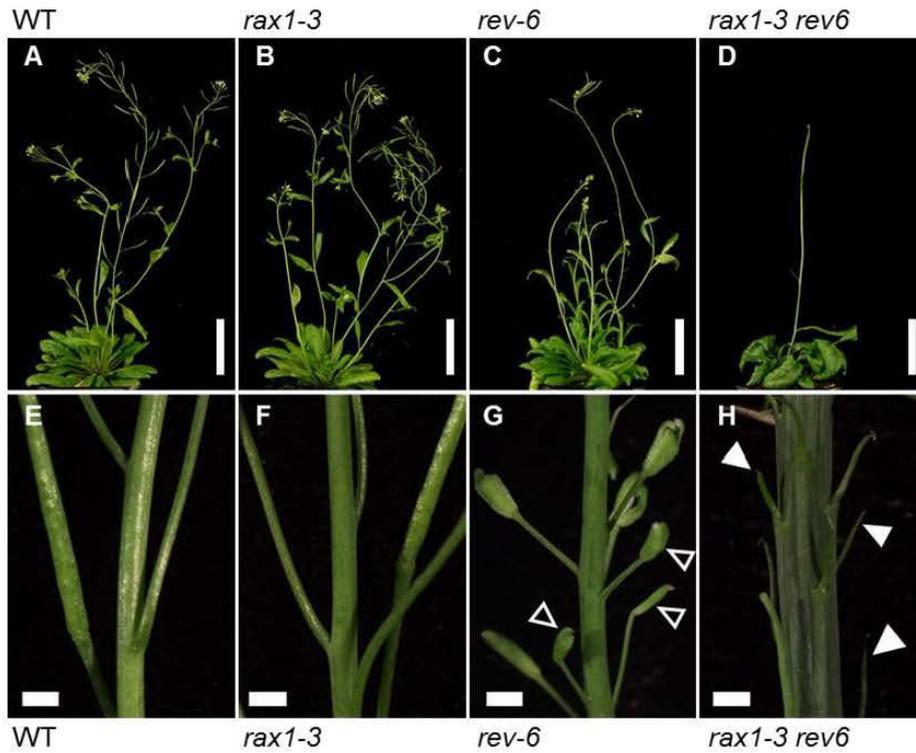
**Figure S1. Disruption of *RAX1* and *REV* coding sequences using a CRISPR/Cas9 system.**

(A-B) Genes models for *RAX1* (A) and *REV* (B) showing intron (lines)/exons (boxes) structures. Exons coding for each protein domains are colour coded (light-red: R2R3 Myb-domain; dark purple: HD-ZIP domain; medium purple: START domain). T-DNA insertion (gray triangle) and EMS mutation are labelled with the name of the corresponding mutant. CRISPR spacer target are localised. (C-F) Coding sequences (C,D) and protein sequences (E,F) of wild-type and CRISPR-mutated lines for *RAX1* (C,E) and *REV* (D,F). Only the relevant part of the sequence are shown. Numbering is relative to the ATG/1<sup>st</sup> methionine. Arrowheads indicate Cas9 cutting site. Non-conserved residues are shown in red.



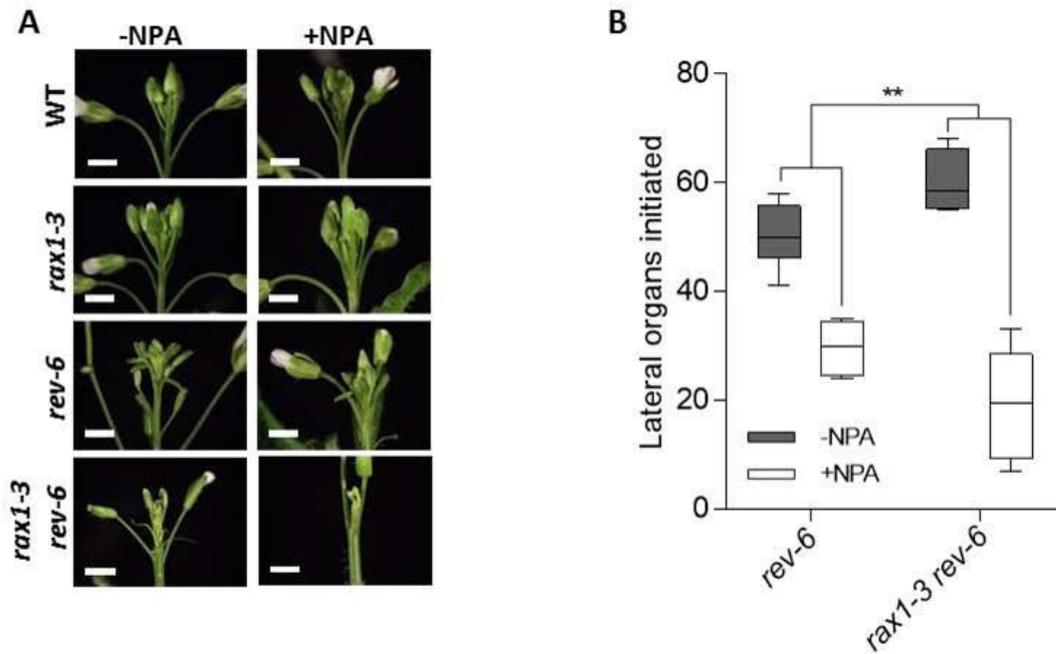
**Figure S2. CRISPR-mutated lines show phenotypes consistent to that of *rax1-3* and *rev-6*.**

(A-D) Habitus of WT (A), *rax1-c1* (B), *rev-c1* (C) and *rax1-c2/c3 rev-c2/c3* (D) plants. (E) Architecture diagram of the single and double mutants three weeks after bolting. Each column represents an individual plant and each square represent a single internode. Squares below the thick black line represent internodes on the rosette and squares above the thick black line, internodes on the main stem. Structures initiated are color-coded as labeled. (F) Proportion of rosettes leaves bearing an axillary bud in the WT and *rax1-c1* mutants in non-inductive short days conditions. \*\*\* $p < 0.001$ , determined by a Wilcoxon-Mann-Whitney test ( $N > 10$ ).



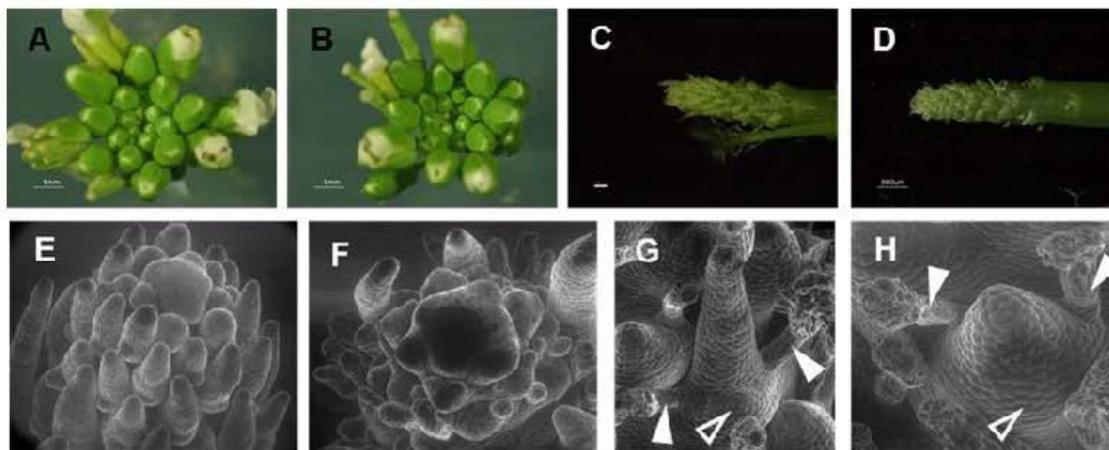
**Figure S3. Growth habit of *rax1-3*, *rev-6* and double mutant in inductive conditions.**

Habitus of WT (A-E), *rax1-3* (B-F), *rev-6* (C-G) and *rax1-3 rev-6* (D-H) plants in long-days conditions, 21 days after bolting (A-D), bar: 5cm. (E-H) Detail of the stems 2.5 cm below the shoot apex, showing flowers lacking internal organ whorls (empty arrowheads) and filamentous structures (full arrowheads). Note the increased diameter and woody aspect of the stem in *rax1-3 rev-6* (H). Bar: 1mm.



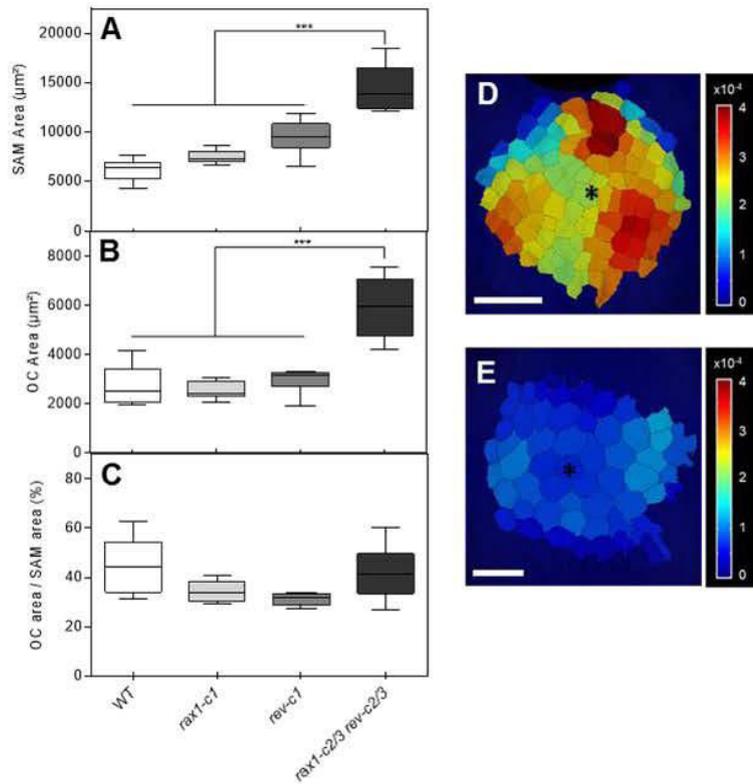
**Figure S4. RAX1 and REV pathways interact with auxin polar transport.**

(A) Inflorescence phenotype of WT, *rax1-3*, *rev-6* and *rax1-3 rev-6* (top to bottom) mock- (left) or NPA- (right) treated plants. (B) Number of lateral organs initiated by mock- or NPA-treated *rev-6* and double mutants four weeks after bolting. \*\* $p < 0.01$ , determined by a two-way ANOVA ( $N > 10$ ).



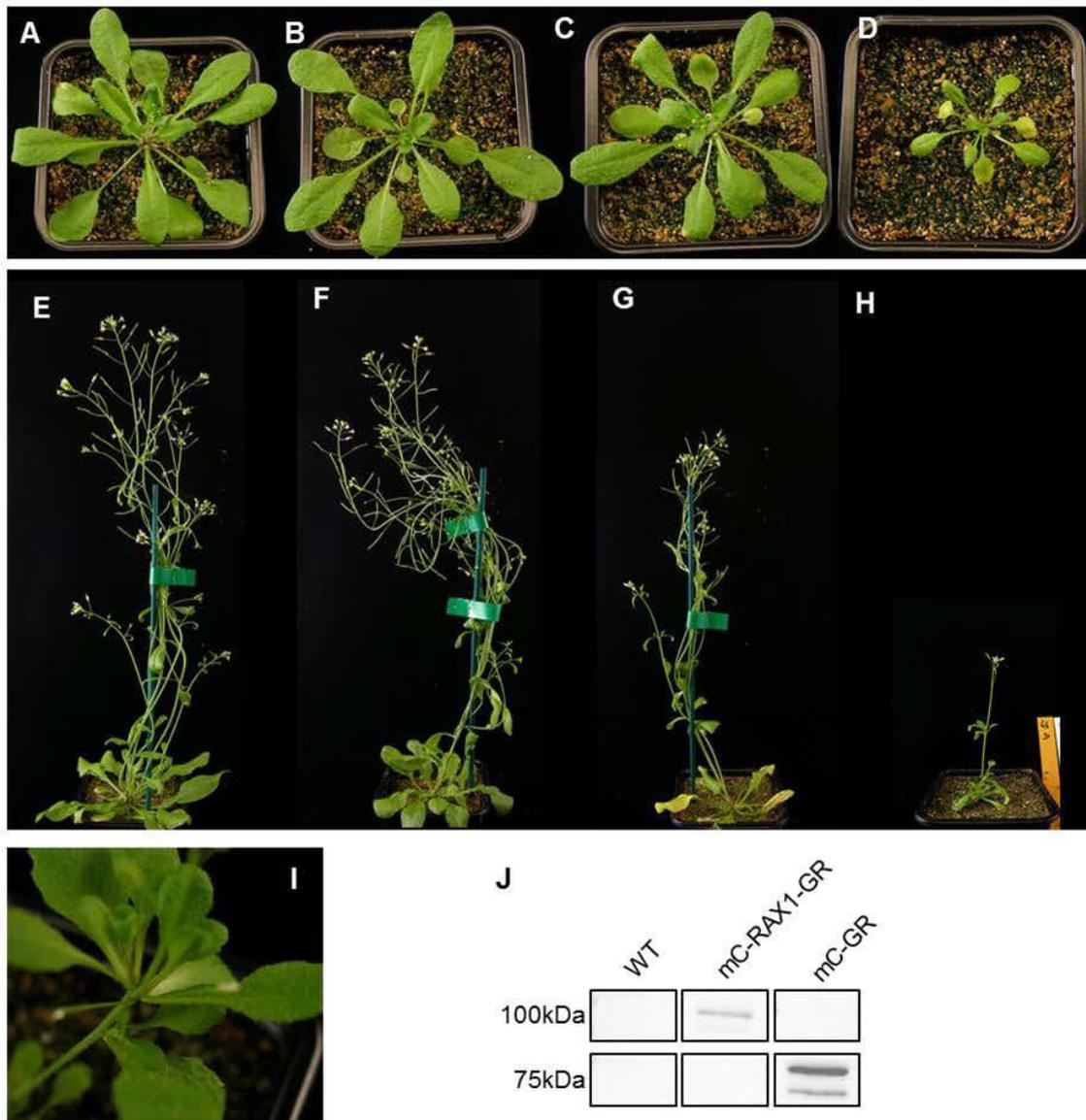
**Figure S5. *rax1-3* and *rev-6* mutants phenotype in non-inductive conditions.**

(A-D) Inflorescences of WT (A), *rax1-3* (B), *rev-6* (C) and double mutants (D) plants. (E, F) Scanning electron microscopy of *rev-6* (E) and *rax1-3 rev-6* (F) inflorescences. (G, H) Details of the structures initiated by the double mutant. The apparent bract is indicated by empty arrowhead and the stipule-like structures by filled arrowheads.



**Figure S6. SAM enlargement in double *rax rev* mutants.**

(A-C) Measure of SAM area (A), WUS-expressing area (B) on z-axis projection of inflorescences and the ratio between these values (C) in WT (white), *rax1-cl* (light grey), *rev-cl* (medium grey) and double mutant (dark grey) plants. \*\*\*\*  $p < 0.001$  determined by a Wilcoxon-Mann-Whitney test ( $N = 8$ ). (D, E) Measure of cellular Gaussian curvature on top of the SAM in WT (D) and double mutants (E). The centre of the SAM is indicated by a star. Scale bar: 20  $\mu\text{m}$ .



**Figure S7. Effects of RAX1 activity induction on growth and development.**

(A-H) 4 week-old plants (A-D) and 6 week-old plants (E-H) expressing either a *mCHERRY-GR* (A-B, E-F) or a *mCHERRY-RAX1-GR* (C-D, G-H) construct and mock- (A, C, E, G) or DEX- (B, D, F, H) treated since they were two weeks old. Notice the retarded growth of DEX-treated plants expressing *RAX1* (D, H). (I) Detail of an aerial rosette in plants in which *RAX1* activity was induced. (J) Western blot showing mCHERRY (mC) expression in the WT and the plants expressing either *mCHERRY-GR* (~75kDa) or *mCHERRY-RAX1-GR* (~100kDa).

	WT	<i>rax1-3</i>	<i>rev-6</i>	<i>rax1-3 rev-6/REV</i>	<i>rax1-3/RAX1 rev-6</i>	<i>rax1-3 rev-6</i>
Fertile flowers (%)	100.0 ± 0.0	97.7 ± 2.3	19.9 ± 5.7	99.0 ± 0.5	9.1 ± 2.4	2.5 ± 0.9
Bare cauline leaf (%)	0.0 ± 0.0	0.0 ± 0.0	3.3 ± 3.3	0 ± 0.0	45.4 ± 8.5	85.7 ± 9.7
Axillary shoot number	4.3 ± 0.3	3.9 ± 0.4	3.8 ± 0.2	3.4 ± 0.2	2.4 ± 0.3	0.6 ± 0.3

**Table S1. Axillary organ phenotype in a *rax1-3 rev-6* F2 segregating population.**

Values indicated are means +/- s.e.m.. n>8.

Identifier	Log2 fold-change	Best score	Probability of occupancy	Gene model
AT3G23240	-3.1045818	14.3718	0.57087673	ERF1
AT5G47220	-1.6378028	12.7267	0.2099967	ERF2
AT5G44790	0.9473666	12.1466	0.25259286	RAN1
AT1G45249	2.10514569	12.7197	0.36296193	ABF2
AT1G21250	-3.2283285	13.7987	0.56945638	WAK1
AT1G21270	-1.7815303	13.7987	0.38298354	WAK2
AT4G28490	-1.653037	14.3718	0.65484614	HAE
AT2G15390	-1.9132121	14.0823	0.63205174	FUT4
AT1G26770	1.28231163	13.7987	0.40058594	EXPA10
AT3G14310	-0.8946615	12.8659	0.27298032	PME3
AT2G47270	-2.1319337	12.7267	0.26848439	UPB1
AT1G75820	-1.2317506	12.7267	0.26726714	CLV1

**Table S2. Non exhaustive list of predicted putative direct RAX1 targets.**

The best score and probability of occupancy by RAX1 were determined from the RAX1 DNA-binding matrix on the genomic locus of the targets. A list of all putative direct targets can be found in Dataset 4.

identifier	Log2FC RAX1	Log2FC REV	Gene model
AT1G26770	1.282	-1.563	EXPA10
AT1G30110	-1.628	0.760	NUDX25
AT1G32090	0.965	2.320	-
AT1G51660	-0.944	-1.034	MKK4
AT1G61380	-2.114	-1.119	SD1-29
AT1G66880	-1.401	-1.618	-
AT2G02800	-0.994	-0.721	APK2B
AT2G20142	-2.143	-1.695	-
AT2G23340	1.559	1.852	DEAR3
AT2G37710	-2.004	-1.158	-
AT2G39705	1.227	1.483	RTFL8
AT2G40330	2.217	1.979	PYL6
AT3G54810	-1.139	-0.724	BME3
AT3G58620	-1.135	1.403	TTL4
AT4G16610	2.893	1.213	-
AT4G21380	-1.881	-1.311	RK3
AT4G23200	-2.357	-2.510	CRK12
AT4G23810	-2.159	1.978	WRKY53
AT4G27730	-1.704	1.410	OPT6
AT4G36410	-1.549	2.753	UBC17
AT4G38560	-4.901	-1.577	-
AT5G14730	-1.189	2.537	-
AT5G47060	1.250	1.211	-

**Table S3. Genes co-regulated by RAX1 and REV.**

Log2 fold-change (Log2FC) in response to DEX treatments in RAX1-GR (this study) and REV-GR (Reinhart et al., 2013) seedlings.

Name	Sequence 5' -> 3'	Binding score
CLV1-1	GTTGGG <u>TACCTAAC</u> TTTCTAA	12.72
RAX1bs	GTTGGG <u>TACCTAAC</u> CTTCTAA	14.37
mRAX1bs	GTTGGG <u><b>TCCATCAC</b></u> CTTCTAA	3.68

**Table S4. EMSA probes used in this study.**

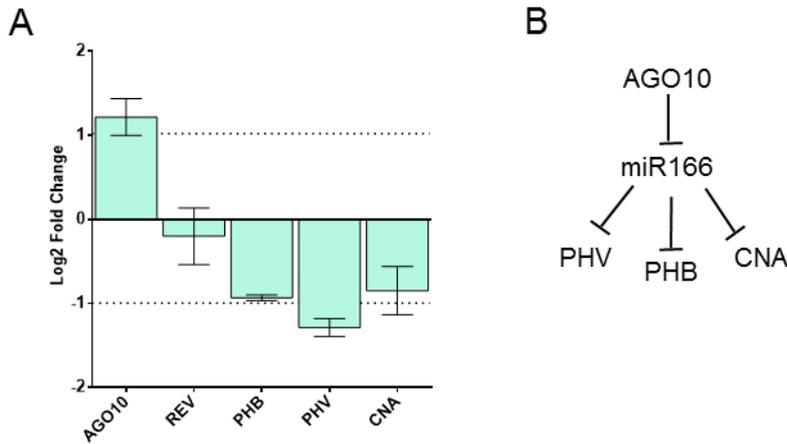
Underlined nucleotides constitute the predicted RAX1 binding-site. Bold nucleotides are mutations aiming to impair binding (in mRAX1bs). Binding score is determined by the DNA binding model of RAX1.

### III. Résultats complémentaires

#### 1. LFY pourrait réguler l'expression des HD-ZIPIII *via* AGO10

Les facteurs de transcription de la famille HD-ZIPIII sont les cibles d'une régulation par les microARN 165 et 166 (miR165/166) (Mallory *et al.*, 2004). miR166, exprimé dans la partie abaxiale des jeunes feuilles est par exemple nécessaire pour restreindre l'activité des HD-ZIPIII à la partie adaxiale de la feuille (Kidner and Martienssen, 2004). Les miR165/166 sont spécifiquement recrutés par ARGONAUTE10/PINHEAD/ZWILLE (AGO10/PNH/ZLL). AGO10 séquestre alors les micro-ARNs et les empêche de dégrader les ARNm spécifiques des HD-ZIPIII (Zhu *et al.*, 2011; Zhou *et al.*, 2015). L'absence d'AGO10 dans les mutants *pnh* entraîne l'inhibition des HD-ZIPIII dans la zone apicale de l'embryon et compromet la formation du SAM (Zhou *et al.*, 2015). Dans les fleurs, les mutations des HD-ZIPIII *PHB PHV* et *CNA* (collectively called *PPC*) entraînent un retard de l'arrêt de l'activité méristème floral, entraînant une augmentation drastique du nombre de carpelles, indiquant un rôle de ces facteurs dans l'inhibition de l'activité du FM (Landau *et al.*, 2015; Prigge *et al.*, 2005).

Il apparaît à l'observation des données ChIP anti-LFY publiées que LFY lie le promoteur de *AGO10* et inhibe son expression, bien que cette régulation ne soit pas commentée dans ces études (Winter *et al.*, 2015, 2011). Cette régulation a potentiellement des implications sur l'étude présenté en article 2. En effet *via* AGO10, LFY pourrait réguler indirectement l'expression de *REV*. De plus les facteurs HD-ZIPIII restent à ce jour très déconnectés du reste du réseau d'initiation florale malgré leur rôle important dans ce processus. Il est donc intéressant de savoir si leur expression est influencée par l'activité de LFY.



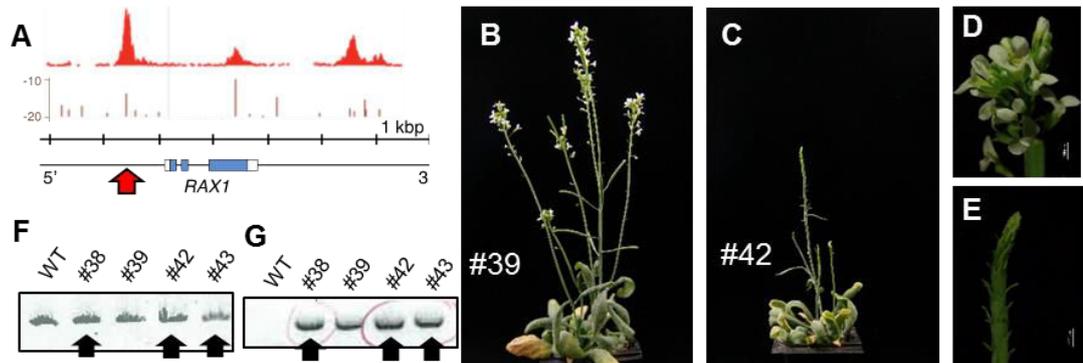
**Figure Chl. 1. Régulation des HD-ZIPIII et AGO10 par LFY.** (A) Abondance relative des transcrits dans des inflorescences du mutant *lfy-12* par rapport au sauvage. (B) Modèle de la régulation des HD-ZIPIII *PHB*, *PHV* et *CNA* par AGO10 qui séquestre miR166, empêchant la dégradation des HD-ZIPIII.

Afin d'étudier la régulation des HD-ZIPIII par LFY, *via* AGO10, nous avons entrepris de mesurer l'abondance de leurs transcrits dans les inflorescences du mutant *lfy-12* (Figure Chl. 1A). Cette analyse montre une augmentation d'environ deux fois des transcrits d'AGO10 dans le mutant *lfy-12* ce qui suggère que LFY réprime AGO10. On observe aussi une baisse du niveau des ARNm *PPC* alors que les transcrits *REV* ne semblent pas affectés (Figure Chl. 1A-B). L'augmentation du niveau d'AGO10 devrait logiquement provoquer une augmentation des transcrits de *PPC*, une inhibition d'AGO10 par LFY serait alors cohérente avec le rôle d'inhibition des *PPC*. Or le contraire est observé ici. Ces résultats contradictoires indiquent que la régulation des HD-ZIPIII dans le contexte des inflorescences *lfy* n'est pas uniquement dépendante d'AGO10. La relation de LFY avec AGO10 et les PPC reste donc à définir par une approche plus complète. En revanche cette analyse indique que LFY ne régule pas *REV*, ces deux FT agissent donc bien dans des voies parallèles.

## 2. Mutation d'un élément *cis* du promoteur de *RAX1*

LFY lie directement le promoteur de *RAX1* et la fusion de LFY au domaine de *trans*-activation VP16 permet de stimuler l'expression de *RAX1* (Chahtane *et al.*, 2013 et Article 2; Moyroud *et al.*, 2011). Cependant les inflorescences des mutants *lfy* expriment toujours *RAX1* (Chahtane *et al.*, 2013). Cela peut être

dû au fait que les inflorescences *lfy* ont une identité végétative et que *RAX1* est régulé par les mêmes facteurs qui contrôlent son expression dans les tissus végétatifs et/ou que d'autres facteurs agissent de manière redondante à LFY dans l'inflorescence.



**Figure Chl. 2. Tentative de mutation d'un site de liaison à l'ADN de LFY dans le promoteur de *RAX1*.** (A) Schéma du locus *RAX1* (bas) avec le signal de ChIP anti-LFY (en rouge, au-dessus) et les sites de fixations prédits de LFY (en brun, au milieu). Le site ciblé par la construction CRISPR/Cas9 est indiqué par une flèche rouge. Adapté de (Chahtane *et al.*, 2013). (B-E) Architecture (B-C) et détail de l'inflorescence (D-E) de plantes au phénotype de type *rev* (#39, B, D) et double mutant (#42, C, E) dans la deuxième génération après transformation. (F) Gels de génotypage (voir Matériel et Méthodes) de mutations induites par une rupture de l'ADN. Aucune mutation n'a pu être détectée. (G) Gel de génotypage pour la présence du T-DNA portant la construction CRISPR/Cas9. Les plantes présentant un phénotype double mutant sont indiquées par une flèche.

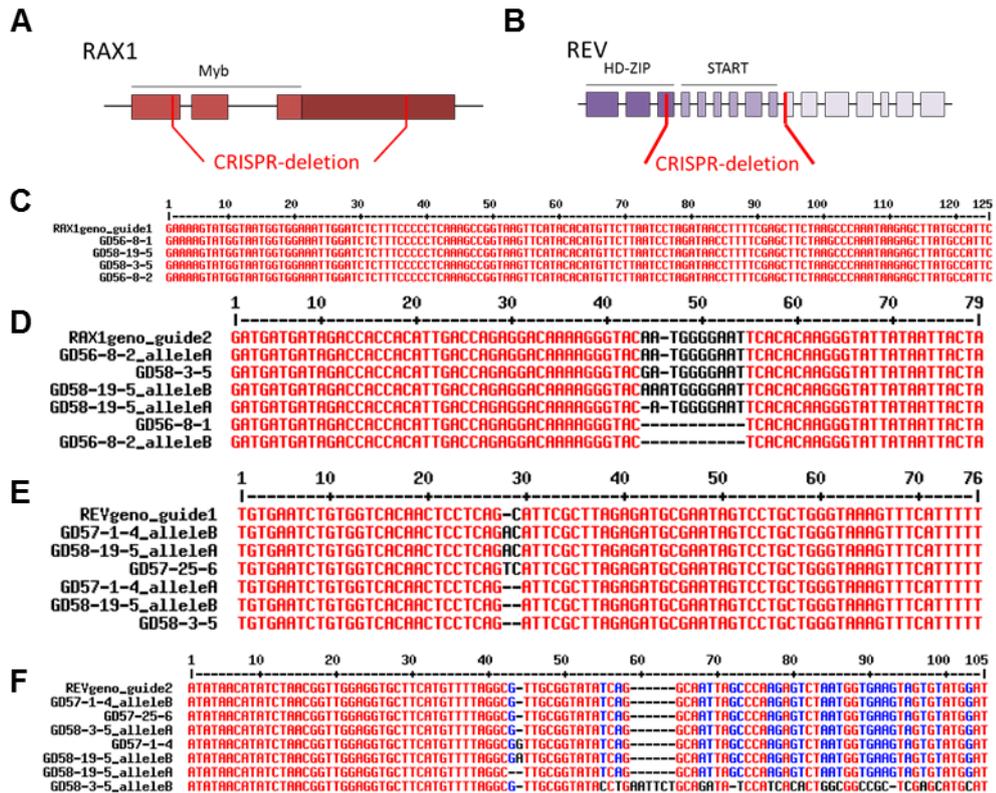
Afin d'estimer l'importance de la régulation de *RAX1* par LFY dans les FM, nous avons tenté de muter le site de liaison à l'ADN de LFY dans le promoteur de *RAX1* grâce à une construction CRISPR/Cas9 (Figure Chl. 2A). Cette approche est facilitée par la présence d'une répétition de guanine dans le site de liaison de LFY (Figure Intro. 11A). Cette répétition de guanine constitue le site PAM (*Protospacer adjacent motif*) nécessaire à la fixation de la Cas9 à l'ADN (Jinek *et al.*, 2012). Afin de pouvoir visualiser facilement le résultat d'une modification de l'expression de *RAX1*, nous avons transformé cette construction dans le fond sensibilisé *rev-6*. Dans la première génération après transformation nous avons pu repérer quelques plantes présentant un phénotype similaire aux mutants *lfy rev* et *rax1 rev* (Figure Chl. 2B-E). Cependant nous n'avons pas détecté de mutations au locus ciblé par la Cas9, malgré nos tentatives répétées sur plusieurs échantillons incluant les inflorescences. Parmi les rares descendants de ces plantes, aucune ne

présentait le phénotype *rax1 rev*. Par contre ce phénotype a été retrouvé dans la descendance d'une plante T1 de phénotype *rev-6*. Là encore parmi ces T2 au phénotype *rax1 rev* aucune mutation n'a pu être détectée au locus ciblé sur de l'ADN d'inflorescence affectée, que ce soit par génotypage PCR ou par séquençage du locus (Figure Chl. 2F). Par contre toutes les plantes présentant un phénotype de fleurs filamenteuses portaient encore le transgène (Figure Chl. 2H).

Nous formulons l'hypothèse que l'effet observé serait dû à la simple présence de la Cas9 dans le promoteur de *RAX1*, bloquant potentiellement la transcription du gène. La relative nouveauté de l'application de la technologie CRISPR à l'édition de génome fait que nous manquons cruellement de données permettant de prédire l'efficacité de la coupure de l'ADN. Il est possible qu'une combinaison de facteurs empêche la coupure du site choisi pour cette étude, induisant une répression du locus.

### 3. Quelques commentaires sur les lignées *CRISPR/Cas9* ciblant *RAX1* et *REV*

Dans l'article 2 nous présentons des lignées mutantes *rax1* et *rev* produites à l'aide de la technologie *CRISPR/Cas9*. Ces lignées ont permis d'accélérer considérablement le déroulement de ce travail pendant le retro-croisement du mutant *rev* dans Col et de combiner des mutations dans deux gènes liés (*REV* et *LFY*). Afin d'identifier plus facilement les plantes chez lesquelles la construction *CRISPR/Cas9* aurait induit des mutations, nous avons utilisé simultanément deux spacers par gènes pour provoquer une délétion d'une partie du locus (Figure Chl. 3A-B). Malgré notre enthousiasme initial cette approche présente des limitations et apparaît rétrospectivement superflue. Ce qui suit est un bref commentaire empirique sur l'usage du CRISPR dans ce travail.



**Figure Chl. 3. Mutations induites par le système *CRISPR/Cas9* dans les loci *RAX1* et *REV*.** (A-B) représentation des loci *RAX1* (A) et *REV* (B). Les exons sont indiqués par des cadres et les introns et les régions non traduites par des lignes. Les sites ciblés par les constructions *CRISPR/Cas9* sont indiqués par des lignes rouges. (C-F) Alignement des séquences des loci ciblés par la construction *CRISPR/Cas9*. (C) *RAX1* guide 1, (D) *RAX1* guide 2, (E) *REV* guide 1, (F) *REV* guide 2. La séquence sauvage est indiquée en premier.

Dans la première génération après transformation (T1) le taux de plantes portant une délétion dans le locus observé dans les dernières feuilles de rosettes était de 3% pour *RAX1* et de 8% pour *REV*. Cependant le nombre de plantes ayant un phénotype de type *rev* (feuilles enroulées et défauts floraux) était bien supérieur à cela avec 37% des plantes (N= 60). Cette différence indique une faible fréquence d'évènements de délétion par rapport aux coupures simples. L'absence de phénotype évident pour *RAX1* ne nous a pas permis d'estimer le taux d'efficacité de la Cas9 dans ces lignées. Afin de maximiser nos chances d'identifier des plantes portant une délétion nous avons génotypé les plantes à partir d'échantillons issus des dernières feuilles caulines. Si le taux de délétion observé était similaire à ce qui avait été observé dans les échantillons de feuilles de rosettes, les plantes identifiées étaient différentes de celles identifiées précédemment, indiquant une importante mosaïcité des plantes de T1.

Pour ces raisons nous avons sélectionné quelques lignées de T1 d'après leur phénotype et/ou l'observation d'une délétion dans un des génotypes (indiquant l'activité du transgène) pour une caractérisation dans la génération suivante (T2). Si nous avons échoué à identifier des plantes portant une délétion dans le locus observé, la grande majorité des plantes portait une mutation. Cela semble indiquer que la coupure simultanée des deux cibles est un phénomène relativement rare, rendant la délétion peu efficace. Dans la T2 nous avons pu observer des plantes mutantes homozygotes, des hémizygotés et quelques plantes mosaïques. L'ADN de ces plantes a été séquencé pour identifier les mutations induites par la coupure de la Cas9 (Figure ChI 3C-F). La gamme de mutations observée comprend principalement des insertions/délétions d'une paire de bases, provoquant un changement du cadre de lecture entraînant la dégénérescence de la séquence protéique et l'apparition d'un codon stop. Nous n'avons pu identifier de mutation induite par le guide ciblant le premier exon de *RAX1* probablement du fait de l'inactivité du guide.

De cette expérience il ressort quelques recommandations sur l'usage du CRISPR. (1) Afin de limiter le travail laborieux de transformation et sélections des mutants il semble important de vérifier *in vitro* l'activité des spacers sélectionnés. (2) L'efficacité d'un guide seul pour induire des mutations est largement supérieure à l'efficacité de deux guides pour induire une délétion. De plus le génotypage par séparation sur gel d'acrylamide (voir Matériel et Méthodes) permet l'identification efficace de mutants portant une insertion/délétion d'une paire de bases. L'utilisation de deux guides par locus apparaît alors comme une perte de temps. (3) Les plantes T1 mutantes étant en grande majorité mosaïque, il est plus simple de sélectionner les T1 en se basant sur le phénotype lorsque cela est possible et de caractériser les mutations en T2.

#### IV. Conclusion et perspectives

Dans l'article 2, nous avons montré que LFY, *via* la régulation de *RAX1* joue un rôle dans l'initiation du méristème floral. Bien que la fonction de REV dans

ce processus semble supplanter en importance celle de la voie contrôlée par LFY, l'utilisation du fond *rev* nous montre que *LFY* et *RAX1* ont un rôle non négligeable dans ce processus. L'action de *RAX1* passe probablement en grande partie par *CLV1*, bien que nos données ne permettent pas de conclure sur le rôle spécifique de l'inhibition de *CLV1* par *RAX1*. Comme discuté dans l'article, nous faisons l'hypothèse que l'inhibition de *CLV1* par *RAX1*, probablement pour moduler son expression puisque *CLV1* est exprimé dans le FM (Clark *et al.*, 1997), est requise pour établir la balance *CLV3/WUS* dans le FM. Démontrer cette hypothèse demandera plus d'expériences visant à déterminer l'expression de *CLV1* dans les doubles mutants *rax1 rev* et l'effet de la quantité de *CLV1* dans le FM sur son initiation.

Une conséquence intéressante de ce travail serait l'intégration de la boucle complète *CLV1/CLV3/WUS* dans les modèles de régulation du méristème. Malgré la relative précision des modèles récents (Gruel *et al.*, 2016; Adibi *et al.*, 2016), la régulation croisée de *WUS* et *CLV3* est simplifiée à ces deux éléments. Il serait particulièrement intéressant de voir s'il est possible d'inférer de nouvelles relations après l'ajout de *CLV1* dans ce système.

Ce travail présente le premier lien mécanistique direct à notre connaissance entre le maximum d'auxine déclenchant la formation d'une fleur et l'établissement du méristème floral. D'une manière remarquable ces processus sont inextricablement liés à la détermination de l'identité florale, grâce à l'implication de *LFY*. L'activité de *LFY* permet probablement de contrôler le développement particulier du FM par rapport au LSM : répression de la croissance de la feuille, formation rapide du méristème et absence de dormance.







# CHAPITRE SECOND

## Etude de l'oligomérisation de LFY et de son rôle dans l'induction d'AGAMOUS

Le facteur de transcription LFY est un régulateur clé du développement floral et est également impliqué dans la formation des méristèmes floraux (*c.f.* Introduction et Chapitre I). En plus de ces rôles LFY est impliqué dans la transition florale (Blázquez et al., 1997), la croissance et la polarité du pédicelle (Yamaguchi et al., 2012), et la régulation des défenses de la plante (Winter et al., 2011). La diversité des rôles de LFY et les différents stades développementaux au cours desquels LFY joue un rôle suggère l'existence de mécanismes permettant la régulation fine de son activité.

LFY régule le développement floral avec l'aide des corégulateurs WUS et UFO qui, lorsqu'ils sont exprimés avec LFY, permettent la formation de fleurs à partir de tissus racinaires ou foliaires respectivement (Gallois et al., 2004; Risseuw et al., 2013). Il a été montré que UFO et LFY interagissent au niveau protéique *via* le DBD de LFY (Chae et al., 2008), cependant l'interaction physique de LFY et WUS n'a jamais été mise en évidence. En plus de

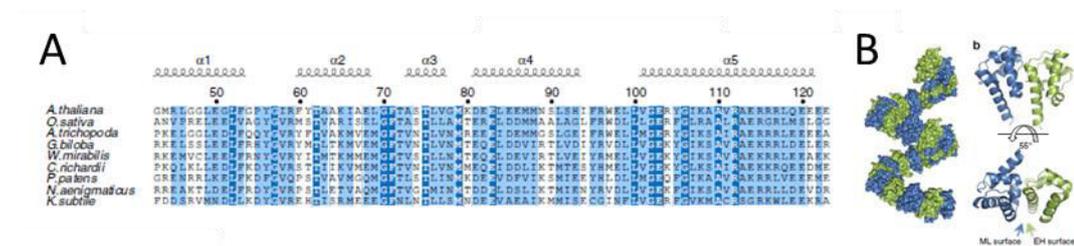
l'interaction avec ses cofacteurs, l'activité de LFY pourrait être régulée par son domaine N-terminal conservé dont la fonction était inconnue au début de ce travail. Dans ce chapitre je présenterai mes travaux sur la caractérisation du domaine N-terminal de LFY, impliqué dans l'oligomérisation de la protéine, et ses implications sur les propriétés régulatrices de LFY. Je me pencherai en particulier sur la régulation du locus *AG* par LFY et WUS, pour laquelle le domaine N-terminal de LFY peut avoir un rôle important.

## I. Rôles du domaine N-terminal de LFY

### 1. Introduction

Il a été montré précédemment que le domaine N-terminal de LFY pourrait être un domaine de type leucine-zipper et permettrait la dimérisation de la protéine d'une manière indépendante de la liaison à l'ADN (Siriwardana and Lamb, 2012a). LFY contient cependant déjà un domaine permettant la dimérisation de la protéine sur l'ADN (Hamès et al., 2008). L'absence de données structurales ou de domaines fonctionnels prédits pour le domaine N-terminal de LFY ne permettant pas de prédire de manière certaine son rôle, une approche structure-fonction a été entreprise avec la résolution de sa structure moléculaire par diffraction aux rayons X (Sayou et al., 2016). Le domaine N-terminal étant très conservé chez les plantes supérieures (Figure ChII.1A), le domaine de la protéine du *Ginkgo biloba* (GbLFY) a été cristallisé (Figure ChII.1B). Cette structure correspond à celle d'un domaine Sterile Alpha Motif (SAM) (Qiao and Bowie, 2005), qui n'avait pu être prédit par comparaison de séquences. Les domaines SAM classiques permettent la formation d'oligomères par interaction tête à queue (Meruelo and Bowie, 2009). Il a été montré que le domaine SAM de LFY a également ces propriétés d'oligomérisation, permettant la formation d'oligomères comportant jusqu'à huit domaines SAM (Sayou et al., 2016). Afin d'étudier les propriétés du domaine SAM une version de LFY affectée spécifiquement pour l'oligomérisation (GbLFY<sub>TERE</sub>) a été créée. Cette double mutation du domaine SAM (T75E, R112E) n'empêche pas la protéine de reconnaître les sites de

liaison de LFY tel que celui présent dans le promoteur d'AP1 (Benlloch et al., 2011; Sayou et al., 2016). Cependant, alors que la protéine sauvage interagit avec la sonde *AP1* uniquement sous forme de dimère, la liaison de monomères de GbLFY<sub>TERE</sub> peut être détectée. Cet effet est d'autant plus marqué en utilisant une sonde *AP1* mutée (*AP1m*), contenant un demi-site de liaison LFY, sur laquelle seul GbLFY<sub>TERE</sub> se fixe (Sayou et al., 2016). Ces résultats indiquent que le domaine SAM empêche la liaison de LFY à l'ADN en tant que monomère et participe à la coopérativité de liaison sur les sites complets (*i.e.* deux demi-sites).



**Figure ChII. 1. Conservation et structure de domaine N-terminal de LFY.** (A) Alignement de séquences du domaine N-terminal dans la lignée verte (B) Modèle structural du domaine N-terminal de GbLFY. Polymérisation du domaine N-terminal en hélice (gauche) et vues de côté (haut droite) et de dessus (bas droite) d'un dimère de domaine N-terminal. Adapté de Sayou *et al.* 2013

Afin d'identifier la fonction du domaine SAM chez *Arabidopsis*, *AtLFY<sub>TERE</sub>* a été exprimée sous contrôle du promoteur LFY endogène dans un fond perte de fonction *lfy-12*, ou sous contrôle du promoteur constitutif 35S (Figure ChII.2). Chez les mutants *lfy-12* l'identité florale n'est pas spécifiée et les fleurs sont converties en tiges ou en structures foliaires (Weigel et al., 1992). Alors que l'expression de *pLFY:LFY* complémente presque complètement la perte de LFY, permettant la formation de fleurs fertiles, l'expression de *pLFY:LFY<sub>TERE</sub>* est bien moins efficace. Ainsi les plantes exprimant *pLFY:LFY<sub>TERE</sub>* produisent uniquement quelques fleurs fertiles puis une multitude de structures mutantes stériles, indiquant un rôle de l'oligomérisation dans la spécification de l'identité florale. De plus, alors que l'expression de *p35S:LFY* entraîne la formation de fleurs à l'aisselle des feuilles de rosettes (Weigel and Nilsson, 1995), les plantes exprimant *p35S:LFY<sub>TERE</sub>* sont

dépourvues de fleurs ectopiques. L'ensemble de ces résultats indique la nécessité de l'oligomérisation de LFY pour son rôle de maître régulateur du développement floral.

A la suite de ces résultats, la question du rôle de l'oligomérisation dans la liaison de LFY à l'ADN et la régulation de ses gènes cibles se pose naturellement. Les résultats ci-après sont issus de ma contribution à ce projet et sont partie intégrante de la publication en Annexe 2 (Sayou et al., 2016).

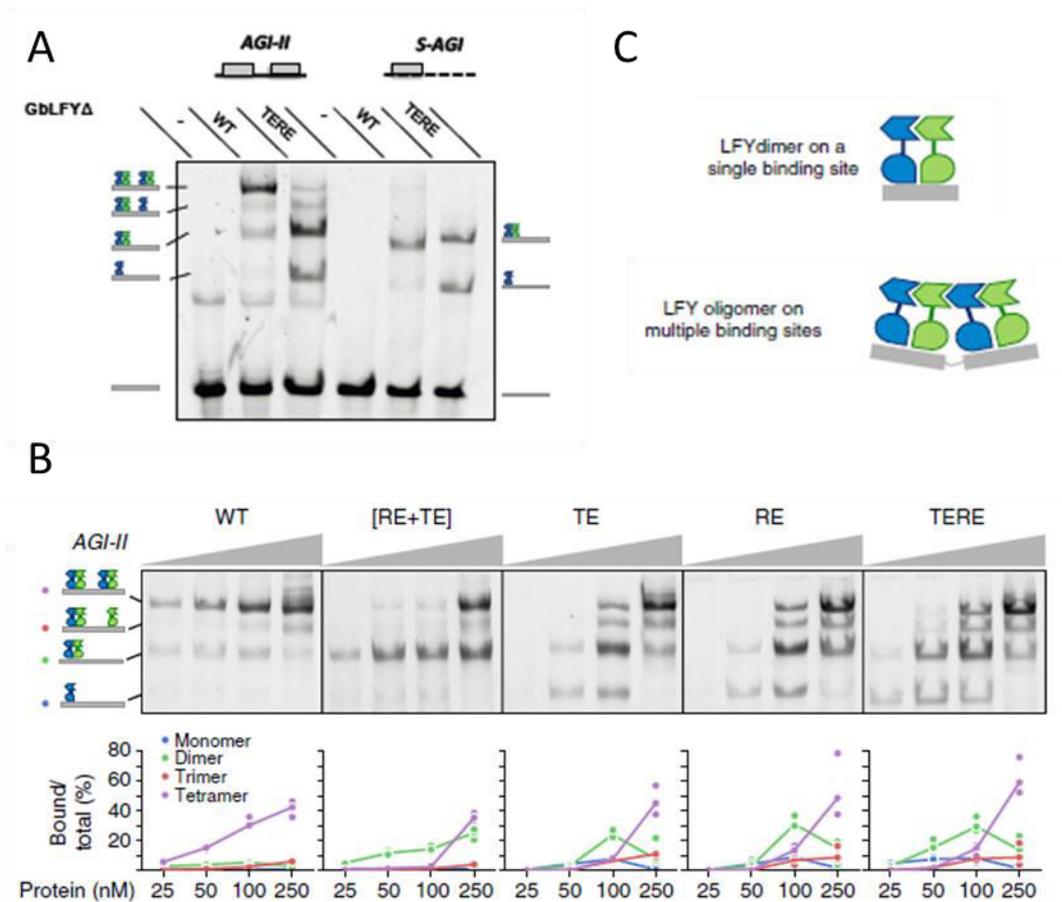


## 2. Résultats

*Le domaine SAM de LFY permet la liaison coopérative de sites distants in vitro*

Il a été proposé que les domaines SAM puissent permettre la liaison coopérative de facteurs de transcription sur plusieurs sites distants (Zhang et al., 2010). Le domaine SAM permettant l'oligomérisation de LFY et étant nécessaire à sa fonction de régulateur clé du développement floral nous avons

testé son rôle pour la liaison *in vitro* de sites multiples. Nous avons utilisé une sonde contenant un ou deux sites de liaison de LFY proches, sur une séquence synthétique ou issue de la région régulatrice du deuxième intron d'*AGAMOUS* (*AGI-II*) (Busch et al., 1999; Lohmann et al., 2001) (Figure ChII.3A-B). De manière cohérente avec les observations faites précédemment sur la sonde *AP1*, GbLFY se lie presque exclusivement en dimère sur la sonde contenant un simple site (*S-AGI*) alors que GbLFY<sub>TERE</sub> se lie avec une affinité similaire sous forme de monomère et de dimère. Sur la sonde *AGI-II*, GbLFY est détecté majoritairement en tant que tétramère et plus faiblement en tant que dimère. GbLFY<sub>TERE</sub> forme majoritairement des complexes correspondant à la liaison de deux monomères et plus faiblement d'un seul monomère, cependant des complexes correspondant à la liaison de trois ou quatre monomères sont également détectés. Nous avons par alors testé la liaison de simples mutants GbLFY<sub>TE</sub> (T75E) et GbLFY<sub>RE</sub> (R112E) qui ne sont plus capables de s'oligomériser mais peuvent former des dimères (Figure ChII. 4). Leurs profils de liaison sur la sonde *AGI-II* est similaire à celui de GbLFY<sub>TERE</sub>. Cependant lorsque ces deux variantes GbLFY<sub>TE</sub> et GbLFY<sub>RE</sub> sont mélangées, les complexes correspondant à la liaison d'un ou trois monomères sont détectés plus faiblement et les complexes correspondants à deux monomères plus fortement. A des concentrations plus fortes en protéines tous les sites présents sur la sonde sont liés pour chacune des variantes de LFY. Comme observé précédemment le domaine SAM empêche la liaison de l'ADN sous forme monomérique, sa mutation résultant en la formation de complexes contenant un, deux, trois ou quatre monomères liés indépendamment. Le mélange [TE+RE] résulte en la formation majoritaire de dimères, probablement due à la capacité de la protéine à former des dimères *via* le domaine SAM (Figure ChII.4). Pour le mélange [TE+RE], les complexes correspondant à la fixation de deux dimères sont observés à partir de 100nM de protéine, alors que pour la protéine sauvage, ces complexes sont observés dès 25nM de protéine. L'ensemble de ces résultats indique que la formation de chaînes oligomériques *via* le domaine SAM permet la liaison coopérative de sites distants *in vitro* (Figure ChII.3C).

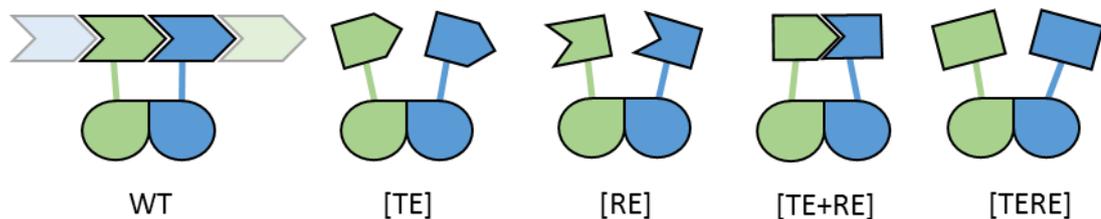


**Figure ChII. 3. L'oligomérisation du domaine SAM de LFY permet la liaison coopérative de sites multiples *in vitro*.** (A) Liaison de GbLFY et GbLFY<sub>TERE</sub> sur des sondes contenant un ou deux sites de liaison. La sonde AGI-II contient deux sites de liaison de LFY extraits de l'intron d'AG. La sonde S-AGI est une sonde synthétique contenant un seul site de liaison issu du même intron. Les complexes supposément formés sont indiqués sur les bords de l'image. (B) Oligomérisation des mutants de GbLFY sur AGI-II. Une concentration croissante de protéine a été utilisée. Les complexes supposément formés sont indiqués à gauche des gels. La quantification de chaque complexe est indiquée sous les gels. (C) Modèle liaison de LFY sur un site unique (haut) et sur un double site (bas). Adapté de Sayou *et al.* 2016.

#### *Le domaine SAM permet la liaison de clusters de sites in vivo*

Nous nous sommes ensuite posé la question du rôle de l'oligomérisation sur la liaison de sites distants *in vivo*. Nous avons analysé les données de liaison à l'ADN issues d'immuno-précipitation de chromatine (ChIP) dans des plantes exprimant *p35S:LFY* ou *p35S:LFY<sub>TERE</sub>* en utilisant un anticorps dirigé contre le DBD de LFY (Sayou *et al.*, 2016). Les lignées présentaient un niveau d'expression similaire de protéine LFY et l'anticorps utilisé reconnaît les protéines avec la même affinité. L'étude a été faite à partir d'échantillons de plantules de 14 jours afin de pouvoir comparer la liaison de LFY et LFY<sub>TERE</sub>

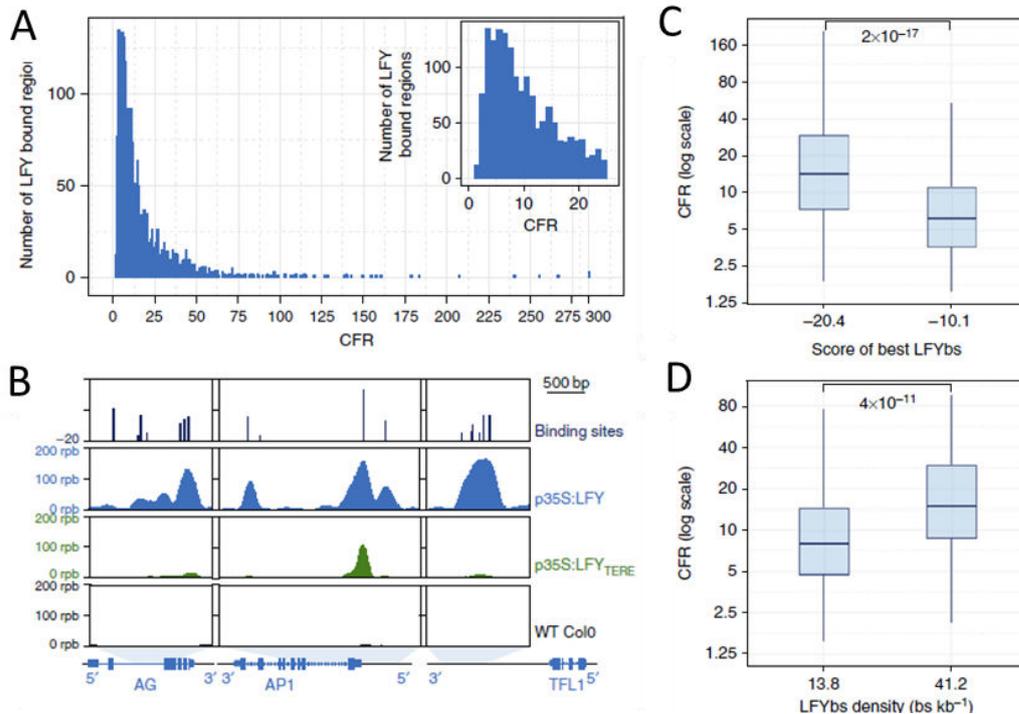
dans les mêmes tissus, ce qui n'aurait pas été possible à des stades de développement plus tardifs, vu les phénotypes de ces plantes. Nous avons comparé l'intensité de liaison de LFY à celle de LFY<sub>TERE</sub> pour chaque région liée par LFY (FDR < 10<sup>-7</sup>). Cette analyse a révélé une réduction globale de l'intensité de liaison (*coverage fold-reduction* CFR) pour LFY<sub>TERE</sub> (Figure ChII.5A). La CFR varie de 1,3 fois à plus de 150 fois, 95% des régions concernées présentant une réduction de plus de 3 fois, et 54% des régions une réduction de plus de 10 fois. Ces résultats indiquent que l'oligomérisation du SAM est requise pour la liaison de LFY à ses gènes cibles *in vivo*.



**Figure ChII. 4. Représentation schématique de LFY et des mutations TE, RE et TERE.** La protéine sauvage dimérise à la fois *via* le DBD (bas) et *via* le domaine SAM (haut). Le domaine SAM médie également l'oligomérisation. Les mutations TE et RE détruisent chacune l'une des interface d'interaction du domaine SAM, empêchant la protéine d'oligomériser. Le mélange de LFY<sub>TE</sub> et de LFY<sub>RE</sub> permet la formation de dimères de SAM.

La CFR peut s'expliquer partiellement par l'étalement des LFY<sub>TERE</sub> sur le génome, les demi-sites LFY étant beaucoup plus nombreux que les sites complets. Cependant une grande hétérogénéité de la CFR a été observée à travers le génome, indiquant la présence de facteurs autre qu'un effet global de dilution de LFY sur le génome (Figure ChII.5A). Afin d'identifier les facteurs jouant sur la CFR nous avons examiné l'influence de la densité et de la qualité des sites de liaison. Pour ce faire nous avons déterminé *in silico* les sites de liaison de LFY dans les régions liées, puis nous avons déterminé la densité des sites de liaison et calculé l'affinité (score) de chaque site (Moyroud et al., 2011; Minguet et al., 2015). Cette analyse montre que les régions de haute CFR ont une densité de sites plus élevée et que ces sites sont de globalement plus faible affinité (Figure ChII.5C-D), comme illustré ici pour certains loci (Figure ChII.5B). Ces résultats indiquent que l'oligomérisation de LFY permet

la liaison de groupes de sites de moyenne ou faible affinité alors que les sites de forte affinité semblent moins dépendants de l'oligomérisation.



**Figure ChII. 5. L'oligomérisation du domaine SAM de LFY permet la liaison coopérative de sites multiples *in vivo*.** (A) Distribution de la CFR pour les régions liées par LFY. (B) Exemples de CFR sur des régions sélectionnées du génome. La première ligne montre les sites de liaison prédits de LFY, les lignes du dessous montrent le signal de ChIP pour LFY, LFY<sub>TERE</sub> et le sauvage respectivement. La structure des loci considérés et la localisation des régions prise en compte sont schématisées en bas du graphe. (C) Distribution de la CFR dans le premier et le dernier décile des régions triées selon le meilleur site de liaison. Les meilleurs sites sont ceux dont le score est le plus proche de zéro. (D) Distribution de la CFR dans le premier et le dernier décile des régions triées par densité de site de liaison. Adapté de Sayou *et al.* 2016.

### Le domaine SAM permet la liaison aux régions fermées de la chromatine

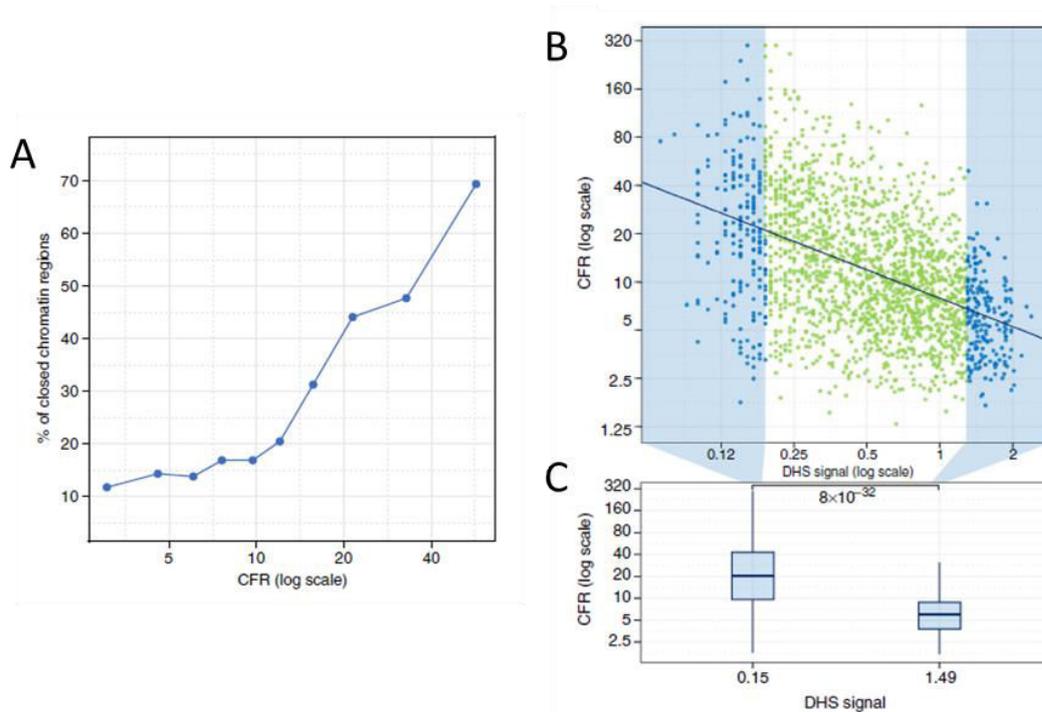
Il a été montré précédemment que LFY interagit avec les remodelleurs chromatinien SPLAYED (SYD) et BRAHMA (BRM) (Wu *et al.*, 2012). Il a également été montré que certaines protéines oligomérisant *via* un domaine SAM pouvaient réguler la chromatine, notamment à travers l'interaction avec des remodelleurs chromatinien (Webber *et al.*, 2013; Isono *et al.*, 2013). Nous avons donc entrepris de croiser les données de CFR avec des données d'accessibilité de la chromatine dans les mêmes tissus, publiées précédemment (Zhang *et al.*, 2012). Nous avons trié les pics de liaison de LFY

par valeur de CFR et groupé ces régions par déciles. Pour chaque décile nous avons déterminé le nombre de régions 'fermées' selon les critères déterminés par Zhang *et al.* Cette analyse montre un enrichissement drastique en régions fermées dans les régions de haute CFR (Figure ChII.6A). Nous avons ensuite déterminé le signal DNaseI-seq (*i.e.* l'accessibilité de la chromatine) pour chaque région liée par LFY. Cette analyse montre là encore un enrichissement du signal DNase dans les régions de faible CFR alors que les régions de forte CFR ont un signal DNase plus faible (Figure ChII.6B-C). Ces résultats indiquent que l'oligomérisation de LFY pourraient lui permettre de lier des régions normalement inaccessibles de la chromatine.

### 3. Discussion

La résolution de la structure moléculaire du domaine N-terminal de GbLFY a révélé qu'il s'agit d'un domaine de type SAM, permettant l'oligomérisation de la protéine *in vitro*. La haute conservation de séquence protéique de ce domaine chez les plantes supérieures permet de penser qu'il est aussi présent et fonctionnel chez *Arabidopsis*. De plus, l'introduction de l'homologue de LFY de *Ceratopteris* (fougère) ou *Welwitschia* (gymnosperme) dans un fond mutant *lfy* chez *Arabidopsis* permet de compenser partiellement la perte de LFY, indiquant que ces protéines relativement distantes permettent de remplir la fonction d'AtLFY (Maizel et al., 2005). De manière cohérente avec cette hypothèse l'expression d'une protéine LFY incapable d'oligomériser chez *Arabidopsis* ne permet pas de remplir la fonction de la protéine sauvage. Il existe plusieurs allèles de LFY affectée dans le domaine SAM : *lfy-22* contient une mutation localisée à la surface du SAM d'après les données structurales présentées ici. Cette mutation ne produit de phénotype qu'en combinaison avec la mutation du cofacteur UFO et ne perturbe donc probablement que légèrement l'oligomérisation (Levin and Meyerowitz, 1995). Par contre la fonction de protéines LFY contenant des mutations dans des leucines à l'intérieur du SAM est fortement perturbée (Siriwardana and Lamb, 2012a). Cependant ces mutations jouent probablement sur le repliement même du SAM, plutôt que sur l'interface d'oligomérisation. L'allèle mutant présenté ici (LFY<sub>TERE</sub>) est perturbé spécifiquement dans l'interface d'oligomérisation et

nous permet donc de poser précisément la fonction de l'oligomérisation d'un point de vue biochimique et génétique.



**Figure ChII. 6. La réduction de liaison du au domaine SAM corrèle avec l'accessibilité de la chromatine.** (A) Proportion de régions annotées comme « chromatine fermées » par décile de régions liées par LFY triées selon la CFR. (B) Signal de DNase (indique l'accessibilité de la chromatine) et CFR pour chaque région liée par LFY. La régression linéaire (noir) à un coefficient de corrélation  $r^2 = 0,2171$ . (C) Distribution de la CFR dans le premier et le dernier décile des régions liées par LFY triées selon leur signal DNase. Adapté de Sayou *et al.* 2016.

Nous avons montré ici que le domaine SAM impose la liaison de LFY en dimère sur deux demi-sites. Le DBD de LFY confère déjà des capacités de dimérisation de la protéine sur l'ADN. Si cette propriété a été montrée pour le DBD seul (Hamès *et al.*, 2008), les propriétés de LFY<sub>TERE</sub> indiquent que domaine SAM est nécessaire à la liaison en dimère de la protéine entière. La manière dont le domaine SAM contrôle la liaison en dimère demeure énigmatique mais pourrait passer par les interactions SAM-SAM ou par une modification de la conformation du DBD. Cette propriété pourrait être requise pour empêcher la dispersion de LFY sur les nombreux demi-sites du génome et concentrer la liaison sur les sites complets. Le domaine SAM pourrait également permettre la liaison coopérative de plusieurs sites distants, *in vitro*

et *in vivo*. Cette propriété permet à LFY de se lier à des sites de faible ou moyenne affinité mais de haute densité et sa perturbation provoque une réduction globale de liaison du facteur sur l'ADN.

L'oligomérisation pourrait également permettre de former des boucles d'ADN entre des sites génomiques plus distants. La formation de boucles d'ADN, par les facteurs de type MADS notamment, est importante pour la régulation de transitions développementales chez *Arabidopsis* (Mendes et al., 2013; Melzer et al., 2009; Cao et al., 2014). De manière intéressante il a été montré récemment que certains facteurs MADS ont un rôle pionnier, liant la chromatine fermée et recrutant les remodelleurs SYD et BRM pour l'ouvrir et permettre la liaison d'autres régulateurs transcriptionnels (Wu et al., 2012; Pajoro et al., 2014a). Il est également intéressant de noter que le facteur ARF5/MP est également capable d'oligomériser *via* le domaine PB1 et il a été montré récemment que sa capacité à recruter SYD et BRM lui permet de modifier l'état chromatinien et d'activer ses gènes cibles (Nanao et al., 2014; Wu et al., 2015). Un rôle similaire a été montré pour des facteurs de type SAM, capables d'oligomériser et de recruter des remodelleurs chromatiniens (Isono et al., 2013; Zhang et al., 2010). LFY interagit également avec SYD et BRM (Wu et al., 2012) et nous montrons ici que le domaine SAM lui permet de lier des régions fermées de la chromatine. Ces propriétés pourraient faire de LFY un facteur pionnier, cependant des expériences complémentaires sont nécessaires pour valider cette hypothèse. En particulier nous avons ici montré une corrélation entre l'état de la chromatine d'une région et l'importance du domaine SAM pour lier cette région, il serait intéressant d'examiner l'état de la chromatine dans les lignées exprimant LFY ou LFY<sub>TERE</sub> afin de déterminer si LFY joue un rôle dans l'ouverture de la chromatine.

Une autre hypothèse intéressante sur le rôle de l'oligomérisation est qu'elle pourrait être un moyen de contrôler la cinétique d'activation des gènes cibles. Le niveau d'expression de LFY croît au court du développement floral (Blázquez et al., 1997). Il est donc possible que les gènes contenant des sites de liaison de LFY de haute affinité (*i.e.* *AP1*) soient liés très tôt, alors que les gènes contenant des sites de moins bonne affinité soient liés plus tardivement dans le développement quand la concentration nucléaire de LFY est plus

importante. Cette hypothèse pourrait être testée par mutagénèse directe des sites de liaison LFY dans les gènes cibles, en utilisant les méthodes de modification de génome de type CRISPR/Cas. Ce mécanisme, associé aux régulations chromatinienne et à l'expression de co-régulateurs pourrait permettre une régulation fine de la cinétique d'expression des gènes cibles lors du développement floral.

## II. Etude de l'interaction de LFY et WUS au locus *AGAMOUS*

### 1. Introduction

Les avancées sur la caractérisation biochimique de LFY et ses propriétés de liaison à l'ADN génomique ouvrent de nouvelles perspectives de travail pour l'étude de ses interactions avec ses co-régulateurs. Dans cette partie, je me penche sur l'étude de la régulation par LFY et WUS d'AG, un FT nécessaire à la formation des carpelles et des étamines. Ce travail a été mené de manière très exploratoire et n'a pas progressé autant que nous l'aurions souhaité. Je présente ici les premiers résultats.

#### *Induction d'AG par LFY et WUS*

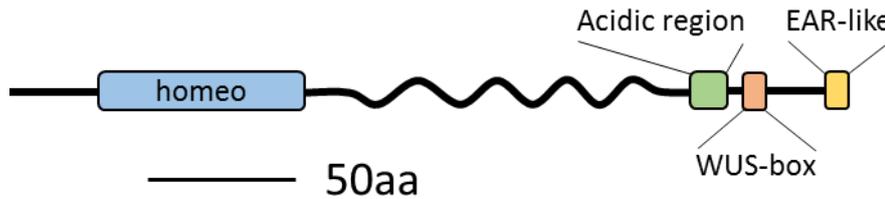
LFY est un régulateur important d'AG : l'expression d'une version activatrice de LFY (*LFY-VP16*) entraîne d'importants défauts dans la morphologie de la fleur, lesquels dépendent largement de l'expression d'AG (Parcy et al., 2002). De plus les mutants *lfy* sur-exprimant *AP1* forment des fleurs qui, dans la plupart des cas, ne portent que des sépales organisés en couronnes consécutives, indiquant que LFY est requis pour l'identité des couronnes internes, et donc pour l'expression d'AG (Liljegren, 1999). LFY lie et régule directement AG *via* deux sites de liaisons dans son deuxième intron (Moyroud et al., 2011; Busch et al., 1999). Le patron d'expression d'AG dans la fleur est en partie contrôlé par les éléments régulateurs présents sur un fragment de ~500 paires de bases (*KB31*) correspondant au site lié par LFY dans ce deuxième intron (Busch et al., 1999). En accord avec ces données, *KB31*

répond aux variations d'expression de *LFY*, cependant lorsque ce fragment est absent, *AG* présente un motif d'expression sauvage dans la fleur mais sa réponse aux variations d'expression de *LFY* est fortement atténuée (Busch et al., 1999). La régulation d'*AG* par *LFY* ne permet donc pas d'expliquer complètement son profil d'expression. *WUS* est un co-régulateur potentiel d'*AG*, la surexpression de *WUS* dans la fleur permettant de réduire en partie le besoin de *LFY* pour l'induction d'*AG* (Lohmann et al., 2001). Des perturbations de l'expression de *WUS* permettent d'influencer l'expression d'*AG*, son expression ectopique entraînant la formation de carpelles et d'étamines surnuméraires (Lohmann et al., 2001). De manière intéressante, *KB31* répond également au niveau d'expression de *WUS* et *WUS* se lie à deux sites de liaison TTAATGG accolés à ceux de *LFY* dans le deuxième intron. Les deux sites de fixation de *WUS* et *LFY* (nommés *AGI* et *AGII* par la suite) sont organisés en tandem et séparés d'une trentaine de paires de bases seulement (Lohmann et al., 2001). L'analyse de l'expression d'un rapporteur contrôlé par un sous fragment de *KB31* contenant les sites *AGI* et *AGII* (*FP50*) montre que *LFY* et *WUS* agissent de manière synergique pour l'expression de gènes rapporteurs dans des systèmes hétérologues, la présence des deux protéines étant requise. Cependant l'étude de la liaison de *LFY* et *WUS* sur le site *AGII* *in vitro* suggère que les deux protéines ne peuvent lier le site simultanément (Lohmann et al., 2001).

Ces données indiquent que *WUS* et *LFY* régulent ensemble l'expression d'*AG* *via* leur liaison à deux sites proches dans le deuxième intron d'*AG*. Si les deux protéines sont requises pour l'activation d'*AG*, elles ne peuvent se lier simultanément sur l'un des sites de fixation. Il n'est cependant pas connu si les deux sites *AGI* et *AGII* sont requis pour l'expression d'*AG* ou si un seul site permet de récapituler son patron d'expression.

#### *WUS, un régulateur insaisissable*

La protéine *WUS* contient plusieurs domaines fonctionnels plus ou moins caractérisés (Figure ChII. 7).



**Figure ChII. 7. Organisation fonctionnelle de la protéine WUS.** La partie N-terminale de la protéine contient le domaine homéobox ou homéodomaine (bleu) permettant la liaison à l'ADN. La partie centrale de la protéine est prédite comme étant désordonnée. La partie C-terminale contient un domaine acide, la WUS-box et le domaine EAR-like. La WUS-box est similaire à un domaine EAR et ces deux domaines pourraient médier l'interaction avec TPL.

Le domaine N-terminal est un domaine homéobox, similaire aux domaines HTH, et permet l'interaction protéine-ADN et la dimérisation de la protéine (van der Graaff et al., 2009; Busch et al., 2010). Le domaine central est prédit comme étant désordonné et régule le mouvement de la protéine dans le SAM (Daum et al., 2014). Le domaine C-terminal contient plusieurs motifs peptidiques conservés : une région acide, la WUS-box et un motif EAR-like. Le motif EAR-like interagit avec le répresseur TPL (Kieffer et al., 2006). En accord avec cela, WUS montre une activité de répresseur transcriptionnel lors d'expériences utilisant des gènes rapporteurs. Cette fonction n'est cependant pas uniquement dépendante du motif EAR-like mais est aussi induite indépendamment par le domaine WUS-box (Ikeda et al., 2009). Il apparaît même que le domaine EAR-like n'est pas absolument nécessaire à la fonction méristématique de WUS, au contraire de la WUS-box, requise pour la répression des ARRs et, étrangement, l'induction de CLV3 (Ikeda et al., 2009). La WUS-box est également requise pour les phénotypes de prolifération de carpelles et étamines obtenus par l'expression ectopique de WUS dans la fleur, et donc pour l'induction d'AG (Ikeda et al., 2009; Lohmann et al., 2001). Cette ambiguïté de l'induction de CLV3 et AG par un domaine répresseur pose la question de savoir si ce domaine a bien un rôle répresseur *in planta*. Lorsqu'elle est exprimée en fusion à d'autres facteurs de transcription, la WUS-box induit bien une activité répressive, similaire au domaine de co-répression SRDX. De plus, sa mutation peut être complétée par la fusion de WUS à SRDX (Ikeda et al., 2009). A l'opposé de ces observations, la mutation de la WUS-box induit un effet dominant-négatif sur l'expression d'AG qui peut être complété par la fusion de WUS au domaine de trans-

activation VP16 (Ikeda et al., 2009). Ces résultats contradictoires indiquent une fonction répressive de WUS dans le cadre de sa fonction méristématique, impliquant la WUS-box, et une fonction activatrice d'AG dans la fleur, impliquant là encore la WUS-box (Ikeda et al., 2009). La WUS-box présentant des similarités de séquences avec les motifs EAR, il n'est pas exclu qu'elle permette le recrutement de corépresseurs de la famille TPL. Cette hypothèse est compatible avec les études d'interactions WUS-TPL (Kieffer et al., 2006). Plus globalement, l'étude de l'expression des cibles directes de WUS dans des lignées permettant l'induction conditionnelle de WUS montre qu'il agit effectivement aussi bien en tant qu'activateur qu'en tant que répresseur transcriptionnel (Lohmann *et al.*, 2001)

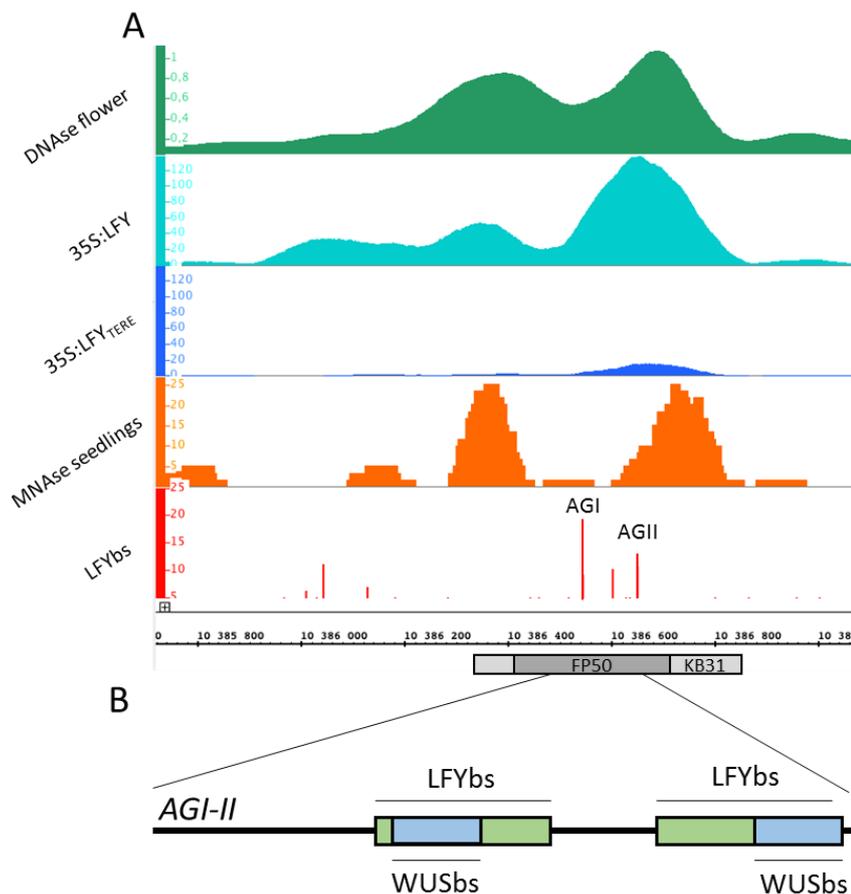
L'identification des sites de liaison de WUS dans de jeunes plantules révèle que WUS lie un motif G-box CACGTG, différent des motifs TTAATGG présents dans l'intron d'AG. On retrouve cependant un enrichissement de ces derniers dans les régions liées par WUS. *In vitro* WUS lie ces deux séquences, avec une légère préférence pour la G-box (Busch et al., 2010)..

#### *Nouvelles données concernant la régulation d'AG par LFY*

L'identification du site de liaison complet de LFY (Moyroud et al., 2011) change considérablement les données sur la fixation de LFY et WUS au locus AG. Au lieu du consensus CCAATG(G/T) utilisé pour les études *in vitro* (Lohmann et al., 2001), LFY lie en fait un motif beaucoup plus large, de 19 paires de bases, englobant les sites de liaison de WUS TTAATGG (Figure ChII. 8B). Cela pourrait expliquer l'impossibilité de liaison simultanée de LFY et WUS sur *AGII* (Lohmann et al., 2001). De plus LFY lie les sites *AGI-II* de manière coopérative *in vitro* grâce à son domaine d'oligomérisation. *In planta*, LFY oligomérisé aussi probablement au niveau du deuxième intron d'AG, la liaison de LFY<sub>TERE</sub> étant considérablement plus faible à ce locus que celle de LFY (Sayou et al., 2016). Il est donc probable que la fixation de LFY au locus AG puisse empêcher la liaison de WUS et *vice versa*.

Le rôle pionnier de LFY pourrait également jouer un rôle important dans la régulation d'AG. Au stade végétatif *FP50*, le fragment d'ADN comportant les sites *AGI-II*, est flanqué par deux nucléosomes dont les positions sont

conservées dans les tissus de plantules (Li et al., 2014) (Figure ChII. 8A). Au stade floral cette région a un signal DNase relativement élevé, indiquant une haute accessibilité de la chromatine (Zhang et al., 2012). Cela indique que le locus *AG* est probablement dans un état chromatinien fermé au stade végétatif et s'ouvre au stade floral, en accord avec les données d'accessibilité de la chromatine au cours du développement floral (Pajoro et al., 2014a). L'ouverture de la chromatine au locus *AG* pourrait être dépendante de l'oligomérisation de LFY, les fleurs des plantes exprimant LFY<sub>TERE</sub> ne formant pas de carpelles ni d'étamines dont la formation est sous contrôle de *AG* (Sayou et al., 2016).



**Figure ChII. 8. Paysage génomique du deuxième intron d'*AG*.** (A) De haut en bas: signal DNase en inflorescence, signal ChIP (anti-LFY) en plantules 35S:LFY de 14 jours, signal ChIP (anti LFY) en plantules 35S:LFY<sub>TERE</sub> de 14 jours, signal MNase en plantules de 14 jours, sites de liaisons prédits de LFY. La numérotation renvoie à la position des nucléotides sur le chromosome 4. Les régions correspondant aux fragments *KB31* et *FP50* sont indiquées en gris clair et foncé respectivement. (B) Représentation schématique de la sonde *AGI-II*. Les sites de liaison de LFY (vert) et WUS (bleu) sont indiqués.

En nous basant sur ces données nous émettons l'hypothèse que l'oligomérisation de LFY au locus d'AG est requise pour sa régulation par LFY et WUS. Nous proposons d'étudier les interactions de LFY et WUS avec *AGI-II* par des approches *in vitro*, en utilisant les protéines complètes. Plusieurs hypothèses permettraient d'expliquer la régulation d'AG par ces deux facteurs : (1) LFY et WUS lient simultanément les deux sites, chaque protéine étant sur un site distinct, (2) les deux protéines lient successivement le locus, cette liaison successive étant nécessaire pour l'activation d'AG, ou (3) les deux protéines interagissent physiquement sur le locus.

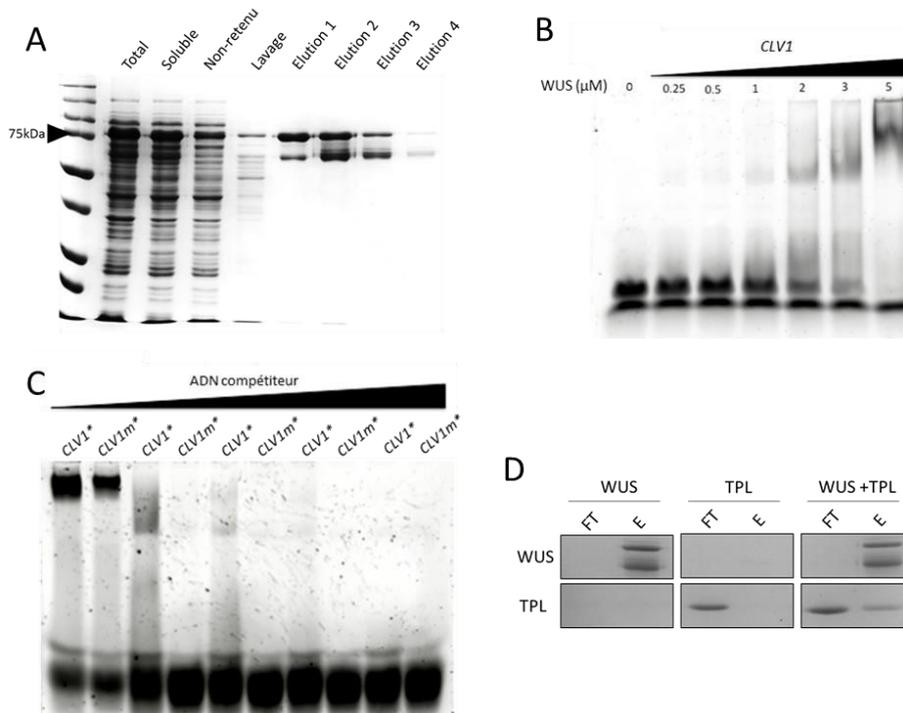
## 2. Résultats

### *Production, purification et activité in vitro de WUS*

Afin d'étudier la liaison de WUS sur *AGI-II* nous avons produit en système hétérologue chez *Escherichia coli* la protéine WUS, fusionnée à la MALTOSE-BINDING PROTEIN (MBP), pour faciliter sa solubilisation, et à une étiquette 6xHis pour simplifier la purification. La protéine est correctement synthétisée par les bactéries et solubilisée dans un tampon optimisé pour la stabiliser. WUS s'adsorbe correctement sur colonne de nickel-sépharose et est éluée par des concentrations croissantes d'imidazole après lavage de la colonne, permettant la purification satisfaisante de la protéine étiquetée sans étape supplémentaire (Figure ChII. 9A). Malgré la présence d'inhibiteurs de protéases et de réducteurs dans le tampon à toutes les étapes de la purification, une bande de poids moléculaire plus faible que la protéine entière est éluée avec la protéine, correspondant probablement à une dégradation partielle de la protéine.

Afin de tester l'activité de liaison à l'ADN de la protéine purifiée nous avons fait interagir WUS avec une sonde marquée constituée du site de liaison de WUS dans le promoteur de *CLV1* (Busch et al., 2010). La séparation des complexes selon leurs poids moléculaires en gel d'acrylamide (gel-retard) montre la formation d'un complexe WUS-sonde (Figure ChII. 9B). De plus, la quantité de sonde liée augmente avec la concentration de WUS. Afin de tester la spécificité de cette liaison nous avons d'une part muté le site de liaison de

WUS sur la sonde utilisée (*CLV1m*) et d'autre part ajouté des concentrations croissantes d'ADN compétiteur aspécifique (Figure ChII. 9C). WUS forme un complexe avec la sonde mutée mais ce complexe WUS-*CLV1m* a été plus sensible à la présence d'ADN compétiteur (*i.e.* est déstabilisé pour de plus faibles concentrations de compétiteur) que le complexe WUS-*CLV1*, indiquant que WUS interagit de manière spécifique avec *CLV1*. Le DBD de la protéine WUS que nous avons purifié est donc fonctionnel. En revanche la concentration relativement haute de protéine requise pour observer la formation d'un complexe protéine-ADN (>1 $\mu$ M) indique la faible affinité de la protéine pour l'ADN cible.



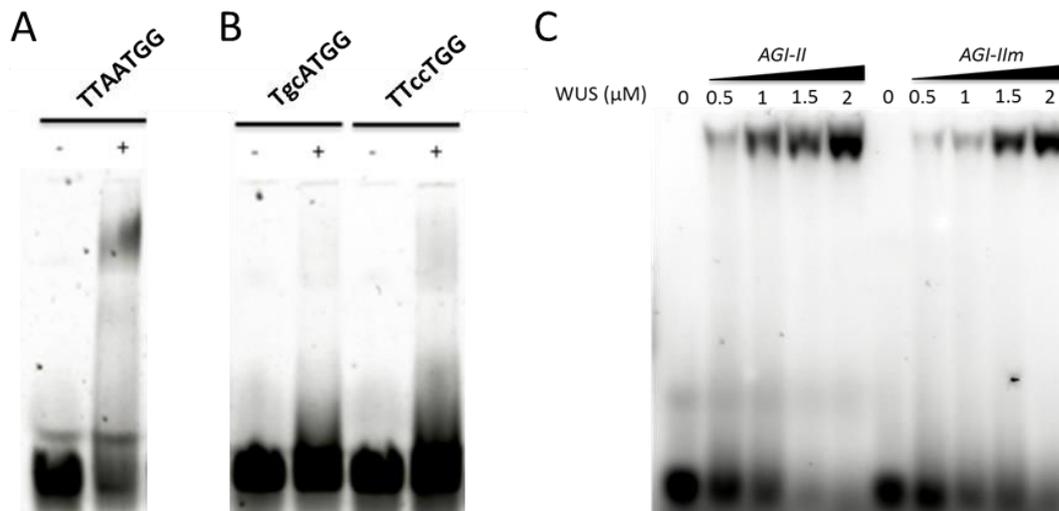
**Figure ChII. 9. Production et activité *in vitro* de WUS.** (A) Production et purification de AtWUS-MBP-6His sur colonne Ni-S. La fraction totale correspond au lysat bactérien, la fraction soluble au surnageant après centrifugation du lysat. LA fraction non retenue correspond aux protéines qui ne s'accrochent pas sur la colonne de Ni-S. Le premier lavage est effectué avec le tampon de solubilisation de la protéine. 20 mM d'imidazole est rajouté au deuxième lavage puis les éluions sont effectuées avec 40, 100, 150 et 300 mM d'imidazole. AtWUS-MBP-6His a un poids moléculaire attendu d'environ 75kDa. (B) Gel retard testant la liaison de WUS sur la sonde *CLV1*. (C) Gel retard testant la spécificité de WUS pour la sonde *CLV1* par rapport à une sonde mutée *CLV1m* avec une concentration croissante d'ADN compétiteur. (D) Co-immunoprécipitations de TPL par WUS-MBP-6His sur colonne d'amylose. FT (*Flow Through*) correspond à la fraction non retenue sur la colonne et E à la fraction éluee.

Afin de vérifier si la faible affinité observée de WUS avec l'ADN est le signe d'un problème de conformation de la protéine, nous avons testé la capacité du domaine C-terminal à interagir avec TPL (Kieffer et al., 2006). L'interaction de la protéine étiquetée WUS-MBP-6His et d'une version tronquée de TPL contenant les 202 premiers acides-aminés (Martin-Arevalillo *et al.*, en préparation) a été testée *in vitro* par co-purification sur colonne d'amylose (Figure ChII. 9D). WUS est adsorbé sur la colonne par sa fusion au domaine MBP. En revanche TPL n'étant pas capable d'interagir avec l'amylose, il n'est pas retenu sur la colonne. Lorsque les deux protéines sont incubées ensemble avant leur passage sur la colonne une fraction de TPL est retenue, indiquant une interaction avec WUS. L'ensemble de ces résultats indique que WUS produit *in vitro* est fonctionnelle. La protéine est capable de lier assez spécifiquement les sites de liaisons du promoteur de *CLV1* et est capable d'interagir avec TPL.

#### *Liaison de WUS sur AGI-II*

Les interactions observées précédemment entre WUS et les motifs TTAATGG de *FP50* reposent sur l'utilisation d'un extrait non-purifié de protéines de levure exprimant WUS (Lohmann et al., 2001). Des cofacteurs également présents dans la levure pourraient être nécessaires à la liaison de WUS sur ces motifs, expliquant que ce motif ne soit que faiblement enrichi dans les SELEX utilisant WUS comme cible (Busch et al., 2010). Pour vérifier la liaison de WUS à ces motifs, nous avons utilisé une sonde contenant un motif de liaison de WUS issu de *FP50* (*AGI-I*). WUS est capable d'interagir avec cette séquence, de plus la mutation du motif perturbe l'interaction, indiquant que WUS est capable de lier le motif TTAATGG sans cofacteur (Figure ChII. 10A-B). Afin de vérifier la liaison de WUS sur les deux sites de liaisons présents dans la séquence de *FP50* nous avons synthétisé une sonde plus longue, contenant les deux sites (*AGI-II*). WUS formant des complexes de haut poids moléculaire sur cette sonde, nous avons séparé les complexes sur gel d'agarose, plus résolutif dans les gammes de poids moléculaires observés. Une seule espèce de très haut poids moléculaire est formée par WUS et *AGI-II*. L'absence de complexes de plus faible poids moléculaire indique que WUS lie la sonde de manière coopérative, similairement à LFY (Figure ChII. 10C). La mutation des motifs

TTAATGG dans cette séquence (*AGI-II*) ne perturbe que faiblement la liaison de WUS, indiquant que d'autres sites de liaison non-prédits pourraient être présent sur la sonde. Cela ajouté à la faible mobilité du complexe formé indique que WUS pourrait oligomériser sur la sonde, ou bien que la liaison observée est relativement peu spécifique d'une séquence particulière.



**Figure ChII. 10. Liaison de WUS aux séquences régulatrices d'AG.** (A) et (B) Gels retards montrant l'interaction de WUS avec une sonde de 20 nucléotides comportant le site TTAATGG (*AGI*) (A) ou des version mutées (B). Puits (-) : sonde seul ; puits (+) : sonde et protéine (+) (C) Gel retard montrant l'interaction de WUS avec des sondes de 85 nucléotides portant deux sites de liaisons (*AGI-II*) ou avec les sites de liaison mutés (*AGI-II*m).

#### *WUS et LFY lient AGI-II de manière compétitive*

Afin de tester la liaison simultanée de LFY et WUS aux séquences régulatrices d'AG, les deux protéines sont incubées à différentes concentrations en présence d'*AGI-II*. Les complexes formés sont séparés sur gel d'agarose (Figure ChII. 11). En présence de la sonde, AtLFY forme un seul complexe, correspondant à la fixation en tétramère, de manière similaire à ce qui a été observé précédemment pour GbLFY (Sayou et al., 2016). AtLFY<sub>TERE</sub> forme lui deux complexes majoritaires, correspondant au dimère et au tétramère, comme observé pour GbLFY<sub>TERE</sub>. A cause du haut bruit de fond et du manque de définition des retards de migrations sur au gel d'agarose, nous ne pouvons nous prononcer sur l'existence de complexes additionnels. Comme précédemment, WUS seul déclenche la formation d'un complexe de très haut

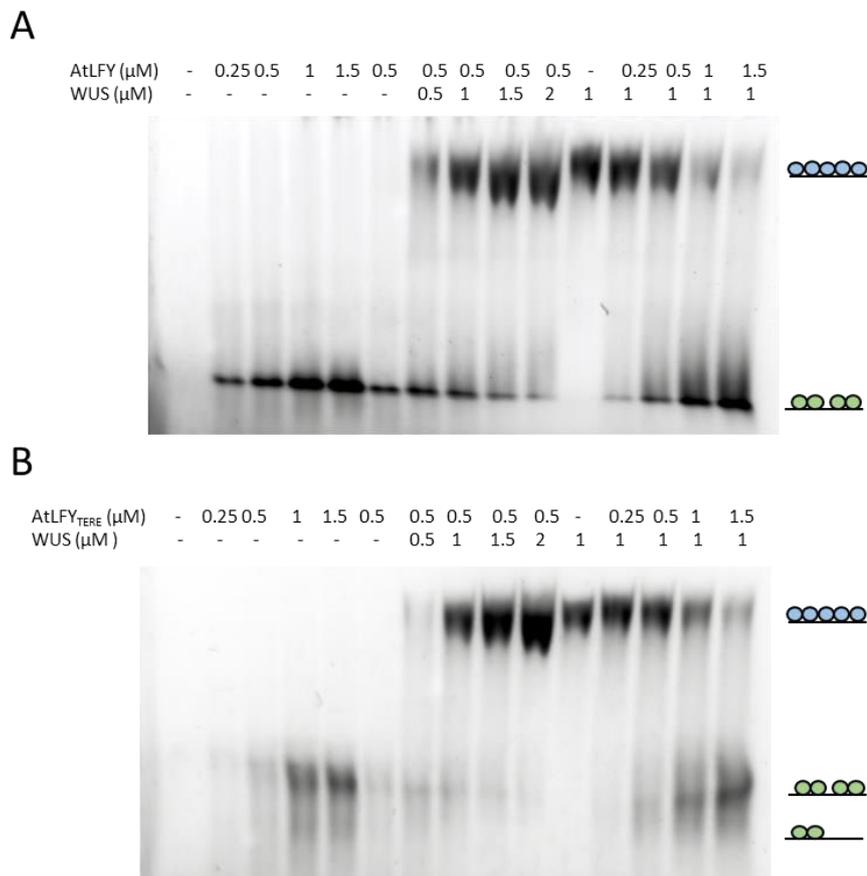
poids moléculaire. En présence de WUS et LFY en même temps on ne détecte pas l'apparition d'une nouvelle espèce : il ne semble pas y avoir de complexe ternaire WUS/LFY/ADN. En revanche une concentration croissante de WUS inhibe proportionnellement la liaison de LFY à la sonde. De façon similaire l'ajout de LFY à des concentrations croissantes inhibe l'interaction entre WUS et la sonde. Ces résultats indiquent que LFY et WUS lient *AGI-II* de manière compétitive et la liaison simultanée des deux protéines sur le même brin d'ADN n'a pas pût être observée. Cela est probablement dû au fait que les deux protéines semblent, indépendamment l'une de l'autre, lier la sonde de manière coopérative, inhibant la liaison de l'autre protéine. De manière surprenante, la même observation est faite pour LFY<sub>TERE</sub> et WUS. Bien que LFY<sub>TERE</sub> ne lie pas la sonde de manière coopérative, aucun nouveau complexe n'a pu être observé lorsque LFY<sub>TERE</sub> et WUS sont incubées simultanément avec la sonde. En revanche LFY<sub>TERE</sub> est plus sensible que LFY à la compétition avec WUS et moins efficace que LFY à déplacer WUS. Cela indique que LFY<sub>TERE</sub> est moins efficace que LFY pour la liaison à *AGI-II*, ce qui n'avait pas été observé pour GbLFY<sub>TERE</sub> (Sayou et al., 2016). Dans l'ensemble ces résultats indiquent que LFY et WUS lient l'ADN de manière compétitive et il est improbable qu'ils soient liés simultanément à des sites de liaison proches.

### 3. Discussion

#### *Liaison de LFY et WUS au locus AG*

Il avait été montré précédemment que LFY et WUS ne peuvent lier simultanément le site *AGII*, probablement dû au fait que les sites de liaison de LFY et WUS se chevauchent (Lohmann et al., 2001; Moyroud et al., 2011). Nous avons montré ici que, comme LFY, WUS oligomériser très probablement sur le site *AGI-II* de manière coopérative. L'oligomérisation de LFY et WUS sur cette séquence empêcherait leur liaison simultanée *in vitro*. Il apparaît donc que si l'activité des deux facteurs de transcriptions est requise pour l'induction d'*AG*, leur liaison à l'ADN doit être transitoire. Cependant, s'il est connu que WUS dimérise *via* le domaine homéobox (Lohmann et al., 2001), il n'a jamais été montré qu'il puisse oligomériser. WUS est encore peu

caractérisé d'un point de vue biochimique. Cependant il est aussi possible que l'oligomérisation observée sur *AGI-II* soit un artefact dû aux hautes concentrations de protéines utilisées ici. Le très haut poids moléculaire apparent des complexes formés malgré l'absence de sites supplémentaires prédits tend vers cette hypothèse. Il avait de plus été montré que la mutation des deux sites de liaison de WUS suffisait à empêcher sa liaison en levure (Lohmann et al., 2001), ce que nous n'avons pas retrouvé *in vitro* et qui renforce l'hypothèse d'une liaison en partie artificielle. Dans l'ensemble, les difficultés présentées par l'étude de liaison à l'ADN de WUS *in vitro* suggèrent que ce système n'est pas le plus approprié pour ce travail.



**Figure ChII. 11. WUS et LFY lient *AGI-II* de manière compétitive.** Expérience de compétition sur gel retard de AtLFY (A) ou AtLFY<sub>TERE</sub> (B) avec WUS sur la sonde *AGI-II*. Les complexes protéines-ADN susceptibles d'être formés sont schématisés à droite avec LFY et LFY<sub>TERE</sub> forment des dimères ou tétramères (vert) et WUS couvre probablement la sonde (bleu). Le nombre de molécules représentées pour WUS est arbitraire. La sonde libre n'est pas visible sur ces gels.

### *Etat chromatinien au locus AG et futures directions*

Le fragment *KB31* utilisé précédemment pour l'étude de la régulation d'AG et dont une sous partie a été utilisée ici est flanqué par deux nucléosomes au stade végétatif (Li et al., 2014). Le fort signal DNase au stade floral, associé à la réduction de liaison de LFY<sub>TERE</sub> par rapport à LFY (Zhang et al., 2012) sont autant d'indices indiquant que ces nucléosomes sont probablement déplacés lors de l'induction d'AG. Cette hypothèse étant cohérente avec les propriétés de LFY<sub>TERE</sub>, incapable d'oligomériser et de lier la chromatine fermée, LFY régule probablement l'ouverture de la chromatine au locus d'AG, en partie *via* l'oligomérisation dans le second intron. Un tel rôle n'a jamais été montré pour WUS et il est possible que l'ouverture de la chromatine par LFY soit un prérequis pour la liaison de WUS. Un mécanisme similaire ("*assisted loading*") a été décrit pour la liaison du récepteur des glucocorticoïdes (GR) et un récepteur de l'estradiol modifié pour reconnaître le même site de liaison (ERpbox) (Voss et al., 2011). Dans ce cas, bien que les deux protéines reconnaissent et lient le même site, la liaison semble facilitée par la présence des deux facteurs simultanément. De plus les deux protéines régulent de manière coopérative l'ouverture de la chromatine aux loci cibles. Le modèle présenté dans cette étude décrit la liaison d'un des facteurs, GR, à la chromatine fermée. Cette liaison permet l'ouverture transitoire de la chromatine, permettant à ERpbox de remplacer GR, de manière compétitive, aux sites de liaison. Le résultat serait une occupation accrue du locus, permettant de maintenir l'ouverture de la chromatine et la transcription des gènes cibles. Il serait intéressant d'étudier l'accessibilité et l'occupation de *KB31* par LFY et WUS seul ou présence de l'autre partenaire en utilisant par exemple des lignées inductibles pour les deux facteurs. Dans ce contexte, la fonction activatrice de WUS pourrait s'expliquer par un rôle dans le maintien de l'accessibilité de la chromatine plutôt que dans la régulation transcriptionnelle, rôle pour lequel WUS est plutôt connu pour agir comme un répresseur (Ikeda et al., 2009).



# MATERIEL & METHODES

## I. Matériel végétal et méthodes de travail sur les plantes

### 1. Matériel végétal

Les souches mutantes et sauvages utilisées dans ce travail sont toutes dans *Arabidopsis thaliana* écotype *Columbia* (Col) sauf si précisé autrement. Les allèles mutants utilisés publiés précédemment sont les suivants : *rax1-3* (Müller et al., 2006; Keller et al., 2006) (ATG23000 ; SALK\_071748) ; *rev-6/vam-1* (Pogany et al., 1998; Otsuga et al., 2001) (AT5G60690) a été d'abord isolé dans l'écotype *Landsberg erecta* (Ler) et a été rétrocroisé au moins trois fois dans Col au cours du présent travail; *lfy-1/lfy-12* (Schultz and Haughn, 1991; Huala and Sussex, 1992) et *lfy-2* (Weigel et al., 1992) (AT5G61850) sont des mutants forts et faibles respectivement.

La lignée rapportrice pWUS:2xVENUS (CW039) a été fournie gracieusement par C. Wenzl (*non publié*) et contient une répétition de deux protéines fluorescentes VENUS sous contrôle du promoteur de *WUS* (~4,4kb) et de son terminateur (Lampropoulos et al., 2013).

## 2. Conditions de cultures

### *Conditions de culture*

Les graines sont stérilisées en phase gazeuse en les plaçant dans un tube eppendorf ouvert dans une boîte hermétique contenant un bécher dans lequel ont été mélangés 20 ml de Javel (NaClO+NaCl, 4,3%) avec 3 ml d'acide chlorhydrique (HCl). Les graines sont incubées en présence du dichlore (Cl<sub>2</sub>) ainsi produit pendant 2 h à 4 h puis semées sur milieu Murashige & Skoog (MS complet 0,5X (Sigma-Aldrich); 1% sucrose ; 0,05% acide 2-(N-morpholino)ethane sulfonique (MES) ; pH 5,7 ajusté au KOH ; 0,8% Agar). Les plantes sont stratifiées à 4°C pendant au moins 48 h puis transférées en chambre de culture pour germer, typiquement à 20°C et 65% d'humidité sous lumière blanche. Après 7 jours les graines sont transférées sur substrat de croissance (terreau de multiplication : vermiculite, 5:1) et cultivées dans les mêmes conditions.

Les conditions de culture jours longs (JL) consistent en cycles de 16 h de lumière et 8 h d'obscurité, les conditions de jours courts (JC) consistent en cycles de 8 h de lumière et 16 h d'obscurité. Lorsqu'une synchronisation de la floraison et/ou un large SAM sont requis, les plantes sont cultivées en JC pendant 6 semaines puis transférées en JL pour induire la floraison. Les SAM sont alors observés ou récoltés 14 jours après transfert.

### *Sélection de transformants sur Hygromycine B*

Les plantes transformées avec un plasmide dérivé du pCAMBIA1300 sont sélectionnées sur Hygromycine B (Sigma-Aldrich). Les graines stérilisées sont semées sur un milieu MS minimal, sans sucre ni vitamines (0,5X MS basal salt (Sigma-Aldrich); pH 5,7 au KOH ; 0,8% Agar) et stratifiées pendant 48 h. Les boîtes sont alors exposées à la lumière pendant environ une heure et transférées dans une chambre de culture à l'obscurité. Après 5 jours les plantules sont exposées à une lumière atténuée pendant une journée puis transférées en chambre de culture en conditions de JL. Les plantules résistantes ont un phénotype d'étiollement avec un hypocotyle très allongé alors que les plantules sensibles ont un hypocotyle court et sont chétives. Les

plantules résistantes sont délicatement transférées en terre et cultivées normalement.

### 3. Extraction d'ADN génomique et génotypage

Un morceau de jeune feuille est récupéré dans un tube eppendorf de 1.5 ml et broyé à l'aide d'un pilon. Alternativement les échantillons sont récupérés en plaques de 96 puits, 2 billes de verres (diamètre 4 mm) sont ajoutées dans chaque tube et les échantillons sont broyés à l'aide du TissueLyser (2 fois 30 s à 30 Hz). Les tissus sont lysés en ajoutant 500 µL de tampon d'Edwards (200 mM Tris-HCl pH 7,5 ; 250 mM NaCl ; 25 mM éthylène diamine tétra-acétique (EDTA) ; 0,5% dodécylsulfate de sodium (SDS) 20%) et vortexés, puis centrifugés 1 min à 16000 g, ou 30 min à 12500 g pour les plaques 96 puits.

400 µl de surnageant sont récupérés dans un nouveau tube et l'ADN génomique est précipité à l'isopropanol (ajouté volume à volume puis centrifugé 5 min à 16000 g (ou 20 min à 12500 g pour les plaques) après 2 min d'incubation à température ambiante). Le culot est nettoyé avec 500 µl d'éthanol 70% puis centrifugé comme précédemment. Le culot est ensuite séché à l'air libre puis suspendu dans 50 µL de Tris-EDTA (TE ; 10 mM Tris-HCl pH 8 ; 1 mM EDTA). Typiquement 2 µL d'ADNg sont utilisés par PCR.

#### *Génotypage par PCR de rax1-3, rev-6 et lfy-2*

Le génotypage de *rax1-3* a été effectué avec les amorces oHC150 et oHC152/SALK Lb1.3 pour l'allèle mutant et oHC150 et oHC151 pour l'allèle sauvage. Le génotypage de *rev-6* a été effectué avec les amorces oGV108 et oGT1025 suivit par une digestion par l'enzyme *TaqI* (New England Biolabs). Le génotypage de *lfy-2* a été effectué avec les amorces oGD136 et oGD137 pour le sauvage et oGD138 et oGD139 pour le mutant. La liste des amorces utilisées pour le génotypage est indiquée dans le Tableau MM 1.

**Tableau MM 1. Amorces de géotypage.**

Nom	Cible	Séquence 5' -> 3'
oHC152 /SALK Lb1.3	T-DNA (Left border)	ATTTTGCCGATTTTCGGAAC
oHC150	RAX1	TGTGAAAAGACCAACCTCACC
oHC151	RAX1	TCGGACATTTTCAGTTTGAAG
oGV108	REV	TAAGGCCTTGTGATGGTGGTGGTTCTATTATTCACATTGTTGATCACCTT
oGT1025	REV	ACAATGCATCGTAGACCACCC
oGD142	FIL	GACAATGAAAGTAGTTTTAATGAG
oGD143	FIL	CAGCTTTGATACGTTGGATCTCCTCC
oGD144	FIL	ACCGTTACGACCGTTTTTCATCC
oGD136	LFY	GGGACAGAGAGACAGAGGGAGCAGCC
oGD137	LFY	GCTGTCCATGATTATAGCTATAGTTCC
oGD138	LFY	ACTTCCCCAGGCTCCGTTACGATAACCA
oGD139	LFY	GGAAGTGGTTACTGGGACGCAGGTCAAG
oGD124	RAX1guide1	GGTGCCTGAAGAAGACTCTAA
oGD126	RAX1guide1	TCACCATGCTTTATGTTTGGTC
oGD125	RAX1guide2	TAATCATTTGCTTCAAATCCCC
oGD134	RAX1guide2	CACACCAGTTTTTCTTCAGACG
oGD127	REVguide1	ACTTCAGGTTAACGATCCAAGC
oGD129	REVguide1	TATAACCTTCATCCCAGGCATC
oGD128	REVguide2	GCTCAAGATCGAAGTACCTGCT
oGD135	REVguide2	AAAAATGACCATTTCCGTGAGT

### Géotypage TaqMan de *lfy-12*

Le géotypage de l'allèle *lfy-12* s'est effectué par chimie TaqMan dans un thermocycleur CFX96™ (Bio-Rad). 1 µL d'ADNg a été mélangé à 10 µL de iTaq™ Universal Probes Supermix (Bio-Rad), 1 µM de chaque amorce de géotypage et 0,2 µM de chaque sonde TaqMan. La sonde s'hybridant sur l'allèle sauvage est : 5' (FAM)-CGCTGCAGCAACAGC-(BHQ) 3' ; la sonde s'hybridant sur l'allèle mutante est : 5' (HEX)-CGCTGCAGTAACAGC-(BHQ) 3'.

L'amplification est effectuée en deux étapes. Après dénaturation de l'ADN à 95°C pendant 5 min, 45 cycles de dénaturation (95°C, 10 s) et élongation (60°C, 30 s) sont effectués. La lecture de fluorescence du FAM et de l'HEX est effectuée pendant l'élongation. La fluorescence finale de l'HEX par rapport au FAM est mesurée pour chaque échantillon et le rapport HEX/FAM permet de déterminer du géotype de chaque individu.

### *Génotypage des lignées CRISPR/Cas9*

Les plantes mutagénisées par l'action de la Cas9 sont identifiées dans un 1<sup>er</sup> temps par PCR, en utilisant des amorces s'hybridant de part et d'autre de la zone à couper (oGD124 et oGD125, et oGD127 et oGD128 pour *RAX1* et *REV* respectivement ; Tableau MM 1). Lorsque cette approche ne permet pas d'identifier des plantes mutées, l'approche élaborée par Zhu *et al.* est utilisée (Zhu et al., 2014). La zone entourant chaque cible de guide est amplifiée par PCR. Les amorces oGD124 et oGD126 ont servi à amplifier la zone autour du guide *RAX1guide1*, les amorces oGD125 et oGD137 la zone autour de *RAX1guide2*, les amorces oGD127 et oGD129 la zone autour de *REVguide1* et les amorces oGD135 et oGD128 la zone autour de *REVguide2*. Le produit de PCR est soumis à une étape de dénaturation de 5 min à 98°C pour permettre la formation d'hétéroduplex d'ADN. Le produit obtenu est ensuite séparé sur gel d'acrylamide non dénaturant à 12% dans du TBE 1X pendant 2 h à 90 V. Les hétéroduplex sont visualisés sur le gel comme des retards de migration par rapport à l'homoduplex d'ADN. En T2 et dans les générations suivantes, où des plantes mutantes homozygotes sont attendues, le produit de PCR est mélangé avant l'étape de fusion à un produit de PCR obtenu par amplification sur l'ADN d'une plante sauvage, dans un ratio équimolaire, afin de permettre la formation d'hétéroduplex détectables sur gel.

Pour les mutants ainsi identifiés, la séquence d'ADN présente entre les séquences cibles des deux guides est amplifiée par PCR et clonée dans pCR®-Blunt (ThermoFisher) avant d'être vérifiée par séquençage (Eurofins Genomics).

#### 4. Génération de lignées transgéniques stables d'*Arabidopsis thaliana*

Tous les clonages sont réalisés en suivant les procédures standard. Les amorces utilisées sont indiquées dans le Tableau MM 2. Tous les clones ont été validés par séquençage (Eurofins Genomics) avant leur utilisation.

### *Clonage GreenGate*

Les plasmides contenant des versions inductibles de RAX1 sont construits en utilisant la méthode de clonage GreenGate (Lampropoulos et al., 2013). La séquence codante pour le domaine de liaison aux glucocorticoïdes du récepteur du rat (GR) a été amplifiée à partir d'ADNg de plantes exprimant un transgène *35S :AP1-GR* (Yu et al., 2004). La séquence codante pour RAX1 a été amplifiée à partir d'ADNc et sous-clonée dans le vecteur d'entrée Gateway® pDONR™221 (ThermoFisher) (pHC03). Une mutation silencieuse a été introduite dans la séquence de *RAX1* avec les amorces oGD122 et oGD123 de façon à éliminer le site de restriction *Bsal* (clone pGD41). L'ADNC a ensuite été réamplifié avec les amorces oGD115 et oGD116 pour le clonage dans les vecteurs d'entrée GreenGate. La cassette de sélection At2S3 :GFP a été amplifiée à partir du plasmide pAlligator1 (Bensmihen et al., 2004) avec les amorces oGD120 et oGD121.

La séquence codant pour RAX1 a été clonée dans pGGC000 pour produire pGD41. Pour le clonage de GR en milieu de construction, la séquence codante a été amplifiée avec les amorces oGD109 et oGD110, clonée dans pGGC000, produisant pGD38. Pour le clonage de GR en C-terminal, un *linker* a été amplifié avec les amorces oGD118 et oGD119 à partir de pGGD001 et GR a été amplifié avec les amorces oGD110 et oGD111. Le linker et GR ont été mélangés lors de la libation, produisant pGD39. La cassette At2S3:GFP a été clonée dans pGGF000 pour produire pGD43. pGD46 et pGD48 ont été produits par assemblage des plasmides pGGZ001, pGGA006, pGGE009 et pGD43 avec pGGB001, pGD41 et pGD39 ou avec pGGB001, pGD38 et pGGD002, respectivement.

**Tableau MM 2. Amorces de clonage.**

Nom	Cible	Sens	Site de clonage	Séquence
oGD01	RAX1	sens	<i>NcoI</i>	AACCATGGGAAGAGCTCCG
oGD02	RAX1	anti-sens	<i>NotI</i>	TAGCGGCCGCTCAGGAGTAGAAATAGGGCAAGC
oGD03	RAX1	anti-sens	Mutagenèse <i>NcoI</i>	TTAGATAAGGAGAAGCCATCGCTGATGATGATGAAT CAG
oGD04	RAX1	sens	Mutagenèse <i>NcoI</i>	CTGATTCATCATCATCAGCGATGGCTTCTCCTTATC TAA
oGD05	RAX1	anti-sens	<i>NotI</i>	TAGCGGCCGCTCAGTTTAGATAAGGAGAAGCCATC GC
oGD109	GR	anti-sens	Overhang "D"	AACAGGTCTCACTGATTTTTGATGAAACAGAAGCTT TTTG
oGD110	GR	sens	Overhang "C"	AACAGGTCTCAGGCTCAGAAGCTCGAAAAACAAAG AAAAA
oGD111	GR	anti-sens	Overhang "E"	AACAGGTCTCAGCAGCTATTTTTGATGAAACAGAAG CTTTTG
oGD115	RAX1	anti-sens	Overhang "D"	AACAGGTCTCACTGAGGAGTAGAAATAGGGCAAGC A
oGD116	RAX1	sens	Overhang "C"	AACAGGTCTCAGGCTCAATGGGAAGAGCTCCGTGT T
oGD118	C-linker	sens	Overhang "D"	AACAGGTCTCATCAGCCGCTAGTGCGATCG
oGD119	C-linker	anti-sens	Overhang "C"	AACAGGTCTCAAGCCCGATCCAGAGCCGCTG
oGD120	AT2S3:GF P	sens	Overhang "F"	AACAGGTCTCAATACGAAACCAAATTAACATAGGGT TTTTA
oGD121	AT2S3:GF P	anti-sens	Overhang "G"	AACAGGTCTCAACTATGCAGATCGTTCAAACATTTG
oGD122	RAX1	sens	Mutagenèse <i>BsaI</i>	GGCTAAACTATTTGAGGCCAAACATAAAGCATGG
oGD123	RAX1	anti-sens	Mutagenèse <i>BsaI</i>	CCATGCTTTATGTTTGCCCTCAAATAGTTTAGCC
RAX1-sg1-F	RAX1guide e1	sens	<i>BbsI</i>	GATTGTATGAACCTACCGGCTTTG
RAX1-sg1-R	RAX1guide e1	anti-sens	<i>BbsI</i>	AAACCAAAGCCGGTAAGTTCATAC
RAX1-sg2-F	RAX1guide e2	sens	<i>BbsI</i>	GATTGAGGACAAAAGGTACAATG
RAX1-sg2-R	RAX1guide e2	anti-sens	<i>BbsI</i>	AAACCATTGTACCCTTTTGTCTC
REV-sg1-F	REVguide 1	sens	<i>BbsI</i>	GATTGCATCTCTAAGCGAATGCTG
REV-sg1-R	REVguide 1	anti-sens	<i>BbsI</i>	AAACCAGCATTGCTTAGAGATGC
REV-sg2-F	REVguide 2	sens	<i>BbsI</i>	GATTGCTTCATGTTTTAGGCGTTG
REV-sg2-R	REVguide 2	anti-sens	<i>BbsI</i>	AAACCAACGCCTAAAACATGAAGC

### *Clonage des constructions CRISPR/Cas9*

Les plasmides utilisés pour le clonage des constructions CRISPR/Cas9 sont un cadeau généreux de W. Yan (Yan et al., 2016). Les guides ARN avec la meilleure spécificité sont choisis à l'aide du logiciel CHOPCHOP (<https://chopchop.rc.fas.harvard.edu/>). Des amorces complémentaires constituées la séquence du guide additionnée d'une extrémité dépassante permettant la libation dans les sites *BbsI* (voir Tableau MM 2) sont mélangées à raison de 10 µM de chaque amorce dans une solution de Tris-HCl pH 7,5 10 µM/ NaCl 150 µM. Le mélange est chauffé à 95°C pendant 3 min et refroidi

lentement à température ambiante afin de permettre l'hybridation des amorces complémentaires.

Les amorces « double brin » sont ensuite liguées dans le site *BbsI* de pBSK-AtU6-26 :guide (Feng et al., 2013), modifié pour permettre la ligation séquentielle de plusieurs guides (Yan et al., 2016). Les guides supplémentaires sont ajoutés successivement par transfert des guides d'un plasmide à l'autre : un guide est isolé d'un plasmide par digestion avec *KpnI* et *XbaI* et transféré un deuxième plasmide contenant un autre guide, ouvert par digestion avec *KpnI* et *SpeI*. Cette ligation élimine les sites *SpeI* et *XbaI*, permettant un autre cycle de transfert. Le plasmide contenant les guides voulus est ensuite digéré par *KpnI* et *SbfI* pour transfert des guides entre les sites *KpnI* et *EcoRI* de pCAMBIA1300 avec la cassette UBQ10:CoCas9:tUBQ10 du pBSK:CoCas9 digéré par *SbfI* et *EcoRI*. Le pBSK:CoCas9 contient une version aux codons optimisés pour l'expression chez *Arabidopsis* de la Cas9 humaine sous contrôle du promoteur *UBQ10* (Yan et al., 2016).

Les plasmides finaux contenant deux guides contre *RAX1*, deux guides contre *REV* ou les quatre guides sont nommés pGD56, pGD57 et pGD58 respectivement.

Le guide ciblant le site de liaison LFY dans le promoteur de *RAX1* a été cloné comme décrit ci-dessus (pGD63).

#### *Transformation stable d'Arabidopsis thaliana*

Les transformations stables sont réalisées en suivant la méthode décrite par Logemann *et al.* (Logemann et al., 2006). Les souches compétentes d'*Agrobacterium tumefaciens* GV3101:pmp90:pSOUP ou C58C1:pmp90 pour les transformations des plasmides dérivés du pGREEN (GreenGate) ou du pCAMBIA respectivement sont transformées par électroporation (2000V) avec ~100 ng de plasmide et cultivées sur milieu LB solide contenant les antibiotiques appropriés pendant 24 à 48 h. Les bactéries sont ensuite ré-étalées uniformément sur milieu YEB solide (Extrait de levure 1%, extrait de bœuf 6%, peptone 6%, sucrose 6%, MgCl<sub>2</sub> 1%, Agar 3g/l) et cultivées en un tapis dense pendant trois jours. Les bactéries sont ensuite resuspendues dans

50 ml de milieu LB liquide puis mélangées au dernier moment à 200 ml d'une solution contenant 5% de sucrose et 0,03% de Silwet L-77®. Les plantes en fleurs sont alors été trempées pendant une dizaine de secondes dans ce mélange et placées sous serre, à l'obscurité, pendant environ 16 h avant d'être cultivées normalement. Les graines sont ensuite récoltées et sélectionnées avec la méthode appropriée.

## 5. Phénotypage

Les plantes à phénotyper sont cultivées soit en JC, soit en JL. La date de floraison est définie par l'élongation de la tige primaire (>0,5 cm). Au moment de la floraison, les feuilles de rosettes sont dénombrées. Le comptage des feuilles, tiges caulinaires et des tiges de rosettes est effectué trois semaines après floraison. Pour le phénotypage floral, les fleurs matures sont observées à la loupe binoculaire tous les deux jours pendant trois semaines après la date de floraison. Pour le comptage des méristèmes axillaires, chaque aisselle de feuille de rosette est observée deux semaines après floraison en détachant la feuille de la plante afin de vérifier la présence ou l'absence de méristème.

Pour les traitements au naptalame (NPA), les plantes sont traitées soit au NPA (10 µM dans 1/10 000 diméthylsulfoxyde (DMSO)) soit au DMSO (1/10 000) (= « *mock* ») par spray tous les deux jours à partir de 14 jours après germination. Le nombre d'organes initiés est déterminé trois semaines après floraison.

Pour les traitements au dexaméthasone (DEX), les plantes ont été traitées soit au DEX (10 µM dans 1/1 000 éthanol) soit à l'éthanol (1/1 000) (= « *mock* ») par spray tous les deux jours à partir de 14 jours.

## 6. Analyse des quantités de protéines par Western Blot

Les tissus sont broyés à l'aide du TissueLyser (2 min à 30 Hz). Les extraits sont ensuite été solubilisés dans 8 µl/mg de masse fraîche de tampon de charge (50 mM Tris-HCl pH 6,8 ; 1% SDS ; 10% Glycerol ; 0,4 % dithiothreitol (DTT)). Après 5 min à 95°C, 10 µl (équivalent à 1,25 mg de matière fraîche) d'extrait est séparé par électrophorèse en gel de polyacrylamide (SDS-PAGE) dans du tampon de Laemmli 1X. Les protéines sont transférées sur membrane

de poly-fluorure de vinylidène (PVDF) dans un tampon Laemmli 0,5X (sans SDS) et éthanol 20% pendant 2 heures à 150 V. La membrane est ensuite incubée en solution de saturation (tampon Tris salin (TBS) 1X ; Tween-20 0,1% ; lait 5%) pendant au moins 1 h. La membrane est ensuite incubée avec l'anticorps primaire (anti-RFP : 6G6 au 1/1000<sup>ème</sup> (ChromoTek)) dans du TBS-T (TBS 1X ; Tween-20 0,1%) pendant 1 h à température ambiante ou la nuit à 4°C. Après 4 lavages de 10 min dans du TBS-T la membrane est incubée avec l'anticorps secondaire dilué au 1/5000<sup>ème</sup> dans du TBS-T pendant 1 h à température ambiante ou bien la nuit à 4°C. Après 4 lavages de 10 min au TBS-T la révélation a été effectuée avec le kit Clarity™ Western ECL Blotting Substrate (BioRad) et imagée par luminescence sur l'imageur ChemiDoc™ (BioRad).

## 7. Analyse de niveau d'expression de gènes par qRT-PCR

### *Extraction des ARNs totaux*

Environ six inflorescences sont disséquées pour ne laisser que des fleurs fermées et récoltées dans des tubes eppendorf de 2 ml contenant 2 billes de verre (diamètre 4 mm) et directement plongées dans l'azote liquide. Après broyage au Tissue Lyser (2X 1 min à 30 Hz), les ARNs totaux sont extraits à l'aide du kit RNeasy Plant Mini Kit™ (Qiagen) en suivant les instructions du fabricant.

L'ADN restant est ensuite dégradé en utilisant la TURBO DNA-free™ DNase en suivant les instructions du fabricant. Les ARNs totaux ainsi obtenus sont quantifiés et leur qualité est contrôlée au NanoDrop.

**Tableau MM 3. Amorces de qRT-PCR.** Les amorces sont choisies pour amplifier un fragment de cDNA contenant un site d'épissage (intron-spanning).

Locus	Nom du gène	Amorce sens 5' -> 3'	Amorce anti-sens 5' -> 3'	Taille d'amplification	Intron spanning
AT2G28390	SAND-domain protein	AACTCTATGCAGCATTTGAT CCACT	TGATTGCATATCTTTATCGC CATC	61	Oui
AT4G34270	TIP41-like family protein	GTGAAAACGTGGAGAGA AGCAA	TCAACTGGATACCCTTTTCG CA	61	Oui
AT4G26410	Uncharacterized protein	GAGCTGAAGTGGCTTCCAT GAC	GGTCCGACATACCCATGAT CC	81	Oui
AT5G43810	AGO10	TGCACTCGGTCGGTCTCTA T	TCGAAATGCTGCAAGATGA G	60	Oui
ATG60690	REV	CGAGCTTGTTTATATGCAGA CG	ATCTCAGGGTCCAGAAATC G	71	Oui
AT2G34710	PHB	TTGGTTTCAGAACCGCAGA	CTGTTTGAAGACGAGCAGC TT	62	Oui
AT1G30490	PHV	TTGGTTCCAGAATCGCAGA	CACTGTCTGAAGACGAGCT GA	64	Oui
AT1G52150	CNA	CCGTCAACATACTCCAATC C	GTCACCACCGATTACAGC	63	Oui
AT1G75820	CLV1	GGATACATCGCCCCAGAGT	TCCAAATTCACCAACAGGTT T	120	Oui
AT5G53950	CUC2	AACTTCCCAGGAGAGCTAA GATGG	TACTTCCGGTCACGGAGGC TAAAG	67	Oui

### *Transcription inverse*

2 µg d'ARNs totaux sont mélangés à 0,5 µg d'amorce oligo(dT) dans un volume total de 10 µl puis chauffés à 70°C pendant 5 min et immédiatement refroidis sur glace pour éviter la formation de structures secondaires. La transcription inverse est ensuite effectuée avec la reverse transcriptase du virus de la leucémie murine Moloney (M-MLV RTase ; Promega) en suivant les instructions du fabricant, à 42°C pendant 60 min. Les ADNc sont dilués au 1/6<sup>ème</sup> immédiatement après réaction.

### *qRT-PCR*

La qRT-PCR est effectuée avec 2 µL d'ADNc, 5 µl SSoAdvance™ Universal SYBR® Green Supermix (Bio-Rad) et 1,2 µM d'amorce sens et anti-sens. La réaction est effectuée dans un thermocycleur CFX96™ (Bio-Rad). Après une dénaturation initiale de 3 min à 98°C, 40 cycles de dénaturation (98°C, 10 s) et d'élongation (60°C, 15 s) sont effectués avec lecture de la fluorescence à chaque cycle pendant l'élongation. Afin de s'assurer de la spécificité de l'amplification, une courbe de fusion est établie après l'amplification en

augmentant graduellement la température des échantillons de 65°C à 95°C (0,5°C/0,05 s ; lecture de fluorescence à chaque palier). Les amorces utilisées pour la qRT-PCR sont indiquées dans le Tableau MM 3.

#### *Détermination de l'expression relative*

Le niveau d'expression de chaque gène a été déterminé par la méthode décrite par Michael W. Pfaffl (Pfaffl, 2001). L'efficacité  $E$  de chaque couple d'amorces est déterminée par la formule :

$$E = 10^{-\frac{1}{\text{pente}}}$$

où *pente* représente la constante liant le  $Ct$  au logarithme népérien de la quantité d'ADNc, déterminé par une dilution géométrique d'ADNc. La quantité relative d'ADNc pour chaque gène est calculée par rapport à la condition contrôle par la formule :

$$RQ = E^{\Delta Ct(\text{contrôle-traitement})} .$$

L'expression relative de chaque gène peut ensuite être calculée comme le ratio de la quantité relative du gène cible par rapport à quantité relative d'un gène de référence :

$$\text{ratio} = \frac{RQ_{\text{cible}}}{RQ_{\text{ref}}}$$

Les gènes de référence pouvant être soumis à des variations d'expression dans certaines conditions, l'expression relative est calculée par rapport à la moyenne géométrique de trois gènes de référence [13] :

$$\text{ratio} = \frac{RQ_{\text{cible}}}{\sqrt[3]{RQ_{\text{ref1}} \cdot RQ_{\text{ref2}} \cdot RQ_{\text{ref3}}}}$$

Chaque réplicat biologique est analysé indépendamment et la moyenne et l'écart-type du ratio d'expression de chaque gène est calculé à travers les réplicats biologiques.

Les gènes de référence utilisés ont été choisis comme étant parmi les gènes les plus stables chez *Arabidopsis thaliana* (Czechowski et al., 2005), et sont AT2G28390, AT4G34270 et AT4G26410.

## 8. Microscopie confocale

### *Préparation et observation des échantillons*

Les plantes sont synchronisées par transition JC/JL. Après deux semaines en JL les SAM sont disséqués en retirant les stades les plus avancés (> stade 3) à l'aide d'une pince aiguisée et plantés verticalement dans une petite boîte de milieu MS.

Pour les plantes traitées, les apex disséqués sont placés dans un milieu MS modifié (MS basal 0,5X ; sucrose 1 % ; pH 5,8 ; agarose 0,8 % ; myo-inositol 5 mg ; acide nicotinique 0,05 mg ; pyridoxine hydrochloride 0,05 mg ; thiamine hydrochloride 0,5 mg ; glycine 0,1 mg ; N6-benzyladenine 200 mM) et incubé en enceinte de culture jusqu'à 4 jours.

Les parois cellulaires sont marquées par incubation 30 min avec du FM4-64 (333 ng/μl ; ThermoFisher). Les tissus sont observés sur un microscope confocal droit TCS SP2 (Leica). L'acquisition des signaux FM4-64 et VENUS sont faits par excitation avec le laser 514 nm et des fenêtres d'acquisition de 630 à 665 nm (PMT 652 V) et de 520 à 555 nm (PMT 465 V) respectivement.

Le signal est enregistré sous forme de « *stacks* » (collection d'images) tridimensionnels composés d'images de résolution 1 024X 1 024 pixels, une gamme de couleurs de 12 bits et une taille des voxels (pixel volumétrique) de 232,51 x 232,51 x 488,44 nm (xyz). Le traitement des images afin d'améliorer la visibilité (luminosité, contrastes et établissement d'un seuil de signal) a été réalisé dans *FIDJI* [14, 15].

### *Analyse d'aire*

Les *stacks* sont ouverts dans *FIDJI* et l'aire du domaine d'expression de *WUS* et le diamètre du SAM sont mesurés à l'aide de l'outil de sélection ovale sur des projections d'intensité maximale du *stack*.

### *Mesure de courbure de la surface du SAM*

Le canal correspondant au FM4-64 est prétraité dans *FIDJI* : le signal est converti en 'Fire' LUT et le bruit de fond est enlevé en relevant les seuils de

valeur minimale avec l'outil 'Lumière/Contrastes'. La mesure de courbure est effectuée dans *MorphoGraphX* (Barbier de Reuille et al., 2015) en suivant les recommandations du manuel d'utilisation en adaptant certains paramètres : la surface du SAM est reconstruite par la méthode de '*marching cube surface*' (taille de cube : 3  $\mu\text{m}$ ). Le signal correspondant aux parois cellulaires est projeté sur la surface. La segmentation des cellules est effectuée par une sélection manuelle des cellules suivie d'une propagation des labels. La courbure gaussienne est ensuite projetée sur la surface (voisinage : 20  $\mu\text{m}$ ) puis représenté et quantifiée à l'échelle cellulaire grâce à l'outil '*Heatmap*'.

## 9. Analyse de l'expression des gènes par hybridation *in situ*

L'hybridation *in situ* est réalisée comme publié précédemment (Porri et al., 2012).

### *Fixation et inclusion des échantillons*

Les plantes sont synchronisées par transition JC/JL. Après deux semaines en JL, les SAM sont disséqués en retirant les stades les plus avancés à l'aide d'une pince aiguisée et placés immédiatement dans la solution de fixation (PBS 1X ; 4% paraformaldéhyde ; pH 7 avec NaOH et H<sub>2</sub>SO<sub>4</sub>), sur glace. Les échantillons sont infiltrés par vide relatif (-0,5 bar ; 20 min) deux fois en changeant la solution de fixation à chaque fois puis incubés sur la nuit en sous agitation à 4°C.

Les échantillons sont ensuite déshydratés par un gradient d'éthanol/solution de fixation (0% ; 50% ; 75% ; 87,5% ; 95% ; 100%). Chaque incubation est effectuée 1 h 30 en rotation à 4°C. Lors de l'étape à 95% d'éthanol, la solution est supplémentée avec 0,1 % d'éosine pour colorer les tissus. Suite au gradient, les tissus sont incubés dans un nouveau bain d'éthanol, sur la nuit, en rotation.

L'éthanol est ensuite progressivement remplacé par de l'Histoclear II par gradient (0% ; 25% ; 43,8% ; 57,8% ; 68,4% ; 76,3% ; 100%) chaque étape durant 1 h, sous rotation, à température ambiante. Les tissus sont finalement incubés 1 h sous rotation dans une solution d'Histoclear fraîche. La moitié du

volume d'Histoclear est ensuite remplacée par des morceaux de Paraplast Plus et l'ensemble est incubé à température ambiante sur la nuit. Le Paraplast est ensuite fondu à 42°C et plus de Paraplast a été ajouté. Après dissolution du Paraplast, les échantillons sont transférés à 56°C et le Paraplast a été changé toutes les 8 à 10 h, au moins 5 fois. Les échantillons dans le Paraplast sont ensuite placés dans des moules et laissés à température ambiante pour solidification.

Les tissus fixés sont découpés au microtome à une épaisseur de 8µm et placés dans un bain d'eau chaude à 42°C. Les sections sont ensuite placées sur des lames de verre prétraitées à la polysine et incubées à 42°C pendant 16 h.

#### *Synthèse des sondes*

L'ADN plasmidique contenant la séquence de la sonde sous contrôle de promoteurs de l'ARN polymérase est linéarisé par digestion sur la nuit puis purifié sur colonne de purification d'acides nucléiques (Nucleospin gel and PCR clean-up, Macherey-Nagel). La transcription est effectuée avec le kit approprié (Promega, T3, T7 ou SP6 RNA-polymerase), selon les recommandations du fournisseur, à partir d'1 µg de plasmide linéarisé. La transcription se déroule à 37°C pendant 1 heure puis les échantillons sont traités à la DNaseI selon les instructions du fournisseur. Les sondes sont ensuite purifiées sur colonne (Chroma SPIN +STE-30 ; ClonTech) puis hydrolysées dans un tampon carbonate (60 mM Na<sub>2</sub>CO<sub>3</sub> ; 40 mM NaHCO<sub>3</sub> ; pH 10,2) afin d'obtenir des fragments d'environ 200 nucléotides selon la formule :

$$t = \frac{L_{sonde} - L_{finale}}{0,11 \cdot L_{sonde} \cdot L_{finale}}$$

A la fin de l'hydrolyse les ARNs sont précipités par ajout de 0,2 mg/ml de glycogène, 1% d'acide acétique, 300 mM NaOAc et 2 volumes d'éthanol suivi d'une incubation 30 min à -80°C et d'une centrifugation à 20 000 g pendant au moins 30 min. Les culots sont lavés à l'éthanol 70% et immédiatement resuspendus dans un tampon de conservation (10 mM Tris-HCl pH 8; 1 mM EDTA pH 8; 50% formamide) et stockés à -20°C.

### *Préparation des tissus*

La cire est retirée des lames par deux bains successifs d'Histoclear pendant 10 min puis les sections sont réhydratées par des bains successifs d'éthanol (100% ; 100% ; 95% ; 90% ; 80% ; 70% ; 50% ; 30% ; 0% ; 0%). Les parois cellulaires sont dégradées par une incubation 20 min dans une solution de HCl 200 mM puis les lames sont lavées (eau puis PBS puis eau) avant d'être traitées 30 min à 37°C par une solution de protéinase K (Tris pH8 100 mM ; EDTA 50 mM ; protéinase K 1 µg/ml). Les lames sont lavées (PBS puis PBS/Glycine (2 mg/ml) puis PBS) pour neutraliser la protéinase K. Les sections sont à nouveau fixées dans une solution de PBS/formaldéhyde 3,7% pendant 20 min avant d'être lavées au PBS. Les sections sont finalement déshydratées par des bains successifs d'éthanol (0% ; 0% ; 30% ; 50% ; 70% ; 80% ; 90% ; 95% ; 100% ; 100%) et séchées sous vide.

### *Hybridation et révélation*

Les sections sont pré-hybridées à 55°C avec le tampon d'hybridation (SSC 1X ; SDS 3% ; 50% formamide ; 0,1 mg/ml yeast tRNA) dans une boîte maintenue à 16 h humide par des papiers absorbants imbibés de formamide 50%. Pour l'hybridation, la sonde est diluée au 1/5<sup>ème</sup> dans une solution de formamide 50% et dénaturée à 75°C pendant 2 min. Ce mélange est ajouté au tampon d'hybridation dans un ratio de 1 µl de sonde pure pour 100 µl de tampon d'hybridation. Le mélange est incubé sur les sections à 55°C pendant 16h (100 µl par section).

Les sections sont enfin lavées dans un tampon de SSC 0.2X / 0.1 % SDS à 55°C puis dans un tampon SSC 2X pendant 2 min, TBS 1X pendant 5 min puis 0.5 % de réactif de Boehringer dans 1X TBS pendant 1 h. Les lames sont ensuite rincées au TBS 1X /BSA 0,5%. L'incubation avec l'anticorps anti-Dig conjugué à la peroxydase alcaline est faite pendant 2 h à température ambiante (TBS 1X ; 0,5% BSA ; 1 :1500 anti-Dig-AP). Les lames sont lavées (TBS 1X / BASA 0,5% / Triton X-100 0,1%, 3 fois pendant 10 min.) puis rincées dans le tampon de détection (100 mM Tris pH 9,6 ; 100 mM NaCl ; 50 mM MgCl<sub>2</sub>) pendant 15 min. La révélation se fait dans le tampon de détection supplémenté par 1,6% v/v de BCPIP et 2,2% v/v de NBT, à l'obscurité pendant

4 à 24 h. La réaction est stoppée par ajout d'eau pure. Les lames sont ensuite montées dans 50% glycérole.

## 10. Microscopie électronique

Les apex sont disséqués juste avant observation et placés rapidement sur un support conducteur adhésif à base de graphite. Les échantillons sont observés sur un microscope électronique à balayage FEI Quanta 250 en mode environnemental avec une tension de 14 kV. La pression est fixée entre 700 et 550 Pa, pour une température entre 1°C et 2°C, en fonction de la résistance des échantillons à la dessiccation.

## II. Production et caractérisation des protéines *in vitro*

### 1. Production de protéines recombinantes

#### *Constructions de plasmides permettant l'expression de protéines recombinantes*

La séquence codante de *RAX1* dans pHC03 a subi une mutation silencieuse avec les amorces oGD03 et oGD04 pour enlever le site de restriction interne *NcoI*, produisant pGD01, à partir duquel *RAX1* a été amplifié par des amorces rajoutant des sites de restriction *NcoI* et *NotI* : la forme complète de *RAX1* (*RAX1<sub>full</sub>*) a été amplifiée avec les amorces oGD01 et oGD02, le domaine Myb (*RAX1<sub>myb</sub>*) a été amplifié avec les amorces oGD01 et oGD05 (Tableau MM 2).

Tous les amplicons ont été sous-clonés dans pCR®-Blunt (ThermoFisher) puis transférés entre les sites *NcoI* et *NotI* des vecteurs d'expression pETM11 ou pETM41 (EMBL). La forme complète de *RAX1* a été clonée dans pETM41, en fusion avec 6 histidines et le domaine de liaison au maltose MBP (pGD19). Le domaine Myb de *RAX1* a été cloné dans pETM11, en fusion avec 6 Histidines (pGD14). La séquence codant pour WUS a été clonée dans pETM41 avant le présent travail (pETH259).

### *Cultures bactériennes et induction de la production de protéines recombinantes*

Des bactéries *Escherichia coli* compétentes de souche Rosetta 2 sont transformées avec ~300 ng de plasmide, étalées sur milieu LB solide avec les antibiotiques nécessaires et cultivées sur la nuit à 37°C. Une colonie est ensuite transférée dans 15 ml de LB liquide et cultivée pendant 3 h avant d'être transférée dans 1 L de milieu. Lorsque la culture atteint une DO de 0,8 à 600 nm, 2 mM de bétaïne sont ajoutés pour prévenir la dégradation des protéines produites et la culture est refroidie à 17°C. La production de protéines est alors induite par ajout de 0,4 mM d'IPTG. Elle se poursuit pendant 16 h à 17°C.

### *Purifications des protéines sur colonne*

500 ml de culture cellulaire sont récoltés par centrifugation à 13 000 g pendant 30 min. Les cellules sont resuspendues dans 50 ml de tampon de solubilisation contenant une tablette de cOmplete™ EDTA-free protease inhibitor (Sigma-Aldrich) puis lysées par sonication. La fraction soluble du lysat est extraite par centrifugation à 13 000 g pendant 40 min et chargée sur colonne de Nickel-Sepharose - permettant l'adsorption des répétitions d'histidine - préalablement équilibrée dans le tampon de solubilisation. La colonne est lavée avec 20 volumes de tampon de solubilisation supplémenté avec 20 mM d'imidazole. L'élution est réalisée dans 3 volumes de tampon de solubilisation supplémenté avec 300 mM d'imidazole. L'éluât est ensuite dialysé dans au moins 60 volumes de tampon de solubilisation pendant 16 h à 4°C. Les tampons de solubilisation utilisés sont : tampon RAX1<sub>myb</sub>-His (Tris-HCl 20 mM ; pH 8 ; TCEP 1 mM), tampon RAX1<sub>full</sub>-MBP (Tris HCl 20 mM ; pH 8 ; NaCl 500 mM; TCEP 1mM) et tampon WUS (Tris HCl 20 mM ; pH 8 ; NaCl 100 mM; TCEP 1mM).

## 2. Protein-binding micro-array

Le *protein-binding micro-array* (PBM) a été réalisé sur la protéine RAX1 entière fusionnée au MBP (pGD19) comme décrit précédemment (Godoy et al., 2011). Cette expérience a été réalisée par des collaborateurs au CNB Madrid. Brièvement, la protéine est hybridée sur une puce ADN double brin présentant toutes les combinaisons de 11 nucléotides. Les brins d'ADN étant

fusionnés au Cy5, ils émettent de la fluorescence à 635 nm. Après trois lavages au tampon phosphate salin - Tween (PBS-T) les lames sont hybridées avec un anticorps primaire polyclonal anti-MBP (Abcam). Après une seconde étape de lavage, les lames sont hybridées avec un anticorps secondaire fusionné à une molécule fluorescente dans le vert (DyLight 549 ; Pierce). Les lames sont ensuite scannées pour la fluorescence à 635 nm et à 549 nm correspondant à la présence d'ADN et de protéine marquée respectivement. Le rapport de fluorescence vert/rouge permet ensuite de déterminer les séquences liées par la protéine. L'alignement des séquences liées a permis de construire une matrice de fréquence représentant un modèle statistique de l'affinité de RAX1 pour l'ADN (pour une revue, voir (Stormo, 2013)).

### 3. Gels retards

#### *Hybridation et marquage des sondes*

Les sondes sont synthétisées à partir d'oligonucléotides complémentaires simple brin. 10  $\mu\text{M}$  de chaque brin est mélangé dans du tampon d'hybridation (10 mM Tris HCl pH 7,5 ; 150 mM NaCl ; 1 mM EDTA) et chauffé à 95°C pendant 3 min puis le mélange est refroidi doucement à température ambiante. Le marquage des sondes est effectué par l'ajout d'une cytosine marquée sur la guanine dépassante d'un des brins de la sonde. 0,2  $\mu\text{M}$  de sonde hybridée sont mélangés à 0,4  $\mu\text{M}$  de Cy5-dCTP (Sigma-Aldrich), la ligation est effectuée avec le fragment de Klenow de l'ADN polymérase (New England Biolabs) dans le tampon approprié pendant 1 h à 37°C à l'obscurité.

#### *Electrophoretic Mobility Gel Shift Assay (EMSA)*

Pour les EMSAs, un mélange est préparé contenant le tampon de liaison, de l'ADN compétiteur (de sperme de hareng) et la sonde marquée à 10 nM dans un volume final de 20  $\mu\text{L}$ . La protéine est ajoutée en dernier au mélange. Le mélange est incubé 15 min sur glace puis chargé dans un gel d'acrylamide non dénaturant de 6 à 8% ou dans un gel d'agarose 2%. L'électrophorèse est typiquement effectuée à 90 V pendant 1 h 15 dans du TBE 0,5 X. Typiquement 28 mg/ml d'ADN compétiteur ont été utilisés. Les tampons de liaison utilisés

sont les suivants : pour RAX1<sub>full</sub>-MBP, tampon Promega1 (Tris-HCl 10 mM, pH 7,5 ; NaCl 50 mM ; MgCl<sub>2</sub> 1 mM ; DTT 1 mM ; EDTA 0,5 mM ; glycérol 4%) ; pour RAX1<sub>myb</sub>-His, tampon Sheng (Tris-HCl pH 7,5 50 mM ; NaCl 10 mM ; ZnCl<sub>2</sub> 0,05 mM ; Triton X-100 0,01% ; albumine sérique bovine (BSA) 20 µg/ml ; DTT 1 mM ; glycérol 10%) et pour WUS-MBP, tampon Eugenio (10 mM acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique (HEPES) pH 7,2 ; spermidine 1 mM ; EDTA 14 mM ; acide 3-[(3-cholamidopropyl)diméthylammonio]-1-propane sulfonique (CHAPS) 0,25% ; tris(2-carboxyethyl)phosphine-HCl (TCEP) 3 mM ; BSA 30 µg/ml ; glycérol 3%). La fluorescence est mesurée sur l'imageur ChemiDoc™ (BioRad).

### III. Méthodes bio-informatiques

#### 1. Analyse des données de transcriptomiques

##### *Librairie et séquençage*

La construction des librairies et le séquençage ont été effectués par la plateforme transcriptomique POPS (IPS2). Les ARNs sont fragmentés à ~300pb, les librairies sont construites en utilisant le kit TruSeq® Stranded (Illumina) et le séquençage a été effectué en *paired-end* sur un HiSeq2000 (Illumina) avec des reads de 100pb. Les *reads* sont nettoyés pour retirer les séquences de basse qualité ou trop courtes et les séquences dupliquées. Les *reads* correspondants aux ARN ribosomiaux sont retirés et les séquences des adaptateurs retirées de l'extrémité des *reads*. Entre 26,4 et 37,3 millions de *reads* propres ont été générés par librairie. Les reads sont localisés sur le génome d'Arabidopsis (TAIR10) en utilisant HISAT2 (v2.0.4) (Kim et al., 2015) en autorisant de 200 à 400 paires de bases entre paires de *read*. Une fois localisés, les *reads* sont filtrés en utilisant un seuil de 30 pour la qualité (correspondant à une confiance de localisation de 0,999) en utilisant SAMTools (Li et al., 2009). Les *reads* sont dénombrés dans chaque intron avec HTSeq-count (Anders et al., 2015).

### *Analyse des gènes différentiellement exprimés*

L'analyse des gènes différentiellement exprimés (DEG) est effectuée avec EdgeR (Annexe 3). Les gènes comptabilisant moins d'un *read* par million de reads totaux dans moins de bibliothèques que de nombre de réplicats sont retirés des jeux de données en raison de leur faible impact sur la variation globale. Le nombre de *reads* dans chaque bibliothèque a été normalisé par la méthode de TMM (*trimmed means of M-values*) et les DEG ont été déterminés par un modèle GLM (*Generalized linear model*) multiparamétrique afin d'identifier les gènes répondant à l'interaction génotype:traitement, en utilisant un test de rapport de probabilités (*Likelihood ratio test* ; LRT) (Robinson and Smyth, 2007; Robinson et al., 2010; McCarthy et al., 2012; Zhou et al., 2014). Les p-values obtenues sont corrigées par la méthode de Benjamini-Hochberg et les gènes ayant un taux de fausse découverte (FDR) inférieur à  $1.10^{-2}$  sont considérés comme différentiellement exprimés.

Les termes d'ontologie de gènes enrichis dans le jeu de données ont été déterminés par un test hypergéométrique sur la plateforme ThaleMine (Krishnakumar et al., 2015).

## 2. Détermination *in silico* des sites de liaison à l'ADN des facteurs de transcription

### *Matrices avec dépendances*

La prédiction des sites de liaison à l'ADN à partir de matrices de fréquences comportant des dépendances entre différentes positions (e.g. LFY) est faite grâce au programme *MORPHEUS* (Minguet et al., 2015). Une matrice de poids-position (PWM) est générée à partir de la matrice de fréquence en attribuant à chaque nucléotide  $n$  à chaque position  $i$ , un poids  $W_{(n,i)}$  calculé comme le logarithme népérien de la fréquence  $f_{(n,i)}$  du nucléotide  $n$  à la position  $i$  par rapport à la fréquence  $f_{(max,i)}$  de la base la plus fréquente à cette position :

$$W_{(n,i)} = \ln \frac{f_{(n,i)}}{f_{(max,i)}}$$

La qualité globale d'un site de liaison  $S_j$  est alors déterminée comme le score correspondant à la somme des contributions de chaque position de la séquence :

$$Score(S_j|W) = \sum_{(n,i)} W_{(n,i)} \cdot S_j$$

Cette méthode de calcul donne un résultat qui est l'opposé de l'énergie relative de liaison décrite par Berg et Von Hippel (Berg and von Hippel, 1987a). Le score ainsi obtenu est compris entre le score minimal de la matrice, soit la séquence la plus coûteuse énergétiquement, et 0, soit la séquence minimisant l'énergie de liaison.

#### *Matrices sans dépendances*

La prédiction des sites de liaison des facteurs dont la matrice ne contient pas de dépendance est faite grâce au module Biopython pour python 2.7 (Cock et al., 2009) (Annexe 4). La matrice construite est une matrice de probabilité calculée comme le logarithme de la fréquence  $f(n, i)$  de la base  $n$  à la position  $i$  sur la probabilité de ce nucléotide dans la séquence:

$$W_{(n,i)} = \ln \frac{f(n,i)}{P_n}$$

Le score est déterminé comme précédemment. Le seuil de score considéré comme significatif est déterminé comme étant le score pour lequel l'information  $I_{(n,i)}$  contenue dans la matrice est égale au négatif du log de l'erreur de type I, comme utilisé pour le programme PATSER (Hertz and Stormo, 1999).

#### *Occupation prédite*

La probabilité pour un facteur de transcription d'occuper une séquence est calculée comme décrit pour le programme GOMER (Granek and Clarke, 2005) (Annexe 4). Le score déterminé par la matrice étant proportionnel à l'énergie libre de liaison ( $\Delta G = R.T.score$ ) (Stormo and Fields, 1998), il est possible de déterminer la constante de dissociation du facteur sur le motif d'ADN comme :

$$K_d = e^{-\frac{\Delta G}{R.T}} \Leftrightarrow K_d = e^{-score}$$

La probabilité que le facteur soit lié à ce motif est donc calculée comme l'équilibre :

$$P_i = \frac{[TF]}{K_{d,i}[TF]}$$

La concentration du facteur de transcription étant inaccessible *in vivo*, elle est fixée comme une constante déterminée comme le  $K_d$  de la protéine pour un site de liaison optimal. Il en résulte une probabilité de liaison de 0,5 pour le site optimal. La probabilité de liaison sur la séquence est calculée comme l'inverse du produit des probabilités du facteur de ne pas être lié à chaque motif rencontré :

$$P_{occ} = 1 - \prod_{i=1} (1 - P_i)$$

Il est à noter que contrairement au calcul de l'occupation prédite (pOcc) effectué par MORPHEUS qui calcule le nombre de molécules liées, cette méthode calcule la probabilité (entre 0 et 1) du facteur de transcription d'être lié à la séquence. Pour MORPHEUS :

$$pOcc = \sum_{i=1} P_i$$

### 3. Analyse des données de ChIP-seq et d'état chromatinien

Les données de ChIP-seq sur LFY et LFY<sub>TERE</sub> (GSE64245 ; Sayou *et al.*, *soumis*) obtenues sur des échantillons de plantules (Col) de 15 jours ont été comparées à des données de DNase-seq obtenues dans des échantillons de plantules (Col) de 14 jours (GSE34318 ; (Zhang *et al.*, 2012)). Les analyses ont été réalisées sous Python 2.7 grâce à la librairie *pybedtools* (Dale *et al.*, 2011), basée sur la suite de *BEDTools* (Quinlan and Hall, 2010). Les opérations de conversion et de transformation des fichiers BED, WIG, bigWig, bigBed et bedGraph ont été réalisées grâce aux outils développés par *UCSC Genome Browser* (<https://genome.ucsc.edu/>).

Les informations contenues dans les fichiers WIG ou bedGraph sont récupérées pour chaque région liée dans le ChIP LFY grâce au sous-programme 'intersect' de *BEDTools* (options `-wa -wb`). La réduction de couverture *CFR* entre le ChIP LFY et LFY<sub>TERE</sub> est calculé comme :

$$CFR = \frac{signal_{LFY}}{signal_{LFY_{TERE}}}$$

Pour les régions où la couverture extrêmement faible du LFY<sub>TERE</sub> donne des valeurs de *CFR* artificiellement hautes ou proches de l'infinie, le *CFR* a été borné arbitrairement à 3000.

Les représentations graphiques sont faites avec les bibliothèques *matplotlib* et *seaborn*, les analyses statistiques ont requis l'utilisation de la bibliothèque *SciPy*.





# CONCLUSION & DISCUSSION

Au cours de ce travail nous avons pu montrer que LFY est impliqué dans la mise en place des FM, en partie *via* la régulation de *RAX1* (Chahtane et al., 2013; Denay et al., 2017 et Article 2). Nous avons montré que RAX1 agissait sans doute à la fois sur les propriétés des parois cellulaires et sur la boucle régulatrice *CLV/WUS*. Dans la deuxième partie de ce travail nous avons mis en évidence le rôle du domaine N-terminal de LFY. Ce domaine permet l'oligomérisation de la protéine et lui permet de lier des régions fermées de la chromatine (Sayou et al., 2016). Puisque LFY interagit avec les remodelleurs chromatinien SYD et BRM, il est probable qu'il ait un rôle de facteur pionnier (Wu et al., 2012). Ces résultats ayant déjà été discutés dans les différentes parties, nous discuterons ici brièvement de leurs implications potentielles dans le développement des fleurs. Dans quelle mesure l'oligomérisation est-elle importante pour la mise en place des FM ? Quel est le rôle spécifique de LFY dans la formation des FM ? Nous discutons ici quelques hypothèses basées sur le travail présenté ici et sur la littérature existante.

### *Rôle de l'oligomérisation de LFY pour l'activation des gènes cibles*

Alors que la sur-expression de LFY induit la formation précoce de nombreux FM à l'aisselle des feuilles de rosette, la sur-expression de LFY<sub>TERE</sub> (incapable d'oligomériser) ne provoque pas ce phénotype. L'oligomérisation de LFY semble donc importante pour l'*initiation* des FM. D'une manière cohérente, des boucles d'ADN peuvent être détectées entre les sites de fixation de LFY au locus *RAX1* (V. Conn, *communication personnelle*). Cela indique que LFY pourrait oligomériser entre les différents sites de liaison à l'ADN et former des boucles, potentiellement importantes pour la dynamique chromatinienne et l'expression des gènes cibles (Mendes et al., 2013; Cao et al., 2014).

### *Rôle de LFY dans la mise en place des FM*

Les gènes impliqués dans la mise en place des FM et des LSM sont très conservés dans les deux processus, indiquant une similarité de mécanismes. La différence majeure entre ces deux processus est l'expression de *LFY* dans les FM. *LFY* joue un rôle dans les processus floraux de répression de la bractée et d'initiation immédiate du méristème non-quiescent (Norberg et al., 2005 et Article 2). De plus, chez les mutants *lfy* les fleurs sont remplacées par des structures rappelant une inflorescence secondaire, sous-tendue par une feuille et à la symétrie spiralée (Norberg et al., 2005; Chandler and Werr, 2014). Ainsi, comme il l'a déjà été proposé il est tentant de spéculer que la formation d'une feuille portant un méristème d'inflorescence est le programme 'par défaut'. *LFY* provoquerait alors la conversion de cette structure en une fleur.

### *La fleur, une suite de transitions développementales*

La formation d'une fleur constitue donc une suite de transitions développementales régulés par des modifications de la chromatine. Le maximum d'auxine déclenche l'activation des ARF et le recrutement des remodelleurs SYD et BRM. Ce processus déclenche l'activation des programmes de développement de l'entre-nœud : croissance de la feuille et développement d'un LSM. Lorsque les conditions externes et internes le permettent, l'activation des ARFs entraîne également l'expression de *LFY*.

LFY peut lui aussi recruter les remodelleurs SYD et BRM et entraîne alors la conversion de l'entre-nœud en structure de type floral par la répression de la croissance de la feuille (bractée) et la modification de la phyllotaxie de l'AM. LFY active également l'expression des MADS-box TF qui permettent la stabilisation de l'identité florale et, toujours *via* le recrutement de SYD et BRM, la détermination de l'identité des différents organes floraux et leur développement.







# RÉFÉRENCES

- Adibi, M., Yoshida, S., Weijers, D., and Fleck, C.** (2016). Centering the Organizing Center in the *Arabidopsis thaliana* Shoot Apical Meristem by a Combination of Cytokinin Signaling and Self-Organization. *PLoS ONE* **11**: e0147830.
- Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., and Tasaka, M.** (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. *Plant Cell* **9**: 841–857.
- Aida, M., Ishida, T., and Tasaka, M.** (1999). Shoot apical meristem and cotyledon formation during *Arabidopsis* embryogenesis: interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes. *Development* **126**: 1563–1570.
- Anders, S., Pyl, P.T., and Huber, W.** (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.
- Aravind, L., Anantharaman, V., Balaji, S., Babu, M.M., and Iyer, L.M.** (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* **29**: 231–262.
- Armengot, L., Marquès-Bueno, M.M., and Jaillais, Y.** (2016). Regulation of polar auxin transport by protein and lipid kinases. *J Exp Bot* **67**: 4015–4037.
- Barbier de Reuille, P. et al.** (2015). MorphoGraphX: A platform for quantifying morphogenesis in 4D. *elife* **4**: 05864.
- Bayer, E.M., Smith, R.S., Mandel, T., Nakayama, N., Sauer, M., Prusinkiewicz, P., and Kuhlemeier, C.** (2009). Integration of transport-based models for phyllotaxis and midvein formation. *Genes Dev* **23**: 373–384.
- Benjamins, R., Quint, A., Weijers, D., Hooykaas, P., and Offringa, R.** (2001). The PINOID protein kinase regulates organ development in *Arabidopsis* by enhancing polar auxin transport. *Development* **128**: 4057–4067.
- Benlloch, R., Kim, M.C., Sayou, C., Thévenon, E., Parcy, F., and Nilsson, O.** (2011). Integrating long-day flowering signals: a LEAFY binding site is essential for proper photoperiodic activation of APETALA1. *Plant J* **67**: 1094–1102.
- Bennett, S.R.M., Alvarez, J., Bossinger, G., and Smyth, D.R.** (1995). Morphogenesis in pinoid mutants of *Arabidopsis thaliana*. *Plant J* **8**: 505–520.

- Bensmihen, S., To, A., Lambert, G., Kroj, T., Giraudat, J., and Parcy, F.** (2004). Analysis of an activated ABI5 allele using a new selection method for transgenic Arabidopsis seeds. *FEBS Lett* **561**: 127–131.
- Berg, O. and von Hippel, P.** (1987a). Selection of DNA binding sites by regulatory proteins. *J Mol Biol* **193**: 723–743.
- Berg, O.G. and von Hippel, P.H.** (1987b). Selection of DNA binding sites by regulatory proteins. Statistical-mechanical theory and application to operators and promoters. *J Mol Biol* **193**: 723–750.
- Besnard, F. et al.** (2014a). Cytokinin signalling inhibitory fields provide robustness to phyllotaxis. *Nature* **505**: 417–421.
- Besnard, F., Rozier, F., and Vernoux, T.** (2014b). The AHP6 cytokinin signaling inhibitor mediates an auxin-cytokinin crosstalk that regulates the timing of organ initiation at the shoot apical meristem. *Plant Signal Behav* **9**.
- Blázquez, M.A., Green, R., Nilsson, O., Sussman, M.R., and Weigel, D.** (1998). Gibberellins promote flowering of Arabidopsis by activating the LEAFY promoter. *Plant Cell* **10**: 791–800.
- Blázquez, M.A., Soowal, L.N., Lee, I., and Weigel, D.** (1997). LEAFY expression and flower initiation in Arabidopsis. *Development* **124**: 3835–3844.
- Bombliès, K., Wang, R.L., Ambrose, B.A., Schmidt, R.J., Meeley, R.B., and Doebley, J.** (2003). Duplicate FLORICAULA/LEAFY homologs zfl1 and zfl2 control inflorescence architecture and flower patterning in maize. *Development* **130**: 2385–2395.
- Bouché, F., Lobet, G., Tocquin, P., and Périlleux, C.** (2016). FLOR-ID: an interactive database of flowering-time gene networks in Arabidopsis thaliana. *Nucleic Acids Res* **44**: D1167–D1171.
- Bowman, J.L., Smyth, D.R., and Meyerowitz, E.M.** (1991). Genetic interactions among floral homeotic genes of Arabidopsis. *Development* **112**: 1–20.
- Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M., and Simon, R.** (2000). Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by CLV3 activity. *Science* **289**: 617–619.
- Brandt, R. et al.** (2012). Genome-wide binding-site analysis of REVOLUTA reveals a link between leaf patterning and light-mediated growth responses. *Plant J* **72**: 31–42.
- Busch, M.A., Bombliès, K., and Weigel, D.** (1999). Activation of a floral homeotic gene in Arabidopsis. *Science* **285**: 585–587.
- Busch, W. et al.** (2010). Transcriptional control of a plant stem cell niche. *Dev Cell* **18**: 849–861.
- Cao, S., Kumimoto, R.W., Gnesutta, N., Calogero, A.M., Mantovani, R., and Holt, B.F.** (2014). A distal CCAAT/NUCLEAR FACTOR Y complex promotes chromatin looping at the FLOWERING LOCUS T promoter and regulates the timing of flowering in Arabidopsis. *Plant Cell* **26**: 1009–1017.
- Chae, E., Tan, Q.K., Hill, T.A., and Irish, V.F.** (2008). An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development* **135**: 1235–1245.
- Chahtane, H. et al.** (2013). A variant of LEAFY reveals its capacity to stimulate meristem development by inducing RAX1. *Plant J* **74**: 678–689.
- Chai, C., Xie, Z., and Grotewold, E.** (2011). SELEX (Systematic Evolution of Ligands by EXponential Enrichment), as a powerful tool for deciphering the protein-DNA interaction space. *Methods Mol Biol* **754**: 249–258.
- Chandler, J.W.** (2012). Floral meristem initiation and emergence in plants. *Cell Mol Life Sci* **69**: 3807–3818.

- Chandler, J.W. and Werr, W.** (2014). Arabidopsis floral phytomer development: auxin response relative to biphasic modes of organ initiation. *J Exp Bot* **65**: 3097–3110.
- Chen, Q., Atkinson, A., Otsuga, D., Christensen, T., Reynolds, L., and Drews, G.N.** (1999). The Arabidopsis FILAMENTOUS FLOWER gene is required for flower formation. *Development* **126**: 2715–2726.
- Cheng, Y., Dai, X., and Zhao, Y.** (2006). Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. *Genes Dev* **20**: 1790–1799.
- Cheng, Y., Dai, X., and Zhao, Y.** (2007a). Auxin synthesized by the YUCCA flavin monooxygenases is essential for embryogenesis and leaf formation in Arabidopsis. *Plant Cell* **19**: 2430–2439.
- Cheng, Y., Qin, G., Dai, X., and Zhao, Y.** (2007b). NPY1, a BTB-NPH3-like protein, plays a critical role in auxin-regulated organogenesis in Arabidopsis. *Proc Natl Acad Sci U S A* **104**: 18825–18829.
- Chickarmane, V.S., Gordon, S.P., Tarr, P.T., Heisler, M.G., and Meyerowitz, E.M.** (2012). Cytokinin signaling as a positional cue for patterning the apical-basal axis of the growing Arabidopsis shoot meristem. *Proc Natl Acad Sci U S A* **109**: 4002–4007.
- Christensen, S.K., Dagenais, N., Chory, J., and Weigel, D.** (2000). Regulation of auxin response by the protein kinase PINOID. *Cell* **100**: 469–478.
- Clark, S.E., Running, M.P., and Meyerowitz, E.M.** (1993). CLAVATA1, a regulator of meristem and flower development in Arabidopsis. *Development* **119**: 397–418.
- Clark, S.E., Williams, R.W., and Meyerowitz, E.M.** (1997). The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell* **89**: 575–585.
- Cock, P.J., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., and de Hoon, M.J.** (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**: 1422–1423.
- Coen, E.S. and Meyerowitz, E.M.** (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* **353**: 31–37.
- Coen, E.S., Romero, J.M., Doyle, S., Elliott, R., Murphy, G., and Carpenter, R.** (1990). floricaula: a homeotic gene required for flower development in antirrhinum majus. *Cell* **63**: 1311–1322.
- Cole, M., Chandler, J., Weijers, D., Jacobs, B., Comelli, P., and Werr, W.** (2009). DORNROSCHEN is a direct target of the auxin response factor MONOPTEROS in the Arabidopsis embryo. *Development* **136**: 1643–1651.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R.** (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* **139**: 5–17.
- Dale, R.K., Pedersen, B.S., and Quinlan, A.R.** (2011). Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. *Bioinformatics* **27**: 3423–3424.
- Daum, G., Medzihradzky, A., Suzaki, T., and Lohmann, J.U.** (2014). A mechanistic framework for noncell autonomous stem cell induction in Arabidopsis. *Proc Natl Acad Sci U S A* **111**: 14619–14624.
- Denay, G., Chahtane, H., Tichtinsky, G., and Parcy, F.** (2017). A flower is born: an update on Arabidopsis floral meristem formation. *Curr Opin Plant Biol*.
- Dharmasiri, N., Dharmasiri, S., Weijers, D., Lechner, E., Yamada, M., Hobbie, L., Ehrismann, J.S., Jürgens, G., and Estelle, M.** (2005). Plant development is regulated by a family of auxin receptor F box proteins. *Dev Cell* **9**: 109–119.

- Dhonukshe, P., Aniento, F., Hwang, I., Robinson, D.G., Mravec, J., Stierhof, Y.D., and Friml, J.** (2007). Clathrin-mediated constitutive endocytosis of PIN auxin efflux carriers in Arabidopsis. *Curr Biol* **17**: 520–527.
- Ditta, G., Pinyopich, A., Robles, P., Pelaz, S., and Yanofsky, M.F.** (2004). The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. *Curr Biol* **14**: 1935–1940.
- Emery, J.F., Floyd, S.K., Alvarez, J., Eshed, Y., Hawker, N.P., Izhaki, A., Baum, S.F., and Bowman, J.L.** (2003). Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. *Curr Biol* **13**: 1768–1774.
- Endrizzi, K., Moussian, B., Haecker, A., Levin, J.Z., and Laux, T.** (1996). The SHOOT MERISTEMLESS gene is required for maintenance of undifferentiated cells in Arabidopsis shoot and floral meristems and acts at a different regulatory level than the meristem genes WUSCHEL and ZWILLE. *Plant J* **10**: 967–979.
- Engelhorn, J., Moreau, F., Fletcher, J.C., and Carles, C.C.** (2014). ULTRAPETALA1 and LEAFY pathways function independently in specifying identity and determinacy at the Arabidopsis floral meristem. *Ann Bot* **114**: 1497–1505.
- Eshed, Y., Baum, S.F., Perea, J.V., and Bowman, J.L.** (2001). Establishment of polarity in lateral organs of plants. *Curr Biol* **11**: 1251–1260.
- Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y., and Zhu, J.K.** (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res* **23**: 1229–1232.
- Ferrándiz, C., Gu, Q., Martienssen, R., and Yanofsky, M.F.** (2000). Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. *Development* **127**: 725–734.
- Fletcher, J.C., Brand, U., Running, M.P., Simon, R., and Meyerowitz, E.M.** (1999). Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science* **283**: 1911–1914.
- Franco-Zorrilla, J.M., López-Vidriero, I., Carrasco, J.L., Godoy, M., Vera, P., and Solano, R.** (2014). DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc Natl Acad Sci U S A* **111**: 2367–2372.
- Franco-Zorrilla, J.M. and Solano, R.** (2014). High-throughput analysis of protein-DNA binding affinity. *Methods Mol Biol* **1062**: 697–709.
- Friml, J. et al.** (2004). A PINOID-dependent binary switch in apical-basal PIN polar targeting directs auxin efflux. *Science* **306**: 862–865.
- Furutani, M., Nakano, Y., and Tasaka, M.** (2014). MAB4-induced auxin sink generates local auxin gradients in Arabidopsis organ formation. *Proc Natl Acad Sci U S A* **111**: 1198–1203.
- Gallois, J.L., Nora, F.R., Mizukami, Y., and Sablowski, R.** (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes Dev* **18**: 375–380.
- Geldner, N., Friml, J., Stierhof, Y.D., Jürgens, G., and Palme, K.** (2001). Auxin transport inhibitors block PIN1 cycling and vesicle trafficking. *Nature* **413**: 425–428.
- Godoy, M., Franco-Zorrilla, J.M., Pérez-Pérez, J., Oliveros, J.C., Lorenzo, O., and Solano, R.** (2011). Improved protein-binding microarrays for the identification of DNA-binding specificities of transcription factors. *Plant J* **66**: 700–711.
- Goldshmidt, A., Alvarez, J.P., Bowman, J.L., and Eshed, Y.** (2008). Signals derived from YABBY gene activities in organ primordia regulate growth and partitioning of Arabidopsis shoot apical meristems. *Plant Cell* **20**: 1217–1230.
- Gordon, S.P., Chickarmane, V.S., Ohno, C., and Meyerowitz, E.M.** (2009). Multiple feedback loops through cytokinin signaling control stem cell number within the Arabidopsis shoot meristem. *Proc Natl Acad Sci U S A* **106**: 16529–16534.

- Van der Graaff, E., Laux, T., and Rensing, S.A.** (2009). The WUS homeobox-containing (WOX) protein family. *Genome Biol* **10**: 248.
- Grandi, V., Gregis, V., and Kater, M.M.** (2012). Uncovering genetic and molecular interactions among floral meristem identity genes in *Arabidopsis thaliana*. *Plant J* **69**: 881–893.
- Granek, J.A. and Clarke, N.D.** (2005). Explicit equilibrium modeling of transcription-factor binding and gene regulation. *Genome Biol* **6**: R87.
- Gray, W.M., Kepinski, S., Rouse, D., Leyser, O., and Estelle, M.** (2001). Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. *Nature* **414**: 271–276.
- Gray, W.M., del Pozo, J.C., Walker, L., Hobbie, L., Risseeuw, E., Banks, T., Crosby, W.L., Yang, M., Ma, H., and Estelle, M.** (1999). Identification of an SCF ubiquitin-ligase complex required for auxin response in *Arabidopsis thaliana*. *Genes Dev* **13**: 1678–1691.
- Greb, T., Clarenz, O., Schafer, E., Muller, D., Herrero, R., Schmitz, G., and Theres, K.** (2003). Molecular analysis of the LATERAL SUPPRESSOR gene in *Arabidopsis* reveals a conserved control mechanism for axillary meristem formation. *Genes Dev* **17**: 1175–1187.
- Gruel, J., Landrein, B., Tarr, P., Schuster, C., Refahi, Y., Sampathkumar, A., Hamant, O., Meyerowitz, E.M., and Jönsson, H.** (2016). An epidermis-driven mechanism positions and scales stem cell niches in plants. *Science advances* **2**: e1500989.
- Guo, D., Zhang, J., Wang, X., Han, X., Wei, B., Wang, J., Li, B., Yu, H., Huang, Q., Gu, H., Qu, L.J., and Qin, G.** (2015). The WRKY Transcription Factor WRKY71/EXB1 Controls Shoot Branching by Transcriptionally Regulating RAX Genes in *Arabidopsis*. *Plant Cell* **27**: 3112–3127.
- Gustafson-Brown, C., Savidge, B., and Yanofsky, M.F.** (1994). Regulation of the *Arabidopsis* floral homeotic gene APETALA1. *Cell* **76**: 131–143.
- Hamann, T., Benkova, E., Bäurle, I., Kientz, M., and Jürgens, G.** (2002). The *Arabidopsis* BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. *Genes Dev* **16**: 1610–1615.
- Hamann, T., Mayer, U., and Jürgens, G.** (1999). The auxin-insensitive bodenlos mutation affects primary root formation and apical-basal patterning in the *Arabidopsis* embryo. *Development* **126**: 1387–1395.
- Hamès, C., Ptchelkine, D., Grimm, C., Thevenon, E., Moyroud, E., Gérard, F., Martiel, J.L., Benlloch, R., Parcy, F., and Müller, C.W.** (2008). Structural basis for LEAFY floral switch function and similarity with helix-turn-helix proteins. *EMBO J* **27**: 2628–2637.
- Heisler, M.G., Ohno, C., Das, P., Sieber, P., Reddy, G.V., Long, J.A., and Meyerowitz, E.M.** (2005). Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the *Arabidopsis* inflorescence meristem. *Curr Biol* **15**: 1899–1911.
- Hertz, G.Z. and Stormo, G.D.** (1999). Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15**: 563–577.
- Honma, T. and Goto, K.** (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* **409**: 525–529.
- Honma, T. and Goto, K.** (2000). The *Arabidopsis* floral homeotic gene PISTILLATA is regulated by discrete cis-elements responsive to induction and maintenance signals. *Development* **127**: 2021–2030.
- Hosoda, K., Imamura, A., Katoh, E., Hatta, T., Tachiki, M., Yamada, H., Mizuno, T., and Yamazaki, T.** (2002). Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the *Arabidopsis* response regulators. *Plant Cell* **14**: 2015–2029.
- Huala, E. and Sussex, I.M.** (1992). LEAFY Interacts with Floral Homeotic Genes to Regulate *Arabidopsis* Floral Development. *Plant Cell* **4**: 901–913.

- Ikeda, M., Mitsuda, N., and Ohme-Takagi, M.** (2009). Arabidopsis WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. *Plant Cell* **21**: 3493–3505.
- Isono, K., Endo, T.A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B.E., and Koseki, H.** (2013). SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. *Dev Cell* **26**: 565–577.
- Jacqumard, A., Gadisseur, I., and Bernier, G.** (2003). Cell division and morphological changes in the shoot apex of Arabidopsis thaliana during floral transition. *Ann Bot* **91**: 571–576.
- Jasinski, S., Piazza, P., Craft, J., Hay, A., Woolley, L., Rieu, I., Phillips, A., Hedden, P., and Tsiantis, M.** (2005). KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. *Curr Biol* **15**: 1560–1565.
- Jeong, S., Trotochaud, A.E., and Clark, S.E.** (1999). The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell* **11**: 1925–1934.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E.** (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**: 816–821.
- Jönsson, H., Heisler, M.G., Shapiro, B.E., Meyerowitz, E.M., and Mjolsness, E.** (2006). An auxin-driven polarized transport model for phyllotaxis. *Proc Natl Acad Sci U S A* **103**: 1633–1638.
- Kayes, J.M. and Clark, S.E.** (1998). CLAVATA2, a regulator of meristem and organ development in Arabidopsis. *Development* **125**: 3843–3851.
- Keller, T., Abbott, J., Moritz, T., and Doerner, P.** (2006). Arabidopsis REGULATOR OF AXILLARY MERISTEMS1 controls a leaf axil stem cell niche and modulates vegetative development. *Plant Cell* **18**: 598–611.
- Kempin, S.A., Savidge, B., and Yanofsky, M.F.** (1995). Molecular basis of the cauliflower phenotype in Arabidopsis. *Science* **267**: 522–525.
- Kidner, C.A. and Martienssen, R.A.** (2004). Spatially restricted microRNA directs leaf polarity through ARGONAUTE1. *Nature* **428**: 81–84.
- Kieffer, M., Stern, Y., Cook, H., Clerici, E., Maulbetsch, C., Laux, T., and Davies, B.** (2006). Analysis of the transcription factor WUSCHEL and its functional homologue in Antirrhinum reveals a potential mechanism for their roles in meristem maintenance. *Plant Cell* **18**: 560–573.
- Kierzkowski, D., Lenhard, M., Smith, R., and Kuhlemeier, C.** (2013). Interaction between meristem tissue layers controls phyllotaxis. *Dev Cell* **26**: 616–628.
- Kim, D., Langmead, B., and Salzberg, S.L.** (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**: 357–360.
- Kitakura, S., Vanneste, S., Robert, S., Löffke, C., Teichmann, T., Tanaka, H., and Friml, J.** (2011). Clathrin mediates endocytosis and polar distribution of PIN auxin transporters in Arabidopsis. *Plant Cell* **23**: 1920–1931.
- Krishnakumar, V. et al.** (2015). Araport: the Arabidopsis information portal. *Nucleic Acids Res* **43**: D1003–9.
- Krizek, B.A. and Fletcher, J.C.** (2005). Molecular mechanisms of flower development: an armchair guide. *Nat Rev Genet* **6**: 688–698.
- Krizek, B.A. and Meyerowitz, E.M.** (1996). The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. *Development* **122**: 11–22.
- Krogan, N.T., Hogan, K., and Long, J.A.** (2012). APETALA2 negatively regulates multiple floral organ identity genes in Arabidopsis by recruiting the co-repressor TOPLESS and the histone deacetylase HDA19. *Development* **139**: 4180–4190.

- Kuroha, T., Tokunaga, H., Kojima, M., Ueda, N., Ishida, T., Nagawa, S., Fukuda, H., Sugimoto, K., and Sakakibara, H.** (2009). Functional analyses of LONELY GUY cytokinin-activating enzymes reveal the importance of the direct activation pathway in Arabidopsis. *Plant Cell* **21**: 3152–3169.
- Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J.U., and Forner, J.** (2013). GreenGate---a novel, versatile, and efficient cloning system for plant transgenesis. *PLoS ONE* **8**: e83043.
- Landau, U., Asis, L., and Eshed Williams, L.** (2015). The ERECTA, CLAVATA and class III HD-ZIP Pathways Display Synergistic Interactions in Regulating Floral Meristem Activities. *PLoS ONE* **10**: e0125408.
- Landrein, B. et al.** (2015). Mechanical stress contributes to the expression of the STM homeobox gene in Arabidopsis shoot meristems. *elife* **4**: e07811.
- Laux, T., Mayer, K.F., Berger, J., and Jürgens, G.** (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development* **122**: 87–96.
- Lee, C. and Clark, S.E.** (2015). A WUSCHEL-Independent Stem Cell Specification Pathway Is Repressed by PHB, PHV and CNA in Arabidopsis. *PLoS ONE* **10**: e0126006.
- Leibfried, A., To, J.P., Busch, W., Stehling, S., Kehle, A., Demar, M., Kieber, J.J., and Lohmann, J.U.** (2005). WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* **438**: 1172–1175.
- Lenhard, M., Bohnert, A., Jürgens, G., and Laux, T.** (2001). Termination of stem cell maintenance in Arabidopsis floral meristems by interactions between WUSCHEL and AGAMOUS. *Cell* **105**: 805–814.
- Lenhard, M., Jürgens, G., and Laux, T.** (2002). The WUSCHEL and SHOOTMERISTEMLESS genes fulfil complementary roles in Arabidopsis shoot meristem regulation. *Development* **129**: 3195–3206.
- Lenhard, M. and Laux, T.** (2003). Stem cell homeostasis in the Arabidopsis shoot meristem is regulated by intercellular movement of CLAVATA3 and its sequestration by CLAVATA1. *Development* **130**: 3163–3173.
- Levin, J.Z. and Meyerowitz, E.M.** (1995). UFO: an Arabidopsis gene involved in both floral meristem and floral organ development. *Plant Cell* **7**: 529–548.
- Li, G., Liu, S., Wang, J., He, J., Huang, H., Zhang, Y., and Xu, L.** (2014). ISWI proteins participate in the genome-wide nucleosome distribution in Arabidopsis. *Plant J* **78**: 706–714.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup** (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- Liljegren, S.J.** (1999). Interactions among APETALA1, LEAFY, and TERMINAL FLOWER1 Specify Meristem Fate. *THE PLANT CELL ONLINE* **11**: 1007–1018.
- Liljegren, S.J., Gustafson-Brown, C., Pinyopich, A., Ditta, G.S., and Yanofsky, M.F.** (1999). Interactions among APETALA1, LEAFY, and TERMINAL FLOWER1 specify meristem fate. *Plant Cell* **11**: 1007–1018.
- Liu, Y., You, S., Taylor-Teeples, M., Li, W.L., Schuetz, M., Brady, S.M., and Douglas, C.J.** (2014). BEL1-LIKE HOMEODOMAIN6 and KNOTTED ARABIDOPSIS THALIANA7 interact and regulate secondary cell wall formation via repression of REVOLUTA. *Plant Cell* **26**: 4843–4861.
- Logemann, E., Birkenbihl, R.P., Ulker, B., and Somssich, I.E.** (2006). An improved method for preparing Agrobacterium cells that simplifies the Arabidopsis transformation protocol. *Plant Methods* **2**: 16.
- Lohmann, J.U., Hong, R.L., Hobe, M., Busch, M.A., Parcy, F., Simon, R., and Weigel, D.** (2001). A molecular link between stem cell regulation and floral patterning in Arabidopsis. *Cell* **105**: 793–803.

- Long, J. and Barton, M.K.** (2000). Initiation of axillary and floral meristems in Arabidopsis. *Dev Biol* **218**: 341–353.
- Long, J.A. and Barton, M.K.** (1998). The development of apical embryonic pattern in Arabidopsis. *Development* **125**: 3027–3035.
- Long, J.A., Moan, E.I., Medford, J.I., and Barton, M.K.** (1996). A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of Arabidopsis. *Nature* **379**: 66–69.
- Long, J.A., Ohno, C., Smith, Z.R., and Meyerowitz, E.M.** (2006). TOPLESS regulates apical embryonic fate in Arabidopsis. *Science* **312**: 1520–1523.
- Maizel, A., Busch, M.A., Tanahashi, T., Perkovic, J., Kato, M., Hasebe, M., and Weigel, D.** (2005). The floral regulator LEAFY evolves by substitutions in the DNA binding domain. *Science* **308**: 260–263.
- Mallory, A.C., Reinhart, B.J., Jones-Rhoades, M.W., Tang, G., Zamore, P.D., Barton, M.K., and Bartel, D.P.** (2004). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J* **23**: 3356–3364.
- Mandel, M.A., Gustafson-Brown, C., Savidge, B., and Yanofsky, M.F.** (1992). Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. *Nature* **360**: 273–277.
- Mandel, T., Candela, H., Landau, U., Asis, L., Zelinger, E., Carles, C.C., and Williams, L.E.** (2016). Differential regulation of meristem size, morphology and organization by the ERECTA, CLAVATA and class III HD-ZIP pathways. *Development* **143**: 1612–1622.
- Mashiguchi, K. et al.** (2011). The main auxin biosynthesis pathway in Arabidopsis. *Proc Natl Acad Sci U S A* **108**: 18512–18517.
- Mathelier, A. et al.** (2016). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **44**: D110–5.
- Mathelier, A. and Wasserman, W.W.** (2013). The next generation of transcription factor binding site prediction. *PLoS Comput Biol* **9**: e1003214.
- Mayer, K.F., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T.** (1998). Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell* **95**: 805–815.
- McCarthy, D.J., Chen, Y., and Smyth, G.K.** (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* **40**: 4288–4297.
- McLeay, R.C. and Bailey, T.L.** (2010). Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* **11**: 165.
- Mellerowicz, E.J., Horgan, K., Walden, A., Coker, A., and Walter, C.** (1998). PRFLL--a *Pinus radiata* homologue of FLORICAULA and LEAFY is expressed in buds containing vegetative shoot and undifferentiated male cone primordia. *Planta* **206**: 619–629.
- Melzer, R., Verelst, W., and Theissen, G.** (2009). The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in “floral quartet”-like complexes in vitro. *Nucleic Acids Res* **37**: 144–157.
- Mendes, M.A., Guerra, R.F., Berns, M.C., Manzo, C., Masiero, S., Finzi, L., Kater, M.M., and Colombo, L.** (2013). MADS domain transcription factors mediate short-range DNA looping that is essential for target gene expression in Arabidopsis. *Plant Cell* **25**: 2560–2572.
- Meruelo, A.D. and Bowie, J.U.** (2009). Identifying polymer-forming SAM domains. *Proteins* **74**: 1–5.

- Michniewicz, M. et al.** (2007). Antagonistic regulation of PIN phosphorylation by PP2A and PINOID directs auxin flux. *Cell* **130**: 1044–1056.
- Minguet, E.G., Segard, S., Charavay, C., and Parcy, F.** (2015). MORPHEUS, a Webtool for Transcription Factor Binding Analysis Using Position Weight Matrices with Dependency. *PLoS ONE* **10**: e0135586.
- Mizukami, Y. and Ma, H.** (1992). Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. *Cell* **71**: 119–131.
- Molinero-Rosales, N., Jamilena, M., Zurita, S., Gómez, P., Capel, J., and Lozano, R.** (1999). FALSIFLORA, the tomato orthologue of FLORICAULA and LEAFY, controls flowering time and floral meristem identity. *Plant J* **20**: 685–693.
- Moyroud, E., Kusters, E., Monniaux, M., Koes, R., and Parcy, F.** (2010). LEAFY blossoms. *Trends Plant Sci* **15**: 346–352.
- Moyroud, E., Minguet, E.G., Ott, F., Yant, L., Posé, D., Monniaux, M., Blanchet, S., Bastien, O., Thévenon, 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, D., Schmitz, G., and Theres, K.** (2006). Blind homologous R2R3 Myb genes control the pattern of lateral meristem initiation in Arabidopsis. *Plant Cell* **18**: 586–597.
- Müller, R., Bleckmann, A., and Simon, R.** (2008). The receptor kinase CORYNE of Arabidopsis transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. *Plant Cell* **20**: 934–946.
- Nanao, M.H. et al.** (2014). Structural basis for oligomerization of auxin transcriptional regulators. *Nat Commun* **5**: 3617.
- Nole-Wilson, S. and Krizek, B.A.** (2000). DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Res* **28**: 4076–4082.
- Norberg, M., Holmlund, M., and Nilsson, O.** (2005). The BLADE ON PETIOLE genes act redundantly to control the growth and development of lateral organs. *Development* **132**: 2203–2213.
- O'Malley, R.C., Huang, S.S., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R.** (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**: 1280–1292.
- Ó'Maoiléidigh, D.S., Graciet, E., and Wellmer, F.** (2014). Gene networks controlling Arabidopsis thaliana flower development. *New Phytol* **201**: 16–30.
- Ogas, J., Cheng, J.C., Sung, Z.R., and Somerville, C.** (1997). Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. *Science* **277**: 91–94.
- Okada, K., Ueda, J., Komaki, M.K., Bell, C.J., and Shimura, Y.** (1991). Requirement of the Auxin Polar Transport System in Early Stages of Arabidopsis Floral Bud Formation. *Plant Cell* **3**: 677–684.
- Okushima, Y. et al.** (2005). Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in Arabidopsis thaliana: unique and overlapping functions of ARF7 and ARF19. *Plant Cell* **17**: 444–463.
- Orenstein, Y. and Shamir, R.** (2014). A comparative analysis of transcription factor binding models learned from PBM, HT-SELEX and ChIP data. *Nucleic Acids Res* **42**: e63.
- Otsuga, D., DeGuzman, B., Prigge, M.J., Drews, G.N., and Clark, S.E.** (2001). REVOLUTA regulates meristem initiation at lateral positions. *Plant J* **25**: 223–236.
- Overvoorde, P.J. et al.** (2005). Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in Arabidopsis thaliana. *Plant Cell* **17**: 3282–3300.
- Pajoro, A. et al.** (2014a). Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome Biol* **15**: R41.

- Pajoro, A., Biewers, S., Dougali, E., Leal Valentim, F., Mendes, M.A., Porri, A., Coupland, G., Van de Peer, Y., van Dijk, A.D., Colombo, L., Davies, B., and Angenent, G.C.** (2014b). The (r)evolution of gene regulatory networks controlling Arabidopsis plant reproduction: a two-decade history. *J Exp Bot* **65**: 4731–4745.
- Parcy, F., Bomblies, K., and Weigel, D.** (2002). Interaction of LEAFY, AGAMOUS and TERMINAL FLOWER1 in maintaining floral meristem identity in Arabidopsis. *Development* **129**: 2519–2527.
- Parcy, F., Nilsson, O., Busch, M.A., Lee, I., and Weigel, D.** (1998). A genetic framework for floral patterning. *Nature* **395**: 561–566.
- Peaucelle, A., Morin, H., Traas, J., and Laufs, P.** (2007). Plants expressing a miR164-resistant CUC2 gene reveal the importance of post-meristematic maintenance of phyllotaxy in Arabidopsis. *Development* **134**: 1045–1050.
- Pelaz, S., Ditta, G.S., Baumann, E., Wisman, E., and Yanofsky, M.F.** (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* **405**: 200–203.
- Pfaffl, M.W.** (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**: e45.
- Pogany, J.A., Simon, E.J., Katzman, R.B., De Guzman, B.M., Yu, L.P., Trotochaud, A.E., and Clark, S.E.** (1998). Identifying novel regulators of shoot meristem development. *J Plant Res* **111**: 307–313.
- Ponting, C.P. and Aravind, L.** (1999). START: a lipid-binding domain in StAR, HD-ZIP and signalling proteins. *Trends Biochem Sci* **24**: 130–132.
- Porri, A., Torti, S., Romera-Branchat, M., and Coupland, G.** (2012). Spatially distinct regulatory roles for gibberellins in the promotion of flowering of Arabidopsis under long photoperiods. *Development* **139**: 2198–2209.
- Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N., and Clark, S.E.** (2005). Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. *Plant Cell* **17**: 61–76.
- Przemeck, G.K., Mattsson, J., Hardtke, C.S., Sung, Z.R., and Berleth, T.** (1996). Studies on the role of the Arabidopsis gene MONOPTEROS in vascular development and plant cell axialization. *Planta* **200**: 229–237.
- Qiao, F. and Bowie, J.U.** (2005). The many faces of SAM. *Sci STKE* **2005**: re7.
- Quinlan, A.R. and Hall, I.M.** (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.
- Raman, S., Greb, T., Peaucelle, A., Blein, T., Laufs, P., and Theres, K.** (2008). Interplay of miR164, CUP-SHAPED COTYLEDON genes and LATERAL SUPPRESSOR controls axillary meristem formation in Arabidopsis thaliana. *Plant J* **55**: 65–76.
- Ratcliffe, O.J., Bradley, D.J., and Coen, E.S.** (1999). Separation of shoot and floral identity in Arabidopsis. *Development* **126**: 1109–1120.
- Reinhardt, D., Mandel, T., and Kuhlemeier, C.** (2000). Auxin regulates the initiation and radial position of plant lateral organs. *Plant Cell* **12**: 507–518.
- Reinhardt, D., Pesce, E.R., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., Traas, J., Friml, J., and Kuhlemeier, C.** (2003). Regulation of phyllotaxis by polar auxin transport. *Nature* **426**: 255–260.
- Reinhart, B.J., Liu, T., Newell, N.R., Magnani, E., Huang, T., Kerstetter, R., Michaels, S., and Barton, M.K.** (2013). Establishing a framework for the Ad/abaxial regulatory network of Arabidopsis: ascertaining targets of class III homeodomain leucine zipper and KANADI regulation. *Plant Cell* **25**: 3228–3249.
- De Reuille, P.B., Bohn-Courseau, I., Ljung, K., Morin, H., Carraro, N., Godin, C., and Traas, J.** (2006). Computer simulations reveal properties of the cell-cell signaling network at the shoot apex in Arabidopsis. *Proc Natl Acad Sci U S A* **103**: 1627–1632.

- Riou-Khamlichi, C., Huntley, R., Jacquard, A., and Murray, J.A.** (1999). Cytokinin activation of Arabidopsis cell division through a D-type cyclin. *Science* **283**: 1541–1544.
- Risseuw, E., Venglat, P., Xiang, D., Komendant, K., Daskalchuk, T., Babic, V., Crosby, W., and Datla, R.** (2013). An activated form of UFO alters leaf development and produces ectopic floral and inflorescence meristems. *PLoS ONE* **8**: e83807.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K.** (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.
- Robinson, M.D. and Smyth, G.K.** (2007). Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* **23**: 2881–2887.
- Roider, H.G., Kanhere, A., Manke, T., and Vingron, M.** (2007). Predicting transcription factor affinities to DNA from a biophysical model. *Bioinformatics* **23**: 134–141.
- Sablowski, R.** (2015). Control of patterning, growth, and differentiation by floral organ identity genes. *J Exp Bot* **66**: 1065–1073.
- Saleh, A., Alvarez-Venegas, R., and Avramova, Z.** (2008). An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. *Nat Protoc* **3**: 1018–1025.
- Sassi, M. and Vernoux, T.** (2013). Auxin and self-organization at the shoot apical meristem. *J Exp Bot* **64**: 2579–2592.
- Sawa, S., Ito, T., Shimura, Y., and Okada, K.** (1999). FILAMENTOUS FLOWER controls the formation and development of arabidopsis inflorescences and floral meristems. *Plant Cell* **11**: 69–86.
- Sayou, C. et al.** (2014). A promiscuous intermediate underlies the evolution of LEAFY DNA binding specificity. *Science* **343**: 645–648.
- Sayou, C. et al.** (2016). A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nat Commun* **7**: 11222.
- Schmitz, G., Tillmann, E., Carriero, F., Fiore, C., Cellini, F., and Theres, K.** (2002). The tomato Blind gene encodes a MYB transcription factor that controls the formation of lateral meristems. *Proc Natl Acad Sci U S A* **99**: 1064–1069.
- Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F., Jürgens, G., and Laux, T.** (2000). The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* **100**: 635–644.
- Schultz, E.A. and Haughn, G.W.** (1991). LEAFY, a Homeotic Gene That Regulates Inflorescence Development in Arabidopsis. *Plant Cell* **3**: 771–781.
- Schuster, C., Gailloch, C., Medzihradzky, A., Busch, W., Daum, G., Krebs, M., Kehle, A., and Lohmann, J.U.** (2014). A regulatory framework for shoot stem cell control integrating metabolic, transcriptional, and phytohormone signals. *Dev Cell* **28**: 438–449.
- Shannon, S. and Meeks-Wagner, D.R.** (1991). A Mutation in the Arabidopsis TFL1 Gene Affects Inflorescence Meristem Development. *Plant Cell* **3**: 877–892.
- Sharon, E., Lubliner, S., and Segal, E.** (2008). A feature-based approach to modeling protein-DNA interactions. *PLoS Comput Biol* **4**: e1000154.
- Shi, B. et al.** (2016). Two-Step Regulation of a Meristematic Cell Population Acting in Shoot Branching in Arabidopsis. *PLoS Genet* **12**: e1006168.
- Siriwardana, N.S. and Lamb, R.S.** (2012a). A conserved domain in the N-terminus is important for LEAFY dimerization and function in Arabidopsis thaliana. *Plant J* **71**: 736–749.
- Siriwardana, N.S. and Lamb, R.S.** (2012b). The poetry of reproduction: the role of LEAFY in Arabidopsis thaliana flower formation. *Int J Dev Biol* **56**: 207–221.

- Smith, R.S., Guyomarc'h, S., Mandel, T., Reinhardt, D., Kuhlemeier, C., and Prusinkiewicz, P.** (2006). A plausible model of phyllotaxis. *Proc Natl Acad Sci U S A* **103**: 1301–1306.
- Smith, Z.R. and Long, J.A.** (2010). Control of Arabidopsis apical-basal embryo polarity by antagonistic transcription factors. *Nature* **464**: 423–426.
- Somssich, M., Ma, Q., Weidtkamp-Peters, S., Stahl, Y., Felekyan, S., Bleckmann, A., Seidel, C.A., and Simon, R.** (2015). Real-time dynamics of peptide ligand-dependent receptor complex formation in planta. *Sci Signal* **8**: ra76.
- Souer, E., van der Krol, A., Kloos, D., Spelt, C., Blied, M., Mol, J., and Koes, R.** (1998). Genetic control of branching pattern and floral identity during *Petunia* inflorescence development. *Development* **125**: 733–742.
- Staden, R.** (1984). Computer methods to locate signals in nucleic acid sequences. *Nucleic Acids Res* **12**: 505–519.
- Stewart, D., Graciet, E., and Wellmer, F.** (2016). Molecular and regulatory mechanisms controlling floral organ development. *FEBS J* **283**: 1823–1830.
- Stieger, P.A., Reinhardt, D., and Kuhlemeier, C.** (2002). The auxin influx carrier is essential for correct leaf positioning. *Plant J* **32**: 509–517.
- Stoma, S., Lucas, M., Chopard, J., Schaedel, M., Traas, J., and Godin, C.** (2008). Flux-based transport enhancement as a plausible unifying mechanism for auxin transport in meristem development. *PLoS Comput Biol* **4**: e1000207.
- Stormo, G.D.** (2000). DNA binding sites: representation and discovery. *Bioinformatics* **16**: 16–23.
- Stormo, G.D.** (2013). Modeling the specificity of protein-DNA interactions. *Cold Spring Harb Symp Quant Biol* **1**: 115–130.
- Stormo, G.D. and Fields, D.S.** (1998). Specificity, free energy and information content in protein-DNA interactions. *Trends Biochem Sci* **23**: 109–113.
- Stormo, G.D. and Zhao, Y.** (2010). Determining the specificity of protein-DNA interactions. *Nat Rev Genet* **11**: 751–760.
- Sun, B., Looi, L.S., Guo, S., He, Z., Gan, E.S., Huang, J., Xu, Y., Wee, W.Y., and Ito, T.** (2014). Timing mechanism dependent on cell division is invoked by Polycomb eviction in plant stem cells. *Science* **343**: 1248559.
- Sun, B., Xu, Y., Ng, K.H., and Ito, T.** (2009). A timing mechanism for stem cell maintenance and differentiation in the Arabidopsis floral meristem. *Genes Dev* **23**: 1791–1804.
- Szemenyei, H., Hannon, M., and Long, J.A.** (2008). TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. *Science* **319**: 1384–1386.
- Talbert, P.B., Adler, H.T., Parks, D.W., and Comai, L.** (1995). The REVOLUTA gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of Arabidopsis thaliana. *Development* **121**: 2723–2735.
- Tanahashi, T., Sumikawa, N., Kato, M., and Hasebe, M.** (2005). Diversification of gene function: homologs of the floral regulator FLO/LFY control the first zygotic cell division in the moss *Physcomitrella patens*. *Development* **132**: 1727–1736.
- Taylor-Teeples, M., Lanctot, A., and Nemhauser, J.L.** (2016). As above, so below: Auxin's role in lateral organ development. *Dev Biol*.
- Theissen, G.** (2001). Development of floral organ identity: stories from the MADS house. *Curr Opin Plant Biol* **4**: 75–85.
- Tiwari, S.B., Wang, X.J., Hagen, G., and Guilfoyle, T.J.** (2001). AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. *Plant Cell* **13**: 2809–2822.

- Ulmasov, T., Murfett, J., Hagen, G., and Guilfoyle, T.J.** (1997). Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9**: 1963–1971.
- Vázquez-Lobo, A., Carlsbecker, A., Vergara-Silva, F., Alvarez-Buylla, E.R., Piñero, D., and Engström, P.** (2007). Characterization of the expression patterns of LEAFY/FLORICAULA and NEEDLY orthologs in female and male cones of the conifer genera *Picea*, *Podocarpus*, and *Taxus*: implications for current evo-devo hypotheses for gymnosperms. *Evol Dev* **9**: 446–459.
- Vernoux, T., Besnard, F., and Traas, J.** (2010). Auxin at the shoot apical meristem. *Cold Spring Harb Perspect Biol* **2**: a001487.
- Voss, T.C., Schiltz, R.L., Sung, M.H., Yen, P.M., Stamatoyannopoulos, J.A., Biddie, S.C., Johnson, T.A., Miranda, T.B., John, S., and Hager, G.L.** (2011). Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. *Cell* **146**: 544–554.
- Wagner, D., Sablowski, R.W., and Meyerowitz, E.M.** (1999). Transcriptional activation of APETALA1 by LEAFY. *Science* **285**: 582–584.
- Wang, Q., Kohlen, W., Rossmann, S., Vernoux, T., and Theres, K.** (2014a). Auxin Depletion from the Leaf Axil Conditions Competence for Axillary Meristem Formation in *Arabidopsis* and *Tomato*. *Plant Cell* **26**: 2068–2079.
- Wang, Y., Wang, J., Shi, B., Yu, T., Qi, J., Meyerowitz, E.M., and Jiao, Y.** (2014b). The Stem Cell Niche in Leaf Axils Is Established by Auxin and Cytokinin in *Arabidopsis*. *Plant Cell* **26**: 2055–2067.
- Wasserman, W.W. and Sandelin, A.** (2004). Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet* **5**: 276–287.
- Webber, J.L., Zhang, J., Cote, L., Vivekanand, P., Ni, X., Zhou, J., Nègre, N., Carthew, R.W., White, K.P., and Rebay, I.** (2013). The relationship between long-range chromatin occupancy and polymerization of the *Drosophila* ETS family transcriptional repressor Yan. *Genetics* **193**: 633–649.
- Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F., and Meyerowitz, E.M.** (1992). LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* **69**: 843–859.
- Weigel, D. and Nilsson, O.** (1995). A developmental switch sufficient for flower initiation in diverse plants. *Nature* **377**: 495–500.
- Weijers, D., Benkova, E., Jäger, K.E., Schlereth, A., Hamann, T., Kientz, M., Wilmoth, J.C., Reed, J.W., and Jürgens, G.** (2005). Developmental specificity of auxin response by pairs of ARF and Aux/IAA transcriptional regulators. *EMBO J* **24**: 1874–1885.
- Weirauch, M.T. et al.** (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**: 1431–1443.
- Whittington, T., Perkins, A.C., and Bailey, T.L.** (2009). High-throughput chromatin information enables accurate tissue-specific prediction of transcription factor binding sites. *Nucleic Acids Res* **37**: 14–25.
- Williams, L., Grigg, S.P., Xie, M., Christensen, S., and Fletcher, J.C.** (2005). Regulation of *Arabidopsis* shoot apical meristem and lateral organ formation by microRNA miR166g and its AtHD-ZIP target genes. *Development* **132**: 3657–3668.
- Winter, C.M. et al.** (2011). LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Dev Cell* **20**: 430–443.
- Winter, C.M., Yamaguchi, N., Wu, M.F., and Wagner, D.** (2015). Transcriptional programs regulated by both LEAFY and APETALA1 at the time of flower formation. *Physiol Plant* **155**: 55–73.
- Won, C., Shen, X., Mashiguchi, K., Zheng, Z., Dai, X., Cheng, Y., Kasahara, H., Kamiya, Y., Chory, J., and Zhao, Y.** (2011). Conversion of tryptophan to indole-3-acetic acid

- by TRYPTOPHAN AMINOTRANSFERASES OF ARABIDOPSIS and YUCCAs in Arabidopsis. *Proc Natl Acad Sci U S A* **108**: 18518–18523.
- Won, K.J., Ren, B., and Wang, W.** (2010). Genome-wide prediction of transcription factor binding sites using an integrated model. *Genome Biol* **11**: R7.
- Wu, M.F., Sang, Y., Bezhani, S., Yamaguchi, N., Han, S.K., Li, Z., Su, Y., Slewinski, T.L., and Wagner, D.** (2012). SWI2/SNF2 chromatin remodeling ATPases overcome polycomb repression and control floral organ identity with the LEAFY and SEPALLATA3 transcription factors. *Proc Natl Acad Sci U S A* **109**: 3576–3581.
- Wu, M.F., Yamaguchi, N., Xiao, J., Bargmann, B., Estelle, M., Sang, Y., and Wagner, D.** (2015). Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate. *elife* **4**: e09269.
- Yadav, R.K., Girke, T., Pasala, S., Xie, M., and Reddy, G.V.** (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. *Proc Natl Acad Sci U S A* **106**: 4941–4946.
- Yadav, R.K., Perales, M., Gruel, J., Ohno, C., Heisler, M., Girke, T., Jönsson, H., and Reddy, G.V.** (2013). Plant stem cell maintenance involves direct transcriptional repression of differentiation program. *Mol Syst Biol* **9**: 654.
- Yadav, R.K. and Reddy, G.V.** (2012). WUSCHEL protein movement and stem cell homeostasis. *Plant Signal Behav* **7**: 592–594.
- Yamada, T., Sasaki, Y., Hashimoto, K., Nakajima, K., and Gasser, C.S.** (2015). CORONA, PHABULOSA and PHAVOLUTA collaborate with BELL 1 to confine WUSCHEL expression to the nucellus in Arabidopsis ovules. *Development*.
- Yamaguchi, N., Winter, C.M., Wu, M.F., Kanno, Y., Yamaguchi, A., Seo, M., and Wagner, D.** (2014a). Gibberellin acts positively then negatively to control onset of flower formation in Arabidopsis. *Science* **344**: 638–641.
- Yamaguchi, N., Winter, C.M., Wu, M.F., Kwon, C.S., William, D.A., and Wagner, D.** (2014b). PROTOCOLS: Chromatin Immunoprecipitation from Arabidopsis Tissues. *Arabidopsis Book* **12**: e0170.
- Yamaguchi, N., Wu, M.F., Winter, C.M., Berns, M.C., Nole-Wilson, S., Yamaguchi, A., Coupland, G., Krizek, B.A., and Wagner, D.** (2013). A molecular framework for auxin-mediated initiation of flower primordia. *Dev Cell* **24**: 271–282.
- Yamaguchi, N., Yamaguchi, A., Abe, M., Wagner, D., and Komeda, Y.** (2012). LEAFY controls Arabidopsis pedicel length and orientation by affecting adaxial-abaxial cell fate. *Plant J* **69**: 844–856.
- Yan, W., Chen, D., and Kaufmann, K.** (2016). Efficient multiplex mutagenesis by RNA-guided Cas9 and its use in the characterization of regulatory elements in the AGAMOUS gene. *Plant Methods* **12**: 23.
- Yanai, O., Shani, E., Dolezal, K., Tarkowski, P., Sablowski, R., Sandberg, G., Samach, A., and Ori, N.** (2005). Arabidopsis KNOX1 proteins activate cytokinin biosynthesis. *Curr Biol* **15**: 1566–1571.
- Yang, F., Wang, Q., Schmitz, G., Müller, D., and Theres, K.** (2012). The bHLH protein ROX acts in concert with RAX1 and LAS to modulate axillary meristem formation in Arabidopsis. *Plant J* **71**: 61–70.
- Yang, L., Zhou, T., Dror, I., Mathelier, A., Wasserman, W.W., Gordân, R., and Rohs, R.** (2014). TFBSshape: a motif database for DNA shape features of transcription factor binding sites. *Nucleic Acids Res* **42**: D148–55.
- Yant, L., Mathieu, J., Dinh, T.T., Ott, F., Lanz, C., Wollmann, H., Chen, X., and Schmid, M.** (2010). Orchestration of the floral transition and floral development in Arabidopsis by the bifunctional transcription factor APETALA2. *Plant Cell* **22**: 2156–2170.
- Yoshida, S., Mandel, T., and Kuhlemeier, C.** (2011). Stem cell activation by light guides plant organogenesis. *Genes Dev* **25**: 1439–1450.

- Yu, H., Ito, T., Wellmer, F., and Meyerowitz, E.M.** (2004). Repression of AGAMOUS-LIKE 24 is a crucial step in promoting flower development. *Nat Genet* **36**: 157–161.
- Yu, Y.T., Wu, Z., Lu, K., Bi, C., Liang, S., Wang, X.F., and Zhang, D.P.** (2016). Overexpression of the MYB37 transcription factor enhances abscisic acid sensitivity, and improves both drought tolerance and seed productivity in *Arabidopsis thaliana*. *Plant Mol Biol* **90**: 267–279.
- Zhang, J., Graham, T.G., Vivekanand, P., Cote, L., Cetera, M., and Rebay, I.** (2010). Sterile alpha motif domain-mediated self-association plays an essential role in modulating the activity of the *Drosophila* ETS family transcriptional repressor Yan. *Mol Cell Biol* **30**: 1158–1170.
- Zhang, W., Zhang, T., Wu, Y., and Jiang, J.** (2012). Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. *Plant Cell* **24**: 2719–2731.
- Zhang, Z., Wang, H., Luo, D., Zeng, M., Huang, H., and Cui, X.** (2011). Convergence of the 26S proteasome and the REVOLUTA pathways in regulating inflorescence and floral meristem functions in *Arabidopsis*. *J Exp Bot* **62**: 359–369.
- Zhao, Y., Christensen, S.K., Fankhauser, C., Cashman, J.R., Cohen, J.D., Weigel, D., and Chory, J.** (2001). A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science* **291**: 306–309.
- Zhao, Y., Ruan, S., Pandey, M., and Stormo, G.D.** (2012). Improved models for transcription factor binding site identification using nonindependent interactions. *Genetics* **191**: 781–790.
- Zhao, Z., Andersen, S.U., Ljung, K., Dolezal, K., Miotk, A., Schultheiss, S.J., and Lohmann, J.U.** (2010). Hormonal control of the shoot stem-cell niche. *Nature* **465**: 1089–1092.
- Zhong, R. and Ye, Z.H.** (2001). Alteration of auxin polar transport in the *Arabidopsis* *ifl1* mutants. *Plant Physiol* **126**: 549–563.
- Zhong, R. and Ye, Z.H.** (1999). IFL1, a gene regulating interfascicular fiber differentiation in *Arabidopsis*, encodes a homeodomain-leucine zipper protein. *Plant Cell* **11**: 2139–2152.
- Zhou, X., Lindsay, H., and Robinson, M.D.** (2014). Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic Acids Res* **42**: e91.
- Zhou, Y., Honda, M., Zhu, H., Zhang, Z., Guo, X., Li, T., Li, Z., Peng, X., Nakajima, K., Duan, L., and Zhang, X.** (2015). Spatiotemporal sequestration of miR165/166 by *Arabidopsis* Argonaute10 promotes shoot apical meristem maintenance. *Cell Rep* **10**: 1819–1827.
- Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S.H., Liou, L.W., Barefoot, A., Dickman, M., and Zhang, X.** (2011). *Arabidopsis* Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. *Cell* **145**: 242–256.
- Zhu, X. et al.** (2014). An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. *Sci Rep* **4**: 6420.







# ANNEXES



## JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles

Anthony Mathelier<sup>1</sup>, Oriol Fornes<sup>1</sup>, David J. Arenillas<sup>1</sup>, Chih-yu Chen<sup>1</sup>, Grégoire Denay<sup>2</sup>, Jessica Lee<sup>1</sup>, Wenqiang Shi<sup>1</sup>, Casper Shyr<sup>1</sup>, Ge Tan<sup>3</sup>, Rebecca Worsley-Hunt<sup>1</sup>, Allen W. Zhang<sup>1</sup>, François Parcy<sup>2</sup>, Boris Lenhard<sup>3,\*</sup>, Albin Sandelin<sup>4,\*</sup> and Wyeth W. Wasserman<sup>1,\*</sup>

<sup>1</sup>Centre for Molecular Medicine and Therapeutics at the Child and Family Research Institute, Department of Medical Genetics, University of British Columbia, Vancouver, V5Z 4H4, BC, Canada, <sup>2</sup>Laboratoire Physiologie Cellulaire & Végétale, Université Grenoble Alpes, CNRS, CEA, IRTSV, INRA, 38054 Grenoble, France, <sup>3</sup>Computational Regulatory Genomics, MRC Clinical Sciences Centre, Imperial College London, Du Cane Road, London W12 0NN, UK and <sup>4</sup>The Bioinformatics Centre, Department of Biology and Biotech Research and Innovation Centre, Copenhagen University, Ole Maaloes Vej 5, DK-2200, Denmark

Received September 12, 2015; Revised October 19, 2015; Accepted October 22, 2015

### ABSTRACT

**JASPAR (<http://jaspar.genereg.net>) is an open-access database storing curated, non-redundant transcription factor (TF) binding profiles representing transcription factor binding preferences as position frequency matrices for multiple species in six taxonomic groups. For this 2016 release, we expanded the JASPAR CORE collection with 494 new TF binding profiles (315 in vertebrates, 11 in nematodes, 3 in insects, 1 in fungi and 164 in plants) and updated 59 profiles (58 in vertebrates and 1 in fungi). The introduced profiles represent an 83% expansion and 10% update when compared to the previous release. We updated the structural annotation of the TF DNA binding domains (DBDs) following a published hierarchical structural classification. In addition, we introduced 130 transcription factor flexible models trained on ChIP-seq data for vertebrates, which capture dinucleotide dependencies within TF binding sites. This new JASPAR release is accompanied by a new web tool to infer JASPAR TF binding profiles recognized by a given TF protein sequence. Moreover, we provide the users with a Ruby module complementing the JASPAR API to ease programmatic access and use of the JASPAR collection of profiles. Finally, we provide the JASPAR2016 R/Bioconductor data package with the data of this release.**

### INTRODUCTION

A key subset of transcription factors (TFs) are involved in the regulation of gene expression at the transcriptional level by binding to DNA regulatory elements. These DNA binding TFs (hereafter referred to only as TFs) can be further divided into classes based on their DNA binding domains (DBDs). Deciphering the DNA sequences bound by TFs is critical for elucidating transcriptional regulation of gene expression, and has been a key focus of large-scale genomics research. Describing the sequence-specific binding preferences of TFs has matured through generations, with the first generation methods consisting of simple consensus sequences. Second generation methods, which remain dominant, quantitatively describe binding preferences with position frequency matrices (PFMs). A PFM is derived from DNA sequences experimentally observed to be bound by a specific TF. The heart of the JASPAR database, the CORE collection, provides non-redundant and manually curated TF binding profiles described as PFMs and associated to TFs from species in six taxonomic groups (vertebrates, nematodes, insects, fungi, urochordates and plants).

Position weight matrices (PWMs, also known as position-specific scoring matrices) are derived from PFMs to predict TF binding sites (TFBSs) within a DNA sequence (see (1) for a review). These matrices represent an additive probabilistic model assuming independence between the TFBS nucleotides.

A third generation of binding models, such as the transcription factor flexible models (TFFMs) (2), have been

\*To whom correspondence should be addressed. Tel: +1 604 875 3812; Fax: +1 604 875 3819; Email: wyeth@cmmt.ubc.ca  
Correspondence may also be addressed to Boris Lenhard. Tel: +44 208 383 8353; Fax: +44 208 383 8577; Email: b.lenhard@csc.mrc.ac.uk  
Correspondence may also be addressed to Albin Sandelin. Tel: +45 353 21285; Fax: +45 3532 5669; Email: albin@binf.ku.dk  
Present address: Rebecca Worsley-Hunt, Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany.

© The Author(s) 2015. Published by Oxford University Press on behalf of Nucleic Acids Research.  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

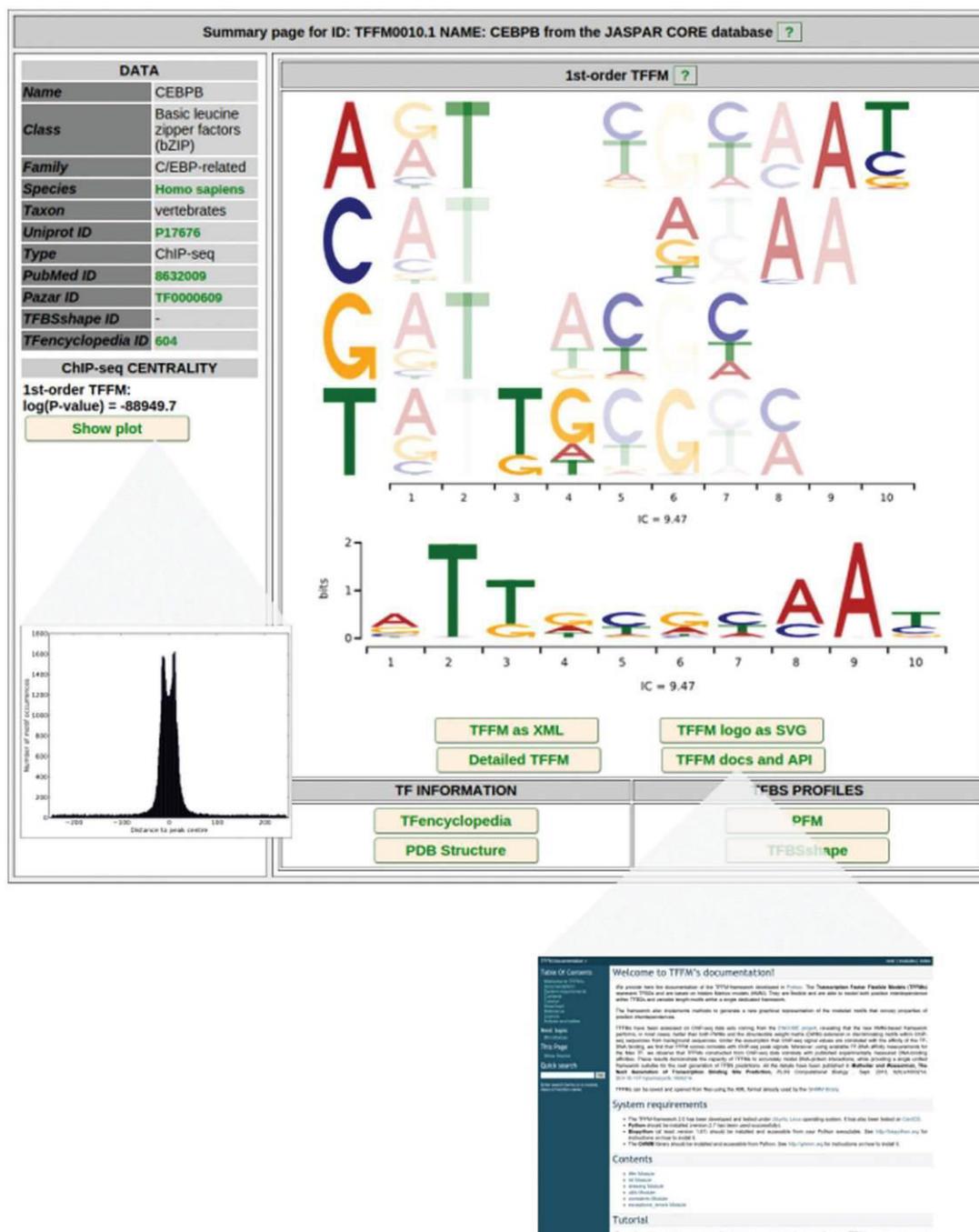
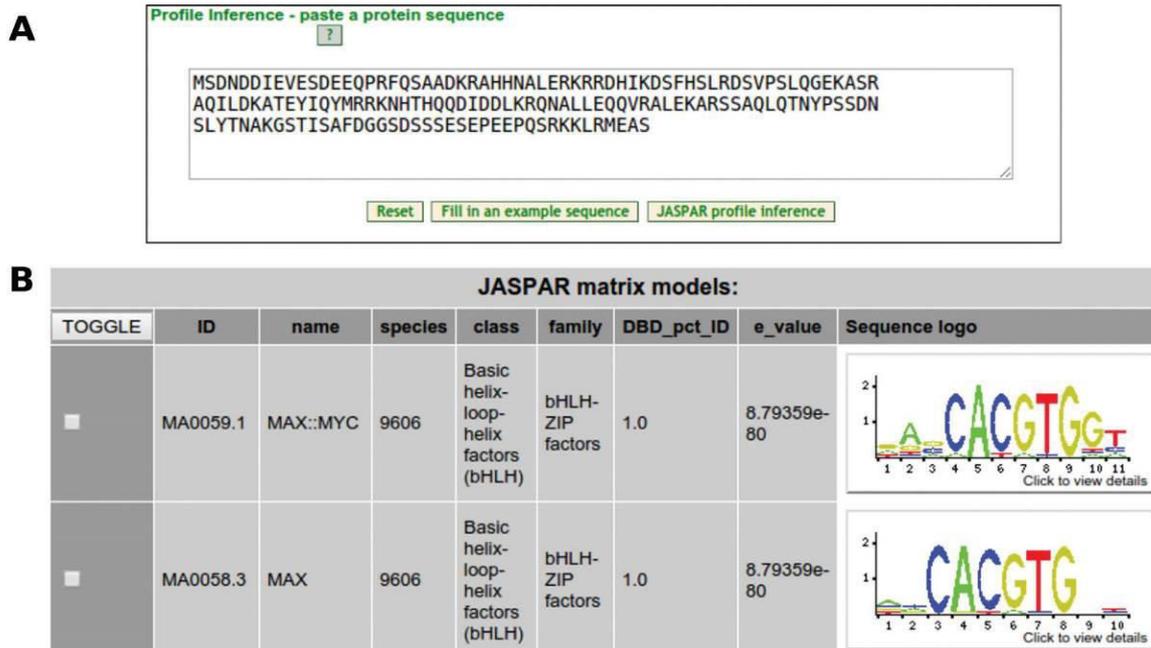


Figure 1. Screenshot of a new TFFM introduced in the new dedicated layout. Refer to (2) for details on TFFMs and interpretation of the dedicated logos describing the dinucleotide dependencies captured by the models.

**Table 1.** Overview of the content growth in JASPAR 2016 (version 6.0) compared to JASPAR 2014 (version 5.0)

Taxonomic group	Number of non-redundant profiles in JASPAR 5.0	New non-redundant profiles in JASPAR 6.0	Updated profiles	Removed profiles	Total profiles (including older versions of profiles)	Total profiles (non-redundant)
Vertebrates	205	315	58	1	635	519
Plants	64	164	0	0	231	227
Insects	131	3	0	0	139	133
Nematodes	15	11	0	0	26	26
Fungi	177	1	1	2	177	176
Urochordata	1	0	0	0	1	1
Total	593	494	59	3	1210	1082

See Supplementary Text for more information.



**Figure 2.** Overview of the JASPAR TF binding profile inference. (A) The user can input a TF protein sequence for which to look for a JASPAR TF binding profile. (B) Binding profiles potentially bound by the TF provided by the user are inferred.

its web interface (<http://jaspar.genereg.net>) or its previous API implemented in several programming languages (R, Perl and Python) (13). Users can refer to the dedicated tutorials and webinar describing how to use these modules (<http://biopython.org/DIST/docs/tutorial/Tutorial.html#htoc172>, [http://www.cisreg.ca/Webinars/JASPAR\\_BioPython\\_MANTA.flv](http://www.cisreg.ca/Webinars/JASPAR_BioPython_MANTA.flv), <http://tfbs.genereg.net/>, <http://bioconductor.org/packages/TFBSTools/>). The current release of JASPAR is accompanied by a new Ruby module (also known as a Ruby gem), based on the BioRuby open-source bioinformatics library (28), at <https://github.com/wassermanlab/jaspar-bioruby>, enabling Ruby users to retrieve the TF binding profiles stored in the database and use them for predicting TFBSs within DNA sequences. It has been implemented to replicate the

functionality of the BioPython module introduced in the 2014 release of JASPAR (13).

**Inferring a JASPAR TF binding profile recognized by a DNA binding domain**

Despite the large expansion of the JASPAR CORE collection, which collects more than 1000 profiles for TFs from six taxonomic groups, the data required for the generation of profiles for many TFs are not yet available. JASPAR users recurrently ask for the most appropriate TF binding profile to use given a TF not present in the database. Recent work has used DBD sequence similarities to infer DNA sequence binding preference (14). Following a similar approach, we provide users with potential profiles to use given a query TF protein sequence (Supplementary Text and Supplementary Figure S1). In brief, the TF binding profile inference fea-

ture compares the DBD sequence of the given TF to those of homologous TFs stored in JASPAR, and infers the TF binding profiles from the best compared JASPAR homologous TFs as potentially recognized by the user's input protein sequence wherever possible (Figure 2).

## CONCLUSIONS AND PERSPECTIVE

The 2016 release of JASPAR maintains the long-term focus on providing high-quality, non-redundant TF binding profiles for the global research community. Consistent with past releases, we have (i) expanded the widely used JASPAR CORE collection, adding 494 profiles; (ii) enhanced usability, incorporating the TFClass structural classification and introducing an associated capacity to select profiles for not yet characterized TFs; (iii) expanded and updated programming tools, highlighted by a Ruby gem for JASPAR access and (iv) introduced a new collection, for the first time incorporating third generation binding profiles.

Looking forward, the introduction of third generation methods may mark a significant transition for JASPAR. As TF binding data continues to expand, and we gain greater insight into each TF, advanced models that address specific TFs or TF-families may become the norm. Determining how best to unite what may be computationally diverse third generation models into a simple-to-use and easy-to-access system will become a focus. Our JASPAR development team looks forward to working with the bioinformatics community as TFBS prediction evolves.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors wish to thank the user community for useful input. We thank Matthew T. Weirauch and Mihai Albu for sharing the TF profile inference code implemented in the CIS-BP website, and Roberto Solano and José Manuel Franco-Zorrilla for providing their PBM data. We thank Miroslav Hatas for systems support and Dora Pak for management support to the WWW lab.

## FUNDING

Genome Canada Large Scaled Applied Research Grant [174CDE] (to WWW lab); Canadian Institute of Health Research (CIHR) Operating Grant [MOP-119586] (to WWW lab); Child and Family Research Institute (CFRI) (to A.M.); British Columbia Children's Hospital Foundation (to A.M.); Postgraduate Scholarships-Doctoral Program from Natural Sciences and Engineering Research of Canada (NSERC); University of British Columbia (UBC) Four Year Doctoral Fellowship (to C.Y.C.); Genome Science And Technology program NSERC-CREATE scholarship; University of Zurich; CFRI Jan M. Friedman Graduate Studentship (to J.L.); China Scholarship Council (to W.S.); UBC Teaching and Learning Enhancement fund (to W.S., A.W.Z.); CIHR Graduate Scholarship [CGSD-GSM to C.S.]; NSERC Discovery Grant [RGPIN 355532-10 to

C.S., R.W.H.]; UBC MD-PhD program (to A.W.Z.); ANR [Blanc-SVSE2-2011-Charmsful to F.P.]; French MRT PhD Fellowship (to G.D.); EU FP7 large scale integrated project ZF HEALTH [HEALTH-F4-2010-242048 to G.T.]; Medical Research Council UK (to B.L.). Funding for open access charge: Genome Canada Large Scaled Applied Research Grant [174CDE].

*Conflict of interest statement.* None declared.

## REFERENCES

- Stormo, G.D. (2013) Modeling the specificity of protein-DNA interactions. *Quant. Biol.*, **1**, 115–130.
- Mathelier, A. and Wasserman, W.W. (2013) The next generation of transcription factor binding site prediction. *PLoS Comput. Biol.*, **9**, e1003214.
- Luscombe, N.M., Laskowski, R.A. and Thornton, J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.*, **29**, 2860–2874.
- Man, T.K. and Stormo, G.D. (2001) Non-independence of Mnt repressor-operator interaction determined by a new quantitative multiple fluorescence relative affinity (QuMFRA) assay. *Nucleic Acids Res.*, **29**, 2471–2478.
- Bulyk, M.L., Johnson, P.L. and Church, G.M. (2002) Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors. *Nucleic Acids Res.*, **30**, 1255–1261.
- Tomovic, A. and Oakeley, E.J. (2007) Position dependencies in transcription factor binding sites. *Bioinformatics*, **23**, 933–941.
- Zhou, Q. and Liu, J.S. (2004) Modeling within-motif dependence for transcription factor binding site predictions. *Bioinformatics*, **20**, 909–916.
- Moyroud, E., Minguet, E.G., Ott, F., Yant, L., Pose, D., Monniaux, M., Blanchet, S., Bastien, O., Thevenon, E., Weigel, D. *et al.* (2011) Prediction of regulatory interactions from genome sequences using a biophysical model for the Arabidopsis LEAFY transcription factor. *Plant Cell*, **23**, 1293–1306.
- Sandelin, A., Alkema, W., Engström, P., Wasserman, W.W. and Lenhard, B. (2004) JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.*, **32**, D91–D94.
- Vlieghe, D., Sandelin, A., De Bleser, P.J., Vlemingck, K., Wasserman, W.W., van Roy, F. and Lenhard, B. (2006) A new generation of JASPAR, the open-access repository for transcription factor binding site profiles. *Nucleic Acids Res.*, **34**, D95–D97.
- Bryne, J.C., Valen, E., Tang, M.H., Marstrand, T., Winther, O., da Piedade, I., Krogh, A., Lenhard, B. and Sandelin, A. (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res.*, **36**, D102–D106.
- Portales-Casamar, E., Thongjuea, S., Kwon, A.T., Arenillas, D., Zhao, X., Valen, E., Yusuf, D., Lenhard, B., Wasserman, W.W. and Sandelin, A. (2010) JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res.*, **38**, D105–D110.
- Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C.Y., Chou, A., Ienasescu, H. *et al.* (2014) JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res.*, **42**, D142–D147.
- Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K. *et al.* (2014) Determination and inference of eukaryotic transcription factor sequence specificity. *Cell*, **158**, 1431–1443.
- Jolma, A., Yan, J., Whittington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G. *et al.* (2013) DNA-binding specificities of human transcription factors. *Cell*, **152**, 327–339.
- Franco-Zorrilla, J.M., Lopez-Vidriero, I., Carrasco, J.L., Godoy, M., Vera, P. and Solano, R. (2014) DNA-binding specificities of plant

- transcription factors and their potential to define target genes. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 2367–2372.
17. Badis, G., Berger, M.F., Philippakis, A.A., Talukder, S., Gehrke, A.R., Jaeger, S.A., Chan, E.T., Metzler, G., Vedenko, A., Chen, X. *et al.* (2009) Diversity and complexity in DNA recognition by transcription factors. *Science*, **324**, 1720–1723.
  18. Wingender, E., Schoeps, T., Haubrock, M. and Donitz, J. (2015) TFClass: a classification of human transcription factors and their rodent orthologs. *Nucleic Acids Res.*, **43**, D97–D102.
  19. Boer, D.R., Freire-Rios, A., van den Berg, W.A., Saaki, T., Manfield, I.W., Kepinski, S., Lopez-Vidriero, I., Franco-Zorrilla, J.M., de Vries, S.C., Solano, R. *et al.* (2014) Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors. *Cell*, **156**, 577–589.
  20. Fonseca, S., Fernandez-Calvo, P., Fernandez, G.M., Diez-Diaz, M., Gimenez-Ibanez, S., Lopez-Vidriero, I., Godoy, M., Fernandez-Barbero, G., Van Leene, J., De Jaeger, G. *et al.* (2014) bHLH003, bHLH013 and bHLH017 are new targets of JAZ repressors negatively regulating JA responses. *PLoS One*, **9**, e86182.
  21. Heyndrickx, K.S., Van de Velde, J., Wang, C., Weigel, D. and Vandepoele, K. (2014) A functional and evolutionary perspective on transcription factor binding in *Arabidopsis thaliana*. *Plant Cell*, **26**, 3894–3910.
  22. Weirauch, M.T., Cote, A., Norel, R., Annala, M., Zhao, Y., Riley, T.R., Saez-Rodriguez, J., Cokelaer, T., Vedenko, A., Talukder, S. *et al.* (2013) Evaluation of methods for modeling transcription factor sequence specificity. *Nat. Biotechnol.*, **31**, 126–134.
  23. Siddharthan, R. (2010) Dinucleotide weight matrices for predicting transcription factor binding sites: generalizing the position weight matrix. *PLoS One*, **5**, e9722.
  24. Zhao, Y., Ruan, S., Pandey, M. and Stormo, G.D. (2012) Improved models for transcription factor binding site identification using nonindependent interactions. *Genetics*, **191**, 781–790.
  25. Bailey, T.L. and Machanick, P. (2012) Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res.*, **40**, e128.
  26. Fulton, D.L., Sundararajan, S., Badis, G., Hughes, T.R., Wasserman, W.W., Roach, J.C. and Sladek, R. (2009) TFCat: the curated catalog of mouse and human transcription factors. *Genome Biol.*, **10**, R29.
  27. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.*, **5**, R80.
  28. Goto, N., Prins, P., Nakao, M., Bonnal, R., Aerts, J. and Katayama, T. (2010) BioRuby: bioinformatics software for the Ruby programming language. *Bioinformatics*, **26**, 2617–2619.





## ARTICLE

Received 18 May 2015 | Accepted 2 Mar 2016 | Published 21 Apr 2016

DOI: 10.1038/ncomms11222

OPEN

# A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor

Camille Sayou<sup>1,†</sup>, Max H. Nanao<sup>2</sup>, Marc Jamin<sup>3</sup>, David Posé<sup>4,5</sup>, Emmanuel Thévenon<sup>1</sup>, Laura Grégoire<sup>1</sup>, Gabrielle Tichtinsky<sup>1</sup>, Grégoire Denay<sup>1</sup>, Felix Ott<sup>4,†</sup>, Marta Peirats Llobet<sup>1,†</sup>, Markus Schmid<sup>4,6</sup>, Renaud Dumas<sup>1</sup> & François Parcy<sup>1,7</sup>

Deciphering the mechanisms directing transcription factors (TFs) to specific genome regions is essential to understand and predict transcriptional regulation. TFs recognize short DNA motifs primarily through their DNA-binding domain. Some TFs also possess an oligomerization domain suspected to potentiate DNA binding but for which the genome-wide influence remains poorly understood. Here we focus on the LEAFY transcription factor, a master regulator of flower development in angiosperms. We have determined the crystal structure of its conserved amino-terminal domain, revealing an unanticipated Sterile Alpha Motif oligomerization domain. We show that this domain is essential to LEAFY floral function. Moreover, combined biochemical and genome-wide assays suggest that oligomerization is required for LEAFY to access regions with low-affinity binding sites or closed chromatin. This finding shows that domains that do not directly contact DNA can nevertheless have a profound impact on the DNA binding landscape of a TF.

<sup>1</sup>Laboratoire de Physiologie Cellulaire et Végétale, Université Grenoble Alpes, CNRS UMR5168, CEA/DRF/BIG, INRA UMR 1417, 17, avenue des Martyrs, 38054 Grenoble, France. <sup>2</sup>European Molecular Biology Laboratory, 71, avenue des Martyrs, 38042 Grenoble, France. <sup>3</sup>Institut de Biologie Structurale CEA/DRF, CNRS, Université Grenoble Alpes, 71, avenue des Martyrs, 38044 Grenoble, France. <sup>4</sup>Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany. <sup>5</sup>Instituto de Hortofruticultura Subtropical y Mediterránea, Universidad de Málaga-Consejo Superior de Investigaciones Científicas, Departamento de Biología Molecular y Bioquímica, Facultad de Ciencias, Universidad de Málaga, 29071 Málaga, Spain. <sup>6</sup>Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden. <sup>7</sup>Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, University of British Columbia, Vancouver, British Columbia, Canada V5Z 4H4. †Present addresses: Wellcome Trust Centre for Cell Biology, University of Edinburgh, Michael Swann Building, Max Born Crescent, Edinburgh EH9 3BF, UK (C.S.); CeGaT GmbH, 72076 Tübingen, Germany (F.O.); Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-Universidad Politécnica de Valencia, ES-46022 Valencia, Spain (M.P.L.). Correspondence and requests for materials should be addressed to M.N. (email: mnanao@embl.fr) or to R.D. (email: renaud.dumas@cea.fr) or to F.P. (email: francois.parcy@cea.fr).

**T**ranscription factors (TFs) play key roles during development and differentiation<sup>1</sup>. By binding to short stretches of DNA called *cis*-elements or TF binding sites (TFBS) and thereby regulating gene expression, they are able to decode the regulatory information present in genomes<sup>2</sup>. Based on structural studies and genome-wide analyses, the rules governing TF DNA interactions are becoming clearer and DNA-binding models with increased prediction power are available<sup>3,4</sup>. TF DNA-binding specificity is primarily determined by their DNA-binding domains (DBDs) through contact with bases and DNA shape readout<sup>4</sup>. DBD are often dimeric or sometimes tetrameric, thereby increasing the TFBS length and the TF specificity<sup>3</sup>. In addition to their DBD, some TF also possess an oligomerization domain<sup>5</sup>. Such a feature is found in several types of factors such as members of the E26 transformation-specific or E-twenty-six (ETS) family<sup>6</sup>, auxin response factors<sup>7</sup>, some zinc-finger TFs<sup>8</sup> or the PRH/Hex<sup>9</sup> TF. How such domains modify the repertoire of TF target regions at the genome-wide scale is largely unknown.

Here we focus on LEAFY (LFY), a plant-specific TF essential for flower development<sup>10</sup>. LFY is one of the few master regulators of flower development, as it integrates environmental and endogenous signals to orchestrate the whole floral network<sup>11–14</sup>. It first triggers the early emergence of a bulge of stem cells on the flanks of the shoot apex<sup>15,16</sup> and subsequently specifies their floral identity by directly inducing floral homeotic genes such as *APETALA1* (*API*), *APETALA3* (*AP3*) or *AGAMOUS* (*AG*) and by repressing the shoot meristem identity gene *TERMINAL FLOWER1* (*TFL1*)<sup>10,11</sup>. Comprehensive lists of genes bound and possibly regulated by LFY have been established by combining genome-wide chromatin immunoprecipitation (ChIP) experiments and transcriptomic approaches<sup>17,18</sup>.

The function of LFY as a master regulator of floral development is most obvious when expressed from a constitutive promoter: it is sufficient to trigger early flowering and ectopic flower production<sup>19</sup>. Together with co-regulators such as the homeodomain TF *WUSCHEL* or the F-Box protein *UNUSUAL FLORAL ORGANS*, it can even induce flowers from root or leaf tissue, respectively<sup>20,21</sup>. This acquisition of floral fate is remarkable, as several LFY targets including *AG* and *AP3* are known to be under the repression of two Polycomb repressive complexes (*PRC1* and *PRC2*) (refs 22,23), which ectopic LFY expression is apparently sufficient to override.

Structure–function studies have demonstrated that LFY binds semi-palindromic 19-bp DNA elements through its highly conserved C-terminal DBD, a unique helix–turn–helix fold that by itself dimerizes on DNA (Fig. 1a)<sup>17,18,24,25</sup>. A biophysical model describing LFY DNA-binding specificity *in vitro* was built and accurately predicts LFY-binding sites (LFYbs) in the *Arabidopsis thaliana* genome<sup>17,26</sup>. In addition to its well-characterized DBD, LFY possesses a second conserved domain at its amino terminus (LFY-N) (Fig. 1a). This domain has been proposed to mediate LFY dimerization by forming a leucine zipper<sup>27</sup>. Here we report the crystallographic structure of LFY-N and elucidate its function using a combination of biochemical and genomic approaches. Our results demonstrate that LFY-N is a Sterile Alpha Motif (SAM) domain that mediates LFY oligomerization. It allows LFY to bind to regions lacking high-affinity LFYbs and confers on LFY the ability to access closed chromatin regions.

## Results

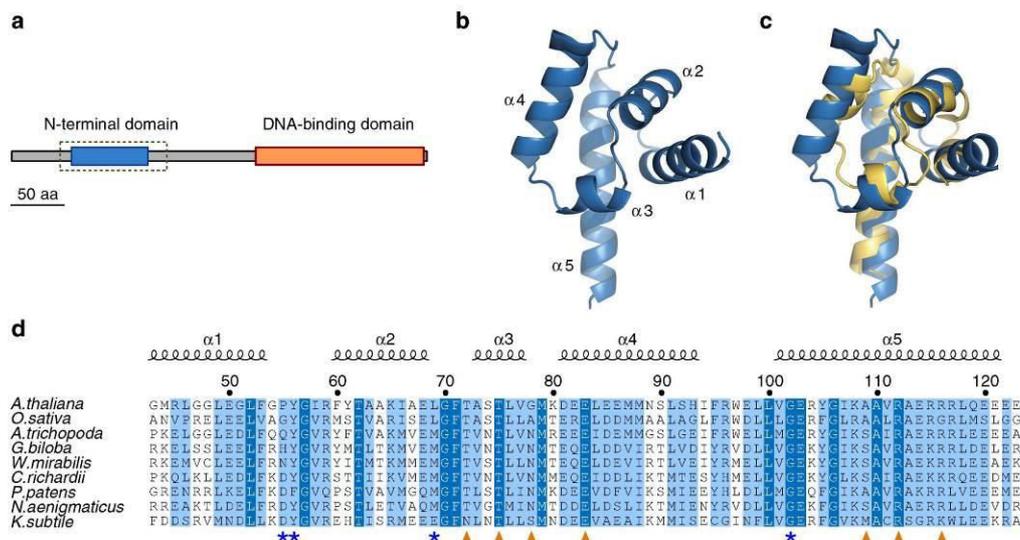
**LFY N-terminal domain is a SAM.** We used a structural approach to unravel the function of the N-terminal domain of LFY. After heterologously overexpressing LFY-N domains from different plant species, we were able to determine the crystal

structure for GbLFY-N from the gymnosperm *Ginkgo biloba* at 2.3 Å resolution (Table 1 and Supplementary Fig. 1). This structure showed that the GbLFY-N monomer is made up of five  $\alpha$ -helices separated by four loops (Fig. 1b). Comparison against the Protein Data Bank revealed a strong structural similarity with the SAM, also called Pointed domain<sup>28</sup> (Fig. 1c), a resemblance that could not be predicted from LFY-N sequences. SAM domains have not been characterized in plants, but they are common protein motifs in other eukaryotes. They interact with proteins, DNA, RNA and lipids, and are involved in many cellular functions including transcriptional regulation and signal transduction<sup>6,28</sup>. Some SAM domains are able to self-associate as oligomers through the so-called mid-loop (ML) and end-helix (EH) surfaces<sup>29,30</sup>. GbLFY-SAM appears to belong to this class: in the crystal, it forms a head-to-tail polymer (Fig. 2a) with monomers contacting their neighbours through two polar surfaces (Fig. 2b). The ML surface comprises negative and polar residues (Fig. 2c), including T72, T75 and E83, which interact with the positively charged EH surface through hydrogen bonds and salt bridges, in particular via R112 and R116 (residue numbering refers to LFY from *A. thaliana*—AtLFY—sequence; Fig. 1d).

**LFY-SAM is an oligomerization domain.** To validate the interactions observed in the crystal in solution, we analysed the oligomerization state of the wild-type (WT) GbLFY-SAM protein and of mutants at the ML surface (T75E substitution, GbLFY-SAM<sub>TE</sub>), EH surface (R112E substitution, GbLFY-SAM<sub>RE</sub>) or at both surfaces (GbLFY-SAM<sub>TERE</sub>) (Fig. 2d). Size-exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS) showed that the WT GbLFY-SAM domain formed oligomers of variable size in solution, depending on protein concentration, containing up to eight monomers (Fig. 2d, Supplementary Fig. 2 and Supplementary Table 1). In comparison, all mutations affecting the interaction interface completely abolished LFY oligomerization and GbLFY-SAM<sub>TE</sub>, GbLFY-SAM<sub>RE</sub> and GbLFY-SAM<sub>TERE</sub> proteins were found to be exclusively monomeric (Fig. 2d and Supplementary Table 1). However, when mixed at an equimolar ratio, the GbLFY-SAM<sub>TE</sub> and GbLFY-SAM<sub>RE</sub> single-face mutants retained the ability to interact, forming a GbLFY-SAM<sub>[RE+TE]}</sub> dimer (Fig. 2d and Supplementary Table 1), thus showing that the different mutations did not alter the overall fold of the SAM domain, and that single face mutant proteins still displayed one functional interaction surface. These results further showed that the head-to-tail arrangement observed in the crystal structure reflects the state of LFY-SAM in solution, which forms an oligomer of limited size in the conditions of this study. The point mutations we designed allowed us to control the oligomerization state, creating either monomeric or dimeric variants of LFY-SAM instead of higher-order oligomers.

Comparing LFY sequences from multiple species throughout the plant kingdom revealed that the key interaction residues of the SAM domain are well conserved (Fig. 1d and Supplementary Fig. 2d), suggesting that the capacity to oligomerize is a general characteristic of LFY-SAM in all species from algae to angiosperms. This hypothesis is supported by the behaviour of the SAM domain of *A. thaliana* LFY (AtLFY-SAM) whose molecular mass in solution increased with its concentration, similar to what was observed for GbLFY-SAM (Supplementary Fig. 2).

**LFY-SAM<sub>TERE</sub> mutants have highly reduced function in planta.** Several mutations that compromise LFY function in rice and *A. thaliana* map to the SAM domain<sup>27,31</sup>. However, according to



**Figure 1 | LFY N-terminus is a SAM domain.** (a) Domain representation of GbLFY protein with the N-terminal domain in blue and DBD in orange. Dashes indicate the region used to obtain the crystal structure. Scale bar, 50 amino acids (aa). (b) Crystal structure of GbLFY-N monomer.  $\alpha$ ,  $\alpha$ -helices. (c) Superimposition of GbLFY-N (blue) and the SAM domain of p63 $\alpha$  (yellow, PDB 2Y9U)<sup>66</sup>. The two structures are very similar and can be superimposed with a root mean square deviation of 1.9 Å for 62 C $\alpha$ . (d) Sequence alignment of LFY-N domain. The residue numbering refers to *A. thaliana* LFY sequence; the secondary structure from GbLFY-SAM crystal structure is indicated above the alignment; orange triangles and blue stars indicate residues involved in the interaction between SAM monomers through the lateral and main chain, respectively.

**Table 1 | Crystallographic data collection and refinement statistics.**

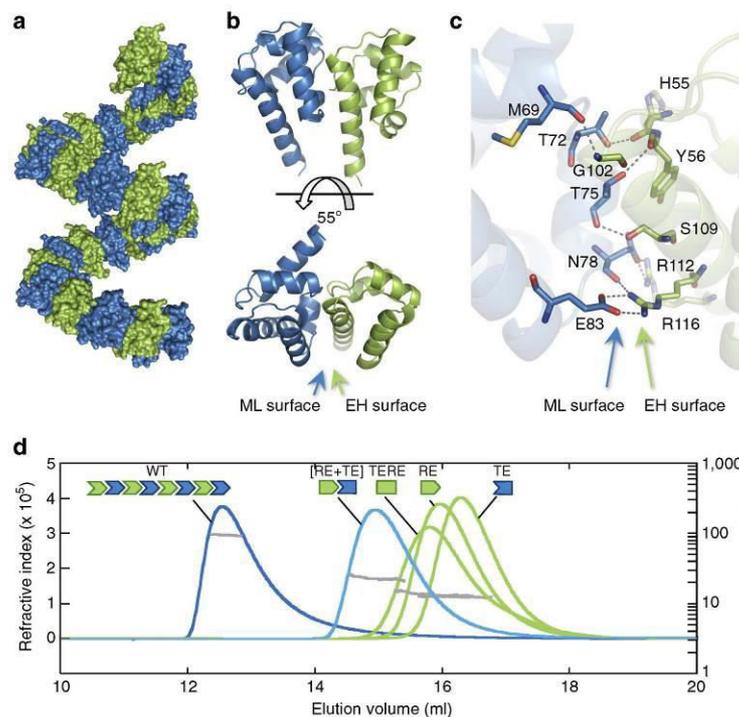
	GbLFY-N (Native)	GbLFY-N (Se-Met)
<i>Data collection</i>		
Wavelength (Å)	0.872	0.872
Resolution (Å) range	52.34–2.25 (2.33–2.25)	100–2.61 (2.77–2.61)
Space group	<i>P</i> 6 <sub>5</sub>	<i>P</i> 6 <sub>5</sub>
Cell dimensions (Å, °)	<i>a</i> = <i>b</i> = 81.1, <i>c</i> = 78.5	<i>a</i> = <i>b</i> = 80.6, <i>c</i> = 78.5
Number of total reflections	57,705	83,685
Number of unique reflections	14,004	8,879
Multiplicity	4.1(3.9)	9.4 (9.2)
Completeness (%)	99.3 (99.0)	99.6 (97.5)
Mean <i>I</i> / $\sigma$ ( <i>I</i> )	15.67 (2.89)	15.28 (2.54)
<i>R</i> -sym	0.069 (0.514)	0.127 (0.915)
<i>Refinement statistics</i>		
Wilson B-factor	44.1	35.85
<i>R</i> -factor	0.1859 (0.2564)	
<i>R</i> -free	0.2298 (0.2767)	
Number of atoms	1,462	
Macromolecule	1,335	
Ligands	25	
Water	102	
Protein residues	163	
Root mean square (bonds)	0.016	
Root mean square (angles)	1.78	
Ramachandran favoured (%)	99	
Ramachandran outliers (%)	0	
Clashscore	13.40	
Average B-factor	46.80	
Macromolecule	46.00	
Solvent	51.80	

Statistics for the highest-resolution shell are shown in parentheses.

our structural data, these positions correspond to buried hydrophobic residues that are unlikely to directly affect oligomerization but are more likely to be important for proper domain folding. Only the very weak *lfy-22* allele in *A. thaliana*<sup>32</sup> carries a mutation (G70D) that localizes to the SAM ML surface and possibly slightly weakens oligomerization. Therefore, to specifically assess the importance of oligomerization for LFY function, we tested whether the TERE mutation affects AtLFY ability to control flower development in the genetically amenable *A. thaliana* plant.

We first tested whether AtLFY<sub>TERE</sub> could complement the *lfy-12*-null mutant, in which flowers are replaced by shoots or sterile shoot/flower intermediate structures lacking petals and stamens (Fig. 3a)<sup>33</sup>. When expressed under the control of the LFY endogenous promoter (*pLFY*), AtLFY<sub>TERE</sub> complemented the *lfy-12* mutant much less efficiently than AtLFY (Fig. 3a,b and Supplementary Table 2, Mann–Whitney rank test, *P* = 0.01006). Whereas most of *pLFY:LFY lfy-12* primary transformants produced fertile flowers, the majority of *pLFY:LFY<sub>TERE</sub> lfy-12* primary transformants produced only a few fertile flowers at the base of the inflorescence. These results indicated that LFY activity was reduced, although not entirely abolished, when its oligomerization was compromised.

To further assess the importance of the SAM domain for LFY master regulator function, we expressed AtLFY from the *p35S* constitutive promoter and monitored its effect at an early developmental stage, when other pathways contributing to flower development have only a minimal impact. When expressed constitutively, AtLFY is known to induce early flowering, and to trigger the precocious termination of the shoots and the formation of ectopic flowers from the axils of rosette leaves<sup>19</sup>. About half of the 61 *p35S:LFY* T1 plants we generated showed this characteristic phenotype. In contrast, ectopic flowers were never observed in 51 *p35S:LFY<sub>TERE</sub>* T1 transformants (Fig. 3c,d, Fisher's exact test, *P*-value =  $1.5 \times 10^{-10}$ ) despite the fact that



**Figure 2 | LFY-SAM is an oligomerization domain.** (a) Helical polymer of GbLFY-SAM in the crystal structure. Twenty-eight monomers are shown, alternatively coloured in blue and green. (b) Side and top views (upper and bottom panels, respectively) of the GbLFY-SAM dimer in the asymmetric unit of the crystal. ML and EH interaction surfaces are indicated. (c) Detailed view of GbLFY-SAM oligomerization interface. Residues mediating contacts between the two monomers are shown as sticks; residues on the ML and EH surfaces are in blue and green, respectively. (d) SEC-MALLS analysis of GbLFY-SAM (WT), GbLFY-SAM<sub>TE</sub> (TE, T75E substitution at the ML surface) and GbLFY-SAM<sub>RE</sub> (RE, R112E substitution at the EH surface), GbLFY-SAM<sub>TERE</sub> (TERE, T75E, R112E) double mutant and GbLFY-SAM<sub>[RE+TE]</sub> ([RE + TE], an equimolar mixture of GbLFY-SAM<sub>TE</sub> and GbLFY-SAM<sub>RE</sub>). 50  $\mu$ l at  $\sim$ 4.5 mg ml<sup>-1</sup> of proteins were used. Elution profiles were monitored by excess refractive index (left ordinate axis). The grey line under each elution peak shows the molecular mass distribution (right ordinate axis). Measured molecular masses (reported in Supplementary Table 1) show that GbLFY-SAM is oligomeric; GbLFY-SAM<sub>TE</sub>, GbLFY-SAM<sub>RE</sub> and GbLFY-SAM<sub>TERE</sub> are monomeric; and GbLFY-SAM<sub>[RE+TE]</sub> is dimeric.

AtLFY and AtLFY<sub>TERE</sub> proteins were expressed at similar levels (Supplementary Fig. 3). Taken together, these results provide evidence that LFY-SAM domain oligomerization is required for the function of LFY as a master switch in flower initiation.

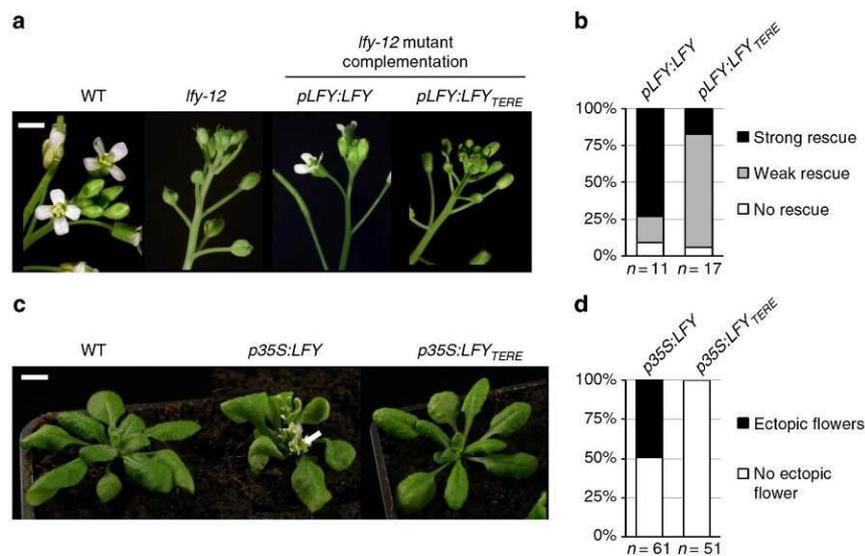
**SAM<sub>TERE</sub> mutations do not alter single LFYbs binding *in vitro*.** Next, we aimed to pinpoint how a functional SAM domain contributes to LFY function at the molecular level. As a TF, LFY specifically binds to *cis*-elements present in the regulatory regions of its target genes. Hence, we tested whether LFY mutants with altered oligomerization status differed in DNA-binding ability. For other SAM proteins, such analyses were hampered by the difficulty in producing recombinant soluble proteins due to the aggregation and precipitation triggered by the SAM domain<sup>34</sup>. By optimizing the purification protocol, we successfully produced a near full-length GbLFY protein (GbLFY $\Delta$ ) lacking only the 53 non-conserved N-terminal amino acids preceding the SAM domain.

As GbLFY and AtLFY have comparable DNA-binding specificities<sup>24</sup>, we tested GbLFY $\Delta$  binding to a DNA fragment carrying a single high-affinity AtLFYbs (*API* probe) from the promoter of the *A. thaliana API* gene<sup>35</sup>. Electrophoretic mobility shift assay (EMSA) and SEC-MALLS experiments showed that GbLFY $\Delta$  and GbLFY $\Delta$ <sub>TERE</sub> bind *API* DNA mostly as dimers (Fig. 4c and Supplementary Table 1) and with the same apparent

affinity (Fig. 4a). Thus, the SAM domain is not essential for the formation of the LFY dimer/DNA complex. This is consistent with previous studies showing that LFY-DBD by itself dimerizes on DNA<sup>25</sup>.

We noticed, however, a slight migration shift for GbLFY $\Delta$ <sub>TE</sub>, GbLFY $\Delta$ <sub>RE</sub> and GbLFY $\Delta$ <sub>TERE</sub> complexes when compared with GbLFY $\Delta$  in EMSA (Fig. 4a,b). This shift disappeared when GbLFY $\Delta$ <sub>RE</sub> and GbLFY $\Delta$ <sub>TE</sub> (hereafter called GbLFY $\Delta$ <sub>[RE+TE]</sub>) were mixed at equimolar concentration (Fig. 4b). This difference in migration behaviour might be due to a slightly different conformation of the complexes when the SAM domain is monomeric. A comparable result was obtained for the AtLFY protein (Supplementary Fig. 4a). Taken together, these findings indicate that LFY-SAM oligomerization is not required to trigger dimer formation on DNA but it appears to only slightly influence the migration of the LFY/DNA complex.

**The SAM domain prevents LFY DNA binding as monomer.** In addition to the main dimeric protein/DNA complex, GbLFY $\Delta$ <sub>TE</sub>, GbLFY $\Delta$ <sub>RE</sub> and GbLFY $\Delta$ <sub>TERE</sub> protein variants showed a weak monomeric complex with the *API* probe in EMSA (Fig. 4b). This was particularly obvious when using a mutated probe, *API*<sub>m</sub>, which contained mutations on each side of the LFYbs pseudo-palindrome and was not bound by WT GbLFY $\Delta$  (Fig. 4a bottom panel). GbLFY $\Delta$ <sub>TERE</sub> was also able to bind to DNA probes that



**Figure 3 | Mutations to the LFY-SAM oligomerization domain impair LFY floral function in *A. thaliana*.** (a) Pictures of WT, *lfy-12* mutant and *lfy-12* expressing *pLFY:LFY* or *pLFY:LFY<sub>TERE</sub>* inflorescences from 40-day-old plants grown under long day conditions (scale bar, 2 mm). (b) Distribution of *pLFY:LFY* and *pLFY:LFY<sub>TERE</sub>* primary transformants in phenotypic complementation classes (see Supplementary Table 2 for details). (c) Pictures of WT and transgenic plants expressing *p35S:LFY* or *p35S:LFY<sub>TERE</sub>* grown for 30 days under long-day conditions (scale bar, 1 cm). The white arrow indicates ectopic rosette flowers. (d) Proportion of the *p35S:LFY* and *p35S:LFY<sub>TERE</sub>* primary transformants showing ectopic rosette flowers.

carried mutations in either half of the palindromic binding site (*API1m1* and *API1m2* probes) as a monomer, whereas GbLFY $\Delta$  showed very little binding to such probes (Supplementary Fig. 4b). These experiments suggest that the LFY-SAM domain might prevent monomeric LFY binding on the numerous half sites present in a genome. To further investigate this possibility, we performed competition EMSA with increasing amounts of nonspecific unlabelled DNA competitor (fish genomic DNA). We found that GbLFY $\Delta_{TE}$ , GbLFY $\Delta_{RE}$  and GbLFY $\Delta_{TERE}$  dissociated from *API* at lower competitor concentrations than GbLFY $\Delta$  and GbLFY $\Delta_{[RE+TE]}$  (Fig. 4d and Supplementary Fig. 4c). Thus, the SAM domain appears to favour LFY dimeric binding at the expense of a less specific monomeric interaction with DNA.

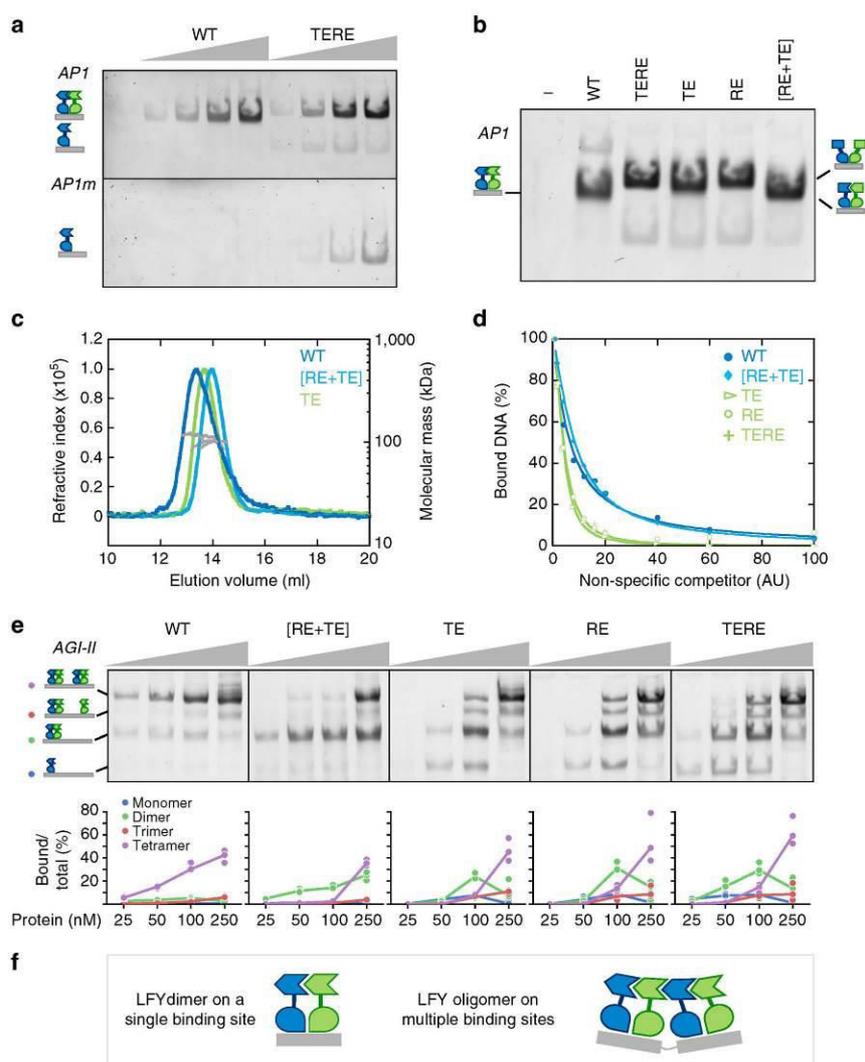
#### LFY-SAM<sub>TERE</sub> mutants show impaired cooperative DNA binding.

Because of their ability to oligomerize, SAM domains have been proposed to allow cooperative DNA binding on regions with multiple TFBS<sup>36</sup>. This property has been inferred from SAM protein variants whose oligomerization potential was limited to dimerization<sup>34</sup> but never established for a WT SAM protein. To test whether LFY-SAM could affect binding on DNA fragments with multiple binding sites *in vitro*, we used probes carrying two nearby LFYbs, either synthetic or from the *AG* regulatory region (*AGI-II*)<sup>17</sup>. With these probes, the GbLFY $\Delta$  protein formed one main complex corresponding to two bound dimers (Fig. 4e and Supplementary Figs 4d and 5). At low concentrations, GbLFY $\Delta_{[RE+TE]}$ , which mimics a WT dimer unable to further oligomerize (Fig. 2d), was not as efficient as GbLFY $\Delta$  at forming a tetrameric complex and the tetramer/dimer ratio was lower for GbLFY $\Delta_{[RE+TE]}$  than for GbLFY $\Delta$  (Fig. 4e and Supplementary Fig. 5). As expected, all monomeric mutants (GbLFY $\Delta_{TE}$ , GbLFY $\Delta_{RE}$  and GbLFY $\Delta_{TERE}$ ) were also impaired in tetramer formation at low concentrations (Fig. 4e and Supplementary Fig. 5). In these assays, the AtLFY protein behaved similarly as GbLFY $\Delta$  (Supplementary Fig. 4a), indicating that mediating cooperative binding to multiple LFYbs is an intrinsic property of

the LFY-SAM domain and is important for the formation of tetrameric complexes (Fig. 4f).

#### A functional SAM domain is required for DNA binding *in vivo*.

To test the importance of a functional SAM domain for LFY DNA binding *in vivo*, we performed biological duplicate ChIP sequencing (ChIP-Seq) experiments with *p35S:LFY*, *p35S:LFY<sub>TERE</sub>* and non-transgenic *A. thaliana* (Col-0) 2-week-old seedlings. At this stage, the expression of the endogenous *LFY* is minimal as compared with that of the transgenes, enabling us to use the WT seedlings as a negative control. We chose to use ectopic expression in seedlings so that we could compare LFY and LFY<sub>TERE</sub> DNA-binding activity in the same tissue, a condition that would not be fulfilled by comparing *lfy* inflorescences expressing LFY or LFY<sub>TERE</sub>. We are aware that, when ectopically expressed, LFY might contact some genomic regions that are not genuine targets. However, it is known that the LFY-bound regions in seedlings and inflorescences significantly overlap and the LFY-binding motifs in the two tissues are extremely similar<sup>17,18</sup>. We therefore compared the intensity of LFY binding in *p35S:LFY* and *p35S:LFY<sub>TERE</sub>* for high confidence bound regions identified in *p35S:LFY*. After signal normalization (Supplementary Fig. 6), we found that the TERE mutation drastically reduced the LFY binding *in vivo* (Fig. 5a,b and Supplementary Fig. 7a). For all subsequent analyses, we define the coverage fold reduction (CFR) as the ratio between the LFY and the LFY<sub>TERE</sub> ChIP-Seq coverages. The CFR values ranged from 1.3 to >150 (Fig. 5a). Ninety-five per cent of the regions have a CFR > 3 and 54% have a CFR > 10. We also performed the converse analysis by selecting the highest confidence bound regions in *p35S:LFY<sub>TERE</sub>* and comparing their binding with that observed in *p35S:LFY*. Except for a few regions for which the binding by LFY<sub>TERE</sub> was slightly stronger, most regions were bound better by LFY (Supplementary Fig. 7b). The binding reduction is unlikely to be due to a less efficient recognition of LFY<sub>TERE</sub> as compared with LFY by the antibody used in the ChIP, as it was raised against LFY-DBD



**Figure 4 | Role of the LFY-SAM oligomerization domain in *in vitro* DNA binding.** (a) EMSA with 10 nM AP1 or AP1m DNA probe and 0, 50, 100, 250 and 500 nM GbLFY $\Delta$  (WT) or GbLFY $\Delta$ <sub>TERE</sub> (TERE) protein. This experiment was performed four times with similar results. (b) EMSA with 10 nM AP1 DNA and 500 nM GbLFY $\Delta$  (WT), GbLFY $\Delta$ <sub>[RE+TE]</sub> [RE + TE], GbLFY $\Delta$ <sub>TE</sub> (TE), GbLFY $\Delta$ <sub>RE</sub> (RE) or GbLFY $\Delta$ <sub>TERE</sub> (TERE) protein. This experiment was performed twice with similar results. (c) SEC-MALLS molecular mass analysis of GbLFY $\Delta$ , GbLFY $\Delta$ <sub>TE</sub> and GbLFY $\Delta$ <sub>[RE+TE]</sub> in complex with the AP1 probe. Fifty microlitres containing 2.6 mg ml<sup>-1</sup> (60  $\mu$ M) protein and 15  $\mu$ M DNA were used. Elution profiles were monitored by excess refractive index (left ordinate axis). The line under each elution peak shows the molecular mass distribution (right ordinate axis). The measured molecular mass for each complex is reported in Supplementary Table 1 and is consistent with that of a protein dimer on DNA. The SEC experiment was performed twice and the MALLS measurement was performed once. (d) Percentage of DNA bound to LFY as a function of nonspecific unlabelled competitor DNA concentration. Quantifications were based on EMSA shown in Supplementary Fig. 4c. This experiment was performed four times with similar results. (e) EMSA with 10 nM AGI-II DNA probe and 25, 50, 100 and 250 nM GbLFY $\Delta$ , GbLFY $\Delta$ <sub>[RE+TE]</sub>, GbLFY $\Delta$ <sub>TE</sub>, GbLFY $\Delta$ <sub>RE</sub> or GbLFY $\Delta$ <sub>TERE</sub> protein. Quantifications show the percentage of total DNA bound by one (monomeric complex, blue), two (dimeric complex, green), three (trimeric complex, red) or four (tetrameric complex, violet) LFY proteins. EMSA have been performed in triplicate and the fit follows the median value. Formation of the tetrameric complex starts at lower concentrations for WT GbLFY as compared with the mutant GbLFY versions. (f) Schematic representation of LFY bound as dimer on a single binding site (left) and as tetramer on two binding sites (right) deduced from the EMSA and SEC-MALLS analysis. DNA is in grey and LFY in blue and green. For all EMSAs, only the protein-DNA complexes are shown.

alone and yielded a comparable signal in western blotting on both proteins (Supplementary Fig. 3). The differences in CFR were apparently neither due to alterations to the sequence specificity of AtLFY<sub>TERE</sub>, as the binding site motifs derived from AtLFY and AtLFY<sub>TERE</sub> ChIP-Seq were very similar (Supplementary Fig. 7c).

This result is consistent with previously published SELEX experiments demonstrating that LFY N-terminal domain deletion does not affect LFY DNA-binding specificity<sup>15,17,24</sup>. Taken together, these findings indicate that LFY SAM domain is crucial for its genome-wide DNA binding *in vivo*.

**LFY-SAM facilitates binding to sites of suboptimal affinity.**

The general reduction in DNA binding of LFY<sub>TERE</sub> is probably partially explained by its capacity to bind as a monomer and scatter over the numerous monomeric sites of the genome. However, as illustrated by three examples (Fig. 5b), the reduction in binding by mutants that are unable to oligomerize varies greatly between genomic regions (Fig. 5a). Gene ontology analysis for genes neighbouring the top high- and low-CFR regions did not detect any specific enrichment. We tried to understand whether the nature and number of LFYbs present in a region might help explaining the binding reduction, as our biochemical analysis suggested. To identify LFYbs, we used a previously validated and highly predictive position weight matrix that computes a score between 0 (highest affinity site) and -56 (lowest affinity) for each 19 pb sequence<sup>17,26</sup>. To avoid using an arbitrary threshold above which a DNA stretch qualifies as LFYbs, we devised a specific procedure to compute the threshold (see methods). The low threshold we obtained (-25) indicates that LFY-bound regions are enriched in LFYbs of very weak affinity (Supplementary Fig. 8a). When applied to regions bound by LFY<sub>TERE</sub>, the same procedure yielded a higher value (-20) (Supplementary Fig. 8b), suggesting that oligomerization promotes binding to low-affinity sites. Consistent with this hypothesis, the bound regions that lack high-affinity LFYbs tend to have higher CFR than regions with high-affinity LFYbs (Fig. 5c). We also analysed the possible influence of LFYbs density and found that regions with high LFYbs density have higher CFR (Fig. 5d) than regions with less LFYbs. Thus, a functional SAM domain helps LFY to bind to DNA regions with either sub-optimal LFYbs or clusters of LFYbs. To investigate whether these clusters display a specific structure, we characterized the distribution of distances between LFYbs in bound regions. We found that binding sites separated by 1 bp were specifically overrepresented (three- to fourfold) in regions with high CFR (Fig. 5e). Reciprocally, the presence of one or more 1-bp distant sites significantly increased the CFR value (Fig. 5f). These analyses suggest that SAM-mediated oligomerization promotes LFY binding to a specific LFYbs configuration. We confirmed this hypothesis biochemically, by showing in EMSA that a distance of 1 bp between LFYbs most favoured LFY tetrameric binding over other spacing (0, 2, 6 and 11 pb) (Supplementary Fig. 5). Similar results were obtained when we limited the same analyses to the subset of genomic regions that are both bound by LFY in seedlings and by endogenous LFY in inflorescences<sup>18</sup> (Supplementary Fig. 9). Except for the effect of binding site density that lost its statistical significance, the conclusions are overall very similar despite the smaller size data set. Altogether, these analyses provide evidence that oligomerization through the LFY SAM domain facilitates LFY binding to regions that lack high-affinity binding sites or display multiple and adjacent sites.

**LFY-SAM allows access to closed chromatin regions.** As other SAM domain proteins have been suggested to regulate chromatin status<sup>37,38</sup>, we also tested whether the CFR could depend on the accessibility of DNA regions. This genomic feature can be monitored using DNaseI hypersensitivity (DHS) coupled to high-throughput sequencing (DNaseI-Seq). Such data were publicly available for 2-week-old WT *A. thaliana* seedlings<sup>39</sup>, the same developmental stage we used for the ChIP-Seq experiments. We therefore calculated the percentage of opened regions (following the criteria defined in ref. 39) in deciles of LFY-bound regions, sorted according to their CFR. In the low CFR regions, only 10–15% of the LFY-bound regions were closed. Strikingly, this number reached 70% in regions with high CFR

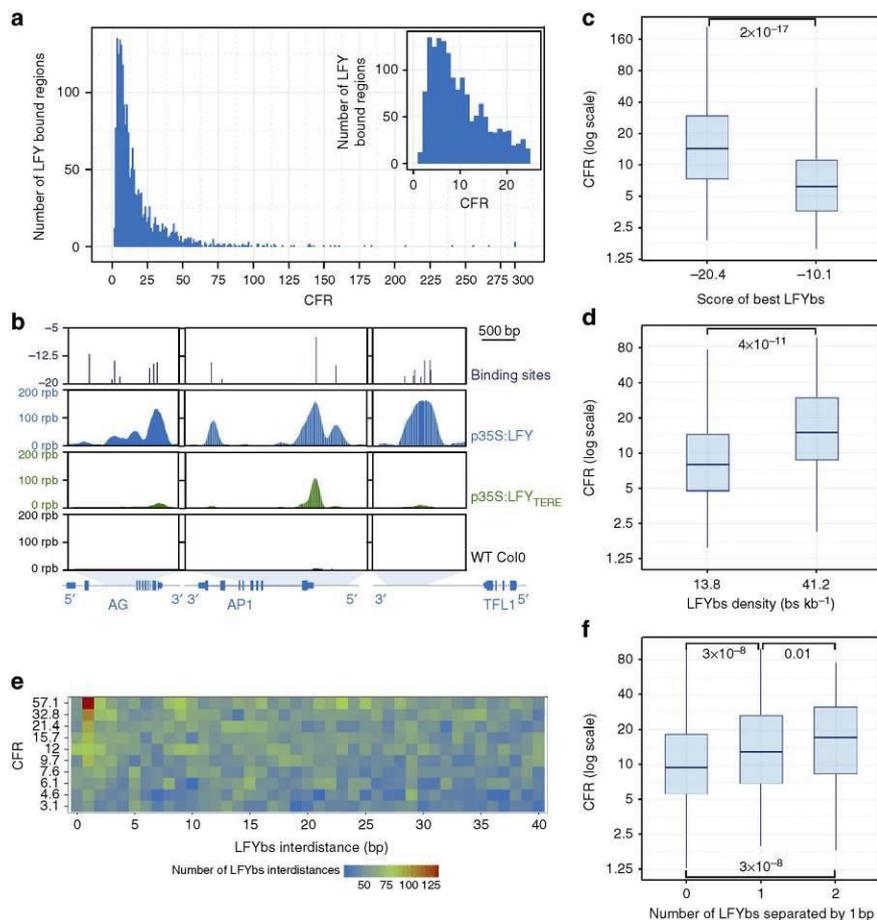
(Fig. 6a), suggesting that the chromatin state of the bound regions has a strong impact on the need for LFY oligomerization. To complement this analysis, we studied the correlation between quantitative DHS levels and the CFR. We found that the CFR values remarkably decreased when the accessibility (that is, the DHS level) increased (Fig. 6b) and poorly accessible regions have higher CFR values than accessible ones (Fig. 6c). An independent DNaseI-Seq data set from slightly younger seedlings (7-day-old) yielded similar results (Supplementary Fig. 7d)<sup>40</sup>. Again, these conclusions held true for the subset of genomic regions for which binding by LFY binding has also been validated in inflorescences (Supplementary Fig. 9g–i). Taken together, these results provide strong evidence that LFY binding to closed chromatin regions depends on the presence of a functional SAM domain.

**Discussion**

The crystal structure of the conserved LFY N-terminus revealed a SAM oligomerization domain. This finding was unexpected not only because the LFY-N primary sequence has no homology with other SAM domains but also because LFY-N was initially proposed to be a putative leucine zipper dimerization domain<sup>27</sup>. The SAM domain appears to be highly conserved throughout LFY evolution (Fig. 1d and Supplementary Fig. 2d) and probably performs the same oligomerization function in all plant species. Both LFY from *G. biloba* (gymnosperm) and *A. thaliana* (angiosperm) oligomerize. Moreover, chimeric proteins bearing the AtLFY-DBD combined with the SAM from the moss *Physcomitrella patens* or the fern *Ceratopteris richardii* complemented the *A. thaliana lfy* mutant, indicating that the biochemical properties of the SAM domain are evolutionary conserved<sup>41</sup>.

*In planta* experiments revealed that altering the capacity of LFY to oligomerize compromised its floral function and drastically reduced its genome-wide DNA binding. The fact that TF oligomerization could contribute to DNA binding was expected based on the biochemical characterization we performed on LFY, as well as previous work on PRH-Hex<sup>9</sup> and on TFs from the ETS family<sup>34</sup>. Specifically, our study shows the extent to which genome-wide DNA binding can be altered when oligomerization is disrupted. The single TF for which this question was addressed so far is YAN, a *Drosophila* ETS TF<sup>38</sup>. The analysis of a non-oligomeric version of its SAM domain indicated that YAN oligomerization did not appear to be the primary determinant for its spreading over extended chromatin regions. Another study focused on PRC1, which is not a TF but a SAM domain member of the polycomb repressive complex in mammals. This work showed that mutation of its oligomerization capacity affected its DNA binding in only 12% of its target regions in mouse embryonic cells<sup>37</sup>. Our study thus offers evidence that TF oligomerization can have a profound impact on its genome-wide DNA-binding landscape.

We identified several reasons why the SAM oligomerization domain might influence LFY DNA binding. First, the SAM domain appears to limit LFY binding as a monomer and therefore might prevent LFY scattering over the genome due to nonspecific binding. How the SAM oligomerization domain reduces monomeric binding is unknown. However, it is tempting to speculate that it might impose constraints on the DBD, reducing its ability to interact with sub-optimal sites as a dimer. Second, combining *in vitro* and *in vivo* experiments, we showed that the SAM oligomerization domain facilitates cooperative higher-order LFY complex formation on regions with multiple or low-affinity LFYbs. Even for LFYbs that appear relatively isolated in the genome, it is conceivable that LFY oligomerization could contribute to cooperative binding through DNA looping.



**Figure 5 | Mutations altering LFY oligomerization affect genome-wide LFY DNA binding in planta.** (a) Histogram showing the distribution of CFR values for the LFY-bound regions in *p35S:LFY*. The inset shows an enlargement for low CFR. (b) AtLFY and AtLFY<sub>TERE</sub> binding profiles on the AG (left), API (middle) and the *TFL1* (right) genomic regions. ChIP-Seq read coverage is shown for *p35S:LFY* (blue), *p35S:LFY<sub>TERE</sub>* (green) and Col-0 (black, background signal). The upper panel shows the LFYbs locations and scores. Blue boxes at the bottom indicate exons. rpb, read per base. (c) Boxplot showing the CFR of the LFY-bound regions with LFYbs of lowest and highest score (first and last deciles). The x axis shows the median LFYbs best score value. *P*-value for Mann-Whitney rank test is indicated on the graph. (d) Boxplot showing the CFR of the LFY-bound regions with lower and higher LFYbs density (first and last deciles of LFY-bound regions). The x axis shows the median density value (bs kb<sup>-1</sup>). *P*-value for Mann-Whitney rank test is indicated on the graph. (e) Heat map representing the distribution of distances between LFYbs. The LFY-bound regions in *p35S:LFY* were pooled in deciles according to their CFR value (y axis shows the groups median CFR value). The x axis shows the distance (in bp) separating two 19-bp LFYbs. Only distances < 40 bp are shown. (f) Boxplot showing the CFR of regions depending on the number of pairs of LFYbs separated by 1 bp (zero, one and two pairs of LFYbs). *P*-values for Mann-Whitney rank test are indicated on the graph.

Cooperative TF DNA binding was previously shown to be important to trigger developmental switches<sup>42,43</sup> and the development of flowers is one of the best examples of a phase transition in plants. When coupled to clusters of low-affinity binding sites, TF oligomerization was shown to increase the sharpness of expression patterns<sup>44</sup>. Therefore, as levels of LFY progressively increase<sup>45</sup>, oligomerization may induce target genes with specific spatio-temporal patterns depending on their promoter *cis*-element topology (LFYbs affinity and density). For example, the early target *API* possesses only a few LFYbs, including one of the highest affinity sites found in the genome (score -7) and might thus be induced early during flower development, whereas genes such as *AP3*, *AG* or *TFL1* that contain a number of lower affinity sites (scores between -10 and -15) are regulated (induced or repressed) later during

flower development. SAM-mediated oligomerization could thus contribute to the timing in the expression of LFY target genes. Finally, although we consider it unlikely, we cannot totally exclude that the TERE mutations could have an indirect effect on LFY DNA binding *in vivo*, independent of oligomerization. These mutations could, for example, affect the interactions with yet unidentified regulators that contribute to the efficiency of LFY DNA binding.

By combining ChIP-Seq data with maps of DNA accessibility<sup>39,40</sup>, we have also provided evidence for a novel and unexpected role for SAM domain-mediated oligomerization: it enables LFY to access closed chromatin regions. Regulatory regions embedded in closed chromatin are usually poorly accessed by TFs<sup>4,46,47</sup>. However, a handful of factors, termed 'pioneer TF', are able to interact with packed chromatin to affect

transcriptional activity<sup>48,49</sup>. Several of those factors are potent developmental regulators involved in cell programming or reprogramming<sup>49</sup>. LFY is known to play a key role in stem cell emergence and floral fate determination<sup>10,15,16</sup>. Whether it qualifies as a pioneer TF will require testing its effect on the chromatin status of bound regions<sup>49</sup>. In recent times, the MADS

TF SEPALLATA3 and AP1 have also been proposed to play a pioneer role during floral organ development<sup>50–52</sup>. Interestingly, these TFs, as LFY<sup>50</sup>, interact with the several chromatin-remodelling factors<sup>51</sup> that, once recruited to closed chromatin regions, could contribute to their opening.

Determining the crystal structure of LFY-N not only revealed its identity as a SAM oligomerization domain but also provided the tools to uncover its function during floral development. We provide evidence that oligomerization can have a profound effect on a TF binding landscape by promoting cooperative binding of LFY to DNA, as was proposed for other oligomeric TFs, and, more unexpectedly, also gives LFY access to closed chromatin regions that are notably refractory to TF binding.

## Methods

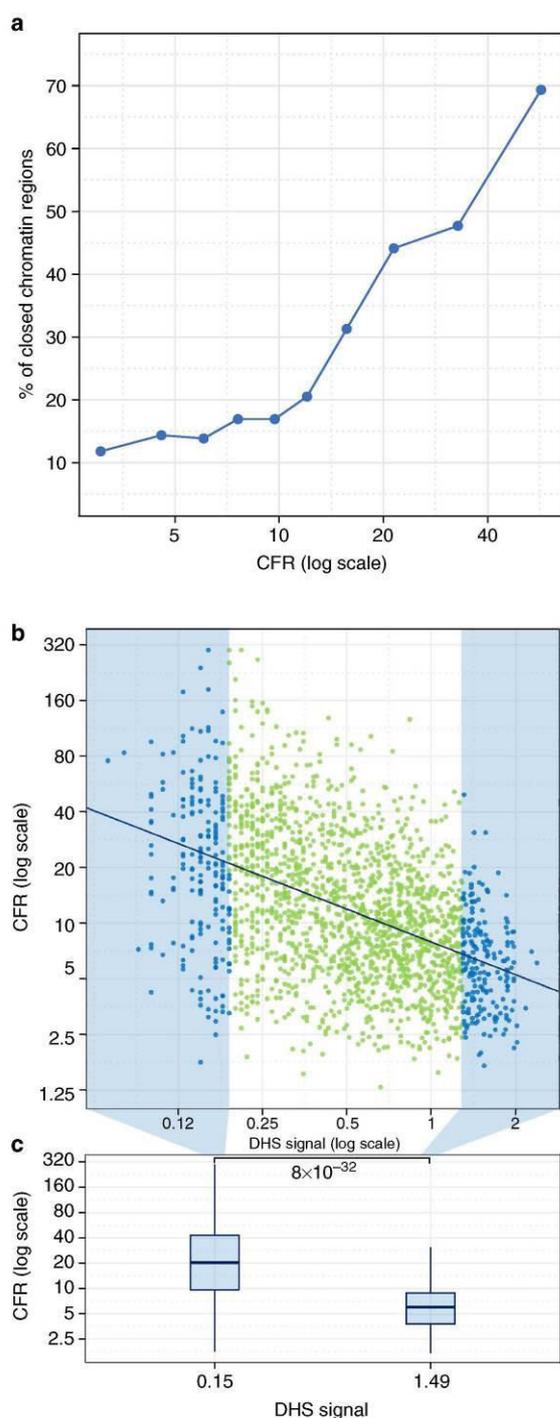
**Plant growth conditions and transformation.** WT, mutants and transgenic lines used are in the *A. thaliana* Columbia-0 background. *A. thaliana* seeds were sown on  $0.5 \times$  Murashige and Skoog (Duchefa biochemie) basal salt mixture medium. Plates were stratified 3 days at 4 °C, grown at 22 °C under long-day conditions and seedlings were transferred to soil. *Lfy-12* mutation genotyping was performed as described<sup>35</sup>. *A. tumefaciens* C58C1 pMP90 was used for stable transformation of *lfy-12* heterozygous plants by the floral dip method<sup>53</sup>.

**Plasmid constructions for protein expression in *Escherichia coli*.** Plasmids were constructed using primers listed in the Supplementary Table 3. GbLFYA (*pETH164*; residues 55–402 from *G. biloba* LFY complementary DNA) was amplified with primers oETH1067 and oETH1068, cloned into the PCR-Blunt vector (Invitrogen) to yield the pCA04 vector and transferred to the *pETM-11* expression vector using NcoI and XhoI restriction sites. GbLFY-SAM (*pETH195*; residues 54–159 from *G. biloba* LFY cDNA) and AtLFY-SAM (*pETH201*; residues 38–151 from *A. thaliana* LFY cDNA) were amplified with the primers oETH1126/oETH1128 and oETH1130/oETH1131, respectively. The PCR products were cloned into the *pETM-11* expression vector<sup>54</sup> using NcoI and XhoI restriction sites, generating fusions with an N-terminal  $6 \times$  His tag, cleavable by the tobacco etch virus (TEV) protease. AtLFY (*pETH94*; residues 1–420 from *A. thaliana* LFY cDNA) was amplified with the primer oETH1031/oETH1032, cloned into the PCR-Blunt vector and transferred into the *pET-30a* vector (Novagen) (C-terminal  $6 \times$  His tag) using NcoI and XhoI restriction sites.

Mutagenesis was done by site-directed mutagenesis with primers listed in the Supplementary Table 4. GbLFY $\Delta_{TE}$  (*pCA21*), GbLFY $\Delta_{RE}$  (*pCA22*) and GbLFY $\Delta_{TERE}$  (*pCA39*) were derived from *pETH164*; AtLFY $\Delta_{TE}$  (*pCA23*), AtLFY $\Delta_{RE}$  (*pCA24*) and AtLFY $\Delta_{TERE}$  (*pCA25*) were derived from *pETH94*; GbLFY-SAM $\Delta_{TE}$  (*pCA15*), GbLFY-SAM $\Delta_{RE}$  (*pCA17*) and GbLFY-SAM $\Delta_{TERE}$  (*pCA20*) were derived from *pETH195*. All plasmids were verified by appropriate digestions or sequencing.

**Binary vector constructions for *A. thaliana* transformation.** *pLFY:LFY<sub>TERE</sub>* (*pCA35*) and *p35S:LFY<sub>TERE</sub>* (*pCA29*) were obtained by inserting a PstI-SalI fragment containing the TERE mutations from *pCA25* in the *pETH29* and *pCA26* (ref. 15), respectively.

**Protein expression and purification.** Proteins were expressed using *E. coli* Rosetta2 (DE3) strain (Novagen). Cells were grown in Luria-Bertani medium supplemented with Kanamycin ( $50 \mu\text{g ml}^{-1}$ ) and Chloramphenicol ( $34 \mu\text{g ml}^{-1}$ ) at 37 °C under agitation up to an optical density of 600 nm of 0.6. Betaine (2 mM) was added and cultures were shifted to 17 °C for 1 h before addition of 0.4 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside. After overnight growth at 17 °C, cells were pelleted. Production of selenomethionine (SeMet) GbLFY-SAM for crystallography was carried out in *E. coli* B834 (DE3) (*met*<sup>-</sup>) strain (Novagen) transformed with the GbLFY-SAM plasmid. The growth of bacteria was initiated in 0.5 l of Luria-Bertani medium supplemented with Kanamycin ( $50 \mu\text{g ml}^{-1}$ ) and Chloramphenicol ( $34 \mu\text{g ml}^{-1}$ ) at 37 °C under agitation up to an optical density of



**Figure 6 | Effect of the chromatin state on the genome-wide LFY DNA binding in planta.** (a) Percentage of closed regions in deciles of *p35S:LFY*-bound regions sorted according to their CFR. (b) Comparison of the CFR of the LFY-bound regions in *p35S:LFY* and the chromatin accessibility in 15-day-old seedlings (DHS signal)<sup>39</sup>. Pearson's correlation test was performed between the CFR and the DHS signal:  $r^2 = 21.71\%$ ,  $P$ -value =  $7.e - 106$ . Data are shown in light blue and green, and fitted values are shown in dark blue. (c) Boxplot representing the CFR signal in the regions of highest and lowest DHS (first and last deciles).  $P$ -value for Mann-Whitney rank test on the two extreme groups is indicated on the graph.

600 nm of 0.5. Bacteria were harvested by centrifugation and washed with M9 minimal medium. The growth was continued at 37 °C until an optical density of 600 of 0.6 in 1 litre of SeMet buffer (2 × M9 minimal medium + MgSO<sub>4</sub> (2 mM), FeSO<sub>4</sub> (25 mg ml<sup>-1</sup>), glucose (4 g l<sup>-1</sup>), vitamins (thiamine, pyridoxine, riboflavin and niacinamide at 1 mg ml<sup>-1</sup>), a mix of all amino acids, except methionine (40 mg ml<sup>-1</sup>), SeMet (40 mg ml<sup>-1</sup>) and antibiotics (34 µg ml<sup>-1</sup> of Chloramphenicol and 50 µg ml<sup>-1</sup> of Kanamycin). The expression of the SeMet protein and growth of the cells at 17 °C were further performed as described above for the GbLFY-SAM protein. All purification steps were performed at 4 °C. All proteins were solubilized and first purified by an affinity chromatography column. The buffer composition used for each protein is given in Supplementary Table 5. Pellets corresponding to 0.5 l culture containing the recombinant protein were sonicated in 50 ml buffer supplemented by one protease inhibitor cocktail tablet Complete EDTA-free (Roche) and centrifuged for 30 min at 20,000 g. The clear supernatant was transferred on a column containing 1 ml Ni-Sepharose High Performance resin (GE Healthcare), washed with buffer containing 20 and 40 mM imidazole, and eluted with buffer containing 300 mM imidazole. Eluted fractions were immediately diluted three times in buffer without imidazole and dialysed overnight.

For GbLFY-SAM, GbLFY-SAM<sub>TE</sub>, GbLFY-SAM<sub>LE</sub>, GbLFY-SAM<sub>TERE</sub> and AtLFY-SAM, after dialysis, the 6 × His tag was cleaved overnight by the TEV protease (5% w/w). The cleavage product was loaded on a Ni-Sepharose High Performance resin to remove the 6 × His tag, the non-cleaved protein and the 6 × His-tagged TEV protease. The protein was eluted with buffer A (Supplementary Table 5) plus 20 mM imidazole. Eluted fractions were applied to a Hi-load Superdex-200 16/60 Prep Grade Column (GE Healthcare) equilibrated with buffer A. After concentration using Amicon Ultra Centrifugal filters (Millipore), the protein concentration was assessed using a NanoDrop-2000 spectrophotometer (Thermo Fisher Scientific Inc.). For crystallography, the SeMet GbLFY-SAM was purified as the native protein.

After dialysis in buffer B, GbLFYA, GbLFYA<sub>TE</sub>, GbLFYA<sub>RE</sub> and GbLFYA<sub>TERE</sub> proteins were concentrated using Amicon Ultra Centrifugal filters and applied to a Hi-load Superdex-200 16/60 Prep Grade Column equilibrated with buffer C. Protein concentration was assessed using a NanoDrop-2000 spectrophotometer.

After dialysis in buffer D, AtLFY, AtLFY<sub>TE</sub>, AtLFY<sub>RE</sub> and AtLFY<sub>TERE</sub> proteins were concentrated using Amicon Ultra Centrifugal filters. Protein concentrations were assessed using the Bradford assay<sup>55</sup>. All proteins were flash frozen in liquid nitrogen and stored at -80 °C.

**Crystal structure determination.** Initial crystallization conditions were identified using the high-throughput crystallization platform at EMBL Grenoble (embl.fr/htdab). The optimum condition was obtained at 20 °C with the hanging drop vapour diffusion method by mixing 1 µl GbLFY-N, native or Se-Met substituted, at 2.5 mg ml<sup>-1</sup> with 1 µl of reservoir solution containing 25 mM Tris-HCl pH 8.8 and 40 mM ammonium sulfate. Crystals were cryoprotected by plunging into liquid nitrogen, after incubation in well conditions to which 20% glycerol had been supplemented. The structure was solved by the single-wavelength anomalous dispersion method from Se-Met substituted protein. A highly redundant data set was collected on the microfocous beamline ID23-2 (ref. 56). The selenium substructure was determined by SHELXC/D<sup>57</sup> and the selenium positions were refined in SHARP<sup>58</sup>. Phases were calculated, solvent-flattened maps were obtained with PIRATE<sup>59</sup> and a partial model was automatically built with BUCCANEER<sup>60</sup>. The model was manually extended into the experimental electron density map and the resultant model was then used as a search model for molecular replacement into the higher-resolution native data (preserving and extending the reflections previously selected for the free set). Successive rounds of model improvement in COOT<sup>61</sup> and refinement in BUSTER<sup>62</sup> were then performed. The built region in the two monomers present in the asymmetric unit comprised residues 54–134 and 56–133, respectively.

**Western blotting.** *A. thaliana* leaves from 15-day-old seedlings grown under long-day conditions were harvested and one leaf from six to eight seedlings were pooled for each transgenic line. The leaves were flash frozen in liquid nitrogen, mechanically crushed, immediately suspended in 80 µl denaturing buffer (50 mM Tris-HCl pH 6.8, 10% w/v glycerol, 1% w/v SDS, 0.0025% w/v Bromophenol blue and 0.4% w/v dithiothreitol) for 10 mg of fresh matter and denatured for 5 min at 95 °C. The samples were run on a SDS-PAGE gel and transferred on a nitrocellulose membrane (Immobilon P transfer membrane, Millipore) for blotting. The equivalent of 18.75 and 12.5 mg of fresh matter were used for anti-LFY and anti-KARI blots, respectively. Five nanograms of recombinant LFY protein (LFYA, produced in *E. coli*) were used as a positive control. Two anti-LFY antibodies were used: JA70 raised in rabbit against the recombinant LFY protein<sup>17</sup> (used at 1/6,000 dilution in Supplementary Fig. 3a) and 4028 (ref. 17) raised in rabbit against the LFY C-terminal amino acids 223–424 (BioGenes). Anti-KARI antibodies (used at 1/3,000) were described in ref. 63. Horseradish peroxidase-conjugated AffiniPure goat anti-rabbit IgG (used at 1/15,000, 111-035-144, Jackson ImmunoResearch) was used as secondary antibody for detection. For Supplementary Fig. 3b, antibodies J70 and 4028 were used at 1/25,000 followed by the secondary antibody at 1/50,000 dilution. Revelation was done using the Pierce ECL2 Western Blotting Substrate kit (Thermo Scientific) and scanned on a Typhoon 9400 scanner

(Molecular Dynamics). Uncropped western blottings and protein ladders are presented as Supplementary Fig. 10.

**Protein sequence and structure analysis.** Three-dimensional structure images were done using PyMOL (www.pymol.org); sequence alignment was done using Multalin (http://multalin.toulouse.inra.fr/multalin/) and visualized using ESPrit (http://espriti.ibcp.fr/ESPrIt/ESPrIt/). Root mean squared deviations were computed using the Dali server (http://ekhidna.biocenter.helsinki.fi/dali\_server/). Surface conservation was computed using the ConSurf server (http://consurf.tau.ac.il/; (ref. 64)).

**Electrophoretic mobility shift assay.** Oligonucleotides used for EMSA are listed in Supplementary Table 6. For *API*, *API1*, *API2*, *AGI-II* and *S-AGI*, complementary single-stranded oligonucleotides were annealed in annealing buffer (10 mM Tris pH 7.5, 150 mM NaCl and 1 mM EDTA). The resulting double-stranded DNA with a protruding G was fluorescently labelled by end filling: 4 pmol of double-stranded DNA was incubated with 1 unit of Klenow fragment polymerase (Ozyme) and 8 pmol Cy3-dCTP or Cy5-dCTP (GE Healthcare) in Klenow buffer during 2 h at 37 °C, followed by 10 min enzyme inactivation at 65 °C. Binding reactions were performed in 20 µl binding buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% glycerol, 0.25 mM EDTA, 2 mM MgCl<sub>2</sub>, 0.01% Tween-20 and 3 mM TCEP) with 10 nM labelled probe, 1 × (28 ng ml<sup>-1</sup>) fish sperm DNA (Roche) as nonspecific competitor and 25–500 nM proteins.

Competition assays were performed in duplicates and 1–100 × fish sperm DNA (Roche) was used in the binding reaction. Signal quantification was performed in using ImageLab v2.0.1 (Bio-Rad Laboratories). Signal of each protein-DNA complex was quantified relatively to total DNA signal. For Fig. 4e, each binding reaction was performed in triplicate. Uncropped gels are presented as Supplementary Fig. 10.

**SEC and SEC-MALLS.** The molecular mass of GbLFY-SAM and AtLFY-SAM was estimated at 4 °C using a Superdex-200 10/300GL column (GE Healthcare), equilibrated with buffer A and calibrated with low- and high-molecular-weight protein standards (gel filtration calibration kit; GE Healthcare). Accurate molecular mass determination using SEC-MALLS was carried out with a Superdex-200 10/300GL column (GE Healthcare). GbLFY-SAM, WT and mutants were analysed in buffer A. Protein-DNA complexes containing GbLFYA, WT or mutants and *API* DNA were analysed in 20 mM Tris-HCl pH 8, 150 mM NaCl, 0.25 mM EDTA, 2 mM MgCl<sub>2</sub> and 1 mM TCEP. Separations were performed at 20 °C with a flow rate of 0.5 ml min<sup>-1</sup>. Elutions were monitored by using a DAWN-EOS detector with a laser emitting at 690 nm for online MALLS measurement (Wyatt Technology Corp., Santa Barbara, CA) and with a RI2000 detector for online refractive index measurements (Schambeck SPD). Molecular mass calculations were performed using the ASTRA software using a refractive index increment (dn/dc) of 0.185 ml g<sup>-1</sup>.

**Chromatin immunoprecipitation sequencing.** Fifteen-day-old seedling from lines CA26 #15 (*p35S:LFY*), CA29 #17 (*p35S:LFY<sub>TERE</sub>*) and Col-0 were grown on 0.5 × Murashige and Skoog medium under long-day conditions. The experiment was performed twice to produce biological replicates. Seedlings (~1 g) were fixed with 1% formaldehyde in MC buffer (10 mM sodium phosphate pH 7.0, 50 mM NaCl and 0.1 M sucrose) for 1 h under vacuum. Fixation was stopped with 0.125 M glycine, followed by three washes with MC buffer. The tissue was ground in liquid nitrogen, the powder was suspended in 15 ml M1 buffer (10 mM sodium phosphate pH 7.0, 0.1 M NaCl, 1 M 2-methyl 2,4-pentanediol, 10 mM β-mercaptoethanol, Complete Protease Inhibitor Cocktail (Roche Diagnostics GmbH, Mannheim, Germany)). The slurry was filtrated three times through 55-µm mesh (Miracloth, Calbiochem) and centrifuged at 1,000 g for 10 min at 4 °C. Filtration and centrifugation were repeated twice. Subsequent steps were at 4 °C, unless indicated otherwise. The pellet was washed five times with 1 ml M2 buffer (M1 buffer with 10 mM MgCl<sub>2</sub> and 0.5% Triton X-100), once with 1 ml M3 buffer (M1 without 2-methyl 2,4-pentanediol) and resuspended in 1 ml Sonic buffer (10 mM sodium phosphate pH 7.0, 0.1 M NaCl, 0.5% Sarkosyl, 10 mM EDTA, Complete Protease Inhibitor Cocktail and 1 mM PEPA BLOC SC (Roche Diagnostics)), and sonicated with a Focused-ultrasonicator S2 (Covaris) (Duty cycle: 20%, intensity: 5, cycles per burst: 200, cycle time: 2 min). After sonication, the suspension was centrifuged (microcentrifuge, top speed) for 5 min and the supernatant was mixed with one volume of IP buffer (50 mM Hepes pH 7.5, 150 mM KCl, 5 mM MgCl<sub>2</sub>, 10 µM ZnSO<sub>4</sub>, 1% Triton X-100 and 0.05% SDS). The solubilized chromatin was incubated overnight with 2.5 µl anti-LFY serum antibody 4028. The immunoprotein-chromatin complexes were captured by incubation with protein A-agarose beads (Santa Cruz Biotechnology) for 1 h on a rotating wheel. The protein A-agarose beads were washed five times with 1 ml IP buffer for 10 min at room temperature. Beads were vortexed for 30 s with 100 µl cold glycine elution buffer (0.1 M glycine, 0.5 M NaCl and 0.05% Tween-20 pH 2.8) and pelleted in a microfuge (room temperature, 1 min, top speed). The supernatant was mixed with 50 µl of 1 M Tris pH 9, to neutralize the eluant. Elution and neutralization were repeated twice. The eluted sample was spun at top speed for 2 min in a microcentrifuge at room temperature. The eluate was treated with 1 µl RNase A/T1

mix 10 mg ml<sup>-1</sup> (Fermentas) and 1.5 µl Proteinase K, recombinant, PCR grade (Roche). After overnight incubation, a second aliquot of Proteinase K was added and incubated at 65 °C for 6 h. DNA was purified using Minelute columns (Qiagen). Libraries for high-throughput sequencing were prepared using standard Illumina protocols. Deep sequencing was performed on an Illumina GAIIx instrument following the manufacturer's instructions.

**Bioinformatics analyses.** Bioinformatics analyses were performed using the SHORE suite (version 0.9.3) (ref. 65) and scripts in python (version 2.7.3) using BioPython (version 1.6.0), pybedtools (version 0.6.4) and Matplotlib (version 1.2.0) modules. All statistical analyses were managed using R (version 3.0.2) and figures were generated using the ggplot2 module (version 1.0.0). All scripts are available on request.

Read mapping and duplicates removal were performed using the SHORE pipeline (preprocess, import, mapflowcell subprocesses) using -H 1,1 -M 0,4 -X 130 -B 1 as parameters. The number of unique reads mapped is indicated in the Supplementary Table 7. Peakfinding was performed with the shore peak subprogramme using -B 1 as parameter. We selected as bound regions all peaks with a BH-FDR- $q < 10^{-7}$  in all four comparisons of the two replicates against the two controls. A python script was used to build bound peak bed files from SUMMARY.txt shore peak output. We obtained 1,954 peaks for *p35S:LFY* and 176 for *p35S:LFY<sub>TERE</sub>*. Corresponding bed files with peak rank and CPR values have been uploaded on the GEO database.

For each genomic position, the read number corresponds to the number of reads extended by 130 bp covering this position. For each bound region (ChIP-Seq peak), the coverage is defined as the area above the read number curve.

The normalization procedure was performed in two steps based on peak coverage comparisons. In *p35S:LFY* and *p35S:LFY<sub>TERE</sub>* we used peaks detected by SHORE, and in the Col control we used a python script as rudimentary peak finder using 20 as a cutoff value for signal (Supplementary Fig. 6). First, we performed intra genotype normalization between ChIP-Seq replicates from the same genotype (Col, *p35S:LFY* or *p35S:LFY<sub>TERE</sub>*). For each genotype, we used the most significant peaks and plotted the coverage values of one replicate against the other. The regression coefficient  $m$  as shown in Supplementary Fig. 6 was used to normalize one replicate against the other.

Replicates were then fused by calculating a normalized read count at each position

$$\text{intraNormReadCount} = \frac{1}{2} * \left( \text{readCount}_{\text{sample1}} + \frac{\text{readCount}_{\text{sample2}}}{m} \right) \quad (1)$$

Second, inter genotype normalization was performed based on the background peaks detected in the Col control sample. For each background peak, coverages were compared between normalized *p35S:LFY* or *p35S:LFY<sub>TERE</sub>* and the Col control sample (Supplementary Fig. 6b), and the regression coefficient  $m'$  was computed and used for final normalization:

$$\text{NormReadCount} = \text{intraNormReadCount} / m' \quad (2)$$

for *p35S:LFY* and *p35S:LFY<sub>TERE</sub>* ChIP-Seq. All coefficients used are listed in Supplementary Table 8.

To calculate read coverages (based on ChIP-Seq or DnaseI-Seq signals), we selected the genomic information included in '.wig' or '.bedgraph files' at bound regions using the bedtools intersect subprogram (-wa -wb option) followed by a python script that calculates the coverage of each of those genomic regions. The CPR due to the TERE mutation was computed with R using equation (3) on normalized wig files for each bound regions.

$$\text{CPR} = \text{coverage}(p35S : LFY) / \text{coverage}(p35S : LFY_{TERE}) \quad (3)$$

When computing CPR for the 1,954 LFY-bound regions (Fig. 5a), the very low coverage in *p35S:LFY<sub>TERE</sub>* for two regions was leading to very high or infinite CPR values. Those two CPR values were arbitrarily set up to 300.

The chromatin information about chromatin state (open versus closed) for each bound region was retrieved with a python script from *dhleaf.txt* and *dhflower.txt* files from ref. 39. A bound region was considered open when its centred half was entirely open.

To find the LFYBs in ChIP-Seq peaks, we used a previously optimized position weight matrix<sup>17</sup> and python script computing a score between -56 and 0 to each 19-pb sequence. Zero corresponds to the site of highest affinity and -56 corresponds to the lowest affinity. To determine what score threshold should be used for binding-site identification, we compared bound regions with a negative set of unbound regions. The negative set was built with a python script that takes a bed file with bound regions as input and randomly selects in the *Arabidopsis* genome regions of same size, similar GC content and with various annotation type (CDS, intron, upstream, downstream or intergenic). The LFYBs scores for each bound and unbound regions were calculated. We then used an R script to compute the density of LFYBs in sliding score windows (going from [0; -1] to [-55; -56]). As shown in Supplementary Fig. 8a, the density of LFYBs is significantly higher in the positive set as compared with the negative one for scores better than -25, as judged using a Mann-Whitney rank test on each window of scores ( $P$ -value < 0.05). The corresponding binding site density might appear surprisingly high but we believe that our score threshold calculation is more valid than

commonly used arbitrary cutoff score values. Moreover, our analyses were robust to threshold change within a reasonable range [-20, -26].

For the heat map of interdistances (Fig. 5e), we used a python script to divide regions in deciles according to their CPR, calculate all distances between LFYBs and plot their number within the decile. To perform the analysis presented in Supplementary Fig. 9, we kept only the 436 regions found as bound in *35S:LFY* seedlings that intersected with regions bound by LFY in inflorescences according to ref. 18 and performed the whole pipeline of analysis. R was used for all figures and statistical analyses, except that Supplementary Fig. 6 was generated by the Matplotlib module of python.

## References

- Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).
- Shlyueva, D., Stampfel, G. & Stark, A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272–286 (2014).
- Weirauch, M. T. & Hughes, T. R. in *A Handbook of Transcription Factors* **52**, 25–73 (2011).
- Slattery, M. *et al.* Absence of a simple code: how transcription factors read the genome. *Trends Biochem. Sci.* **39**, 381–399 (2014).
- Bienz, M. Signalosome assembly by domains undergoing dynamic head-to-tail polymerization. *Trends Biochem. Sci.* **39**, 487–495 (2014).
- Hollenhorst, P. C., McIntosh, L. P. & Graves, B. J. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu. Rev. Biochem.* **80**, 437–471 (2011).
- Nanao, M. H. *et al.* Structural basis for oligomerization of auxin transcriptional regulators. *Nat. Commun.* **5**, 3617 (2014).
- Katsani, K. R., Hajibagheri, M. A. N. & Verrijzer, C. P. Co-operative DNA binding by GAGA transcription factor requires the conserved BTB / POZ domain and reorganizes promoter topology. *EMBO J.* **18**, 698–708 (1999).
- Williams, H., Jayaraman, P.-S. & Gaston, K. DNA wrapping and distortion by an oligomeric homeodomain protein. *J. Mol. Biol.* **383**, 10–23 (2008).
- Moyroud, E., Tichtinsky, G. & Parcy, F. The LEAFY floral regulators in angiosperms: conserved proteins with diverse roles. *J. Plant Biol.* **52**, 177–185 (2009).
- Moyroud, E., Kusters, E., Monniaux, M., Koes, R. & Parcy, F. LEAFY blossoms. *Trends Plant Sci.* **15**, 346–352 (2010).
- Irish, V. P. The flowering of *Arabidopsis* flower development. *Plant J.* **61**, 1014–1028 (2010).
- Chandler, J. W. Floral meristem initiation and emergence in plants. *Cell. Mol. Life Sci.* **69**, 3807–3818 (2012).
- Posé, D., Yant, L. & Schmid, M. The end of innocence: flowering networks explode in complexity. *Curr. Opin. Plant Biol.* **15**, 45–50 (2012).
- Chahtane, H. *et al.* A variant of LEAFY reveals its capacity to stimulate meristem development by inducing RAX1. *Plant J.* **74**, 678–689 (2013).
- Yamaguchi, N. *et al.* A molecular framework for auxin-mediated initiation of flower primordia. *Dev. Cell* **24**, 271–282 (2013).
- Moyroud, E. *et al.* Prediction of regulatory interactions from genome sequences using a biophysical model for the *Arabidopsis* LEAFY transcription factor. *Plant Cell* **23**, 1293–1306 (2011).
- Winter, C. M. *et al.* LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Dev. Cell* **20**, 430–443 (2011).
- Weigel, D. & Nilsson, O. A developmental switch sufficient for flower initiation in diverse plants. *Nature* **377**, 495–500 (1995).
- Gallois, J.-L., Nora, P. R., Mizukami, Y. & Sablowski, R. WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes Dev.* **18**, 375–380 (2004).
- Risseuw, E. *et al.* An activated form of UFO alters leaf development and produces ectopic floral and inflorescence meristems. *PLoS ONE* **8**, e83807 (2013).
- Kaufmann, K., Pajoro, A. & Angenent, G. C. Regulation of transcription in plants: mechanisms controlling developmental switches. *Nat. Rev. Genet.* **11**, 830–842 (2010).
- Engelhorn, J., Bianvillain, R. & Carles, C. C. Gene activation and cell fate control in plants: a chromatin perspective. *Cell. Mol. Life Sci.* **71**, 3119–3137 (2014).
- Sayou, C. *et al.* A promiscuous intermediate underlies the evolution of LEAFY DNA binding specificity. *Science* **343**, 645–648 (2014).
- Hamès, C. *et al.* Structural basis for LEAFY floral switch function and similarity with helix-turn-helix proteins. *EMBO J.* **27**, 2628–2637 (2008).
- Minguet, E. G., Segard, S., Charavay, C. & Parcy, F. MORPHEUS, a webtool for transcription factor binding analysis using position weight matrices with dependency. *PLoS ONE* **10**, e0135586 (2015).
- Siriwardana, N. S. & Lamb, R. S. A conserved domain in the N-terminus is important for LEAFY dimerization and function in *Arabidopsis thaliana*. *Plant J.* **71**, 736–749 (2012).
- Qiao, P. & Bowie, J. U. The many faces of SAM. *Sci. STKE* **2005**, re7 (2005).

29. Meruelo, A. D. & Bowie, J. U. Identifying polymer-forming SAM domains. *Proteins* **74**, 1–5 (2009).
30. Kim, C. a *et al.* Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *EMBO J.* **20**, 4173–4182 (2001).
31. Ikeda-Kawakatsu, K., Maekawa, M., Izawa, T., Itoh, J.-I. & Nagato, Y. ABERRANT PANICLE ORGANIZATION 2/RPL, the rice ortholog of *Arabidopsis* LEAFY, suppresses the transition from inflorescence meristem to floral meristem through interaction with APO1. *Plant J.* **69**, 168–180 (2012).
32. Levin, J. Z. & Meyerowitz, E. M. UFO: an *Arabidopsis* gene involved in both floral meristem and floral organ development. *Plant Cell* **7**, 529–548 (1995).
33. Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F. & Meyerowitz, E. M. LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* **69**, 843–859 (1992).
34. Green, S. M., Coyne, H. J., McIntosh, L. P. & Graves, B. J. DNA binding by the ETS protein TEL (ETV6) is regulated by autoinhibition and self-association. *J. Biol. Chem.* **285**, 18496–18504 (2010).
35. Benlloch, R. *et al.* Integrating long-day flowering signals: a LEAFY binding site is essential for proper photoperiodic activation of APETALA1. *Plant J.* **67**, 1094–1102 (2011).
36. Zhang, J. *et al.* Sterile alpha motif domain-mediated self-association plays an essential role in modulating the activity of the *Drosophila* ETS family transcriptional repressor Yan. *Mol. Cell Biol.* **30**, 1158–1170 (2010).
37. Isono, K. *et al.* SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. *Dev. Cell* **26**, 565–577 (2013).
38. Webber, J. L. *et al.* The relationship between long-range chromatin occupancy and polymerization of the *Drosophila* ETS family transcriptional repressor Yan. *Genetics* **193**, 633–649 (2013).
39. Zhang, W., Zhang, T., Wu, Y. & Jiang, J. Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. *Plant Cell* **24**, 2719–2731 (2012).
40. Sullivan, A. M. *et al.* Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana*. *Cell Rep.* **8**, 2015–2030 (2014).
41. Maizel, A. *et al.* The floral regulator LEAFY evolves by substitutions in the DNA binding domain. *Science* **308**, 260–263 (2005).
42. Cherry, J. L. & Adler, P. R. How to make a biological switch. *J. Theor. Biol.* **203**, 117–133 (2000).
43. Lebrecht, D. *et al.* Bicoid cooperative DNA binding is critical for embryonic patterning in *Drosophila*. *Proc. Natl Acad. Sci. USA* **102**, 13176–13181 (2005).
44. Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. & Gaul, U. Predicting expression patterns from regulatory sequence in *Drosophila* segmentation. *Nature* **451**, 535–540 (2008).
45. Blázquez, M. A., Soowal, L. N., Lee, I. & Weigel, D. LEAFY expression and flower initiation in *Arabidopsis*. *Development* **124**, 3835–3844 (1997).
46. Sherwood, R. I. *et al.* Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* **32**, 171–178 (2014).
47. Voss, T. C. & Hager, G. L. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat. Rev. Genet.* **15**, 69–81 (2014).
48. Guertin, M. J. & Lis, J. T. Mechanisms by which transcription factors gain access to target sequence elements in chromatin. *Curr. Opin. Genet. Dev.* **23**, 116–123 (2013).
49. Iwafuchi-doi, M. & Zaret, K. S. Pioneer transcription factors in cell reprogramming. *Cell* **158**, 2679–2692 (2014).
50. Wu, M.-P. *et al.* SWI2/SNF2 chromatin remodeling ATPases overcome polycomb repression and control floral organ identity with the LEAFY and SEPALLATA3 transcription factors. *Proc. Natl Acad. Sci. USA* **109**, 3576–3581 (2012).
51. Smaczniak, C. *et al.* Characterization of MADS-domain transcription factor complexes in *Arabidopsis* flower development. *Proc. Natl Acad. Sci. USA* **109**, 1560–1565 (2012).
52. Pajoro, A. *et al.* Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome Biol.* **15**, R41 (2014).
53. Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W. & Chua, N.-H. Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* **1**, 641–646 (2006).
54. Dümmler, A., Lawrence, A.-M. & de Marco, A. Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb. Cell Fact.* **4**, 34 (2005).
55. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
56. Plot, D. *et al.* The ID23-2 structural biology microfocus beamline at the ESRF. *J. Synchrotron Radiat.* **17**, 107–118 (2010).
57. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 479–485 (2010).
58. Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 2023–2030 (2003).
59. Cowtan, K. Generic representation and evaluation of properties as a function of position in reciprocal space. *J. Appl. Crystallogr.* **35**, 655–663 (2002).
60. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
61. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
62. Bricogne, G. *et al.* BUSTER. Available at <http://www.globalphasing.com> (2011).
63. Dumas, R., Joyard, J. & Douce, R. Purification and characterization of acetoxyhydroxyacid reductoisomerase from spinach chloroplasts. *Biochem. J.* **262**, 971–976 (1989).
64. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533 (2010).
65. Ossowski, S. *et al.* Sequencing of natural strains of *Arabidopsis thaliana* with short reads. *Genome Res.* **18**, 2024–2033 (2008).
66. Sathyamurthy, A., Freund, S. M. V., Johnson, C. M., Allen, M. D. & Bycroft, M. Structural basis of p63 $\alpha$  SAM domain mutants involved in AEC syndrome. *FEBS J.* **278**, 2680–2688 (2011).

### Acknowledgements

We thank E. Moyroud for the gift of GbLFY cDNA; A. Mathelier, A. Fejes, R. Worsley-Hunt, W.W. Wasserman, C. Petosa, B. Ycart and P. Lauß for helpful discussions; and C. Zubieta, J. Zubieta, J. Engelhorn, C. Carles and J. Govin for critical reading of the manuscript. This work was supported by a DFG-funded ERA-PG grant (BloomNet; SCHM1560/7-1) to M.S., the ANR (Blanc – SVSE2 – 2011- Charmful) and CEA NanoProt3D programme to F.P. and R.D., the iRTeLIS CEA PhD programme to C.S. and the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement numbers 227764 (P-CUBE) and 283570 (BioStruct-X) to R.D. This work used the platforms of the Grenoble Instruct Center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

### Author contributions

C.S., R.D., M.S. and F.P. designed the experiments. C.S., M.P.L., E.T., G.D., M.J. and R.D. performed the biochemical experiments. M.H.N., C.S. and R.D. performed the crystallographic experiments. C.S. and G.T. performed plant experiments. D.P. and M.S. performed the ChIP experiment. G.D., L.G., F.O. and F.P. performed the bioinformatics analyses. All authors interpreted the results. F.P. wrote the paper with the help of C.S., R.D., M.H.N. and M.S., and input from all authors.

### Additional information

Accession codes: The GbLFY-N structure has been deposited in the RCSB with code 4UDE. The ChIP-Seq data have been submitted to the NCBI GEO database with accession number GSE64245.

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Sayou, C. *et al.* A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nat. Commun.* **7**:11222 doi: 10.1038/ncomms11222 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>





## ANNEXE 3 : EDGER SCRIPT

---

```
#!/usr/bin/env Rscript

# Script is largely taken from the EdgeR training made by Etienne Delannoy and Marie-
Laure Martin-Magniette

library(edgeR)

setwd("D:/Charmfull_RAX1_RNAseq_project/")

# threshold for FDR control
alpha=0.001

# ReadDGE to import count table form mutiple files
dFull <- readDGE(c("46_M1", "46_M2", "46_M3", "46_D2", "46_D3", "46_D4", "48_M1", "48_M2", "48_M3", "48
_D2", "48_D4", "48_D5"),
                path="./Count_tables",
                #s=c(1,2),

group=c("46M", "46M", "46M", "46D", "46D", "46D", "48M", "48M", "48M", "48D", "48D", "48D"))

write.table(dFull$samples, "./DE_analysis/tag_number.csv", sep="\t") # outputs library
size

# Filtering genes that have less than 1 count per million and are expressed in less
than 3 librairies (because 3 rep)
keep <- rowSums(cpm(dFull)>1) >= 3
d <- dFull[keep, , keep.lib.sizes=FALSE]
write.table(dFull$samples, "./DE_analysis/filtered_tag_number.csv", sep="\t") # prints
lib.size table after filtering

# TMM normalization
d <- calcNormFactors(d) # Pairwise TMM normalization, each library is weighted for
total counts
write.table(dFull$samples, "./DE_analysis/normalized_tag_number.csv", sep="\t") #
prints normalization factors

# Plotting multi dimensional scaling plot to see sample similiraties
pdf("./DE_analysis/MDS.pdf")
plotMDS(d, method="bcv", col=as.numeric(d$samples$group))
legend("bottomright", as.character(unique(d$samples$group)), col=1:3, pch=20)
dev.off()

# Design matrix
target <- read.table("./Count_tables/design_matrix.txt", header=TRUE) #import design
matrix
target$factor1 <- relevel(target$factor1, ref="GD48") # set GD48 as ref
target$factor2 <- relevel(target$factor2, ref="control") # set control as ref
d$design <- model.matrix(~ factor1 + factor2 + factor1:factor2, data=target) # setting
multifactor design

# estimating dispersions
d <- estimateGLMCommonDisp(d, d$design)
d <- estimateGLMTrendedDisp(d, d$design)
d <- estimateGLMTagwiseDisp(d, d$design)

# BCV plot
pdf("./DE_analysis/BCV.pdf")
plotBCV(d, "GLM multifactor design")
dev.off()

# statistical testing using GLM test
# fit genewise glms
fit <- glmFit(d, d$design)
colnames(fit$design)

## conduct likelihood ratio test for the treatment effect
##(coefficient = 4) :
lrt <- glmLRT(fit, coef=4)

# display the top ten differentially expressed tags
cat("\nthe top ten differentially expressed tags :\n")
print(topTags(lrt, n=10, adjust.method="BH"))
```

```

# display the total number of differentially expressed genes at FDR < alpha (here alpha
= 0.05)
de <- decideTestsDGE(lrt, adjust.method="BH", p.value=alpha)
cat("\nResults of the differential analysis for a FDR control at 0.05\n")
cat(paste("Number of differentially expressed genes: ",length(which(de[,1]!=0))),"\n")
cat(paste("Number of down-regulated differentially expressed genes:
",length(which(de[,1]==-1))),"\n")
cat(paste("Number of up-regulated differentially expressed genes:
",length(which(de[,1]==1))),"\n")

# p-values histogram
pdf("./DE_analysis/pval.pdf")
hist(lrt$table$PValue, breaks=150, col="skyblue", border="slateblue",
main="Distribution of raw p-values_GLMTest_edgeR",xlab="raw p-value", las=1)
dev.off()

# Plot fold changes
pdf("./DE_analysis/DE_plot.pdf")
detags <- rownames(d)[as.logical(de)]
plotSmear(lrt, de.tags=detags)
abline(h = c(-2, 2), col = "blue")
dev.off()

# Saving results as table
result <- topTags(lrt, n=nrow(lrt$table), adjust.method = "BH", sort.by="logFC")
write.table(result, "./DE_analysis/DE_analysis.txt", sep="\t")

```

## ANNEXE 4 : BIOPYTHON SCRIPT

---

```
#!/usr/bin/env python
# coding: utf-8

"""
Wrapper script for BioPython allowing scanning of sequences with a matrix,
Pocc calculation and standard results output.
Can also runs as a script.
REQUIRES Biopython and seaborn
Outputs a summary file for each sequence
    a summary file with best score and pOcc for every AGI
    a bedgraph file with all the motifs found
"""

### IMPORTS #####
from Bio import SeqIO, motifs, Seq
from Bio.Alphabet import IUPAC
import re, copy, argparse, os
from math import exp
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import random

##### USER CONSTANTS### #####
# Background frequencies
#These were inferred from the total sequences of TAIR10:
# upstream_3000bp + genomic + downstream_3000bp
BASE_FREQUENCY = {'A': 0.310, 'C': 0.190, 'G': 0.190, 'T': 0.310} # seq base frequency

class Score(object):
    """Stores information relative to a single score"""
    def __init__(self, val, posInSeq, posInGenome, strand, motifMatch):
        self.score = val
        self.pos = posInSeq
        self.genomePos = posInGenome
        self.strand = strand
        self.motif = motifMatch

    def toString(self):
        """called by ScoreRecord.__str__"""
        return "%s\t%s\t%s\t%s" % (self.score, self.strand, self.genomePos, self.motif)

    def __str__(self):
        return("""score:                %s\nposition:                %s\nstrand:                %s\nmotif:                %s""" % (self.score, self.pos, self.genomePos, self.strand, self.motif))

class ScoreRecord(dict):
    """Simple class to score a sequence and keep the results.
    Provide link to the PSSM and the SeqRecord.
    Inherits from dict, individual scores can be accessed by positions.
    """
    def __init__(self, pssm, seqRecord):
        """Constructor method.
        Takes a Bio.motifs.PSSM object and a Bio.SeqRecord object as arguments.
        Computes scores with forward and reverse complement matrix and keeps the best
        score at each position."""
        # Checking that the arguments are valid
        try:
            assert isinstance(pssm, motifs.matrix.PositionSpecificScoringMatrix)
            assert isinstance(seqRecord, SeqIO.SeqRecord)
        except AssertionError:
            raise ValueError('ScoreRecord constructor was called with illegal
arguments')
        # Initializing attributes
        self.pssm = pssm
        self.seqRecord = seqRecord
        # changing Seq alphabet format
        seqRecord.seq.alphabet = IUPAC.unambiguous_dna
        # calculating scores
        self.fScores = pssm.calculate(seqRecord.seq)
```

```

# calculating reverse complement scores
self.rScores = pssm.reverse_complement().calculate(seqRecord.seq)
for i in xrange(len(self.fScores)):
# Selecting max score for each position and getting strandness
    score, strand = max((self.fScores[i], '+'), (self.rScores[i], '-'))
    # getting genomic position
    genPos = ScoreRecord.getGenomicPosition(seqRecord, i+1)
    self[i+1] = Score(score, i+1, genPos, strand,
seqRecord.seq[i:i+pssm.length])
    self.threshold = None

def __str__(self):
    out = 'pos\tscore\tstrand\tgenomicPos\tmotif\n'
    for key in self:
        out += str(key) + '\t' + self[key].toString() + '\n'
    return out

def write(self, path):
    """Writes itself to a file.
    returns path"""
    makeFolders(os.path.dirname(os.path.abspath(path)))
    with open(path, 'w') as out:
        out.write('# Sequence: %s\n# Score threshold: %s\n' %(self.seqRecord.id,
self.threshold))
        out.write(self.__str__())
    return path

def filterScore(self, threshold, inplace = False):
    """Filter raw scores based on score threshold.
    returns a new instance of ScoreRecord unless inplace is True
    """
    if inplace:
        for key, val in self.items():
            if val.score < threshold:
                del(self[key])
        self.threshold = threshold
        return
    else:
        new = copy.copy(self)
        new.filterScore(threshold, inplace = True)
        return new

@property
def bestScore(self):
    """returns the best score"""
    bestScore = Score(self.pssm.min, 0, '', '', '')
    try:
        for score in self.values():
            if score.score > bestScore.score:
                bestScore = score
        return bestScore
    except ValueError:
        return bestScore

@property
def pOcc(self, concentration = None):
    """Calculates the probability of occupancy of the sequence form calculated
scores
Based on GOMER algorithm, does not account for competition/cooperation of the
TF.
concentration : (float) defaults to the Kd for optimal site (as in GOMER)
return probability of binding (float)"""
    # get [TF]
    if not concentration:
        concentration = ScoreRecord.calcKd(self.pssm.max)
    p = 1
    # for every score in record
    for score in self.values():
        # increment pNotBound
        p *= ScoreRecord.pNotBound(concentration, score.score)
    # pBound is 1-pNotBound
    return 1 - p

@staticmethod
def pNotBound(concentration, score):
    """Calculates probability of the factor not being bound on a motif
concentration: float, TF concentration

```

```

        score: float, score for the motif
        return p: float"""
        return 1 - (concentration / (concentration + ScoreRecord.calcKd(score)))

    @staticmethod
    def calcKd(score):
        """Calculates Kd from a score
        return Kd, float"""
        return exp(-score)

    @staticmethod
    def getGenomicPosition(seqRecord, posInSeq):
        """Gets the genomic position corresponding to the position in sequence form the
description of SeqRecord.
        return string chrN:startPos-endPos. The length is always 1.
        """
        try:
            chr, start = re.split('\W+', seqRecord.description)[1:3]
            pos = int(start) + posInSeq
            return '%s:%s-%s' %(chr, pos, pos+1)
        except ValueError:
            return

class ScoreCollection(dict):
    """A collection of ScoreRecords, built with a matrix and a sequence dictionary"""
    def __init__(self, pssm, seqDict, threshold = None):
        """Constructor method
        Creates a ScoreRecord for each sequence in the dictionary
        Calculates pOcc and BestScore automatically"""
        for seq in seqGenerator(seqDict):
            self[seq.id] = ScoreRecord(pssm, seq)

    def applyThreshold(self, threshold):
        """Filter scores under threshold across all ScoreRecords
        Filters in place !"""
        for sc in self.values():
            sc.filterScore(threshold, inplace = True)

    def writeSummary(self, path, bestScore = True, pOcc = True):
        """Writes the best score and pOcc of each sequence in a file
        collection: iterable of ScoreRecords
        can toggle bestScore and pOcc on or off
        path: path output file
        return path"""
        makeFolders(os.path.dirname(os.path.abspath(path)))
        header = "id"
        if bestScore:
            header += "\tbestScore\tpos"
        if pOcc:
            header += "\tpOcc"
        header += "\n"
        with open(path, 'w') as out:
            out.write(header)
            for sc in self.values():
                line = sc.seqRecord.id
                if bestScore :
                    line += "\t" + str(sc.bestScore.score) + "\t" +
str(sc.bestScore.pos)
                if pOcc :
                    line += "\t" + str(sc.pOcc)
                out.write(line + "\n")
        return path

    def writeBedGraph(self, path):
        """Writes an unsorted bedGraph file with all sites
        path: path output file
        return path"""
        makeFolders(os.path.dirname(os.path.abspath(path)))
        with open(path, 'w') as out:
            out.write("track type=bedGraph\n")
            for sc in self.values():
                for score in sc.values():
                    if score.genomePos:
                        chrom, start, stop = re.split('\W+', score.genomePos)
                        out.write(chrom + '\t' + start + '\t' + stop + '\t' +
str(score.score) + '\n')
        return path

```

```

def pOccDistribution(self, fileName):
    """Plots a kernel density estimate of pOcc distribution"""
    ScoreCollection._kdePlot(self._pOccList, fileName)

def scoreDistribution(self, fileName):
    """Plots a kernel density estimate of pOcc distribution"""
    ScoreCollection._kdePlot(self._scoreList, fileName)

def writeScoreFiles(self, folder):
    """Write one file per sequence with all the scores"""
    makeFolders(os.path.dirname(os.path.abspath(folder)))
    for sc in self.values():
        path = os.path.join(folder, sc.seqRecord.id + '.txt')
        sc.write(path)

@property
def _pOccList(self):
    l = [sc.pOcc for sc in self.values()]
    return np.nan_to_num(np.array(l))

@property
def _scoreList(self):
    l = [s.score for sc in self.values() for s in sc.values()]
    return np.array(l)

@staticmethod
def _kdePlot(data, fileName):
    # create folders
    makeFolders(os.path.dirname(os.path.abspath(fileName)))
    # plot
    plt.clf()
    sns.set(style="white", palette="muted", color_codes=True)
    ax = sns.kdeplot(data, shade = True, color = "g", linestyle = '-')
    ax.set(xlabel = 'pOcc', ylabel = 'density', yticklabels=[])
    sns.despine(left=True)
    plt.savefig(fileName)

##### MATRIX PROCEDURES #####
def loadMatrix(file):
    """Load a matrix from a JASPAR formatted file.
    Returns PSSM object."""
    # Loading matrix from file
    with open(file) as f:
        m = motifs.parse(f, "jaspar")[0]
    # Conversion to pwm and then to pssm
    pwm = m.counts.normalize(pseudocounts = BASE_FREQUENCY)
    pssm = pwm.log_odds(BASE_FREQUENCY)
    return pssm

def getThreshold(pssm, method = 'fpr', val = 0.01):
    """Gets the threshold for a given matrix.
    method: str of fpr, fnr, balanced or patser, see BioPython doc
    val: float or int, value associated with the method chosen
    return threshold (float)"""
    # calculate score distribution
    distribution = pssm.distribution(background = BASE_FREQUENCY, precision = 10**4)
    # return threshold
    if method == 'patser':
        return distribution.threshold_patser()
    elif method == 'fpr':
        return distribution.threshold_fpr(val)
    elif method == 'fnr':
        return distribution.threshold_fnr(val)
    elif method == 'balanced':
        return distribution.threshold_balanced(val)
    else:
        raise ValueError

### FASTA PROCEDURES #####
def loadFasta(fileList, identifiers, shuffle = False):
    """Reconstruct sequences for identifiers across multiple files.
    search for the Agi in the files provided and append the sequences in the order of
    the list.
    fileList: ordered list of fasta files with sequences. Can provide single file as
    string

```

```

        identifiers: identifier list. Only identifiers in this list will be extracted from
fasta files
return SeqRecord dictionary with identifiers as keys
"""
# Checking file input format
if type(fileList) == str:
    fileList = [fileList]
# Initializing output dictionary
newDict = {}
# Load all sequences as a dictionary
for f in fileList:
    with open(f, 'rU') as handle:
        seqDict = SeqIO.to_dict(SeqIO.parse(handle, 'fasta'))
        for agi in identifiers:
            try:
                if newDict.has_key(agi):
                    newDict[agi].seq += seqDict[agi].seq
            else:
                newDict[agi] = copy.deepcopy(seqDict[agi])
        except KeyError:
            print "Warning: identifier %s not in file %s" %(agi, f)
# Creating shuffled sequences
if shuffle:
    for agi, seq in newDict.items():
        newDict[agi] = shuffleSeq(seq)
return newDict

def seqGenerator(fastaDict):
    """Yields seqRecords contained in a dict"""
    for agi in fastaDict:
        yield fastaDict[agi]

def loadAgi(agiFile):
    """Loads a list of identifiers from a line separated file
return list of agi"""
    return [agi.strip() for agi in open(agiFile, 'rU').readlines()]

def shuffleSeq(seq):
    """Shuffles nucleotides in a SeqRecord object
return shuffled SeqRecord object"""
    l = list(seq.seq)
    random.shuffle(l)
    newSeq = Seq.Seq(''.join(l), IUPAC.unambiguous_dna)
    return SeqIO.SeqRecord(newSeq, id= seq.id + '_shuffled', name = seq.name +
'_shuffled', description = seq.description)

### MAIN PROCEDURE #####
def makeFolders(directory):
    """Creates folders if needed"""
    if not os.path.exists(directory):
        os.makedirs(directory)

def main(matrixFile, fastaFile, targetFile, outputFolder, thresholdMethod = 'patser',
thresholdValue = None):
    """Analyze mutiple sequence

matrixFile: path to the file containing the matrix information in JASPAR format
fastaFile: path to the file containing sequences in fasta format, a list of files
can be provided
targetFile: line delimited file with target identifiers
outputFile: path to output folder, not file !
thresholdMethod: {'patser', 'fpr', 'fnr', 'balanced', 'value', 'none'}
                see biopython documentation
                value allows user-defined threshold
                none does not filter scores
thresholdValue: value associated to threshold calculation method (see biopython)
"""
    # Asserting threshold arguments
    if thresholdMethod not in ['patser', 'fpr', 'fnr', 'balanced', 'value', 'none']:
        raise ValueError('Illegal threshold method: %s' %thresholdMethod)
    if thresholdMethod in ['patser', 'fpr', 'fnr', 'balanced', 'value']:
        if not thresholdValue:
            raise ValueError('Please provide a threshold for %s method'
%thresholdMethod)
        elif type(thresholdValue) not in [int, float]:
            raise TypeError('Threshold value must be of type int or float')
    # Opening files

```

```

try:
    pssm = loadMatrix(matrixFile)
    agiList = loadAgi(targetFile)
    seqDict = loadFasta(fastaFile, agiList)
except IOError, e:
    raise IOError('Unable to locate file %s' %e)
# Calculating threshold
if thresholdMethod != 'none':
    if thresholdMethod == 'value':
        threshold = thresholdValue
    else:
        threshold = getThreshold(pssm, method = thresholdMethod, val =
thresholdValue)
else:
    threshold = None
# Making the collection and outputting results
coll = ScoreCollection(pssm, seqDict)
if threshold:
    coll.applyThreshold(threshold)
coll.writeScoreFiles(outputFolder)
coll.writeSummary(os.path.join(outputFolder, 'summary.txt'))
coll.writeBedGraph(os.path.join(outputFolder, 'binding_sites.bedgraph'))

### SCRIPT #####
if __name__ == '__main__':
    parser = argparse.ArgumentParser(description = "Finds TFBS in a set of sequences
for a set of AGI")
    parser.add_argument('-m', '--matrix', required = True, help = 'Matrix file in JASPAR
format')
    parser.add_argument('-f', '--fasta', nargs = '+', required = True, help = 'Fasta
file containing sequences. Multiple file can be given, sequences will be concatenated
in the order of the files')
    parser.add_argument('-t', '--targets', required = True, help = 'Line delimited file
containing AGI of targets')
    parser.add_argument('-o', '--output', nargs = '?', default = 'TFBS_output', help =
'Folder to output the analysis')
    parser.add_argument('--threshold_method', nargs = '?', default = 'patser', type =
str, choices = ['fpr', 'fnr', 'balanced', 'patser', 'defined', 'none'], help = 'Method
for threshold determination (see BioPython for ref), defined allow a user-defined
threshold and none will find all scores')
    parser.add_argument('--threshold_value', type = float, help = 'Value associated to
threshold method (see BioPython doc)')
    args = parser.parse_args()
    main(args.matrix, args.fasta, args.targets, args.output, args.threshold_method,
args.threshold_value)

```