Numéro National de Thèse : 2016LYSEN049

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée par
l'Ecole Normale Supérieure de Lyon

Ecole Doctorale N° 340
Biologie Moléculaire, Intégrative et Cellulaire

Spécialité de doctorat : Développement végétal
Discipline : Sciences de la vie

Soutenue publiquement le 02/12/2016, par :
Samuel COLLAUDIN

Exploration de l'origine de la robustesse de la dynamique d'expression d'*AGAMOUS* pendant le développement de la fleur en utilisant une approche pluridisciplinaire

Exploring the basis of robust *AGAMOUS* expression dynamics during flower development using a pluridisciplinary approach

---

Devant le jury composé de :

Pr. Henrik JÖNSSON, Professeur, SLCU (Cambridge), Rapporteur
Dr. Cristina FERRANDIZ, Chargée de recherche, IBMCP (Valence), Rapporteure
Dr. Olivier GANDRILLON (DR), Directeur de recherche, LBMC – ENS de Lyon (Lyon), Examinateur
Dr. Christophe GODIN (DR), Directeur de recherche, INRIA (Montpellier), Examinateur
Dr. François PARCY (DR), Directeur de recherche, CEA (Grenoble), Examinateur
Pr. Anna MARCINIAK-CZOCHRA, Professeure, IWR (Heidelberg), Examinatrice
Dr. Pradeep DAS, Maitre de conference, RDP – ENS de Lyon (Lyon), Directeur

Résumé

Les plantes terrestres ont une biomasse estimée 1000 fois plus grande que celle des animaux. Les angiospermes, ou plantes à fleurs, est le groupe de plantes prédominant avec 90% des espèces de plantes sur terre. Avec leur architecture fortement conservée et leur rôle crucial dans divers aspects de l'existence humaine, ils forment un important sujet de recherche. L'identité des organes floraux est définie par l'expression de gènes homéotiques appartenant à la famille des MADS-box au début du développement floral. Un de ces gènes, *AGAMOUS* (*AG*), est responsable de l'identité des étamines et des carpelles chez *Arabidopsis thaliana*. Dans ce manuscrit, je tente de comprendre les propriétés spatiales et temporelles de l'expression d'AG en cherchant à connaître les mécanismes impliqués dans le bon établissement de la dynamique d'expression d'*AG* pendant les jeunes stades du développement floral.

Je débute par développer un modèle de réaction-diffusion qui prend en compte la croissance de la fleur pendant les stades d'intérêt, ainsi que quelques facteurs de transcriptions clefs impliqués dans la régulation d'*AG*, tel que APETALA1, APETALA2, LEAFY et WUSCHEL. Utilisant ce modèle avec plusieurs ensembles de paramètres, j'ai pu reproduire la dynamique d'expression d'*AG* connue. Ensuite, dans le but d'avoir une description détaillée de la dynamique d'*AG in vivo*, j'ai imagé en direct et en 4D la croissance des fleurs pour quantifier l'activation de l'expression d'*AG* de son initiation à son patron d'expression stable. En couplant l'analyse du modèle, ses simulations et la quantification *in vivo* de la dynamique d'*AG*, je montre que son expression se déroule en deux phases: une phase de faible expression durant laquelle quelques cellules expriment stochastiquement *AG*, et une phase de forte expression qui correspond au patron d'expression classique. Bien que toutes les cellules du dôme central de la fleur présentent un profil d'activation d'AG similaire, le temps précis au cours du développement où *AG* est activé est différent pour chacunes d'entre elles et est à l'origine de la stochasticité du patron d'expression. Avec l'aide du modèle, je propose quatres nouvelles hypothèses relatives à la régulation d'*AG* :

- AG est capable de maintenir sa propre activation en se liant directement à son second intron au travers d'un complexe protéique. Le modèle suggère que ce complexe comprend au moins deux molécules d'AG, ce qui produit un seuil d'auto-activation pour AG qui retarde son activation et permet une augmentation rapide de l'expression

quand ce seuil est dépassé.

- AP2 influence la valeur de ce seuil, restreint l'expression d'*AG* dans le dôme central de la fleur et produit un retard dans l'activation complète d'*AG*.

- LFY et WUS sont nécessaire à l'accumulation des protéines d'AG dans les cellules pour pouvoir atteindre le seuil d'auto-activation et obtenir une expression complète d'*AG*.

- Le mouvement d'AG est nécessaire pour obtenir l'expression d'*AG* dans toutes les cellules du dôme central, ce mouvement peut expliquer le patron d'expression d'*AG*, spécifiquement la présence d'un anneau avec une expression plus faible à la frontière de son domaine d'expression.

Pour prouver ces hypothèses, j'ai réalisé différentes expériences. En premier, utilisant une expérience de FRET-FLIM dans les protoplastes, nous proposons qu'AG est capable de s'associer en homodimer dans les cellules végétales. Néanmoins, sur-exprimer *AG* pour aider les cellules à atteindre le seuil d'auto-activation plus tôt que dans la plante sauvage ne semble pas modifier la dynamique d'expression de l'*AG* endogène. En deuxième, j'ai testé le rôle précis de LFY au cours des différentes phases et transitions de la dynamique d'expression d'*AG* en mutant les sites d'interactions spécifiques pour LFY au sein des séquences de régulation d'*AG*. Ces mutations retardent l'expression l'expression d'*AG* et modifient légèrement son patron d'expression. Je montre que seulement d'important retards dans l'activation d'*AG* induit des modifications phénotypiques. Ensuite, pour tester le rôle de la répression par AP2 dans la dynamique d'expression d'*AG*, j'analyse le rapporteur d'*AG* dans le contexte d'un mutant fort d'*ap2*. Dans ce mutant, l'expression d'*AG* s'étend à une région plus large et le retard entre l'initiation de l'expression d'*AG* et la transition entre les phases de faible et forte expressions est diminué. Ces résultats correspondent aux simulations du modèle. Finalement, pour comprendre l'importance du mouvement d'AG d'une cellule à l'autre dans sa propre dynamique, je bloque cette capacité de bouger en utilisant un tag de localisation nucléaire. Bien que cela induit un retard dans l'activation de quelques cellules au stade 3 au moment où toutes les cellules du dôme centrale de la fleur expriment *AG* dans la plante sauvage, ce retard n'a pas d'effets visible sur le phénotype.

Ce travail montre comment l'association de la modélisation avec l'analyse quantitative de la dynamique d'expression d'un gène peut être utilisée pour élaborer de nouvelles hypothèses

à propos des mécanismes impliqués dans la formation de patron d'expression, que l'on peut expérimentalement tester par la suite.

Mots clefs

Développement floral, AGAMOUS, modèles de réaction-diffusion, régulation génétique, imagerie 4D, dynamique d'expression, stochasticité, patron d'expression

Abstract

Land plants are estimated to have a biomass 1000 times greater than animals. Angiosperms, or flowering plants, are in turn the most predominant group, making up 90% of all plant species on earth. Presenting a highly conserved architecture, they form an important focus of research, due to their crucial role in diverse aspects of human existence. The identity of flower organs is defined by the expression of homeotic genes during early development that belongs to the MADS-box family. One of these genes, *AGAMOUS* (*AG*), is responsible for the identity of the stamens and the carpels in *Arabidopsis thaliana*. In this manuscript, I attempt to fully understand the spatial and temporal properties of *AG* expression by investigating the mechanisms underlying the proper establishment of *AG* expression dynamics during the early stages of flower development.

I start by developing a reaction-diffusion model that takes into account the growth of the flower at the relevant stages, as well as the few key transcription factors involved in *AG* regulation, such as APETALA1, APETALA2, LEAFY and WUSCHEL. Using this model and several different parameter sets, I was able to reproduce the known dynamics of *AG* expression. Next, in order to have a very detailed description of *AG* dynamics *in vivo*, I used real-time 4D imaging on growing flowers to quantify the activation of *AG* expression from its onset to the stable pattern. By coupling analysis of the model, its simulations, and the *in vivo* quantification of *AG* dynamics, I show that the *AG* expression occurs in two phases: a low-expression phase, during which a few cells stochastically express *AG*, and a high-expression phase, which encompases it's typical pattern of expression. Thus although all cells of the central dome of the flower present similar profiles of *AG* activation, the precise developmental time at which *AG* is activated is different in each case, and is the origin of the initial stochastic pattern. With the aid of the model, I also propose four new hypotheses to explain *AG* regulation:

- AG is able to maintain its own activation by directly binding its own second intron through a protein complex. The model suggests that this complex contains at least two molecules of AG, which produces a threshold in AG auto-activation that in turn delays activation and permits a rapid increase in expression when the threshold is reached.

- AP2 influences the value of this threshold, restraining *AG* expression to the central dome of the flower and producing a delay in complete *AG* activation.

- LFY and WUS are necessary to accumulate AG proteins in cells in order to reach the auto-activation threshold and obtain a full expression of *AG*.
- AG movement is necessary to obtain expression of *AG* in every cell of the central dome, this movement could explain the pattern of *AG* expression, especially the presence of a ring of lower *AG* expression at the border of its domain of expression.

To prove these hypotheses, I have carried out various experiments. Firstly, using FRET-FLIM experiment in protoplast cells, we suggest that AG is able to form homo-dimers in plant cells. However, overexpressing *AG* to help cells reach the auto-activation threshold earlier than in the wild-type does not appear to alter the endogenous *AG* dynamics of expression. Secondly, I test the precise role of LFY in the different phases and transitions in the *AG* expression dynamics by mutating specific interaction sites for LFY within *AG* regulatory sequences. These mutations appear to delay *AG* expression and slightly modify its pattern of expression. I show that only important delays in *AG* activation induce phenotypic differences. Then, to test the role of AP2 repression in *AG* expression dynamics, I analyse the *AG* reporter in the context of a strong *ap2* mutant. In these mutants, *AG* expression spreads to a wider region and reduces the delay between the onset of *AG* expression and the transition from low- to high-expression. These results match with simulations of the model. Lastly, to understand the importance of AG cell-to-cell movement in *AG* dynamics, I block its ability to move using a nuclear localisation tag. Although this induces a delay in the activation of few cells at stage 3, when all cells of the central dome of the flower express *AG* in the WT. This delay has no visible effects on the phenotype.

This work shows how the combination of modelling and of a quantified analysis of gene expression dynamics can be used to elaborate new hypotheses on the mechanisms underlying pattern formation, which can be subsequently tested experimentally.

Keywords

Flower development, AGAMOUS, reaction-diffusion modelling, gene regulation, 4D imaging, expression dynamics, stochasticity, pattern formation

Table of contents

Résumé & Mots clefs i
Abstract & Keywords iv
Table of contents ix

1 Introduction **1**
1.1 The flower: a highly successful evolutionary adaptation in angiosperms . . . 2
1.2 *AGAMOUS*, a key determinant of reproductive organ identity 6
1.2.1 AG is a MADS box protein 6
1.2.2 *AG* is part of a complex gene regulatory network 7
1.2.3 *AG* 2nd intron is the location of most of its regulation 11
1.2.4 Epigenetic has a role in *AG* regulation 12
1.2.5 Discussion 13
1.3 Interactions and structure in the MADS family 15
1.3.1 Structural biology in plant development 15
1.3.2 Study of the MADS-box proteins through SEPATALA3 15
1.4 How can we use a modelling approach to a better understanding of *AG* regulation 18
1.4.1 Genetic Regulatory Networks 21
1.4.2 Reaction-diffusion model with bistability 22
1.4.3 Example of models used in developmental biology 23
1.4.4 Applications of GRN computing methods on the flower patterning . . 27
1.4.5 Attempt to create dynamic models 28
1.4.6 Discussion 30
1.5 Methods to detect and quantify gene expression 30
1.5.1 Methods to localise mRNA 30
1.5.2 Confocal imaging to locate the protein in the tissue 31
1.5.3 4D imaging of the flower development 33
1.5.4 Discussion 35
1.6 Stochasticity and robustness in development, two fields that are often linked 35
1.6.1 Robustness in plant development 35
1.6.2 Link between stochasticity and robustness 36

2 A diffusion/reaction equations system to model dynamic of *AGAMOUS* expression **45**
Abstract 46
2.1 Introduction 46
2.2 Mathematical model 48
2.3 Biological justification of the model 50
2.3.1 Choice of the domain 50
2.3.2 Model kinetics 50

2.4 Model reduction . . . 52
2.4.1 Spatial transformations . . . 53
2.4.2 Reduction of the parameter space . . . 54
2.5 Numerical simulations . . . 54
2.5.1 Model discretization . . . 54
2.5.2 Visualization of the results . . . 55
2.6 Results of the model analysis and simulations . . . 56
2.6.1 The model reproduces qualitatively the dynamics of the WT . . . 56
2.6.2 The pattern of *AP1* and *LFY* expressions depend mainly on AG . . . 58
2.6.3 Parameters of *AG* auto-activation play the key role in the dynamics of *AG* expression . . . 59
2.6.4 Role of AP2 in the spatial and temporal restriction of AG . . . 61
2.6.5 LFY and WUS are necessary for *AG* activation . . . 62
2.6.6 Diffusion of AG is important to obtain a proper pattern . . . 64
2.6.7 External peak of AG expression accelerates the AG activation . . . 64
2.6.8 Analysis of the key parameters in robustness of the model . . . 65
2.7 Discussion, next steps . . . 67
2.8 Author contribution . . . 67
2.9 Supplementary data . . . 68

3 Analysis of the stochastic activation of *AGAMOUS* expression during flower development 71
Abstract . . . 71
3.1 Introduction . . . 72
3.2 Results . . . 73
3.2.1 Stochastic *AG* expression is observed at stage 2 . . . 73
3.2.2 Peripheral and central cells in the *AG* expression domain present a lower concentration of AG proteins . . . 74
3.2.3 The activation of *AG* expression is stochastic . . . 78
3.2.4 *AG* expression spreads from one cell to its adjacent cells . . . 79
3.2.5 *AG* activation profiles correspond to an auto-activation of AG itself as an homodimer . . . 79
3.3 Discussion and perspectives . . . 81
3.4 Methods . . . 85
3.5 Author contributions . . . 87
3.6 Acknowledgement . . . 87
3.7 Supplementary data . . . 88

4 The origins of the robust AGAMOUS pattern of expression 95
Abstract . . . 95
4.1 Introduction . . . 96
4.2 Results . . . 98
4.2.1 Over-expression of *AG* have an impact on its localization . . . 98
4.2.2 Activation by LFY and WUS is necessary for proper *AG* expression and a WT phenotype . . . 99
4.2.3 Auto-activation of *AG* is repressed until stage 3 of flower development 103
4.2.4 AG likely forms stable homo-dimers . . . 104
4.2.5 The role of cell-to-cell movement in establishing the *AG* pattern . . . 106

4.2.6 AP2 is one of the factors that restrain the activation of *AG* in the early stages of flower development 108
4.3 Discussion and perspectives 108
4.4 Methods 114
4.5 Acknowledgment 116
4.6 Author contributions 116
4.7 Supplementary data 116

5 General discussion and perspectives **121**

Bibliography **129**

List of figures **152**

List of tables **153**

Chapter 1

Introduction

Contents

1.1 The flower: a highly successful evolutionary adaptation in angiosperms 2

1.2 *AGAMOUS*, a key determinant of reproductive organ identity 6

1.2.1 AG is a MADS box protein 6

1.2.2 *AG* is part of a complex gene regulatory network 7

1.2.3 *AG* 2nd intron is the location of most of its regulation 11

1.2.4 Epigenetic has a role in *AG* regulation 12

1.2.5 Discussion 13

1.3 Interactions and structure in the MADS family 15

1.3.1 Structural biology in plant development 15

1.3.2 Study of the MADS-box proteins through SEPATALA3 15

1.4 How can we use a modelling approach to a better understanding of *AG* regulation 18

1.4.1 Genetic Regulatory Networks 21

1.4.2 Reaction-diffusion model with bistability 22

1.4.3 Example of models used in developmental biology 23

1.4.4 Applications of GRN computing methods on the flower patterning 27

1.4.5 Attempt to create dynamic models 28

1.4.6 Discussion 30

1.5 Methods to detect and quantify gene expression 30

1.5.1 Methods to localise mRNA 30

1.5.2 Confocal imaging to locate the protein in the tissue 31

1.5.3 4D imaging of the flower development 33

1.5.4 Discussion 35

1.6 Stochasticity and robustness in development, two fields that are often linked 35

1.6.1 Robustness in plant development 35
1.6.2 Link between stochasticity and robustness 36

1.1 The flower: a highly successful evolutionary adaptation in angiosperms

The Plant kingdom has evolved over millions of years to comprise more than 375,000 living species [Christenhusz and Byng, 2016]. The vast majority of these are seed plants, which are separated in two classes, the gymnosperms and the angiosperms. The presence of seeds within an enclosure is the defining characteristic of angiosperms, implying the presence of a fruit and flowers from which those fruit form. The angiosperms first appeared in the mid Cretaceous (120-90 Mya) and, within 20-30 million years, became the predominant plants on the planet [Friis et al., 2010], and currently comprise about 300,000 species [Christenhusz and Byng, 2016]. The primary traits of the flower are quite conserved between these species, despite the incredible diversity of sizes and colours.

One of the main plant model systems that has been used to carry out a detailed study of the mechanisms underlying flower development is *Arabidopsis thaliana*. This species is native to Eurasia, has a short life cycle of less than two months in normal growth conditions, measures around 20-25 cm with hundreds of flowers that are capable of self-pollination. Its relative small genome has been fully sequenced and a vast array of tools allow us to precisely study developmental mechanisms.

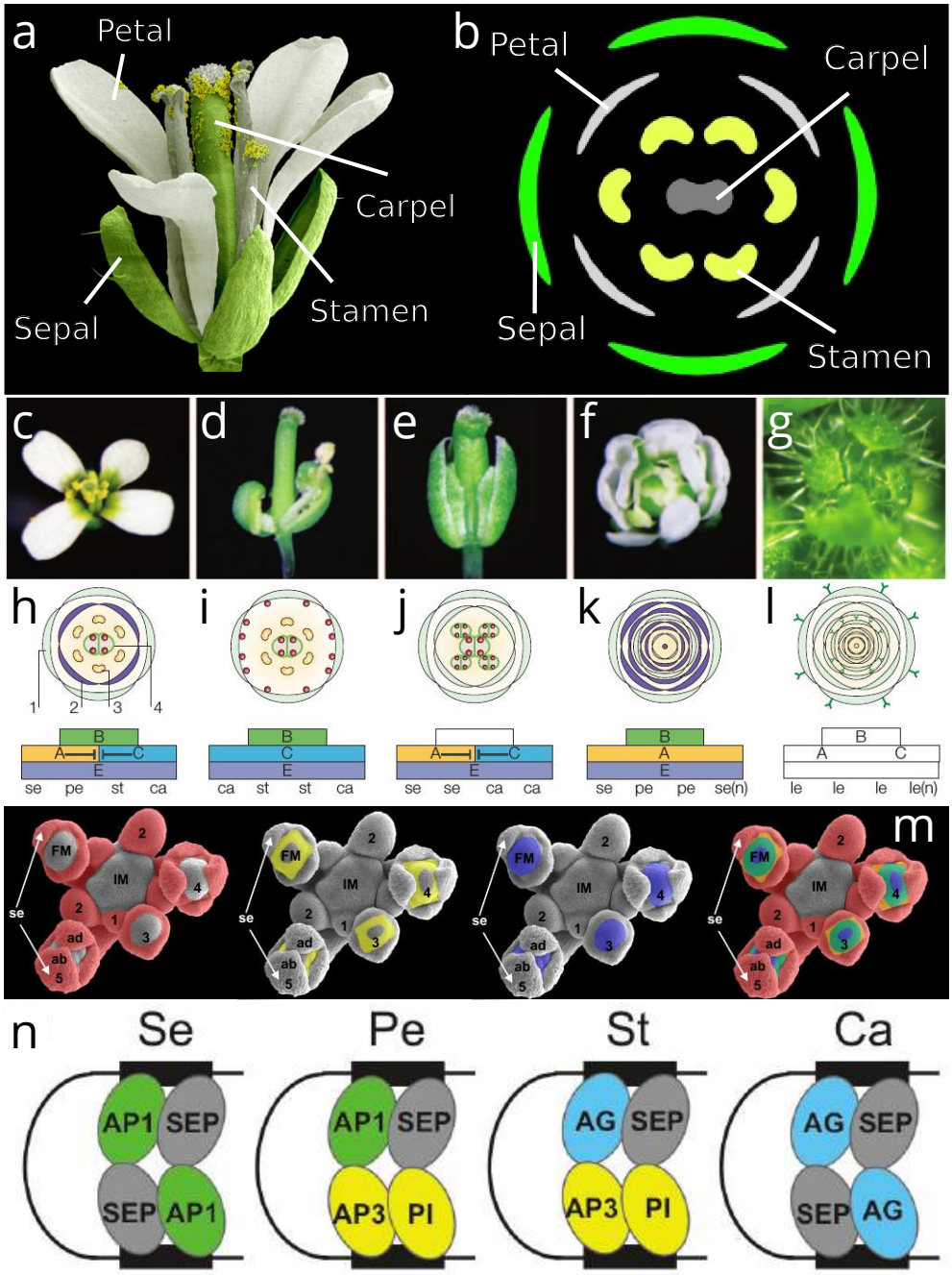
In this manuscript, I will describe work that relates to how flowers gain their specific shapes. Flowers form at the growing tip of the plant, around a structure known as the shoot apical meristem (SAM). The SAM is an indeterminate structure that produces lateral organs (first leaves, then flowers) throughout the life of the plant, and is maintained by a pool of stem cells at its centre. These stem cells continuously divide and produce cells that migrate at the periphery and are able to differentiate in response to hormonal fluxes within the SAM [Barton, 2010]. Flowers form as a small bump of cells at the flanks of the meristem that will become a flower after few days and then grows themselves through stages of development to give a fully functional flower and then seeds. Flower development has been dissected in different stages from bud initiation to seeds fall [Smyth et al., 1990]. *Arabidopsis* flowers are composed of four different kinds of organs (Figure 1.1 a.) – 4 sepals, 4 petals, 6 stamens

and 2 congenitally fused carpels. These organs appear in concentric discs, or whorls, within the flower (Figure 1.1 b.). Three classes of genes, expressed in overlapping domains, act combinatorially to define the identities of these organs, in what is known as the ABC model of flower development [Coen and Meyerowitz, 1991]. Thus the A class genes *APETALA1* (*AP1*) and *AP2* are expressed in the outer two whorls, and mutants in these genes display transformed sepals and petals (Figure 1.1 d & i). The B class genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) are expressed in the second and third whorls, and the loss of their activities give rise to defective petals and stamens (Figure 1.1 e & j). The C class gene *AGAMOUS* (*AG*) is expressed in the third and fourth whorls and mutations in this gene lead to defects in the reproductive structures, with stamens replaced by petals and carpels replaced by a new flower (Figure 1.1 f & k). The latter phenotype of a new flower, rather than a simple change in organ identity, is possibly due to the role of AG in repressing the stem cell organiser gene *WUSCHEL* (*WUS*).

WUS activity has been shown to be crucial for the maintenance of functional meristem, via a direct positive regulation of the stem cell identity gene *CLAVATA3* (*CLV3*). It is expressed in a small group of cells underlying the stem cell zone. The repression of *WUS* expression at stage 6 leads to the arrest of stem cell proliferation in the central dome of the flower, and allows proper flower development to occur [Lohmann et al., 2001, Lenhard et al., 2001].

The ABC model also includes the activity of a fourth class of genes, called the E class, which is composed of the four *SEPALATA* genes (*SEP1*, *SEP2*, *SEP3* and *SEP4*) with partially overlapping expression patterns and redundant functions. Indeed this class of genes, called the 'glue proteins' because they act as scaffolding proteins, are necessary to allow the formation complexes of ABC proteins (the so-called quartets) that act to confer organ identity in the different whorls (Figure 1.1 m. & n.) [Immink et al., 2009]. Quadruple mutants of the *SEP* genes results in floral organs that resemble leaves (Figure 1.1 g & l.) [Ditta et al., 2004]. The ABCE model thus stipulates that sepal formation requires the activities of A and E class genes; petals require A, B and E activities; stamens require B, C and E activities and carpels require C and E activities.

The architecture of the flower in well conserved between species. The ABCE model has been shown to be present in other species with sometimes small modifications. The original comparison was performed with *Antirrhinum* that present the same structure [Coen and



(Caption next page.)

Figure 1.1: Flower architecture and the ABCE model: (a.) An *Arabidopsis thaliana* flower with the following four organ types labelled: sepals, petals, stamens and carpels. **(b.)** Floral diagram showing the numbers and relative positions of floral organs. **(c.-m.)** Summary of the ABCE model showing WT and mutant flowers **(c.-g.)**, floral diagrams **(h.-l.)** and the relative expression domains of patterning genes (m.). **(c. & h.)** The ABCE model in the WT, with a WT flower (c) its architecture and the description of the ABCE model. Superposition of A and E gives sepal identity; A, B and E gives petal identity B, C and E gives stamen identity; and C and E gives carpel identity. **(d. & i.)** A class A mutant, where sepals are transformed into carpels and petals into stamens. **(e. & j.)** A class B mutant, with petals transformed into sepals and stamens into carpels. **(f. & k.)** A class C mutant, with stamens are transformed into petals and a new flower emerges from the center. **(g. & l.)** A class E mutant, where all organs of the flower are replaced by leaf-like structures. **(m.)** Expression patterns of the ABC genes in young flowers. The numbers 1 to 4 indicate the stage of flower development. The A class genes are in red, the B in yellow and the C in blue. The fourth panel provides the superposed image in the four whorls of the flower. **(n.)** The quartet model that leads to the identity of organs in the flower. Two molecules of AP1 and two of SEP act to give sepals; one molecule each of AP1, PI, AP3 and SEP give petals; one each of AG, PI, AP3 and SEP give stamens and two AG and two SEP give carpels. Panel a was adapted fromwww.weigelworld.org/, panel b was produced by P. Das, panel c - l are adapted from [Krizek and Fletcher, 2005], panel m from [Alvarez-Buylla et al., 2010] and panel n from [Wellmer et al., 2014].

Meyerowitz, 1991]. But, for instance for *Amborella* or *Nuphar*, the fusion of sepals and petals produces tepals present mainly in spirals instead of whorls [Soltis et al., 2007]. They can also contain larger domain of expression of the B-class genes.

Interactions between genes of the ABCE model, their pattern of expression, their evolution, their rules in the flower development have been extensively studied over the last 25-30 years. Nevertheless many questions still remain unaddressed, especially with respect to how the very particular spatio-temporal dynamics of gene expression are established within the context of the rapid growth flower that occurs during early flower development. A clear example of such specific dynamics is for *AGAMOUS*, the C class gene that appears very suddenly at stage 3 in the central dome of the flower. In this manuscript, I will focus my efforts on this gene, to precisely study its pattern of expression to understand the key biological parameters required for this patterning to occur. I will focus my efforts on the first few stages of flower development, during the period when this special expression pattern is established.

1.2 *AGAMOUS*, a key determinant of reproductive organ identity

AGAMOUS (*AG*), the only C class gene in *Arabidopsis*, has been intensively studied during the last decades for its role in flower development. It is a transcription factor that plays a role in the identity and development of the carpels and the stamens [Bowman et al., 1989, Yanofsky et al., 1990, Drews et al., 1991] and in the termination of stem cell identity in the flower [Lenhard et al., 2001]. AG is an homeotic gene encoding for a 252 residue long MADS domain transcription factor. This family is thought to act by binding specific promoter elements called CArG boxes.

1.2.1 AG is a MADS box protein

Like most of the factors of the ABCE model excepting AP2, AG is a member of the MADS family of transcription factors. The name MADS is derived from four proteins in four different species: MINICHROMOSOME MAINTENANCE 1 (MCM1) from the budding yeast, AGAMOUS (AG) from *Arabidopsis thaliana*, DEFICIENS (DEF) from *Antirrhinum majus* and SERUM RESPONSE FACTOR (SRF) from humans, which share the ability to

bind to the 'MADS' domain in gene regulatory sequences. Although the MADS-box family is present in all eukaryotes species, they are nevertheless present in greater numbers in plants than in animals. In animals these proteins are mainly involved in muscle development and cell proliferation and differentiation [Shore and Sharrocks, 1995]. In plants, MADS-box proteins are involved in many developmental process at all stage of plant development [Smaczniak et al., 2012] (Figure 1.2). There are 107 MADS-box genes in *Arabidopsis*, which are divided into two classes the type I MADS TFs bear a MADS DNA binding domain and a C-terminal domain, while the type II MADS TFs are comprised of four domains, including a MADS domain, an I (intervening) domain important in the dimerisation process, a K (keratin-like coiled-coil) domain important for dimerisation and tetramerisation and a C-terminal domain involved in transactivation and formation of higher order complexes (Figure 1.3).

AG, like the other MADS-box proteins involved in flower development, is a type II MADS TF. Both types of MADS TF have been shown to bind as dimers to DNA sequences ($CC[A/T]_6GG$ or $CTA[A/T]_4TAG$) called CArG-boxes [Pollock and Treisman, 1991, Riechmann et al., 1996a, Riechmann et al., 1996b], although differences can be observed in the number of A/T residues in different target gene promoters [Nurrish and Treisman, 1995]. MADS dimers, as well as tetramer one can also play a role in the regulation of a transcription factor by MADS-box proteins [Puranik et al., 2014]. I will address the structural properties of these proteins later in this manuscript.

1.2.2 *AG* is part of a complex gene regulatory network

AG expression has been shown to commence at early stage 3 of flower development with an homogeneous pattern of expression in the two central whorls and remains uniformly expressed until stage 7 in the presumptive stamens and carpels [Yanofsky et al., 1990, Drews et al., 1991]. Thereafter, *AG* becomes restricted to specific tissues within the differentiating stamens and carpels, such as the filaments and connectives anther walls of developing stamens [Bowman et al., 1991, Sieburth and Meyerowitz, 1997].

AG has four principal functions in flower development. First, it interacts with other MADS-box proteins, such as SEP3, PI and AP3 in the third whorl, or SEP3 in the fourth, to confer stamen and carpel identity to those cells [Coen and Meyerowitz, 1991, Gómez-Mena et al., 2005, Smaczniak et al., 2012]. The identity of these organs is due to direct interactions of these complexes to DNA binding sites. Second, AG restrains the expression of the A class

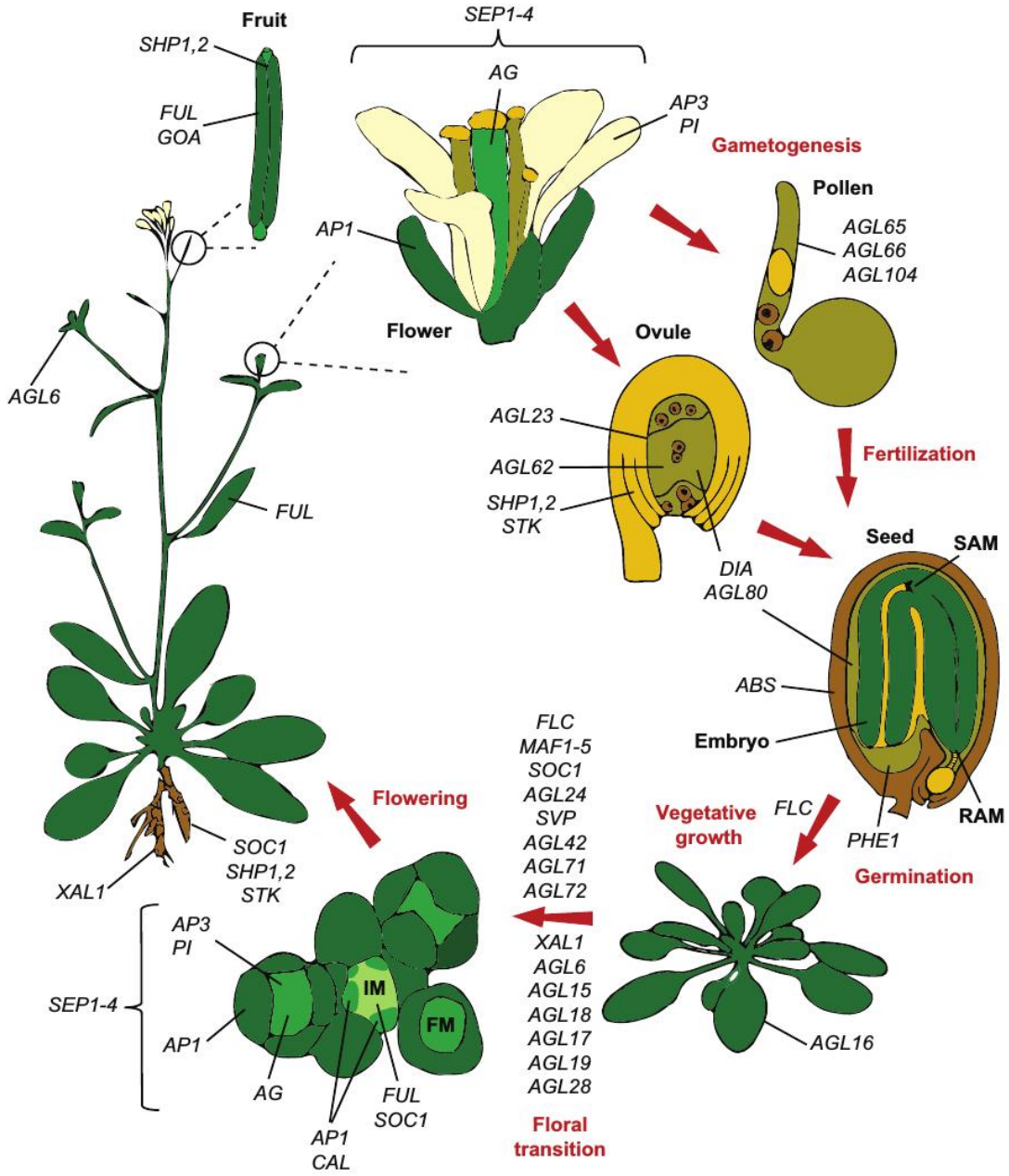


Figure 1.2: Role of MADS-box proteins in the *Arabidopsis thaliana* development Adapted from [Smaczniak et al., 2012].

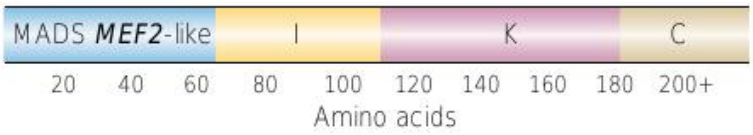


Figure 1.3: Type II MADS-box protein The type II MADS-box proteins have an N-terminal MADS DNA binding domain that resembles the animal myocyte enhancer factor 2 (MEF2), followed by a weakly conserved intervening region (I) and a Keratin-like coiled-coil domain (K) that are both involved in the dimerisation and the tetramerisation of these proteins. These are followed by a C-terminal region that is important for the transactivation process and to stabilize interactions with other proteins. [Ng and Yanofsky, 2001]

gene *AP1* in the two outer whorls by directly binding to the *AP1* promoter [Gustafson-Brown et al., 1994]. Third, AG plays a role in repressing *WUS* in the centre of the flower to arrest stem cell proliferation and ensure floral determinacy [Lohmann et al., 2001, Lenhard et al., 2001]. Finally, AG has various specific roles in late stamen and carpel development, such as in anther morphogenesis, dehiscence and filament formation and elongation [Ito et al., 2007]. In carrying out these various functions, AG activates a number of genes some of which are summarised here (Figure 1.4 A. & B.).

AG is itself regulated by several transcription factors during flower development. It is activated by LEAFY (LFY) and WUS, which together bind regulatory sequences at the *AG* locus [Parcy et al., 1998, Lohmann et al., 2001, Lenhard et al., 2001]. LFY is a master regulator of floral fate that is expressed throughout the flower at early stages, whereas *WUS* begins to be expressed in the stem cell niche at late stage 1 and is required for the maintenance of stem cell fate. AG is also capable of binding its own promoter to activate its own expression [Gómez-Mena et al., 2005]. *AG* expression is repressed in the outer whorls by two different pathways – one including AP2 and the micro RNA miR172 that in turn regulates the expression of *AP2* [Drews et al., 1991, Chen, 2004, Wollmann et al., 2010], the other with the complex of LEUNIG (LUG) and SEUSS (SEU). *LUG* and *SEU* are two cadastral genes that work together as a transcriptional co-repressor complex that restrain *AG* in the two inner whorls [Liu and Meyerowitz, 1995, Sridhar et al., 2004]. The third one involve epigenetic factors as ULTRAPETALA1 & 2 (ULT1 & 2) and will be detailed later. AP2, LEU and SEU are present early in flower development in the outer whorls. It has also been shown that SEP3 can have a positive role in *AG* regulation [Castillejo et al., 2005]. SEP3 interacts with AG (see below) and has a pattern of expression similar to *AG*,

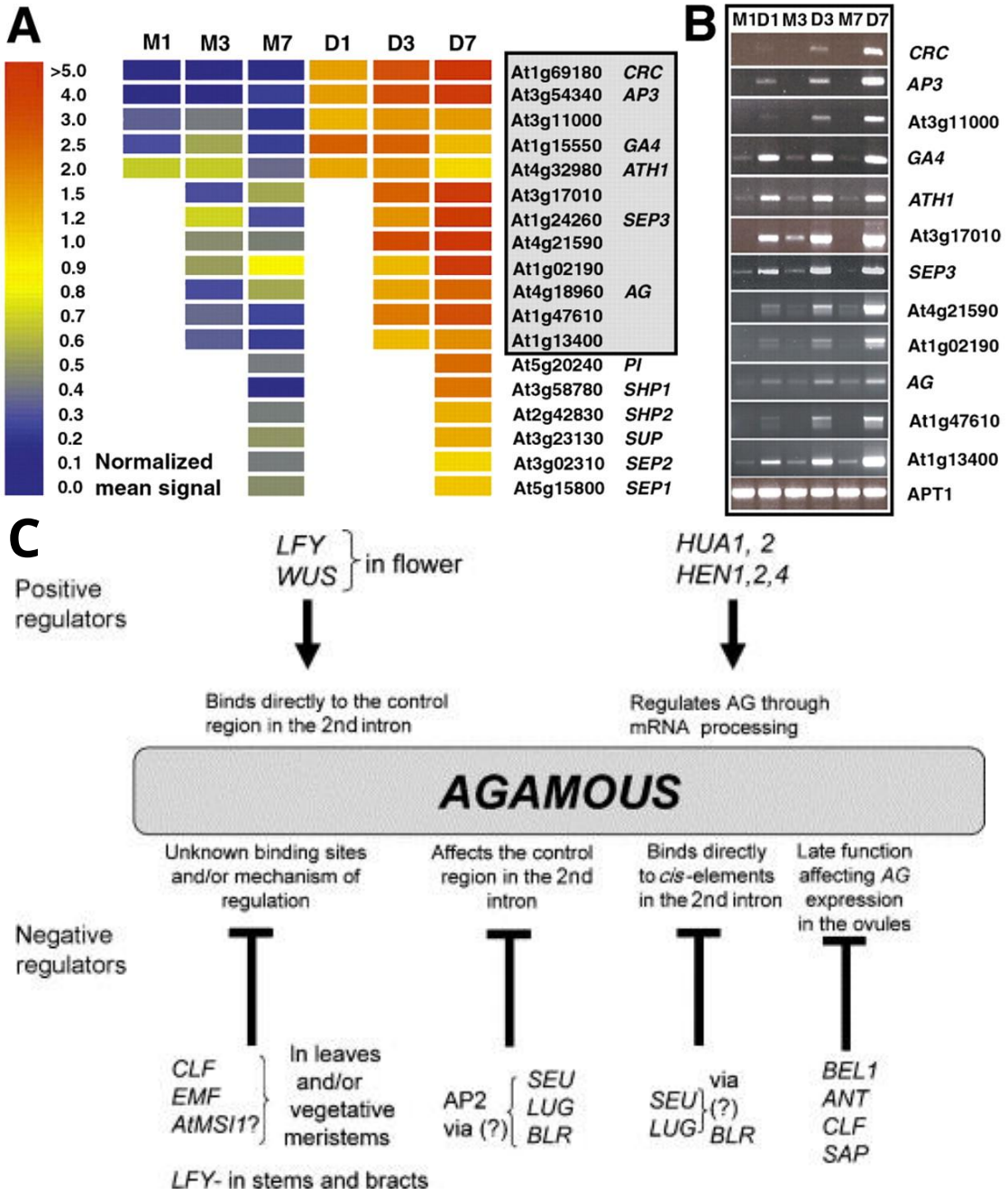


Figure 1.4: Regulation by and of *AGAMOUS*: (A & B.) Activation of transcription factors by AG during flower development. Expression levels of selected genes by an inducible *35S::AG-GR* construct, where a rat glucocorticoid receptor (GR) moiety is fused in frame to the AG protein, rendering it inactive in the absence of dexamethasone (DEX). **(A.)** Expression detected by oligonucleotide arrays. M1 to M7 and D1 to D7 indicate the number of days after mock or DEX treatment respectively (1, 3 and 7 days). Genes listed in the grey box are targets, including AG itself, that show sustained activation. **(B.)** Validation of the 12 selected targets by RT-PCR. *APT1* is a constitutive control. **(C.)** Principal regulators of *AG* during early and late flower development. Panels A and B are adapted from [Gómez-Mena et al., 2005], and panel C is adapted from [Zahn et al., 2006].

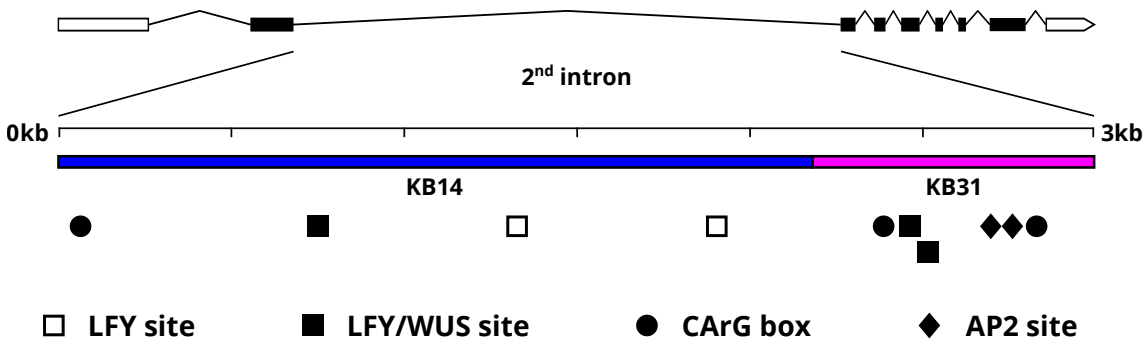


Figure 1.5: Regulatory sequences at the *AG* locus: Schematic of the *Arabidopsis thaliana AG* 5.68-kb coding strand (top row). Empty rectangles are the 5' and 3' UTR regions, black rectangles are exons, solid lines are introns. Zoomed view of the 3-kb second intron where many key regulatory sequences are to be found (lower panel). The coloured bars represent regions that are thought to contain redundant regulatory elements[Busch et al., 1999]. Evolutionarily conserved binding sites for LFY, WUS and MADS family protein are represented as solid or empty shapes and are labelled [Deyholos and Sieburth, 2000, Hong and Hamaguchi, 2003, Moyroud et al., 2011] and AP2 binding sites based on Yant and Dinh [Yant et al., 2010, Dinh et al., 2012]

although at a lower level [Urbanus et al., 2009]. The regulation of *AG* at later floral stages is much more complex and involves direct or indirect activities of various transcription factors (reviewed in [de Folter et al., 2005, Zahn et al., 2006, Ferrándiz et al., 1999]) (Figure 1.4. C).

1.2.3 *AG* 2nd intron is the location of most of its regulation

The key regulatory sequences that control the expression of *AG* lie within the 5.7 kb coding region, and more specifically within the 3-kb second intron (Figure 1.5). Indeed, This intron has been extensively studied, both to identify the specific sequences bound by various known regulators, as well as to determine how conserved these sequences are over evolution [Busch et al., 1999, Deyholos and Sieburth, 2000, Hong and Hamaguchi, 2003, Yant et al., 2010, Moyroud et al., 2011]. Constructs wherein the GUS reporter gene was fused to this intron reproduces a similar pattern of expression to that determined by *in situ* hybridisation experiments for *AG* RNA [Busch et al., 1999]. Furthermore, a 2.2-kb 5' and a 0.8-kb 3' region of this intron can independently recapitulate the proper expression pattern (Figure 1.5) [Busch et al., 1999, Deyholos and Sieburth, 2000].

A phylogenetic footprinting analysis of this intron from the *AG* locus in 29 different species showed that several small regions are highly conserved and correspond either to

sequences bound by known factors, such as the LFY and WUS consensus binding sites or to unidentified complexes, such as the CCAAT boxes [Hong and Hamaguchi, 2003]. The five conserved LFY binding sites have since been experimentally confirmed [Moyroud et al., 2011]. Three consensus MADS-domain TF binding sites, the CArG-boxes have been identified in the second intron, though only one has been shown to be important for the auto-activation and maintenance of *AG* expression [Gómez-Mena et al., 2005]. AP2 is also able to directly bind the second intron of *AG* [Yant et al., 2010]. Two consensus sequences have been found in the corresponding region [Dinh et al., 2012] (Figure 1.5).

1.2.4 Epigenetic has a role in *AG* regulation

The regulation of *AG* expression is not affected only by transcription factors. Proteins with epigenetic functions have been described as having a strong effect on *AG* expression. While I have not directly focussed on the epigenetic aspects of *AG* regulation in my thesis work, it is definitely an important factor that can influence the results of our experiments and that need to be taken into account while interpreting those results. Below, I describe this regulation in some detail, so as to have a clear insight into its importance in *AG* regulation.

Phenotypic traits can be inherited without DNA sequence specificity but through chromatin structures. Epigenetics is the study of this chromatin conformation during transcription [Moore, 2015]. The role of chromatin in regulating the expression of a gene is a main mechanism in biology shared by all eukaryotic species [Li et al., 2007]. The DNA is structured by proteins called histones that wrap the DNA around them to form chromatin. These histones may be modified via the addition or removal of chemical tags such as methyl or acetyl groups on different residues of the histones. Depending on the methylation or acetylation states of these histones, the DNA may take on different conformations and thus have an altered capacity to bind proteins and for DNA transcription to be activated or repressed.

A few genes have been shown to regulate *AG* epigenetically. *ULT1* and *ULT2* are two such genes, which have been identified in *Arabidopsis* with a role in flower development. ULT1 represses the accumulation of meristematic cells during development. In comparison to the wild type, loss of function *ult1* mutants display a larger inflorescence meristem, more flowers, supernumerary floral organs, and a decrease in floral meristem determinacy [Fletcher, 2001, Carles et al., 2004]. In the *ult1* mutant, a reduced level of *AG* mRNA is observed in the central dome of the flower at stage 6. This reduction correlates with a delayed

repression of *WUS*, which confers some of the phenotypes. It has also been shown that the ubiquitous expression of *ULT1* under the *35S* promoter leads to a higher level of *AG* expression throughout the flower [Carles and Fletcher, 2009]. ULT2 has a similar, albeit lower, effect as ULT1 in regulating shoot stem cell maintenance [Monfared et al., 2013].

CURLY LEAF (CLF) has the opposite effect to ULT1 in *AG* regulation. CLF represses *AG* expression in leaves, in the inflorescence meristem and in the outer whorls of the flower. Strong *clf* alleles show *AG* mRNA accumulation in the whole flower and a phenotype similar to the overexpression of *ULT1* [Saleh et al., 2007, Carles and Fletcher, 2009]. It has been proposed that these genes act in different pathways to regulate chromatin [Li et al., 2007]. CLF is a Polycomb group (PcG) factor, which is part of the Polycomb Repressive Complex 2 (PRC2), which mediates Histone 3 Lysine 27 trimethylation (H3K27me3) and results in the repression of transcription of the targeted gene. This repression is detected on both upstream *AG* sequences and in the 2nd intron. On the other hand, ULT1 is a trithorax group (trxG) factor that mediates the action of ATX1 (ARABIDOPSIS TRITHORAX1) in a histone methyltransferase in the trxG complex, on the Histone 3 Lys 4 residue (H3K4me3), an activating mark that results in high levels of *AG* transcription [Saleh et al., 2007, Carles and Fletcher, 2009]. One model is thus that ULT1 competes with CLF and the PcG to methylate the H3K4 and to promote, along with ATX1 and other co-activators, transcription at the *AG* locus (Figure 1.6).

This mechanism of *AG* regulation acts independently of, and in addition to, the regulatory effects of other transcription factor pathways. It has previously been shown that the activation of *AG* by LFY occurs separately from the positive action of ULT1 [Engelhorn et al., 2014]. Furthermore, little is understood about the upstream regulators of these epigenetic factors. For these reasons, I will focus specifically on the actions of transcription factors like LFY, WUS etc. in the *AG* regulation and not on the actions of the epigenetic factors.

1.2.5 Discussion

AG is an important protein in flower development that is regulated via many different pathways. Much is known about the mechanisms by which these pathways act to establish the pattern of AG expression. These pathways must exert their individual effects during stage 2 in order for AG expression to appear very suddenly at the transition to stage 3.

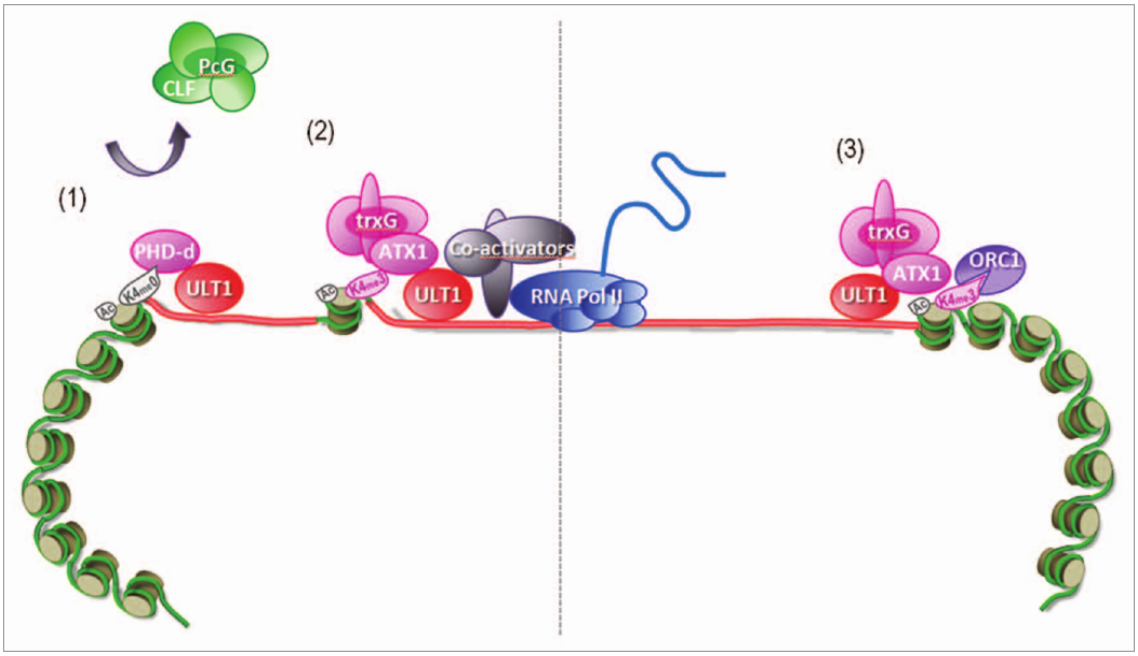


Figure 1.6: Epigenetic regulation of *AG*: (1) PcG factors, such as CLF, are evicted from DNA to allow a PHD-domain containing protein to bind unmethylated H3K4 chromatin. PHD proteins interact with ULT1 to recruit **(2)** co-activators and transcription complexes. ULT1 also interacts with the trxG complex containing ATX1 that deposits active histone marks, such as H3K4me3. **(3)** ORC1 directly bind these marks to increase transcriptional activity. This figure was adapted from [Carles and Fletcher, 2010].

Additionally, the flower undergoes immense growth during stage 2 (up to an eight-fold change in volume), implying the all these regulatory pathways must also account for growth. t is not easy to understand how these numerous mechanisms work together to robustly establish the spatio-temporal dynamics of AG expression. In the next section, I discuss the new tools that allow to us to more precisely delineate the pattern of AG expression and the role of these regulators in establishing this pattern.

1.3 Interactions and structure in the MADS family

1.3.1 Structural biology in plant development

Structural biology addresses how various biological molecules within the cell are folded in space and its interactions with other proteins at the molecular and the atomic levels. It allows us to gather precise information about molecular interactions that are observed by genetics or other means. These techniques are important to gain a better understanding of the systems under study. For instance, the structure of LEAFY and its capacity to bind the DNA as a monomer, a homodimer or a multimer, provides insight into the evolution of this protein between species and its role in development [Sayou et al., 2014]. It also allows us to understand the link between LFY multimerisation and its capacity to access theoretically low-affinity binding sites or region with closed chromatin [Moyroud et al., 2011, Sayou et al., 2016].

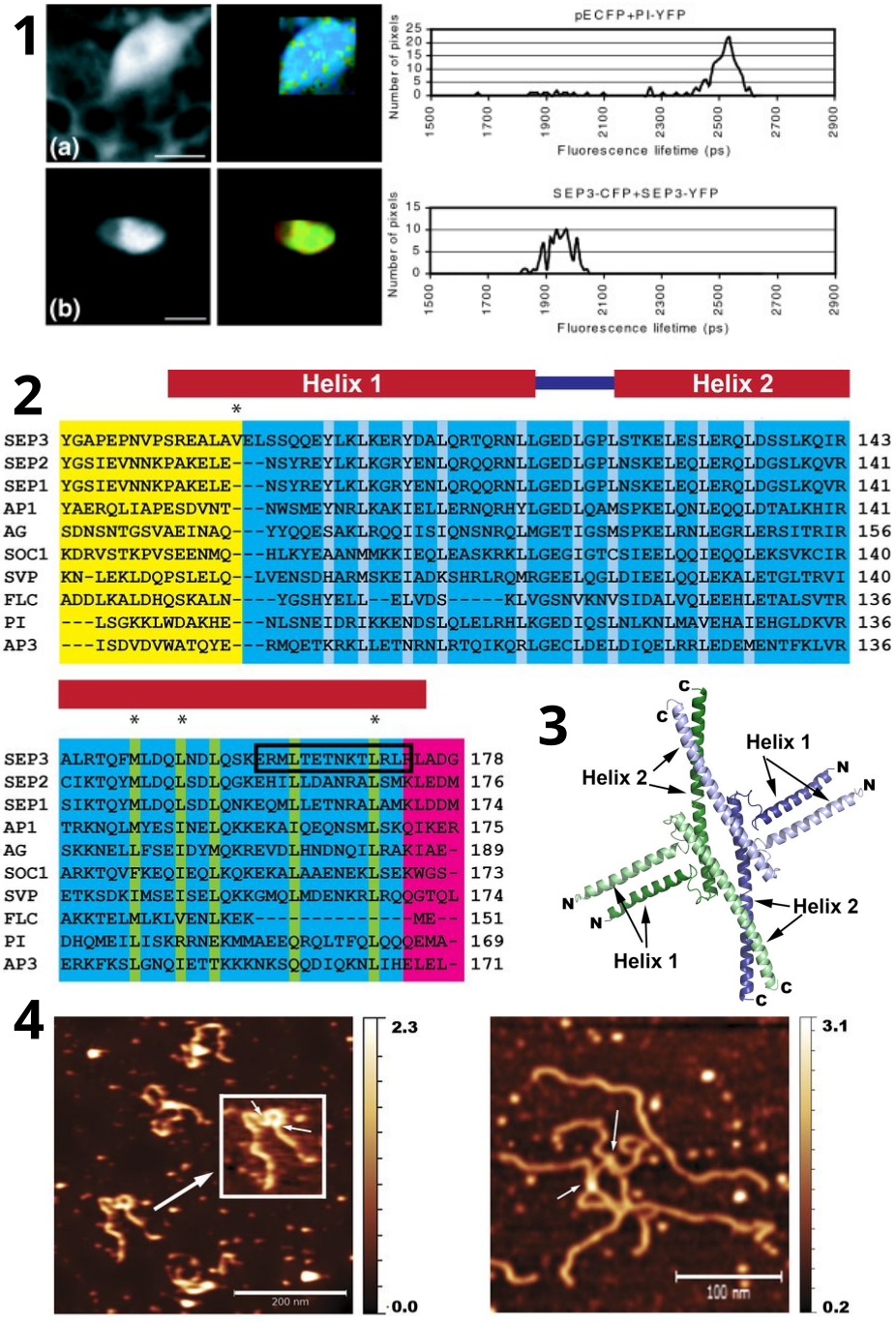
1.3.2 Study of the MADS-box proteins through SEPATALA3

The determination of the identity of flower organs is based on the presence of different MADS-box proteins in different whorls. MADS-box proteins have a conserved structure, as described previously (see section 2). The different members of this family are known to bind DNA sequences called CArG boxes as multimers, specifically tetramers. This capacity to form multimers has been widely studied. Yeast two hybrids assays have been performed between all the 107 *Arabidopsis* MADS-box proteins, revealing numerous 272 interactions that provide insight into the quartet model [de Folter et al., 2005]. One of the MADS proteins that are able to bind AG is SEP3. This protein has a comparable sequence to AG. Results are known about the structure and the interaction of SEP3 that can give us some ideas

concerning AG. One interesting result emerging from this study is that SEP3 is not able to form homodimers. Two recent studies show that this is likely not the case and that SEP3 is indeed able to form homodimers and homotetramers. First FRET-FLIM experiments in protoplasts clearly show the capacity of two molecules of SEP3 to bind each other (Figure 1.7.1) [Immink et al., 2009]. Additionally, a large part of the SEP3 protein (residues 75 – 178), including the entire K domain, which carries the capacity for multimerisation, as well as parts of the I and the C domains have been crystallized. Its structure reveals the presence of two α-helices that are involved in dimer and tetramer formation (Figure 1.7 3.) [Puranik et al., 2014]. The sequence similarities between members of the MADS family suggests that the residues involved in dimerisation and tetramerisation are well-conserved (Figure 1.7.2). Further experiments show that these tetramers are able to bind the DNA via the formation of a loop that corresponds to the binding of the tetramer to two CArG boxes located seperated by 93bp of each other (Figure 1.7.4) [Puranik et al., 2014]. Concerning AG, three CArG boxes are present in the second intron, the first one at its beginning, the two other are separated by 429bp. This is enough to allow creation of a loop as for SEP3.

The crystal structure of AG has not yet been resolved. The yeast two hybrid interaction assays suggests that AG cannot form homodimers [de Folter et al., 2005], but as noted above, this was also the case for SEP3. However sequence similarity between *AG* and *SEP3* (Figure 1.7.2), as well as the quartet model would, on the other hand, suggest that AG is able to form a tetramer with two proteins of SEP3 in the central whorl to confer carpel identity to the floral organs growing here. Recently, EMSA experiments aimed at examining the capacity of SEP3 and AG to form hetero or homodimers to bind the *SUPPRESSOR OF CONSTANS1* (*SOC1*) promoter [Silva et al., 2016] showed that AG can homodimers and to hetero-tetramerise with SEP3. It also showed that the hetero-tetramer of SEP3 and AG is easily able to bind the DNA.

We previously seen that the interaction between different proteins with DNA can have an impact on the expression of a target gene. While some insights into the capacity of AG to form homodimers and to tetramerise with SEP3 do exist, the role of this interaction in the ABC model and in its dynamics are very unclear. The composition of the MADS complexe that bind CArG boxes of AG is a parameter that may play an important role in the establishment of the correct pattern of expression of *AG*, and will need to be taken in account in the future.



(Caption next page.)

Figure 1.7: Structural properties of SEP3: (1.) FRET-FLIM experiment in protoplasts shows the capacity of SEP3 to form dimers. **(a.)** Control with pECFP and PI-YFP. **(b.)** SEP3-CFP with SEP3-YFP showing that SEP3 is able to form stable homodimers [Immink et al., 2009]. **(2.)** Sequence of residues 75-178 of SEP3 showing its homology with other MADS box proteins. The I domain is indicated in yellow and the K domain in blue. The various structural motifs are indicated above: The two main alpha helices are represented based on the crystal structure of the SEP3 protein. Residues involved in dimerisation and tetramerisation are highlighted in light blue and light green respectively. **(3.)** Tetramer of $SEP3^{75-178}$. The two helices are at the base of the two dimers in blue and green and the tretramer formation. **(4.)** Atomic Force Microscopy of SEP3 tetramers bound to DNA. Full length SEP3 protein forms complexes with 1-kb of a target promoter DNA at 10 to 15nM protein and 5nM DNA before dilution to 1nM DNA for imaging. Arrows indicate DNA looping due to SEP3 interactions. Scale bars indicate 200nm for the right and 100nm for the left. [Puranik et al., 2014]

1.4 How can we use a modelling approach to a better understanding of AG regulation

Mathematical modelling is a powerful technique to address key questions and paradigms in model systems and to provide quantitative mechanistic understanding of dynamics and control of intra- and intercellular signalling and the resulting self-organisation and spatio-temporal structure formation at the level of tissues and organisms.

First used in physics, partial differential equations models have been introduced in biology by pioneers such as Allan Turing who proposed reaction-diffusion equations to model morphogenesis in animal and plant development [Turing, 1952]. The Turing hypothesis can be stated as follows: *When two chemical species with different diffusion rates react with each other, the spatially homogeneous state may become unstable, thereby leading to a nontrivial spatial structure.* The idea looks counter-intuitive, since diffusion is expected to yield a uniform distribution. However, mathematical analysis of reaction-diffusion equations provides explanation for the phenomenon postulated by Turing. The mechanism is related to a local behaviour of solutions of the system in the neighbourhood of a constant stationary solution that is destabilised through diffusion. Patterns arise through a bifurcation, called *diffusion-driven instability (DDI).*

The original concept was presented by Turing on the example of two linear reaction-diffusion equations describing dynamics of two chemical components with a significantly

different diffusion rates [Turing, 1952]. Changing the ratio of diffusion coefficients may lead to emergence of a variety of patterns described by the same equations (Figure 1.8). The concept became a paradigm for pattern formation and led to development of numerous theoretical models describing natural patterns such as stripes, spirals, spots emerging from a homogeneous steady state [Murray, 2002, Murray, 2003]. The most famous realisation of the Turing's idea in a mathematical model and the first numerical simulation of a system exhibiting Turing patterns is the activator-inhibitor model proposed by Gierer and Meinhardt in 1972 [Gierer and Meinhardt, 1972]. That abstract model has provided some hints how complex intracellular pathways coupled to cell-to-cell communication through diffusive molecules may lead to symmetry breaking and emergence of spatially heterogeneous patterns of gene expression [Meinhardt, 1982, Gordon and Beloussov, 2006]. Nowadays, an enormous amount of information on a molecular and cellular scale has become available and much has been learned about the molecular components involved in signal transduction during development. Information on how these components are integrated into networks may be translated into *bottom-up* models allowing to link signalling networks to processes observed on a macroscopic scale of the tissue [Asthagiri and Lauffenburger, 2000, Savageau, 1979].

In other areas of biology, such as neurophysiology or ecology, mathematical modelling has led to many discoveries and insights through a process of synthesis and integration of experimental data, see [Murray, 2002, Murray, 2003]. Also in developmental biology many different morphologies have been the subject of mathematical modelling. Different models are able to produce similar patterns. The question is how to distinguish between them so as to determine which may be the relevant mechanisms. The first necessary condition is reproducibility of the observed patterns. But then it is important to design new experiments to validate the models and consequently verify model hypotheses and biological theories. This may allow, in turn, to provide alternative explanations for the observed phenomena.

The ABC system has already been a subject of mathematical modelling addressing different questions [de Jong, 2002]. Building a new model we have to select its ingredients. The complexity of the model is a key question here. Do we need to take care of all known interactions and environmental factors? We might be tempted to use a detailed knowledge of the system. However, in many cases a simple model may provide better understanding of the underlying mechanisms [Editors, 2000, Tomlin and Axelrod, 2007]. The complexity of a model does not allow a precise analysis that may lead to a mechanistic understanding of

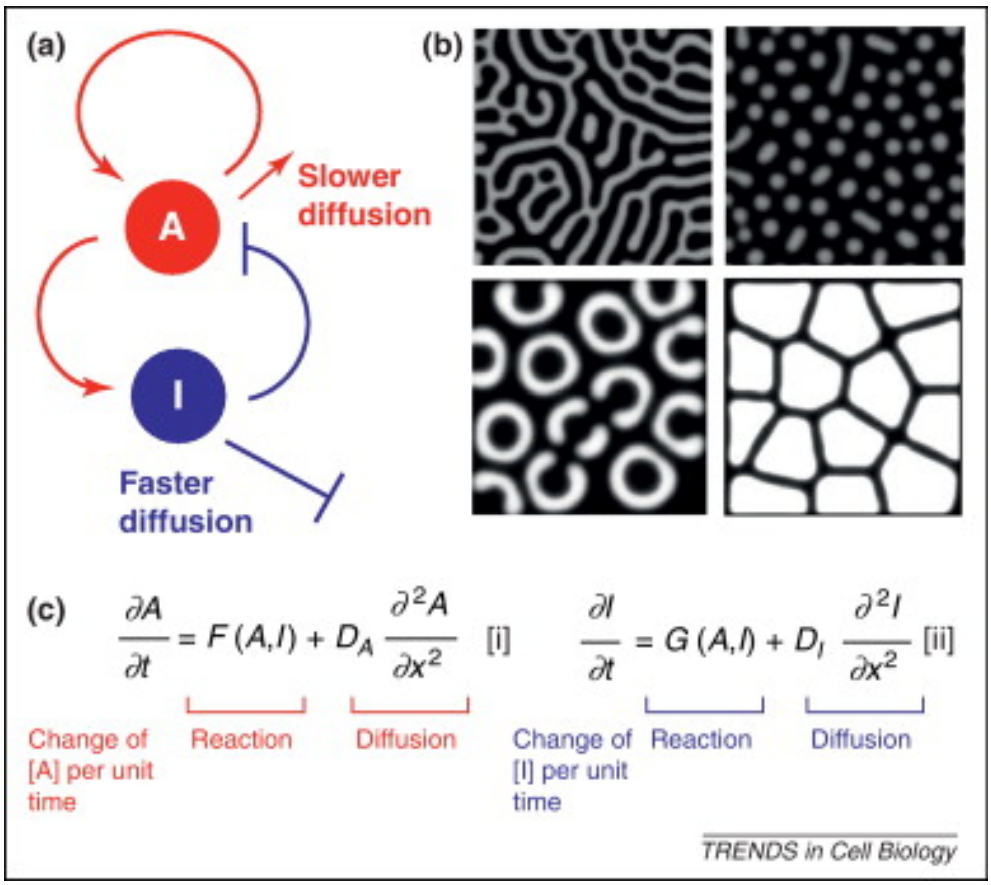


Figure 1.8: Turing models : **(a.)** a Turing model is composed of at least two components, an inhibitor I that diffuse faster than an activator A. **(b.)** Depending of the parameters, obtained patterns can be different. In these simulations we can observe stripes like spots described by the same system of equations. **(c.)** The model can be described by two reaction-diffusion equations that summarize interactions between the components and its capacity to move. [Torii, 2012]

the processes. The complexity is also linked to the number of parameters. A huge number of parameters is usually causing difficulties in model calibration and may lead to existence of different parameter sets providing the same pattern, and hence requiring additional data to validate the model [Giurumescu et al., 2006].

Different kind of models have been created during the past decades [de Jong, 2002]. In the following, we present two of them, first a Boolean model of a genetic regulatory network and then, a reaction-diffusion system applied in the remainder of this thesis. In the latter we follow a parsimonious approach to modelling in which comprehensive models are better understood in view of simpler models.

1.4.1 Genetic Regulatory Networks

To take into account the complexity of certain genetic pathways, Kauffman introduced in 1969 a model of Genetic Regulatory Network (GRN) [Kauffman, 1969]. The approach has been then followed by Aldana [Aldana, 2003]. It is based on a Boolean net of N elements $\{\sigma_1, \sigma_2, \ldots, \sigma_N\}$, each of them being a binary variable $\sigma_i \in \{0, 1\}, i = 1, 2, \ldots, N$ that corresponds to N genes expression of our system. The parameter $\sigma_i = 1$ if the ith gene is expressed and $\sigma_i = 0$ if it is not. At fixed time, each of these elements is given by a function of other elements depending on interactions between the different genes. Then, the value of σ_i at time $t+1$ is given by the value of its K_i controlling elements $\sigma_{j_1(i)}(t), \sigma_{j_2(i)}(t), \ldots, \sigma_{j_{K_i}(i)}(t)$ at time t:

$$\sigma_i(t+1) = f_i(\sigma_{j_1(i)}(t), \sigma_{j_2(i)}(t), \ldots, \sigma_{j_{K_i}(i)}(t)) \tag{1.1}$$

where f_i is a Boolean function associated with the ith element that depends on its controlling elements. The f_i functions can be fixed during all iterations. In that case the function is given by a table of interactions. Nevertheless, these functions can change during the system evolution. After a few iterations with some given initial conditions, fixed points or fixed orbits that are generally called attractors can be obtained. Different attractors can be compared by measuring their differences using, for example, the Hamming distance. It allows to understand possible transitions between different attractors. Evolution of these differences allows to know if the system is chaotic (without stable attractors) or stable.

Similar theoretical approach has been applied to several biological systems, Belleza, [Balleza et al., 2008], including the flower pattering in *Arabidopsis*. The application of

such methods shows that the model should be robust to have stable attractors, but it also should be flexible to allow the creation of different attractors such as, for example, the four whorls of the flower. Modelling needs to find a balance between order and chaos. It is the critical area studied by Aldana [Aldana, 2003]. This convergence near the critical area can be explained by the evolution of the system near a robust and stable system that can be easily conserved between species [Hasty et al., 2001].

1.4.2 Reaction-diffusion model with bistability

Beside the described above Turing patterns located around a destabilised constant equilibrium (*close to equilibrium* patterns), reaction-diffusion systems may lead to emergence of *far from equilibrium* patterns due to bistability and hysteresis in nonlinear interactions.

Such mechanism has been proposed by Marciniak-Czochra in case of receptor-based models, i.e. systems of a single reaction-diffusion equation coupled to ordinary differential equations (ODE) in context of pattern formation in hydra [Marciniak-Czochra, 2006]. In general, equations of such models can be represented by the following initial-boundary value problem,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u,v)\frac{\partial v}{\partial t} = g(u,v) \tag{1.2}$$

with zero-flux boundary conditions.

The variable u describes diffusing extracellular molecules or enzymes, which provide cell-to-cell communication, while v is a vector of variables localised on cells, eg. cell surface receptors or intracellular signalling molecules.

In such models no Turing patterns are stable [Marciniak-Czochra et al., 2016]. Consequently, bistability is necessary to provide stable spatially heterogeneous stationary patterns. Bistability and hysteresis are properties of the kinetic function (Figure 1.9 A). Bistable systems are the systems that have two stable steady states, meaning two attractors that can be reached depending on initial conditions. They correspond to two stable spatially homogeneous equilibria of the full reaction-diffusion system. Diffusion tries to average different states and is the cause of spatio-temporal patterns. As shown in numerical simulations existence of travelling waves excludes formation of stable patterns and the solutions converge to spatially homogenous states. In the case when travelling waves do not exist, stable spatially heterogenous patterns can be observed. The latter is possible in case of a hysteresis-based relation in the quasi-stationary state in the ODEs subsystem of (1.2), i.e. $g(u,v) = 0$, as

depicted in Fig. 1.9 A. It has been shown that the system with bistability but no hysteresis effect cannot exhibit stable spatially heterogeneous patterns, [Köthe and Marciniak-Czochra, 2013]. Hysteresis is necessary to obtain stable patterns.

Properties of hysteresis-based mechanism of pattern formation have been recently studied in a minimal version of the model consisting of one reaction-diffusion equation and one ODE [Köthe and Marciniak-Czochra, 2013]. In such model different stationary solutions can be constructed. The solutions are jump-discontinuous in the non-diffusing variables and may have a form of a transition layer or a boundary layer as shown in [Marciniak-Czochra et al., 2013]. They may also include several jump points, see Figure 1.9 C. Stability of such patterns has been recently investigated in [Härting et al., 2016].

Reaction-diffusion-ODE models with hysteresis have been applied for instance to explain co-existence of different patterns in grafting experiments in hydra [Marciniak-Czochra, 2006]. The model is based on the observation that when transplanting a piece of tissue from a higher position to a lower position along the body axis may lead to formation of an additional head. The process of head formation in hydra is linked to patterns in wnt expression which high concentration leads to cell differentiation and head creation (establishes the identity of the organ), see Figure 1.9 B. The hysteresis model allows to explain co-existence of different patterns depending on the initial condition (transplantation).

1.4.3 Example of models used in developmental biology

Mathematical and computational models have proved to be a powerful technique to address key questions and paradigms in diverse model systems and to provide quantitative insights into mechanisms of development both in animals and plants [Tomlin and Axelrod, 2007, Prusinkiewicz and Runions, 2012].

One of the most famous models was developed to understand the segmentation in drosophila embryo [von Dassow et al., 2000, Von Dassow and Odell, 2002, Ingolia, 2004, Albert and Othmer, 2003]. Segments polarity is defined by the expression of two homeotic genes *wingless* and *Engrailed*. The expression of this two genes is very precise and controls segmentation, as depicted in Figure 1.10a. These genes are regulated by a complex pathway, see Figure 1.10c. The model proposed by von Dassow is an example of *bottom-up* approach. It allows better understanding of which interactions are essential to produce the patterning and how sharp boundaries of different genes expression domains are formed.

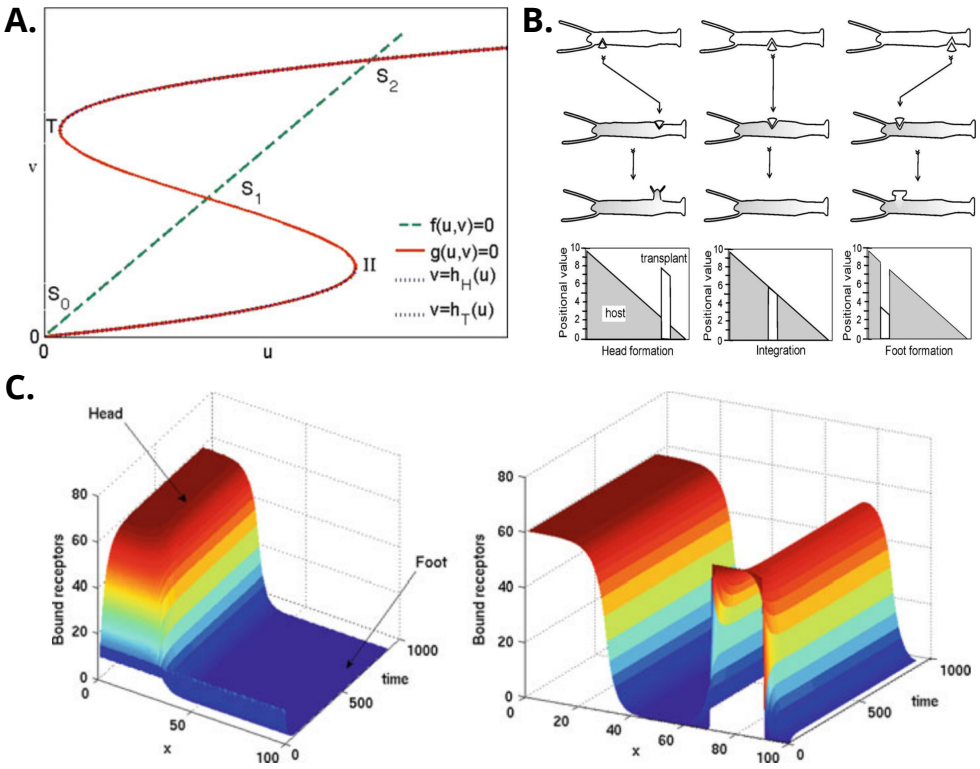


Figure 1.9: Model of wnt concentration in the *hydra*: (A.) Kinetic functions of a model (Equation 1.2) with hysteresis. **(B.)** Representation of the experiment performed on hydra that show that organ definition depend on the initial concentration of a morphogen. **(C.)** Simulation of the morphogen concentration that reproduce the experiments. [Torii, 2012]

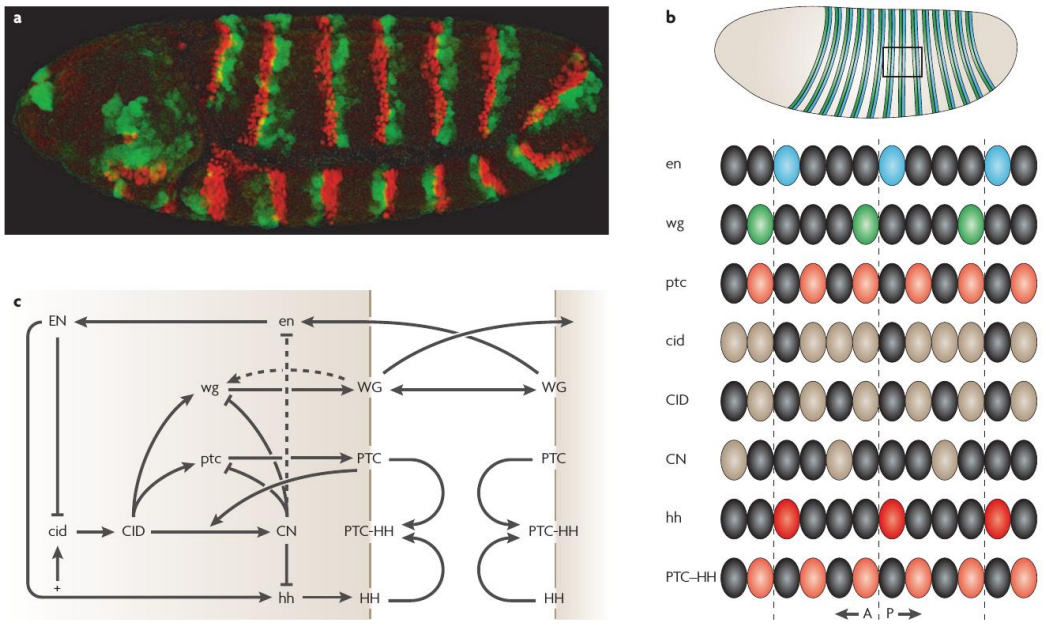


Figure 1.10: Model of the drosophila segment polarity network: (a.) Expression of the segment polarity genes *wingless* (*wg*) in green and *Engrailed* (*EN*) in red. **(b.)** Segment polarity that result on models of von Dassow et al.. The model reproduce segments by modelling the interactions between genes of the network c. The proteins are in upper-case and the mRNA in lower-case. The anterior (A) posterior (P) axis is represented by the arrow. **(c.)** The network used in the model of segment polarity. [Tomlin and Axelrod, 2007]

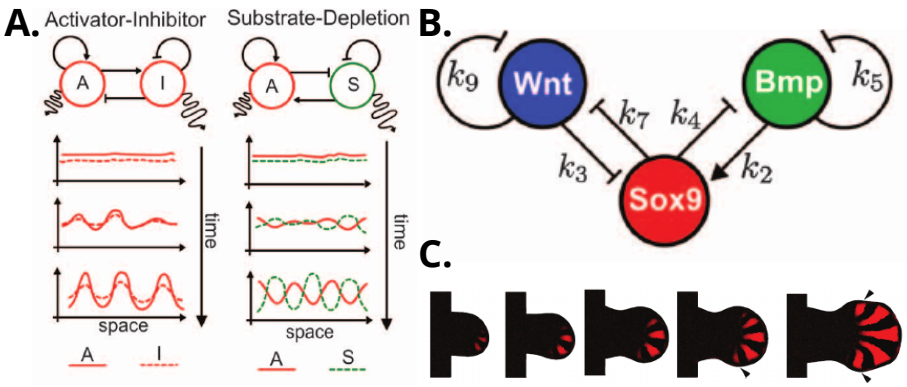


Figure 1.11: Model of the limb development: (A.) Study of interactions necessary to be able to reproduce the pattern formation that lead to digits development. **(B.)** Once the model was selected, protein have been find to correspond of the components of the model. **(C.)** Simulation of the *sox9* protein expression in the limb. [Raspopovic et al., 2014]

The other developmental process for studying pattern formation is the limb development [Raspopovic et al., 2014, Uzkudun et al., 2015]. It is an example of *top-down* models with unknown components to reproduce the observed patterning process (Figure 1.11 A.). Once the pattern was reproduced, the abstract model ingredients were linked to gene expression observed in experiments (Figure 1.11 B.). Construction of the model was based on a simple growth system which was not sufficient to explain the pattern formation. The model was then extended to account for additional components showing that a gradient of two proteins *fgf* and *hoxd13* was necessary to govern a proper digit formation (Figure 1.11 C.).

Plant development provides also interesting modeling questions. An advantage of the plant system for modelling is a fixed structure of the tissue, based on adjacent cells which do not migrate. It allows to model the tissue growth without a necessity to account for individual cell dynamics. One such example is a model of phyllotaxie pattern observed in *Arabidopsis thaliana* during the growth of the plant apex [Smith et al., 2006]. The other processes addressed by mathematical modelling is the maintenance of stem cells in the apex. Interactions between homeotic proteins like WUS or CLV3 in the shoot apex meristem (SAM) and their patterns of expression are well understood, what allowed to build a model of reaction-diffusion equations to reproduce these interactions [Yadav et al., 2011]. The model of signalling pathway was integrated in a simulation of a growing meristem with the capacity of proteins to move between cells. The resulting model proved to be able to reproduce the pattern of expression of the genes involved in stem cell maintenance. It allowed to draw hypotheses explaining genes apparition in the flower [Gruel et al., 2016]. However,

due to the model complexity its quantitative calibration to the data is still an open problem.

1.4.4 Applications of GRN computing methods on the flower patterning

This theoretical approach can be applied to study the establishment of the ABC system. The most prominent model based on the GRN methods is called a FOS-GRN (floral organ specification GRN). Taking into account different gene interactions allows explaining structure of the ABC system. The first approach is based on a binary system of eleven genes (the ABC genes *AP1, AP3, PI, AG* and other transcription factors responsible for the identity of the floral meristem and the activation or repression of the ABC genes: *LFY, EMBRYONIC FLOWER 1 (EMF1), TERMINAL FLOWER 1 (TFL1), CAULIFLOWER (CAL), LUG, UFO* and *SUPERMAN (SUP)*). Based on a table of interactions, iterations of the model show existence of six attractors, four of them corresponding to the four whorls of the flower, one representing non-floral cells and one observed in mutant experiments [Mendoza and Alvarez-Buylla, 1998]. These six attractors can also be deduced from a reduced system of ten genes without *CAL* with three levels of expression 0, s and 1. This system explains that certain genes have less importances in the existence of the four flower attractors. *EMF1* is a floral repressor gene only expressed in non-floral attractors, and UFO, LUG and SUP regulate the fine spatial expression patterns of organ identity genes, so they only play a role in the activation of genes such as *AP3, PI* and *AG*. Thus, *TFL1, LFY, AP1, AG, AP3* and *PI* are sufficient to explain the existence of the four attractors [Mendoza et al., 1999].

Analysis of the fifteen genes system explains the robustness of the four attractors. This system with three levels of gene expression is used to explain different mutant patterns such as *ap2, lug, ap3...* or *p35S::AG* transgenic lines, see Figure 1.12. Interactions among the different mutants are resumed in tables and their analysis provides patterns observed in experiments [Espinosa-Soto et al., 2004]. It is the first model providing explanation of the ABC system. Similar results have been obtained by a GR with only two levels of expression, [Chaos et al., 2006]. This has indicated that a Boolean system is sufficient to explain the existence of four attractors. The results show robustness of the system.

More recently, the FOS-GRN has been applied to analyse transitions between different attractors, leading to an epigenetic landscape. Using stochastic simulations [Alvarez-Buylla et al., 2008] and a continuous GRN model [Davila-Velderrain et al., 2015], evolution of the

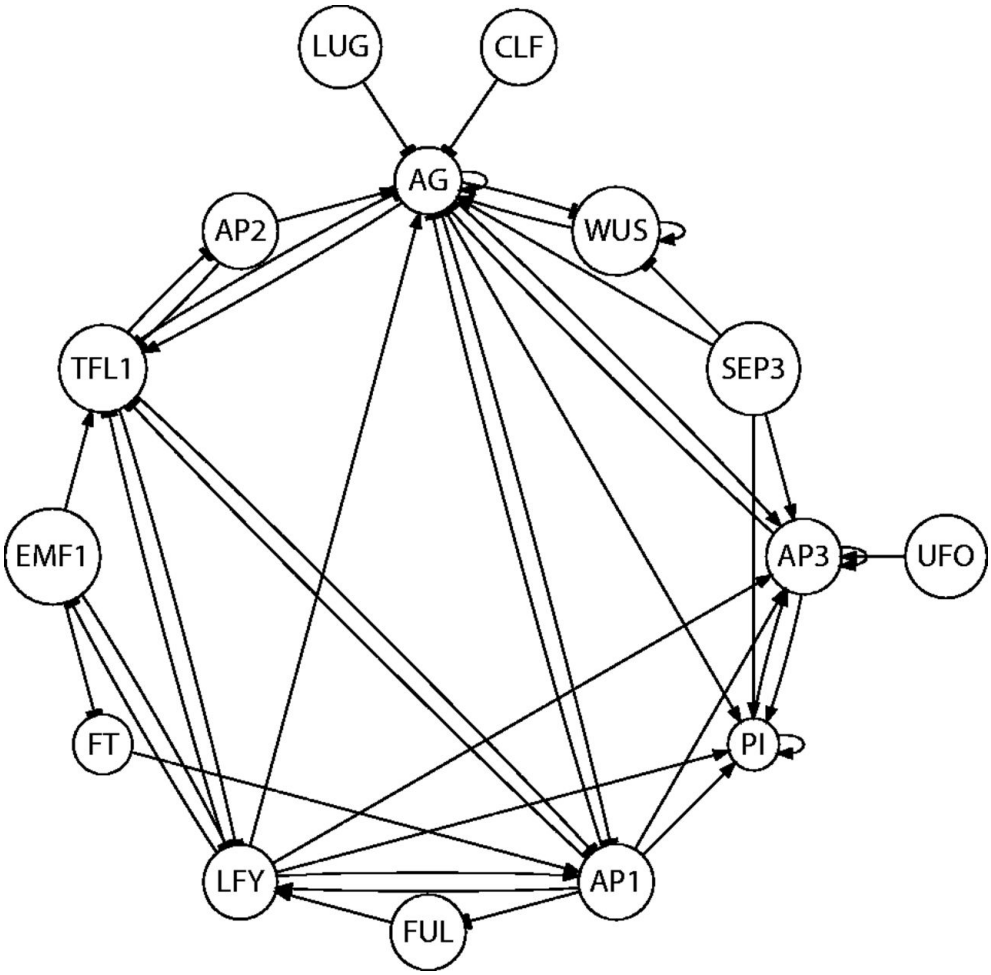


Figure 1.12: Floral organ specification gene regulatory network (FOS-GRN (reproduced from [Davila-Velderrain et al., 2015]) based mainly on mRNA expression patterns. It controls the early differentiation of inflorescences and floral organs in *Arabidopsis Thaliana*. Positive and negative regulatory interactions are represented by arrows and perpendicular lines respectively. This figure is adapted from [Espinosa-Soto et al., 2004].

sepal attractor to the petal one, the petal to the stamen and the stamen to the carpel have been shown. The studies have contributed to the understanding of the geometrical pattern of the ABC genes.

1.4.5 Attempt to create dynamic models

The different models present in the previous section describe a structure of stationary states resembling those found in the four different whorls. However, these models do not explain the organisation of these steady states into four concentric circles. A spatio-temporal model has been proposed recently by Barrio [Barrio et al., 2010]. The model is based on a simplification of the floral meristem into a spherical shape that does not change in time. The spatio-

temporal dynamics should account for two aspects: growing domain (physical field) leading to a symmetry breaking in its shape and evolution of gene interaction network on such a domain. The physical growth is modelled using a phase-field approach with a 'phase-field' ϕ and a 'spontaneous curvature-like' u that follow differential equations. The equations are coupled to a simple GRN accounting for four genes: *AP1, AP3, AG* and *WUS* following the rules described in the previous sections. Numerical simulations of such coupled model show establishment of the four whorls, which is robust to parameter variations. This model has also a capability to reproduce patterns of several mutants. Another model for the four whorls pattern formation has been proposed by Ming [Ming, 2009]. It is based on a GRN of three transcription factors (A, B and C) with auto-regulation of A and C and antagonism of these two transcription factors. B expression is activated in presence of A and C and auxin activates A and C expression. These transcription factors are modelled by three ordinary differential equations based on a function f of activation probability of the transcription of the gene X by the protein Y:

$$f(Y) = \frac{Y^m}{Y^m + \theta_X^m}$$

where θ_X^m is the concentration of Y at which $f = 1/2$ and m is the number of Y binding sites. If Y represses X, then $1 - f(Y)$ function is used. A fourth ODE explains the transport of auxin by PIN FORMED1 (PIN1), one of the auxin efflux facilitators. These equations are integrated in a system regulating cell division. Numerical simulation show that the model is able to capture aspects of the wild-type pattern described in previous sections. However this system stays inconsistent, possibly due to the small number of transcription factors considered [Ming, 2009].

These different models aim to describe spatio-temporal dynamics of proteins production in cells as a consequence of the network of protein interactions. However, they neglect effects of cell-to-cell communication through a transport of the molecules between the cells. There are only few studies explain the importance of diffusion in plant models, for example by using diffusion in the GRN system [Benítez et al., 2008]. A gradient of protein concentrations can also be obtained without diffusion equations, as it was shown in a growing tissue by Chisholm [Chisholm et al., 2010].

1.4.6 Discussion

The existing mathematical models of the ABC system are mainly based on the FOS-GRN approach. They approach the questions of existence and the stability of different whorls of the flower. Our biological question of which mechanisms regulate expression of *AG* in the central dome of the flower at the stage three of development has not been addressed by mathematical models so far. Taking into account the complexity of the *AG* regulation, we develop a mathematical model to test different hypotheses on the regulation of *AG* expression. We consider factors such as a regulation of gene expression by key proteins, ability of a protein to move from one cell to another and domain growth. Different examples show that diffusion of a protein can facilitate formation of spatial patterns [Crampin et al., 1999]. A tissue growth can also have an impact on the patterns of genes expression. Altogether, to investigate the role of these parameters and the outcome of their interplay, we propose a new model of reaction-diffusion equations on a growing domain.

1.5 Methods to detect and quantify gene expression

One of the striking challenges in developmental biology is precise description the expression level of a gene or a protein in a given cell or tissue. During the last decades, researchers have developed methods that allow us to determine where a gene is expressed and, more recently, that allow us to quantify expression at cellular resolution. Below, I will review two methods to measure the transcriptional activity of a gene – by *in situ* hybridisation and the *in vivo* quantification of proteins with confocal imaging. These methods have enabled a proper characterisation of *AG* expression.

1.5.1 Methods to localise mRNA

in situ hybridisation methods were first developed most than thirty years ago to localise mRNA within a given tissue [Cox et al., 1984]. The principle is to use a thin section of fixed tissue and then detect the presence of a desired mRNA with antisense sequence of RNA linked to a moiety that can be revealed with different methods of colouration. This allows a qualitative estimation of the region where a particular mRNA is present in a given tissue, but not the quantification of the degree of expression, as the length of the revelation (colouring) step may dramatically change the final aspect of the *in situ*. This method also

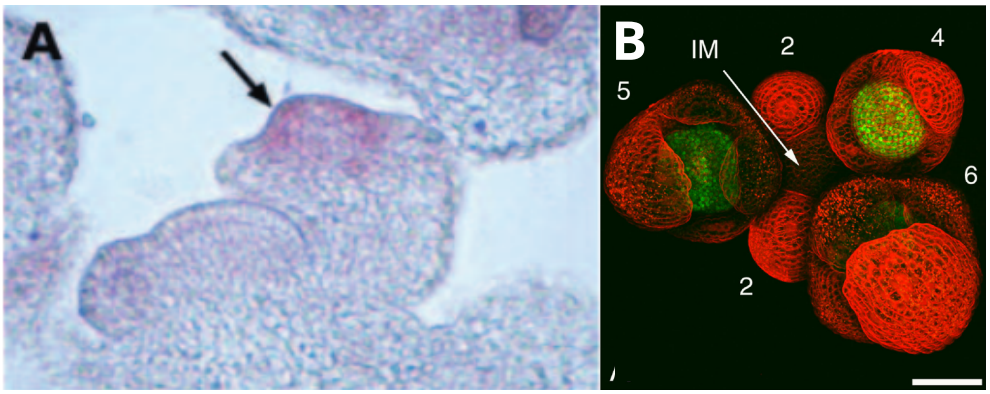


Figure 1.13: Determination of AG expression patterns by different approaches: **(A.)** *in situ* hybridisation of AG showing the localisation of AG mRNA in the central dome at stage 3 flowers . **(B.)** Projection of a confocal image of the translational fusion of AG with a Green Fluorescent Protein (GFP) reporter, counterstained with FM4-64 to mark plasma membranes showing a protein pattern similar to the mRNA. Panel A is adapted from [Gómez-Mena et al., 2005] and panel B from [Urbanus et al., 2009]

requires thin sections to be generated and so doesn't easily allow the detection of the pattern of expression in whole tissues. It is using this method that the pattern of *AG* expression was first described, showing that mRNA is present in the central dome of stage three flowers [Drews et al., 1991, Gómez-Mena et al., 2005] (Figure 1.13 A.).

Because this method provides data on sections through the meristem, gaining an understanding of the spatio-temporal expression dynamics of a gene of interest requires an examination of a large number of samples, as well as some level of subjective interpretation, as it is impossible to know the exact plane of the section through the tissue. Furthermore, it is not clear what the minimal size of a domain must be in order for it to be detected through such colorimetric stainings on sectioned tissues. To get around these limitations in order to gain a more holistic view of gene expression dynamics, a method was developed in the lab to do *in situ* hybridisations on whole meristems [Rozier et al., 2014].

1.5.2 Confocal imaging to locate the protein in the tissue

While colorimetric whole-mount *in situ* hybridisation experiments provide tissue-level data on expression dynamics, it is still insufficient to generate detailed cellular quantifications, as well as to examine how the gene behaves in single cells over time. To these ends, the preferred option is to generate a fusion of the protein of interest with a fluorescent protein like GFP and then to integrate this construct into the genome of the plant. Fluorescent proteins like GFP, which was discovered from the jellyfish *Aequora victoria* [Shimomura et al., 1962] allow

the visualisation of the fused protein within an *in vivo* tissue using confocal laser scanning microscopy (CLSM). Here, the fluorescent protein (FP) is excited with a laser of a very specific wavelength and the emission are recorded with the help of a sophisticated detector. Different FP with different optical properties have been developed, either by mutating known FP or by investigating different organisms, thus permitting the simultaneous observation of several proteins in the same tissue [Cranfill et al., 2016]. However, adding a fluorescent protein at the C- or N-terminal ends of the protein of interest can alter its properties, by changing the conformation of the protein, disturbing its binding capacity to DNA or to another protein, resulting in a non-functional or partially-functional protein and a potentially incorrect pattern of expression. It is thus important to test the functionality of the chimeric protein by rescuing a mutant, and to confirm the observed pattern with *in situ* hybridisation experiments. In a similar vein, we can also stain specific components of the tissue to have a better visualisation and measurement of our protein, for instance by marking the plasma membrane with a lypophillic vital stain such as FM4-64. This technique allows AG to be observed in the flower and confirms the expression of *AG* at stage 3 (Figure 1.13 B.).

CLSM observations provide excellent precision with respect to the tissular dynamics of protein expression, and even to the relative intensities of protein levels in different regions, but is still insufficient for a thorough quantification of those levels and a complete analysis of dynamics at the cell level. One issue in this respect is the z resolution, which is usually much lower than the x or y resolutions. Our experiments, for example, typically use a resolution of $0.2\mu m$ along the x and y axes and a resolution of $1\mu m$ in the z axis. This difference is due to the properties of the objective lenses on the microscope. To reduce this problem and to obtain isotropic images with identical resolutions in all axes, our group has collaborated with bio-informaticians in Montpellier developed a method called MARS, wherein a given sample is imaged from 2-4 different angles, which are then fused to generate a single higher-resolution image that can be segmented [Fernandez et al., 2010] (Figure 1.14). This treatment can be performed on cell outlines, as well as on any fusion reporter present in the plant. This technique allows us to measure different properties of the cell, such as the volume, the anisotropy or the curvature of the cell and to precisely quantify and characterise cell behaviours at the different stages of flower development.

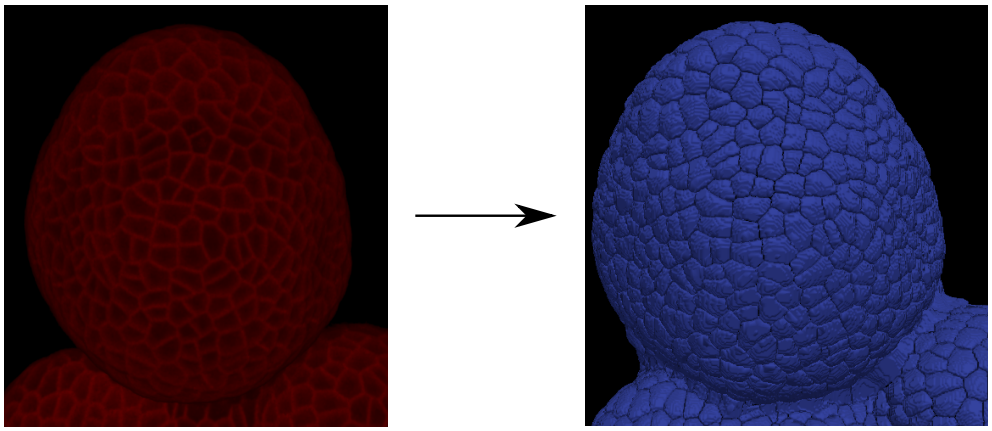


Figure 1.14: 3D reconstruction of a stage 2 flower. Projection of a stage 2 flower with a membrane marker (on the left). After imaging from multiple angles and the fusion of those image stacks, the resultant high-resolution image is segmented to provide an image where each cell is assigned a number (on the right). This data can be used to determine cellular metrics such as volume, anisotropy, curvature etc.

1.5.3 4D imaging of the flower development

One of the advantages of confocal imaging is that it can be carried out *in vivo*. Furthermore, with the improvement of confocal technologies (such as better detectors), the lasers may be used at a much lower power, thus causing less damage to the images tissue. This in turn permits us to image the same sample several times at regular intervals during development and to study the growth of an organ or the dynamics of gene expression. This method has been widely used in animal and in plant development [Roeder et al., 2012, Faure et al., 2016]. Using the tools described above, we are able to analyse development (Figure 1.15), by imaging the same flower every 4-18 hours, segmenting those images and then determining the lineages of the cells in the sample. Based on this analysis, tools have been developed in the group to create a temporal graph that summarizes the cellular properties with their evolution over time (Figure 1.16. Subsequently, statistical analyses may be carried out on these data in order to find correlations between different properties, such as between gene expression and cell morphometry. This work is unpublished but is described in the thesis of Jonathan Legrand and will be used later in next chapters.

One of the key limitations in the published protocol for MARS and 4-D real-time growth analysis [Fernandez et al., 2010] is the time interval between two serial image acquisitions. Indeed, our protocol used the lipophilic dye FM4-64 to stain cell membranes in order to be able to segment the tissue. Because this dye is internalised via endocytosis within a few

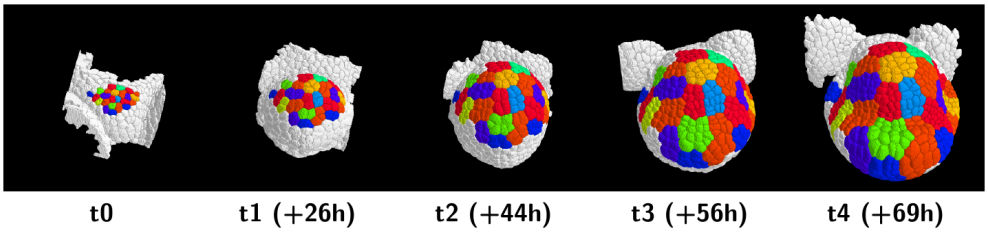
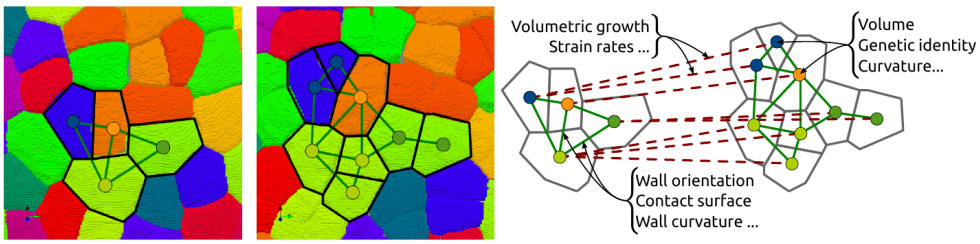


Figure 1.15: Time course of a flower during early stages of development: Time course of a single flower stained with FM4-64 and image at multiple time points from multiple angles. The images for each time point are fused and segmented and the lineage is then determined across the entire time course. The colors represent these lineages. The time interval between images is indicated below the projections.



a. Topological graph **b.** Spatio-temporal graph

Figure 1.16: Definition of the temporal graph: The topological graph summarizes the geometrical properties of the cells of a single image. It can include all available data for each cell, such as protein signal or other properties. Using the lineage, a spatio-temporal graph is created to link properties of mother cells to daughter cells.

hours, it has to reapplied to the sample just prior to every image acquisition. Additionally, because it is slightly toxic to the plant, it cannot be applied very frequently, thus precluding the possibility of short time intervals. One way to get around this is by using an FP fused to a protein or protein fragment with a membrane localisation signal [Reddy et al., 2004].

1.5.4 Discussion

Later in this manuscript, I will show how I have used these techniques to observe the dynamics of *AG* expression. Furthermore, I will analyse these dynamics using quantified protein fluorescence at cellular resolution using a time series with short intervals between acquisitions. This has allowed us to gain a better understanding of precisely how this gene is activated in a robust spatio-temporal manner within the growing flower.

1.6 Stochasticity and robustness in development, two fields that are often linked

Stochasticity and robustness are two processes that are widely studied in developmental biology. Behind these two concepts are different principles that I will briefly review here. The study of stochasticity and robustness are sometimes linked and I will try to explain why. I have previously explored what is behind stochasticity in development, especially in plants and its link with robustness (Article below this section).

1.6.1 Robustness in plant development

Robustness can be defined as the absence, or a low level, of variation in response to a specific environmental or genetic perturbation [Félix and Barkoulas, 2015]. The concept is used in evolution to try to understand why phenotypic traits are conserved between species. Robustness can be studied in two main contexts: firstly to understand the evolution of phenotypic traits and assess the variation caused by environmental changes; and secondly, to analyse the stability of a pattern in complex systems, in order to determine which interactions produce redundancy in a system.

Robustness to environmental studies was first studied by Waddington who said that “the wild-type of an organism, that is to say, the form which occurs in nature under the influence

of natural selection, is much less variable in appearance than the majority of mutant races" [Waddington, 1942]. However, detailed studies of robustness are mainly much more recent. One example is the attempt to measure and quantify variations between species, or at the tissular and cellular levels within an organism [Mestek Boukhibar and Barkoulas, 2015]. In plants, the flower has a structure that is more or less conserved between species, though individual characteristics, such as flower size or petal color can be very different [Krizek and Fletcher, 2005]. At the tissue level, one study shows that meristem size size depends on light or temperature conditions [Landrein et al., 2015]. With respect to AG, one study analysed the conservation of its second intron, showing that a few sites are very well conserved between species, and that these sites correspond to regulatory elements bound by known key floral patterning factors [Hong and Hamaguchi, 2003, Causier et al., 2009].

In complex systems, robustness is a key concept. It addresses the key question of whether small variations in developmental parameters could give rise to similar patterns. This question is important because it is difficult to precisely reproduce a biological model, implying that a model is only an approximation of our system and to provide meaning to it, the result should be robust to small variations present in the biological system. With respect to the ABC model of flower development, boolean model were used to study the stability of gene expression in the four whorls. This showed that the system is stable with small variations in each whorls, but also that pattern of expression in the adjacent whorls are closer, meanings that smaller variations in gene expression are sufficient to reach the identity of the adjacent whorls than the one to reach a further one, so that significant variation can lead to the expression of TF that correspond to the identity of the adjacent whorls. This model was also used to study the conservation of the ABC model between species [Sánchez-Corrales et al., 2010]. These results show how robust can be the final pattern of expression of TF that are necessary for flower organs identity in *Arabidopsis thaliana* but also in other species, nevertheless, the mechanism that lead to this robust pattern are not fully understood.

1.6.2 Link between stochasticity and robustness

Stochasticity and robustness are often linked [Macneil and Walhout, 2011]. It has been shown that there are mechanisms that try to buffer stochasticity in order to obtain robustness [**?**, Holloway et al., 2011]. We also see that stochasticity can produce robust patterns like for the cone of the drosophila eyes [Wernet et al., 2006]. Another example is the reprogramming

of somatic cells into induced pluripotency cells. In this case, stochasticity is useful to produce a switch to a state that is robust. The activation is due to the expression of a group of transcription factors that permits dedifferentiation. Nevertheless, only few cells that are treated with this transcription factors are dedifferentiated (1% to 20% after 4 weeks) but this process is very robust after a longer time (92% after 18 weeks) [Hanna et al., 2009].

Previous studies show that the ABC system and AG expression are robust in *Arabidopsis* but also in other species. The pattern of AG expression is well known and allows us to judge the importance of a switch leading to a rapid activation in the central dome at stage 3.However, no precise data exists for AG expression during flower development. The tools presented in previous subsections will allow us to generate such data, and to test the robustness of the AG pattern at tissular and cellular levels, as well as to determine the presence of stochasticity.

frontiers in
PLANT SCIENCE

OPINION ARTICLE
published: 14 November 2014
doi: 10.3389/fpls.2014.00643

Models to reconcile plant science and stochasticity

Sam Collaudin[1,2] and Vincent Mirabet[1,2]*

[1] Reproduction et Développement des Plantes, INRA, CNRS, Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, Lyon, France
[2] Laboratoire Joliot-Curie, CNRS, Ecole Normale Supérieure de Lyon, Lyon, France
*Correspondence: vincent.mirabet@ens-lyon.fr

Edited by:
Naomi Nakayama, University of Edinburgh, UK

Reviewed by:
James Locke, University of Cambridge, UK
Jose Teles, University of Cambridge, UK

Keywords: stochasticity, plant development, modeling, pattern, morphogenesis, tissue

Plants are modular organisms that exhibit diverse adaptations to variability. This variability can be intrinsic in nature, as in the case of cell shape or division plane stochasticity, protein distribution in a cell, variations in internal mechanical properties etc... (Altschuler et al., 2008; Besson and Dumais, 2011). It can also be extrinsic, as with variations in environmental conditions at different time scales (Wolpert et al., 1998; Sultan, 2000; Franklin, 2009; Leyser and Day, 2009). When it comes to rationalizing data acquisition and interpretation, one has the tendency to define what part of the variability is arguably unhelpful stochasticity and what part does in fact contain meaningful information.

Systems biology, which combines methodologies from various disciplines, can be used to understand the mechanisms of development. For example, complex network analysis (Lucas et al., 2011), computer simulations (Band et al., 2012) or physical measurements through atomic force microscopy (Milani et al., 2014) can be combined with biological experiments. For instance, such an approach has been able to produce reasonable explanations for how patterning at the meristem level can lead to the stem structure (Prusinkiewicz et al., 1995). Stochasticity in models as a variable or as a methodological tool has been a subject of interest for many years in physics and mathematics (Saguès et al., 2007; Friedrich et al., 2011; Wilkinson, 2011). Studies have already been published in biology but only a few focused on plant development, and are often more recent (for a review of this aspect, see Meyer and Roeder, 2014). Along with a better understanding of growth processes, those studies have also illustrated how our vision of stochasticity was previously too derogatory (Kliebenstein, 2012). Those new methodologies illustrate how stochasticity can be both a consequence and an origin of core mechanisms in development.

Here we use specific examples to illustrate how mathematical or computational models are well-suited to the study of stochasticity in plant functions. Moreover, models enable the use of measured phenotypic stochasticity at multiple scales to elucidate the underlying processes. We suggest that models used for such purposes do not need to be overly complex, and various complex models of the same process will in fact converge toward similar conclusions. We will focus our attention on apical meristems and the growth that they generate, where cell–cell interactions underlie the emergence of various interesting properties of the tissues and organs.

STOCHASTICITY CAN BE BUFFERED BY GENETIC NETWORKS AT CELLULAR AND TISSULAR SCALES

It has been shown that low levels of a protein or a chemical component induce a high level of noise that can impact on pattern formation (Shnerb et al., 2000). For instance, in *Drosophila*, the *Hunchback* (*Hb*) gene is crucial for the proper segmentation of the embryo, and is regulated by *Bicoid* (*Bcd*). The Bcd gradient can cause intense noise due to the small numbers of molecules. Nevertheless, the definition of segmentation and boundary position are well-conserved between embryos despite Bcd stochasticity. Holloway proposed in his study that noise can be affected by the number and the strength of binding sites on a promoter (Holloway et al., 2011). He develops computational models to confirm that the high number of Bcd binding sites on the Hb promoter reduce stochastic noise of Bcd gradient.

Complex gene regulatory networks drive morphogenesis in all multicellular contexts. Feedback and redundancy within these networks compensate for intrinsic or extrinsic stochasticity. For instance flower formation is driven by the activation of a small network mainly composed by *LEAFY (LFY)*, *APETALA1 (AP1)*, *CAULIFLOWER (CAL)*, and *TERMINAL FLOWER1 (TFL1)*. This network is regulated by environmental and physiological inputs to start flower initiation at the appropriate time for reproduction. This network contains many feedback loops and mutual activations, such as the induction by LFY of AP1 and CAL, which themselves positively regulate LFY. These interactions can buffer the environmental noise to obtain the formation of a robust pattern and to avoid the reversal of flowering (Blazquez et al., 2006).

MODELS EXPLAIN THE ROBUSTNESS OF PATTERNING

Plant tissues, even those as little differentiated as meristems, exhibit strong self-organizational properties. Intense local cell–cell interactions through diverse exchanges (Murray et al., 2012; Landrein and Vernoux, 2014) contribute to the emergence of patterns. Those patterns may be in the form of either simple genetic differentiation or more complex morphogenetic events. Among the various properties of such self-organized

patterning, robustness is crucial for the principal meristematic properties: their ability for self-renewal or to produce various lateral organs. Models are well-suited to predict how a set of linked cells can generate shape and differentiate following emergent processes.

Auxin signaling processes are amongst the best-studied cases of tissue patterning in plants, and furthermore, they have also been extensively modeled and linked to numerous biological observations. In the auxin flux models, simple cell-to-cell communication occurs via the local amplification of auxin flux by the PIN-FORMED proteins. The auxin response patterns observed in the meristematic tissues of the stem and at the vascular generation zones, are an emergent property of those molecular interactions (Sassi and Vernoux, 2013). This system clearly illustrates the robustness of emergent patterns to external noise. External noise can be as intense as multicellular injuries. It has been shown that an injury to the meristem can be compensated. The patterning system is robust enough and can maintain the activity of organogenesis (Snow and Snow, 1932; Reinhardt et al., 2005). Computational Models of phyllotaxis can predict how the plant might cope with such ablations; the pattern is very quickly deformed around the ablation, but re-emerges naturally as growth continues to produce healthy new cells. This resistance to local injuries is also observed in vasculature development. Cutting a part of the provasculature induces its spontaneous reconfiguration, such that the new vasculature is reshaped around the ablation (Sauer et al., 2006). Models predict that such reconfiguration does not need any specific change in cell behavior, and that cell–cell communication itself is sufficient to enable such changes (Wabnik et al., 2010).

STOCHASTICITY IN PATTERNING IS FILTERED IN PLANT TISSUES

If spatial robustness is a natural outcome of the self-organization described above, this kind of patterning is itself sometimes stochastic. The phyllotactic angle between successive lateral organs forming on the shoot apical meristem (SAM) is approximately 137 degrees in *Arabidopsis thaliana*. For a long time, research has focused on predicting mechanisms behind the astonishing regularity of phyllotaxis in various plant systems (Adler, 1997). More recently, the close examination of plants phyllotaxis has led to the discovery of strange phenotypic alterations. The *histidine phosphotransfer protein 6 (ahp6)* mutant presents curious alterations called M-shaped successions, where three successive lateral organs display altered angles (Besnard et al., 2013). Other, more complex, successions are visible with lower frequencies. Strikingly, these types of alterations also occur in wild type plants, though less frequently. In order to understand the source of this stochasticity and the specific pattern of alterations, both statistical (Refahi et al., 2011; Guédon et al., 2013) and agent (Mirabet et al., 2012) models have been used to study phyllotaxis in *Arabidopsis*. These models have predicted that the auxin system, under the influence of stochasticity, can spontaneously generate the alterations seen along the stem. Indeed in mutants some organs can be generated simultaneously, whereas a delay (called the plastochron) occurs in a typical normal situation. Because the organs appear simultaneously, the way they are arranged along the stem may be inverted, thus producing the characteristic M-shaped structure. This stochasticity in timing would thus appear to be a spontaneous outcome of the spatial patterning of the auxin system.

These studies have helped clarify that the *Arabidopsis* SAM possesses a second patterning system, based on the AHP6 protein, that partially overlaps the auxin system and ensures that new primordia will emerge successively through time. Thus, the temporal stochasticity of the auxin system is compensated for by a second patterning process that filters it. Without the use of a "systemic" view of the entire patterning process, it would have been difficult to decipher the role of the AHP6 system.

STOCHASTICITY AS A SOURCE OF PATTERNING AND MORPHOGENESIS

In developmental biology, stochastic gene expression can lead to the formation of coherent patterns. An example is in the ommatidium of the *Drosophila* eye, which consists of eight photoreceptor cells. Two of them (R7 and R8) express rhodopsin, which is responsible for the detection of color. It has been shown that the separation of "yellow" and "pale" ommatidia determined by rhodopsin regulation in R7 and R8 is due to the stochastic expression of the *SPINELESS* receptor (Wernet et al., 2006). This stochasticity is both necessary and sufficient for proper ommatidial development. In this example, stochastic gene expression at the cell level can become instructional at the tissue level.

Through the use of simple activator-inhibitor model systems, Turing managed to describe the self-organization of various spatial patterns (Turing, 1952). These patterns mainly depend on the strength of molecular interactions and on the geometry of the domains where the activators and inhibitors are expressed. In these computational models, stochasticity is necessary to trigger the dynamics that leads to the final stable pattern. Stochasticity of cell behaviors becomes the motor of patterning. Nevertheless, this stochasticity is in a way buffered by the interactions, as its intensity has only a small effect on the final pattern. In plants, examples of such systems exist in trichomes positioning in leaves (Benítez et al., 2007; Greese et al., 2012). Interactions can be summarized into an activator complex (that consists of *WEREWOLF*, *GLABRA1*, *GLABRA3*, *ENHANCER OF GLABRA3,* and *TRANSPARENT TESTA GLABRA*) as well as the redundant inhibitory activity of *CAPRICE* and *TRIPTYCHON*. With Turing-like models applied to those components, the authors were able to reproduce the experimentally observed patterns.

Stochasticity is present not only in gene expression, but is an inherent property of cells, notably with respect to cell growth. A recent study showed that cells are able to interact mechanically to adapt their growth depending on the behaviors of their neighbors (Uyttewaal et al., 2012). Interestingly, this function seems to increase variability instead of compensating for it. In turn this positive feedback is necessary for correct morphogenesis of new primordia. Models predict that an optimum exists between variability of cell growth and feedback between cells. Depending upon the relative strength of both parameters, the tissue can grow more or less efficiently. This intricate interplay between stochasticity and cell–cell

communication is a fundamental aspect of tissue morphogenesis, and would appear to be regulated. Models can help predict the optimal ratio between stochasticity and feedback necessary for proper morphogenesis. Interestingly, it is not this theoretical optimum that seems to be generated in meristems, a fact that may allow the tissue to undergo growth bursts, which may in turn lead to primordia emergence (Alim et al., 2012).

SIMPLE MODELS TRANSLATE VARIABLE PHENOTYPES INTO VALUABLE INFORMATION

Complex systems can be modeled quite simply. An example is human crowds being modeled as simple interacting agents with very basic properties. Such models can efficiently predict the behavior of these groups (Helbing et al., 2000). Similarly, plant cells and tissues can also be modeled using such approaches. With a simple model such as that of Turing (with less than 10 parameters), it is possible to add noise measured at the cell scale, and study its consequences at an higher (tissue or plant) level. Thus, phenotypic variability at this higher level can be interpreted through the model, that gives the ability to search for the cellular parameters leading to the mutant phenotype of interest.

In the example of phyllotaxis described above, the types and frequencies of alterations may be interpreted through the use of the model. They are predicted to be an outcome either of alterations of the meristem structure or the auxin system. This scenario may be easily tested with further experimentation, for example searching for defects in the pin network or meristematic size.

This reasoning is in fact multiscale, each conclusion at one scale providing the data for models that focus on the link to the next level of organization. Our ability to measure variability and stochasticity at various scales has recently been increased. Experimental techniques and analysis tools have immensely improved the precision of measurements both spatially and temporally, and at various scales (Fernandez et al., 2010; de Reuille et al., 2014). Those new techniques point out the importance of heterogeneity and stochasticity in biological systems. Modeling approaches will be more and more helpful in this new context to explain those data.

It is time to switch from seeing biology as clockwork perfection to looking at its natural variations more thoroughly. That will undoubtedly help us decipher where plants' real beauty is hidden: behind those so-called imperfections.

ACKNOWLEDGMENTS

We thank Pradeep Das and Arezki Boudaoud for fruitful comments.

REFERENCES

Adler, I. (1997). A history of the study of phyllotaxis. *Ann. Bot.* 80, 231–244. doi: 10.1006/anbo.1997.0422

Alim, K., Hamant, O., and Boudaoud, A. (2012). Regulatory role of cell division rules on tissue growth heterogeneity. *Front. Plant Sci.* 3:174. doi: 10.3389/fpls.2012.00174

Altschuler, S. J., Angenent, S. B., Wang, Y., and Wu, L. F. (2008). On the spontaneous emergence of cell polarity. *Nature* 454, 886–889. doi: 10.1038/nature07119

Band, L. R., Fozard, J. A., Godin, C., Jensen, O. E., Pridmore, T., Bennett, M. J., et al. (2012). Multiscale systems analysis of root growth and development: modeling beyond the network and cellular scales. *Plant Cell* 24, 3892–3906. doi: 10.1105/tpc.112.101550

Benítez, M., Espinosa-Soto, C., Padilla-Longoria, P., Díaz, J., and Alvarez-Buylla, E. R. (2007). Equivalent genetic regulatory networks in different contexts recover contrasting spatial cell patterns that resemble those in Arabidopsis root and leaf epidermis: a dynamic model. *Int. J. Dev. Biol.* 51, 139–155. doi: 10.1387/ijdb.062183mb

Besnard, F., Refahi, Y., Morin, V., Marteaux, B., Brunoud, G., Chambrier, P., et al. (2013). Cytokinin signalling inhibitory fields provide robustness to phyllotaxis. *Nature* 505, 417–421. doi: 10.1038/nature12791

Besson, S., and Dumais, J. (2011). Universal rule for the symmetric division of plant cells. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6294–6299. doi: 10.1073/pnas.1011866108

Blazquez, M., Ferrandiz, C., Madueno, F., and Parcy, F. (2006). How floral meristems are built. *Plant Mol. Biol.* 60, 855–870. doi: 10.1007/s11103-006-0013-z

de Reuille, P. B., Robinson, S., and Smith, R. S. (2014). Quantifying cell shape and gene expression in the shoot apical meristem using MorphoGraphX. *Methods Mol. Biol.* 1080, 121–134. doi: 10.1007/978-1-62703-643-6_10

Fernandez, R., Das, P., Mirabet, V., Moscardi, E., Traas, J., Verdeil, J.-L., et al. (2010). Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. *Nat. Methods* 7, 547–553. doi: 10.1038/nmeth.1472

Franklin, K. A. (2009). Light and temperature signal crosstalk in plant development. *Curr. Opin. Plant Biol.* 12, 63–68. doi: 10.1016/j.pbi.2008.09.007

Friedrich, R., Peinke, J., Sahimi, M., and Reza Rahimi Tabar, M. (2011). Approaching complexity by stochastic methods: from biological systems to turbulence. *Phys. Rep.* 506, 87–162. doi: 10.1016/j.physrep.2011.05.003

Greese, B., Wester, K., Bensch, R., Ronneberger, O., Timmer, J., Huulskamp, M., et al. (2012). Influence of cell-to-cell variability on spatial pattern formation. *IET Syst. Biol.* 6, 143–153. doi: 10.1049/iet-syb.2011.0050

Guédon, Y., Refahi, Y., Besnard, F., Farcot, E., Godin, C., and Vernoux, T. (2013). Pattern identification and characterization reveal permutations of organs as a key genetically controlled property of post-meristematic phyllotaxis. *J. Theor. Biol.* 338, 94–110. doi: 10.1016/j.jtbi.2013.07.026

Helbing, D., Farkas, I., and Vicsek, T. (2000). Simulating dynamical features of escape panic. *Nature* 407, 487–490. doi: 10.1038/35035023

Holloway, D. M., Lopes, F. J., da Fontoura Costa, L., Travençolo, B. A., Golyandina, N., Usevich, K., et al. (2011). Gene expression noise in spatial patterning: hunchback promoter structure affects noise amplitude and distribution in *Drosophila* segmentation. *PLoS Comput. Biol.* 7:e1001069. doi: 10.1371/journal.pcbi.1001069

Kliebenstein, D. J. (2012). Model misinterpretation within biology: phenotypes, statistics, networks, and inference. *Front. Plant Sci.* 3:13 doi: 10.3389/fpls.2012.00013

Landrein, B., and Vernoux, T. (2014). "Auxin, chief architect of the shoot apex," in Auxin and *its* Role in Plant Development, eds E. ZaŽímalová, J. Petrášek, and E. Benková (Wien: Springer-Verlag), 191–212. doi: 10.1007/978-3-7091-1526-8_10

Leyser, O., and Day, S. (2009). *Mechanisms in Plant Development*. Oxford: John Wiley & Sons.

Lucas, M., Laplaze, L., and Bennett, M. J. (2011). Plant systems biology: network matters. *Plant Cell Environ.* 34, 535–553. doi: 10.1111/j.1365-3040.2010.02273.x

Meyer, H. M., and Roeder, A. H. K. (2014). Stochasticity in plant cellular growth and patterning. *Front. Plant Sci.* 5:420. doi: 10.3389/fpls.2014.00420

Milani, P., Mirabet, V., Cellier, C., Rozier, F., Hamant, O., Das, P., et al. (2014). Matching patterns of gene expression to mechanical stiffness at cell resolution through quantitative tandem epifluorescence and nanoindentation. *Plant Physiol.* 165, 1399–1408. doi: 10.1104/pp.114.237115

Mirabet, V., Besnard, F., Vernoux, T., and Boudaoud, A. (2012). Noise and robustness in phyllotaxis. *PLoS Comput. Biol.* 8:e1002389. doi: 10.1371/journal.pcbi.1002389

Murray, J. A. H., Jones, A., Godin, C., and Traas, J. (2012). Systems analysis of shoot apical meristem growth and development: integrating hormonal and mechanical signaling. *Plant Cell* 24, 3907–3919. doi: 10.1105/tpc.112.102194

Prusinkiewicz, P., Hammel, M., Mech, R., and Hanan, J. (1995). "The artificial life of plants," in *Artificial Life for Graphics, Animation, and Virtual Reality* (ACM SIGGRAPH), 7, 1–1, 1–38.

Refahi, Y., Farcot, E., Guédon, Y., Besnard, F., Vernoux, T., and Godin, C. (2011). "A combinatorial model of phyllotaxis perturbations in *Arabidopsis thaliana*," in *Combinatorial Pattern Matching,* eds R. Giancarlo and G. Manzini (Berlin; Heidelberg: Springer), 323–335. doi: 10.1007/978-3-642-21458-5_28

Reinhardt, D., Frenz, M., Mandel, T., and Kuhlemeier, C. (2005). Microsurgical and laser ablation analysis of leaf positioning and dorsoventral patterning in tomato. *Development* 132, 15–26. doi: 10.1242/dev.01544

Saguès, F., Sancho, J., and García-Ojalvo, J. (2007). Spatiotemporal order out of noise. *Rev. Mod. Phys.* 79, 829–882. doi: 10.1103/RevModPhys.79.829

Sassi, M., and Vernoux, T. (2013). Auxin and self-organization at the shoot apical meristem. *J. Exp. Bot.* 64, 2579–2592. doi: 10.1093/jxb/ert101

Sauer, M., Balla, J., Luschnig, C., Wisniewska, J., Reinöhl, V., Friml, J., et al. (2006). Canalization of auxin flow by Aux/IAA-ARF-dependent feedback regulation of PIN polarity. *Genes Dev.* 20, 2902–2911. doi: 10.1101/gad.390806

Shnerb, N. M., Louzoun, Y., Bettelheim, E., and Solomon, S. (2000). The importance of being discrete: life always wins on the surface. *Proc. Natl. Acad. Sci. U.S.A.* 97, 10322–10324. doi: 10.1073/pnas.180263697

Snow, M., and Snow, R. (1932). Experiments on phyllotaxis. II. The effect of displacing a primordium. *Phil. Trans. R. Soc. Lond. B* 222, 353–400. doi: 10.1098/rstb.1932.0019

Sultan, S. E. (2000). Phenotypic plasticity for plant development, function and life history. *Trends Plant Sci.* 5, 537–542. doi: 10.1016/S1360-1385(00)01797-0

Turing, A. M. (1952). The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond. B* 237, 37–72. doi: 10.1098/rstb.1952.0012

Uyttewaal, M., Burian, A., Alim, K., Landrein, B., Borowska-Wykrt, D., Dedieu, A., et al. (2012). Mechanical stress acts via Katanin to amplify differences in growth rate between adjacent cells in Arabidopsis. *Cell* 149, 439–451. doi: 10.1016/j.cell.2012.02.048

Wabnik, K., Kleine-Vehn, J., Balla, J., Sauer, M., Naramoto, S., Reinöhl, V., et al. (2010). Emergence of tissue polarization from synergy of intracellular and extracellular auxin signaling. *Mol. Syst. Biol.* 6, 447. doi: 10.1038/msb.2010.103

Wernet, M. F., Mazzoni, E. O., Çelik, A., Duncan, D. M., Duncan, I., and Desplan, C. (2006). Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature* 440, 174–180. doi: 10.1038/nature04615

Wilkinson, D. J. (2011). *Stochastic Modelling for Systems Biology*. Boca Raton, FL: CRC Press.

Wolpert, L., Beddington, R., Brockes, J., Jessell, T., Lawrence, P., and Meyerowitz, E. (1998). *Principles of Development*. London: Current Biology Ltd.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 July 2014; accepted: 30 October 2014; published online: 14 November 2014.
Citation: Collaudin S and Mirabet V (2014) Models to reconcile plant science and stochasticity. Front. Plant Sci. ***5****:643. doi: 10.3389/fpls.2014.00643*
This article was submitted to Plant Evolution and Development, a section of the journal Frontiers in Plant Science.
Copyright © 2014 Collaudin and Mirabet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The establishment of the correct *AG* expression pattern during early flower development is a key mechanism to ensure not only proper organ identities in the third and fourth whorls, but also the repression of *WUS* at the appropriate time to generate flowers with the correct numbers of organs. Despite many articles that have addressed the issue, it is not fully clear how this pattern arises in such a robust manner even as the flower grows. One of the important questions that remains unanswered is precisely how the various regulatory input coordinate with each other and within the growing flower for AG to appear in a very specific domain and at a very specific developmental stage. In this manuscript, I try to address this question using a variety of approaches.

Earlier in this chapter, I discuss how a pluridisciplinary approach, combining different scienctific disciplines, can help gain a better understanding of any given system. In the past, attempts have been made to use boolean models to address the mechanistics of ABC patterning in general, as well as the existence of whorls within the flower. However, such models are not suitable for addressing the question of how the spatial and temporal dynamics of *AG* expression are put in place. The expression patterns of the main transcription factors that regulate *AG* expression are well known. The two main activators of *AG* expression, *LFY* and *WUS* are expressed at early stage 1, whereas protein of the main repressor AP2 is not visible prior to stage 3. Nevertheless, the literature has always maintained that AG itself only appears in the central dome of the flower at stage 3. To tackles this discrepancy, my first step was to use a reaction-diffusion approach to model *AG* regulation to determine whether the regulation by these transcription factors was sufficient to explain this *AG* dynamics, and then to determine whether, by proposing hypotheses to make the model work better, I could gain a better understanding of the mechanisms underlying this dynamic.

Then it is interesting to study precisely the dynamic of *AG* pattern of expression. Indeed we see previously that this pattern is described at the tissue level on fixed time point. Tools allow today to study the dynamic of expression at a cellular level in an *in vivo* tissue during time. Studying the profile of *AG* expression in different cells can give us a better understanding of the process of activation and regulation of *AG*.

Once we will have analyse precisely the dynamic of expression of *AG* at a cellular level, based on the hypotheses proposed by the reaction-diffusion model, we can performed biological experiments to test these hypotheses and prove the explanation provided about robustness of *AG* expression. We can summarized the questions we will ask in the activation

process:

- How the capacity of AG to form higher complexes can influence the process of auto-activation by itself?
- How AP2 can either create a threshold to delay AG activation and to restrain AG in the centre dome?
- How AG movement can help to have a robust and fast homogeneous pattern of expression?
- Do we observe stochasticity in AG activation, which stochasticity and what is its origin?

We will propose in this thesis answers to these questions and further hypotheses for future research work.

Chapter 2

A diffusion/reaction equations system to model dynamic of *AGAMOUS* expression

Contents

Abstract **46**

2.1 Introduction **46**

2.2 Mathematical model **48**

2.3 Biological justification of the model **50**

2.3.1 Choice of the domain 50

2.3.2 Model kinetics 50

2.4 Model reduction **52**

2.4.1 Spatial transformations 53

2.4.2 Reduction of the parameter space 54

2.5 Numerical simulations **54**

2.5.1 Model discretization 54

2.5.2 Visualization of the results 55

2.6 Results of the model analysis and simulations **56**

2.6.1 The model reproduces qualitatively the dynamics of the WT . . . 56

2.6.2 The pattern of *AP1* and *LFY* expressions depend mainly on AG . 58

2.6.3 Parameters of *AG* auto-activation play the key role in the dynamics of *AG* expression 59

2.6.4 Role of AP2 in the spatial and temporal restriction of AG 61

2.6.5 LFY and WUS are necessary for *AG* activation 62

2.6.6 Diffusion of AG is important to obtain a proper pattern 64

2.6.7 External peak of AG expression accelerates the AG activation . . . 64

2.6.8 Analysis of the key parameters in robustness of the model 65
2.7 Discussion, next steps 67
2.8 Author contribution 67
2.9 Supplementary data 68

Abstract

Flower architecture is determined during its development through the expression of transcription factors. In *Arabidopsis thaliana*, a plant model to study morphogenesis, expression of three class of homeotic genes occur early in flower development in specific patterns that determine the identity of the four floral organ types: sepals, petals, stamens and carpels. The existence and the stability of these four whorls was previously studied with boolean models. However, the spatial and dynamic expression patterns of these genes are difficult to understand. In this chapter, I have focused on the regulation of the C-class gene, *AGAMOUS* (*AG*), which is involved in conferring identity to stamen and carpels. In order to elucidate the mechanisms underlying the expression pattern of this gene, I create a diffusion-reaction model, which successfully reproduces the known AG pattern and suggests new hypotheses about certain underlying mechanisms. It also suggests that the delay between the appearance of the genes that are known to activate AG and its actual activation is due to the existence of a threshold defined by the nature of autoregulation of AG and by the levels of its main repressor, APETALA2 (AP2), involve either on the spatial and temporal regulation of AG. I also propose that the capacity of AG to move between cells is important to obtain a proper pattern of expression.

2.1 Introduction

Flowers are present in a great number of different plants in nature, including over 200,000 species of dicotyledonous plants. They are meristems with time limited organogenesis that produce 4 kinds of organs: sepal, petal, carpel and stamen situated on 4 different concentric circles called whorls (Figure 2.1 A & B). The most common model system to study flower development is *Arabidopsis thaliana*. In this plant, organ position is determined by expression of 3 classes of genes (A, B and C) that are expressed at stage 3 of flower development in 3 different whorls (Figure 2.1 C.). With superposition, we obtain 4 whorls that determine

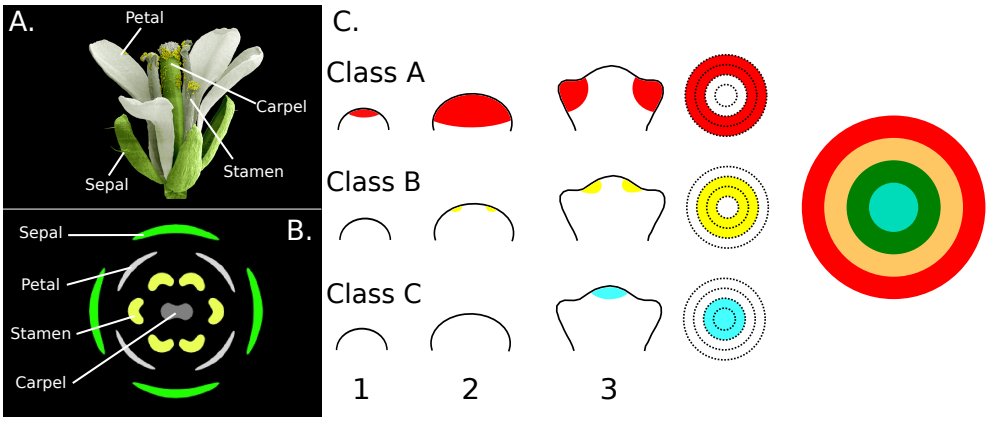


Figure 2.1: The ABC model: (A.) *Arabidopsis thaliana* flower that show the 4 kinds of organs: sepals, petals, carpels and stamens. **(B.)** Representation of the 4 whorls of flower organs. **(C.)** representation of the dynamics of expression of the A, B and C-class genes during the 3 first stages of flower development. The superposition of these 3 class of genes gives the 4 whorls of different gene expression that induce the organ identity

the 4 kind of organs. This system, discovered by Coen and Meyerowitz in 1991 [Coen and Meyerowitz, 1991] is called the ABC model.

In the 25 years since the ABC model was first proposed, the field has come to understand a lot about how the three classes of genes are regulated, and how they interact with each other, both genetically and molecularly [O'Maoileidigh et al., 2014, Mendoza and Alvarez-Buylla, 2000]. However, one aspect that is still poorly understood is exactly how these genes become expressed in very specific spatial domains and at very specific times.

To be able to understand the existence of the 4 whorls and their stability, Alvarez-Buylla [Alvarez-Buylla et al., 2010] proposed boolean models on the meristem and on the flower that give attractors for the 4 organs identity. Two spatial models have been created on the ABC model, the first [Barrio et al., 2010] that use a macroscopic field to obtain whorls formation. A second simple model proposed by Ming [Ming, 2009] study briefly the spatial problem.

To be able to more precisely understand the mechanisms underlying the dynamics of ABC gene expression patterns, I decided to focus on the only C-class gene, *AGAMOUS* (*AG*). AG is a MADS-box protein that is required to confer proper identity to carpels and stamens and is thus important in the reproduction of the plants [Yanofsky et al., 1990]. It has been shown that *AG* expression first begins to appear in the central dome of stage 3 flowers [Hong and Hamaguchi, 2003]. However, the key activators of *AG*, are present days

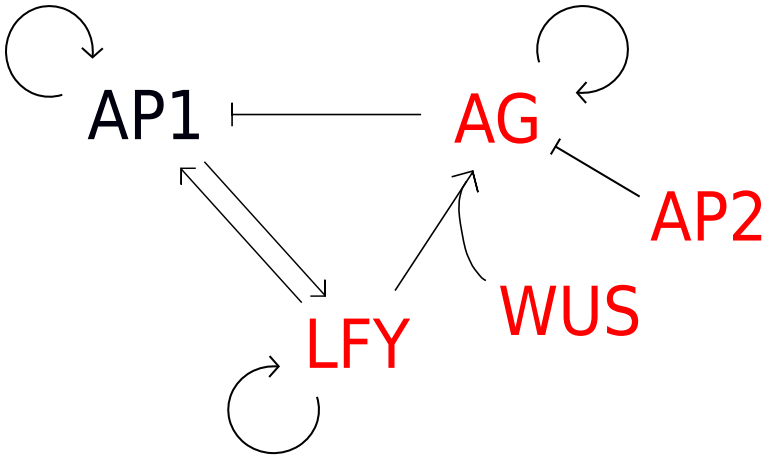


Figure 2.2: A simple model to understand *AG* dynamic of expression: The scheme summarize the known regulations of gene transcription (positive or negative) by direct or undirect binding of the protein to the DNA. The red proteins are known to be able to move between cells.

before the onset of *AG* expression. To understand how these regulators bring about the precise spatio-temporal regulation of *AG*, we developed a simple model based on reaction-diffusion equations that reflected the principal known interactions of these regulators, and combined this with a method that recapitulates floral growth during the relevant stages.

To build the model, we decide to focus on the A-class gene *APETALA1* (*AP1*) and *APETALA2* (*AP2*) that are known to be the antagonist genes of *AG* [Krizek and Fletcher, 2005], and on two activators that work together, *LEAFY* (*LFY*) and *WUSCHEL* (*WUS*) [Lohmann et al., 2001, Lenhard et al., 2001, Engelhorn et al., 2014] . This simple system (Figure 2.2) allows us to reproduce the various gene expression patterns observed in the literature, and to propose hypotheses to better understand the system and that can be tested with biological experiments. A mathematical analysis of the model suggested that the auto-activation of *AG* coupled with the repression by AP2 defines a threshold in *AG* expression. It also suggested that AG diffusion play an important role in its capacity to be highly expressed in the central dome just after activation.

2.2 Mathematical model

To describe the spatio-temporal dynamics of the C-class gene, *AG*, we have built a model based on partial differential equations that accounts for interactions between the main components of the ABC system, as well as cell-to-cell communication through protein movement. For simplicity, we have reduced the three dimensional geometry of the flower by considering

only the L1 cell layer, and by further treating it as a flat surface.

The model accounts only for the protein concentration. Simulations of an extended version of the model, involving mRNA concentrations as separate variables, predict the same dynamics, and hence we focus on a reduced model. A detailed biological justification of the model nonlinearities is provided in the next section.

The model variables, denoted by AP1, LFY, AG, WUS and AP2, are functions of time $t \in \mathbb{R}^+$ and space $x \in \overline{\Omega}(t) \subset \mathbb{R}^2$. We focus on the time interval $[0, T]$, with $t = 0$ corresponding to a young stage 1 of the flower and $t = T$ corresponding to a late stage 3 of the flower. The domain Ω is growing with time as explained further on. Moreover, due to the rotational symmetry of the flower, we may assume $\Omega(t) = B_{r(t)}(0)$ with a radius $r : [0, T] \to \mathbb{R}$ being a growing function of time with a given initial $r(0) > 0$.

The domain grows due to the cell division, which is assumed to be homogeneous in time and space. However, the domain of WUS production is fixed to a certain number of cells, which corresponds roughly to a 1/4 of the initial domain, i.e. $\Omega_W = B_{\frac{r_0}{4}}(0)$. On the other hand the ring of expression of AP2, which is controlled by an independent mechanism and hence prescribed in the model, is changing in time and corresponds to a half of the domain size, i.e. $\Omega_A(t) = \Omega(t) \setminus B_{\frac{r}{2}}(0)$.

The model consists of the following five reaction-diffusion equations

$$\partial_t AP1 = D_1 \Delta_x AP1 + a_1 \frac{AP1^{n_1}}{M_1^{n_1} + AP1^{n_1}} \cdot \frac{LFY^{n_2}}{M_2^{n_2} + LFY^{n_2}} \cdot \frac{M_3^{n_3}}{M_3^{n_3} + AG^{n_3}} - \delta_1 AP1 \tag{2.1}$$

$$\partial_t LFY = D_2 \Delta_x LFY + a_4 \frac{AP1^{n_4}}{M_4^{n_4} + AP1^{n_4}} + a_5 \frac{LFY^{n_5}}{M_5^{n_5} + LFY^{n_5}} - \delta_2 LFY \tag{2.2}$$

$$\partial_t AG = D_3 \Delta_x AG + \left(a_6 \frac{AG^{n_6}}{M_6^{n_6} + AG^{n_6}} + a_7 \frac{LFY \cdot WUS}{M_7 + LFY \cdot WUS} \right) \cdot \frac{M_8^{n_8}}{M_8^{n_8} + AP2^{n_8}} - \delta_3 AG \tag{2.3}$$

$$\partial_t WUS = D_4 \Delta_x WUS + a_9 \cdot \mathbb{1}_{\Omega_W \times [t_1, T]}(x, t) - \delta_4 WUS \tag{2.4}$$

$$\partial_t AP2 = D_5 \Delta_x AP2 + a_{10} \cdot \mathbb{1}_{\Omega_A(t) \times [t_2, T]}(x, t) - \delta_5 AP2, \tag{2.5}$$

supplemented with initial and boundary conditions. The latter should allow free flow of a protein on the boundary $\partial\Omega$, the tissue is not closed at the periphery of the domain and proteins can diffuse in the tissue. The domain is restricted to the part of the tissue where the AC system is active. Extending the domain to the whole structure would require taking into account additional signaling factors, a different geometry and different growth conditions. To streamline our study, we focus on the restricted domain. To cope with this restriction we assume the free flow condition on the boundary.

Cell-to-cell transport of the proteins is modeled by a diffusion operator with a constant coefficient D_i, $i = 1, ..., 5$, which is set to zero in case of the non-diffusing protein AP1.

2.3 Biological justification of the model

2.3.1 Choice of the domain

The first layer of cells of the flower, called L1, is structured in a way that cell division planes are always anticlinal and that growth is horizontal [Traas and Vernoux, 2002]. Although the components of the AC model are found in different cell layers, the experimental data are extracted at the surface of the flower. Further experiments show that the inside layers behave in a similar way that the first one. Cells in the plant are stick together due to the cell wall, so protein are able to diffuse inside the cell and to be passively transported directly from one cell to the other. To describe the pattern of WUS which is not produced in L1 but diffuses from a lower cell layer [Yadav et al., 2011, Daum et al., 2014], we project the source of WUS to the L1 domain. Furthermore, we model the L1 layer as a flat surface. Considering more realistic geometry is postponed to a further work using agent-based models.

2.3.2 Model kinetics

The interactions between the model components are based on the available information on the regulation of the transcription of respective genes. We use Hill functions to describe activation and inhibition processes and linear functions for a protein degradation.

AP1 is a MADS-Box protein and belongs to the A-class gene. AP1 is important to determine the identity of the sepals and petals [Coen and Meyerowitz, 1991, Gustafson-Brown et al., 1994]. It is expressed in the whole flower at the initial stage of development and disappears in the two inner whorls at the stage 3 [Urbanus et al., 2009]. It has been shown that AP1 does not move between cells so $D_1 = 0$ [Sessions et al., 1999, Urbanus et al., 2010]. No direct self-regulation of *AP1* is known, nevertheless, AP1 has an indirect effect by repressing its own repressors, as the bZIP protein FD or SQUAMOZA PROMORTER BINDING PROTEIN-LIKE 9 (SPL9) [Kaufmann et al., 2010]. To account for this effect, we use a Hill function with an exponent $n_1 = 1$. LFY activates *AP1* by direct and indirect regulation (via repression of *TERMINAL FLOWER1* (*TFL1*)) [Mandel and Yanofsky, 1995,

Liljegren et al., 1999]. LFY can bind directly the promoter of *AP1* [Moyroud et al., 2011] and is known to bind the DNA as a multimer [Sayou et al., 2014]. Hence, we model this interaction with a Hill function with $n_2 = 2$. Numerical simulations indicate that the choice of the n_2 has no significant influence on the model dynamics. Nevertheless, LFY is also able to activate *AP1* indirectly [Pastore et al., 2011]. To summarize these multiple pathways of *AP1* activation by LFY and AP1 itself, we decide to use one activation term that is the combination of the two Hill equations above. Furthermore, the experiments showing AP1 present everywhere in the *ag* loss of function mutants [Gustafson-Brown et al., 1994] suggest that *AP1* is being repressed by AG. To model this inhibition, we apply a Hill function under a hypothesis that the capacity to bind the DNA in complex is similar to the capacity to bind its own second intron, what yields $n_3 = n_6$. The choice of specific number of these exponents cannot be decided based on experimental evidence and hence we test numerically consequence of different assumptions, as shown in subsection 2.6.3.

LFY is a transcription factor defining the flower identity [Weigel and Nilsson, 1995]. It is present at the initiation of a new flower and stay strongly expressed during the first stages of the development [Blázquez et al., 2006]. LFY is able to move between cells [Sessions et al., 1999, Wu et al., 2003]. It is activated by AP1 through a direct binding [Liljegren et al., 1999, Kaufmann et al., 2010], what we model using a Hill function of $n_4 = 1$. The results of model simulations do not change for higher values of n_4. LFY is also known to self-enhance by binding its own regulatory sequence by forming multimers [Moyroud et al., 2011]. We model it with a Hill function with $n_5 = 2$. Higher values of n_5 lead to similar results. We observe than in *ap1-1* mutant, LFY is still present [Weigel et al., 1992] so we choose to add the two activation terms in this equation.

AG is a MAD-Box protein and the only member of the C-class genes [Yanofsky et al., 1990, Coen and Meyerowitz, 1991]. It is important to determine the identity of the carpels and stamens. The stage 3 of the flower development is linked to a rapid expression of *AG* in the two inner whorls [Deyholos and Sieburth, 2000, Gómez-Mena et al., 2005]. The protein is able to relocate between cells but only at a distance of few cell diameters [Urbanus et al., 2010]. AG self-enhance its own production by a direct binding of its second intron [Yanofsky et al., 1990, Gómez-Mena et al., 2005]. Since the exact the binding process is not known, we model it with a Hill nonlinearity with an arbitrary n_6, which we then investigate using numerical simulations. *AG* is also known to be activated by the combination of LFY and

WUS that can bind *AG* 2nd intron on different site. This activation is not necessary to maintain a proper AG expression but is important to initiate its activation [Busch et al., 1999, Lohmann et al., 2001, Lenhard et al., 2001]. Hence we choose to add these two terms for the activation part of this equation. The spatial restriction of *AG* is due to AP2 that can also bind the 2nd intron of *AG* [Drews et al., 1991, Yant et al., 2010, Krogan et al., 2012]. Here we choose to use an Hill equation to inhibit both *AG* auto-activation and activation by LFY and WUS.

WUS is transcription factor that is involved in the maintenance of the stem cells in the meristem and in the flower [Yadav et al., 2011, Daum et al., 2014]. It is regulated by a different pathway. It is known to be expressed in few cells in inner layers and to diffuse inside the L1 layer of cells. As mentioned before, restricting the model domain to the L1 layer, we project the expression domain of WUS to Ω_W subdomain at L1. It allows to reproduce the patterning observed in experiments and provides a model of diffusion of the protein in the whole domain.

AP2 is the second gene from the A-class [Kunst et al., 1989]. It represses *AG* strongly at the periphery of the flower. Regulation of *AP2* is maintained through additional signaling pathway which is independent on the AC-system and its production is mainly repressed by miR172 [Wollmann et al., 2010]. Hence, we reproduce this known pattern of expression of *AP2* and at the periphery starting at the beginning of the stage 2 of flower development. This allows us to model further influence of the prescribed AP2 pattern on the AC-system. Following experimental evidence of an AP2 gradient on the domain part which does not exhibit *AP2* production. Since the exact mechanism of this process is unknown, we model the gradient using a diffusion with the coefficient D_5.

We also add some hypothesis to reduce parameters variations. In a first time, we consider that diffusion terms, when they are different of 0 are equal to the same value D. We also consider that the degradation rate of the different proteins are equal to δ. These conditions are not applied to AP2 that is regulated by a microRNA miR172.

2.4 Model reduction

To cope with the growing domain we transform the model to a fixed domain by changing the variables (proof of this reduction is in the supplementary data).

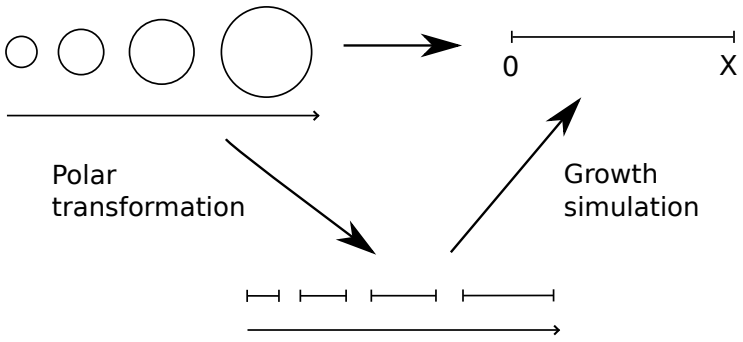


Figure 2.3: Reduction of the spatial domain

2.4.1 Spatial transformations

To be able to change variables, we can write each of our equation as:

$$\partial_t Y(x,t) = D.\Delta_x Y(x,t) + f(Y(x,t))$$

Because of the rotational symmetry of the flower, we reduce the model to a one dimensional domain $[0, \rho(t)]$ defined along a radius of $\Omega(t)$, see Figure 2.3. Polar transformation of the coordinates yields to the following equation:

$$\partial_t Y(\rho,t) = D(\Delta_\rho Y(\rho,t) + \frac{1}{\rho}\partial_\rho \cdot Y(\rho,t)) + f(Y(\rho,t))$$

The growth of the domain is modeled by the simplest linear growth model

$$\frac{d}{dt}\rho = c \cdot \rho$$

with a constant growth rate $c > 0$ and the initial condition $\rho(0) = r_0$. The coefficient c can be estimated based on experimental data showing a three-fold growth of the time during the considered time span $[0, T]$. Hence $c = \frac{ln3}{T}$ and $\rho = r \cdot e^{\frac{ln3 \cdot t}{T}}$.

Changing variables, we obtain a model on a fixed domain $[0, r_0]$ with the equation given by

$$\partial_t y(r,t) = De^{-2c \cdot t}(\Delta_r y(r,t) + \frac{1}{r}\partial_r \cdot y(r,t)) + cr \cdot \partial_r y(r,t) + f(y(r,t)) \tag{2.6}$$

2.4.2 Reduction of the parameter space

In the next step we simplify the model by rescaling the variables. Taking $r' = r/r_0$ and $t' = t/T$ and

$$\begin{aligned} AP1' &= \frac{AP1}{M_1} \\ LFY' &= \frac{LFY}{M_2} \\ AG' &= \frac{AG}{M_6} \\ WUS' &= \frac{WUS}{a_9 T} \\ AP2' &= \frac{AP2}{a_{10} T}, \end{aligned}$$

and re-defining the parameters $D_i = TD_i/r_0^2$, $\delta_i = T\delta_i$, for $i = 1, ..., 5$, $a_1 = a_1 T/M_1$, $M_3 = M_3/M_6$, $a_4 = a_4 T/M_2$, $M_4 = M_4/M_1$, $a_5 = a_5 T/M_2$, $M_5 = M_5/M_2$, $a_6 = a_6 T/M_6$, $a_7 = a_7 T/M_6$, $M_7 = M_7/M_2 T a_9$, $M_8 = M_8/T a_{10}$, $t_1 = t_1/T$, $t_2 = t_2/T$, $w_0 = w_0/r_0$ and $m_0 = m_0/r_0$, we obtain a new system (previous parameters are replaced by the new one without the apostrophe):

$$\begin{aligned} \partial_t AP1 &= D_1 \Delta_r AP1 + a_1 \frac{AP1^{n_1}}{1 + AP1^{n_1}} \cdot \frac{LFY^{n_2}}{1 + LFY^{n_2}} \cdot \frac{M_3^{n_3}}{M_3^{n_3} + AG^{n_3}} - \delta_1 AP1 \\ \partial_t LFY &= D_2 \Delta_r LFY + a_4 \frac{AP1^{n_4}}{M_4^{n_4} + AP1^{n_4}} + a_5 \frac{LFY^{n_5}}{M_5^{n_5} + LFY^{n_5}} - \delta_2 LFY \\ \partial_t AG &= D_3 \Delta_r AG + \left(a_6 \frac{AG^{n_6}}{1 + AG^{n_6}} + a_7 \frac{LFY \cdot WUS}{M_7 + LFY \cdot WUS} \right) \frac{M_8^{n_8}}{M_8^{n_8} + AP2^{n_8}} - \delta_3 AG \\ \partial_t WUS &= D_4 \Delta_r WUS + \mathbb{1}_{(0, w_0(t)) \times [t_1, 1]}(r, t) - \delta_4 WUS \\ \partial_t AP2 &= D_5 \Delta_r AP2 + \mathbb{1}_{(0.5, 1) \times [t_2, 1]}(r, t) - \delta_5 AP2 \end{aligned}$$

with $r \in (0, 1)$ and $t \in [0, 1]$ and $w_0(t) = \frac{1}{4} e^{-c \cdot t}$.

2.5 Numerical simulations

2.5.1 Model discretization

The model is solved numerically using a C script with CVODE library to simulate the resolution of ordinary differential equations. To use this library, we discretize the spatial operator by using Taylor expansion. The diffusion operation in polar coordinates is composed of a term of the second order derivative and two of the first order (Equation 2.6). We

discretize the coordinate r with a fixed time step τ. The first second order term is discretized by

$$\frac{\partial^2 Y}{\partial r^2} = \frac{1}{\tau^2}(Y(r+\tau) + Y(r-\tau) - 2Y(r)) + o(\tau^2) \tag{2.7}$$

For the first term of the first order part we expand the two terms $Y(r+\tau)/(r+\tau)$ and $Y(r-\tau)/(r-\tau)$ to obtain by additions:

$$\frac{1}{r}\frac{\partial Y}{\partial r} = \frac{1}{2\tau}\left(\frac{Y(r+\tau)}{r+\tau} - \frac{Y(r-\tau)}{r-\tau} + \frac{1}{r^2}Y(r)\right) + o(\tau) \tag{2.8}$$

The second one is obtained by the sum of $Y(r+\tau)$ and $Y(r-\tau)$:

$$r \cdot \frac{\partial Y}{\partial r} = \frac{r}{2\tau}\left(Y(r+\tau) - (r-\tau)\right) + o(\tau) \tag{2.9}$$

Hence, with combination of (2.7), (2.8) and (2.9) we obtain:

$$\begin{aligned}\Delta Y(r) \approx \frac{1}{\tau}\Bigg[&\left(e^{-2c\cdot t}\left(\frac{1}{\tau} + \frac{1}{2(r+\tau)}\right) + \frac{cr}{2D}\right)\cdot Y(r+\tau) \\ &+ \left(e^{-2c\cdot t}\left(\frac{1}{\tau} - \frac{1}{2(r-\tau)}\right) - \frac{cr}{2D}\right)\cdot Y(r-\tau) + e^{-2c\cdot t}\left(\frac{-2}{\tau} + \frac{1}{2r^2}\right)\cdot Y(r)\Bigg]\end{aligned} \tag{2.10}$$

About boundary conditions, we choose for $r = 0$ to have a 0-flux border condition, then $Y(0-\tau) = Y(0+\tau)$ so:

$$\Delta Y(0) \approx \frac{1}{\tau}\left[\frac{2e^{-2c\cdot t}}{\tau}\cdot Y(\tau) - \frac{2e^{-2c\cdot t}}{\tau}\cdot Y(0)\right] \tag{2.11}$$

At the peripheral boundary ($r = 1$), we use a periodic condition, so $Y(1+\tau) - Y(1) = Y(1) - Y(1-\tau)$. Then we obtain:

$$\Delta Y(1) \approx 0 \tag{2.12}$$

2.5.2 Visualization of the results

Following the one-dimensional model reduction, the solutions represent a distribution of protein concentration along the radius of the L1 disc at different time points. Simulations are visualized by a heat map of the concentration of proteins in a spatio-temporal coordinates of the space variable shown in the abscissa and the time variable in the ordinate. They present

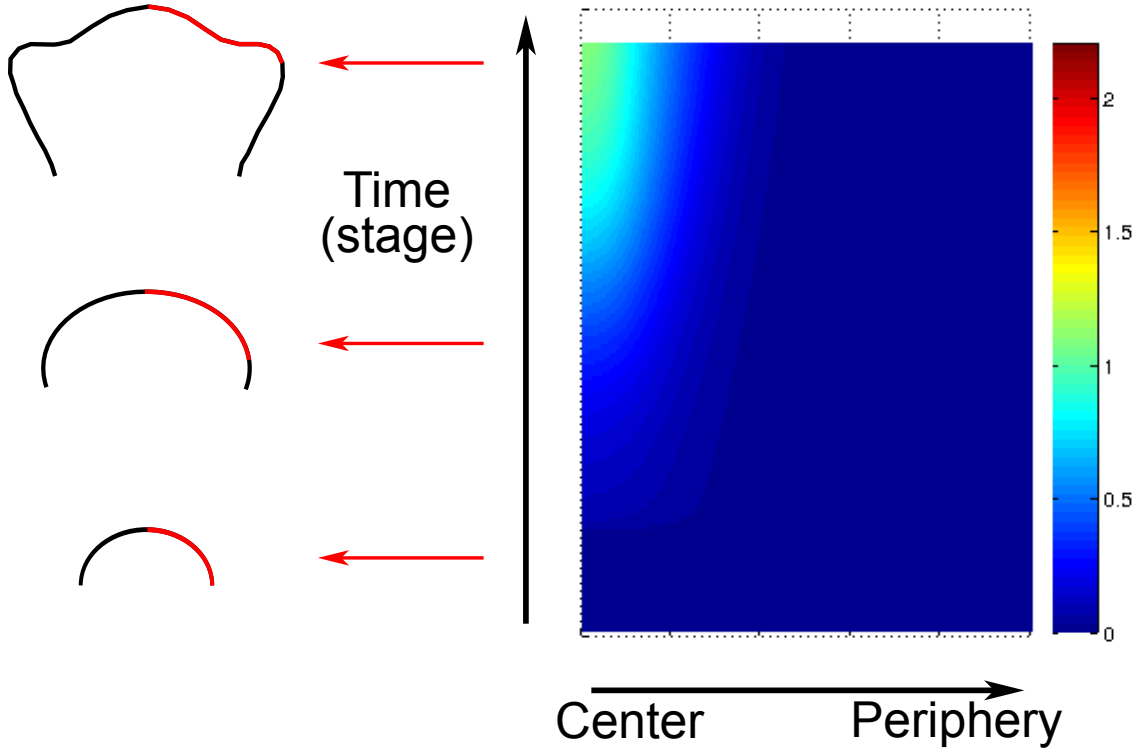


Figure 2.4: Visualization of the results: Simulations are represents by an heat map that correspond to protein concentrations in function of the time and the space variables.

solutions of the model in transformed coordinates (fix frame) and hence with a fixed radius of the domain. Simulations represent the evolution of the system between the early stage 1and a middle stage 3 (Figure 2.4). The heat map is the same for all concentrations of all simulations.

2.6 Results of the model analysis and simulations

2.6.1 The model reproduces qualitatively the dynamics of the WT

Values of the parameters have been chosen based on the the available experimental knowledge about the interactions, and on mathematical analysis and simulations to obtain results that qualitatively reproduce observations in literature and of recent experiments (Figure 2.5) [Drews et al., 1991, Gómez-Mena et al., 2005, Urbanus et al., 2009]. We set the following basic set of parameters:

$d_1 = 0$, $a_1 = 1.6$, $n_1 = 1$, $n_2 = 2$, $M_3 = 3$, $n_3 = n_6$, $\delta_1 = 0.7$, $d_2 = 0.01$, $a_4 = 0.2$, $n_4 = 1$, $a_5 = 1.6$, $n_5 = 2$, $\delta_2 = 0.7$, $d_3 = 0.01$, $a_6 = 0.72$, $n_6 = 2$, $a_7 = 0.05$, $M_7 = 1$, $M_8 = 0.15$, $n_8 = 1$, $\delta_3 = 0.2$, $d_4 = 0.01$, $w_0 = 0.25$, $t_1 = 0.05$, $\delta_4 = 0.7$, $d_5 = 0.05$, $m_0 = 0.6$, $t_2 = 0.3$ and

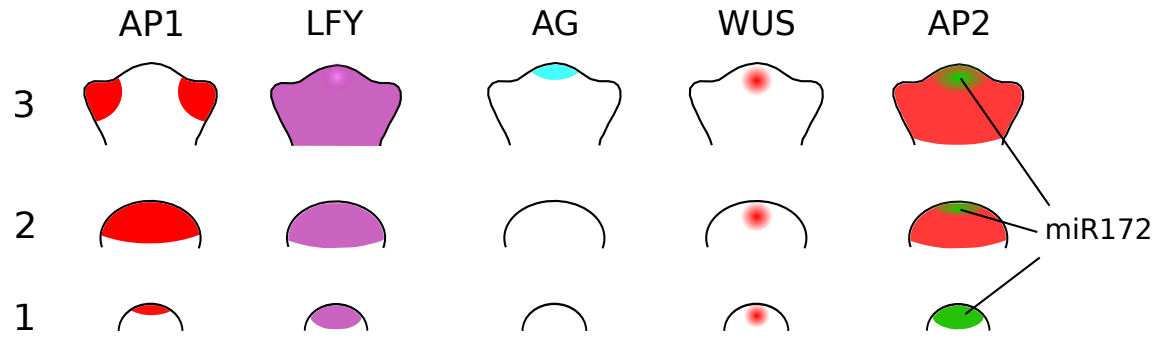


Figure 2.5: Expression pattern of the AC model based on literature and recent experiments

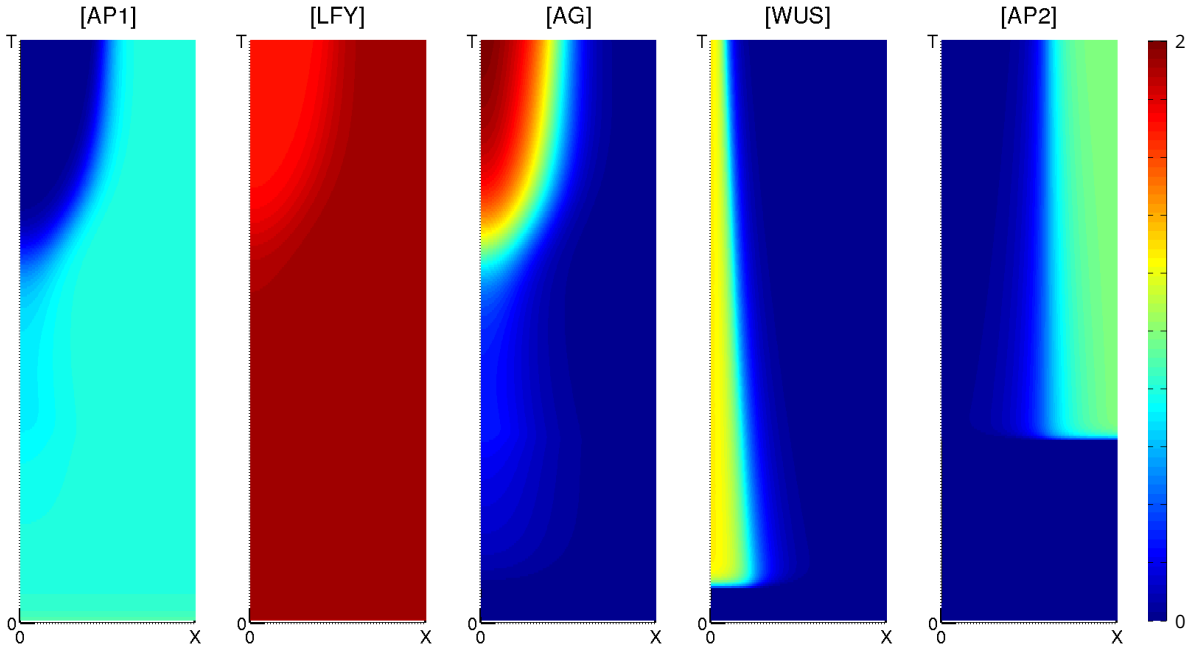


Figure 2.6: Simulation in the WT condition: The graph represents the level of expression of the 5 components of the AC model

$\delta_5 = 1$

The model simulation reproduce qualitatively the dynamics observed experimentally:

- The expression of AG is delayed comparing to the presence of its activators.
- AG appears rapidly in the central domain of the flower (Figure 2.6).
- A very low level of AG expression in the initial phase of the process can be observed in simulations. This low expression will be discussed in next chapters.

Based on these observations, we can define 3 phases in the dynamics of AG expression: the first when AG is expressed at a very low level in the center, the second when there is a transition from low to high expression level of AG due to auto-activation, and the third

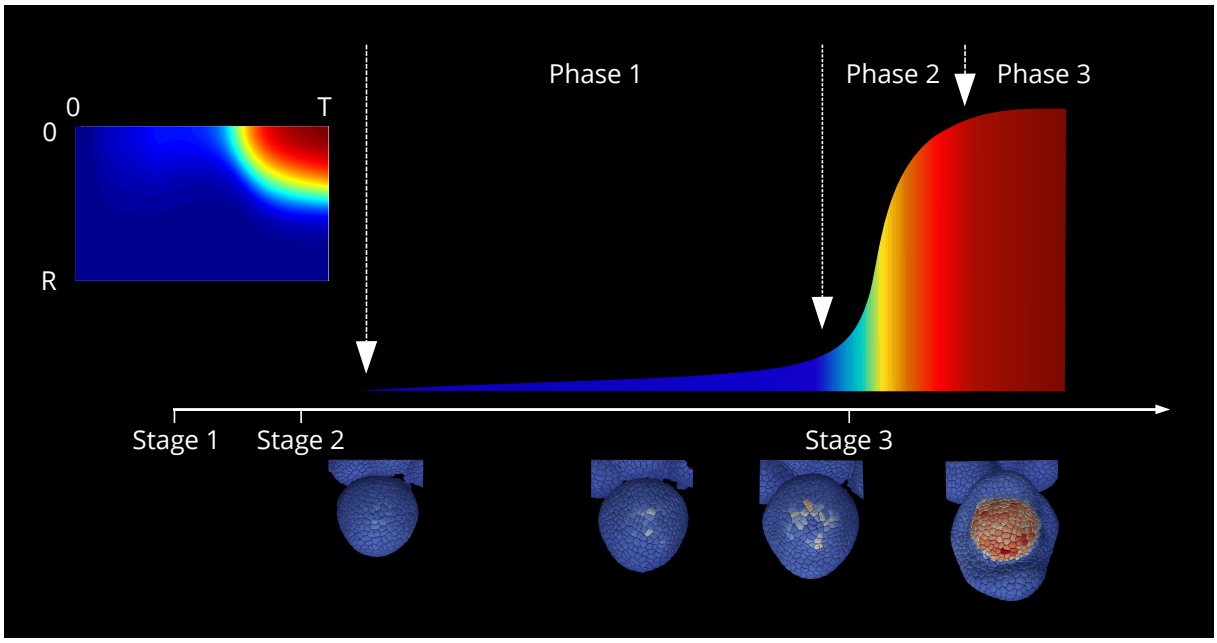


Figure 2.7: Description of the phases of *AG* activation: The scheme represent the 3 phases of *AG* activation. *AG* expression level is represented during the time of the 3 first stages of flower development and is also represented with the same color ode than in our simulations. Correspondence with experiments are present below the scheme.

phase when the final pattern is established and remains stable (Figure 2.7).

2.6.2 The pattern of *AP1* and *LFY* expressions depend mainly on AG

AP1 and *LFY* expressions are determined by the first two equations which depend directly also on AG concentration. To have a better understanding of the system dynamics, we analyze the structure and stability of spatially homogenous stationary solutions. In other words, we look at the steady states in the gene regulation system on the level of a single cell. Thus, we set diffusion to zero and solve the following stationary problem corresponding to the first two equations depending on AG:

$$\begin{aligned} 0 &= a_1 \frac{AP1}{1+AP1} \cdot \frac{LFY^2}{1+LFY^2} \cdot \frac{M_3^2}{M_3^2+AG^2} - \delta_1 AP1 = f_1(AP1, LFY, AG) \\ 0 &= a_4 \frac{AP1}{M_4+AP1} + a_5 \frac{LFY^2}{M_5^2+LFY^2} - \delta_2 LFY = f_2(AP1, LFY) \end{aligned} \tag{2.13}$$

We analyze the system depending on the AG value. Using numerical computation approach of linearized stability, we check that for small values of $AG \geq 0$ the system has two

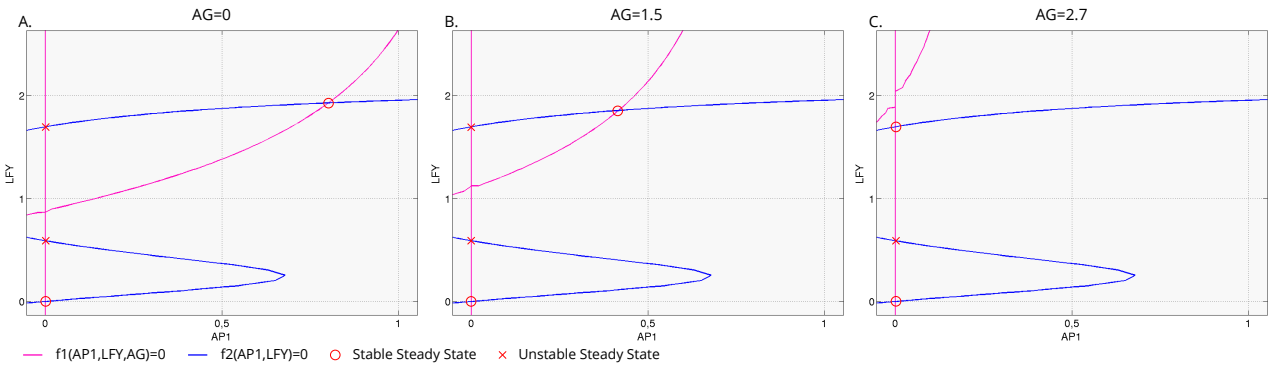


Figure 2.8: Nullclines and steady states of the AP1/LFY system: the three graphs represent the nullclines of the simplified system 2.13 with different values of AG. The stability of the steady states has been calculated numerically with pplane8 module of matlab

stable and two unstable steady states (see Figure 2.13). The dynamics depends on the initial conditions an the trivial equilibrium $(0, 0)$ cannot be reached for the realistic values of model initial data. Indeed, the initiation of a new flower is correlated with a high level of LFY and AP1 [Gustafson-Brown et al., 1994, Parcy et al., 1998]. However, for increasing AG values, the AP1 value of the positive stationary solution decreases and when it crosses $AP1 = 0$ a change of stability appear and the semi-trivial equilibrium with $AP1 = 0$ becomes stable. These two cases correspond of the known pattern of expression in the two inner whorls (high level of AG and no AP1) and in the two outer whorls (no AG and high level of AP1) during the third phase of development [Krizek and Fletcher, 2005].

2.6.3 Parameters of *AG* auto-activation play the key role in the dynamics of *AG* expression

AG is regulated by three factors, its own auto-regulation, the activation by the combination of LFY and WUS and the repression by AP2. We can first focus on the auto-activation in the case of no AP2.

First we decide to focus on the auto-activation part of the third equation:

$$\frac{d}{dt}AG = a_6 \frac{AG^{n_6}}{1 + AG^{n_6}} - \delta_3 AG \tag{2.14}$$

The dynamics of this equation depends significantly on the parameter n_6, which reflects the way how AG binds to its own promoter. For $n_6 = 1$ there exists only one positive stable

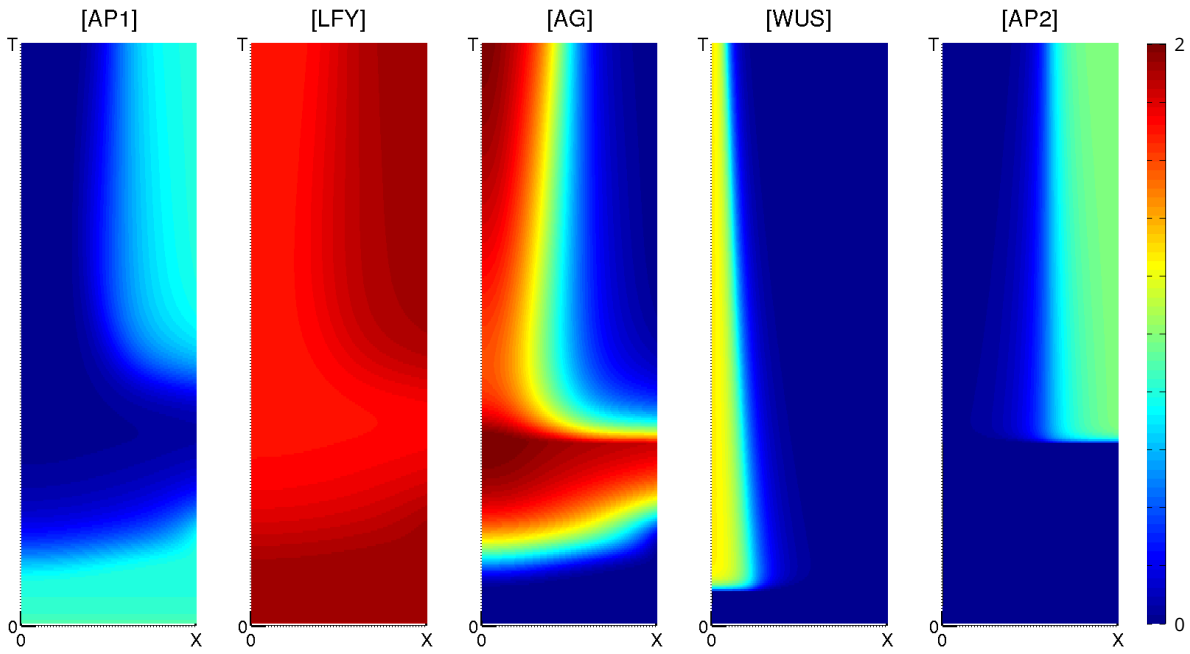


Figure 2.9: Simulation of the auto-activation with a single protein: we use $n_6 = 1$ and we rescale $a_6 = 0.78$ to obtain a similar final pattern of expression. We see that *AG* activation appears earlier than expected and in the full domain before *AP2* expression.

steady state ($\frac{a_6-\delta_3}{\delta_3}$) if and only if $a_6 > \delta_3$ and one unstable steady state equal to 0. It yields a convergence of the solution to this unique positive equilibrium. Consequently, a model describes emergence of a wide pattern of AG expression, which takes place relatively fast, i.e. before the appearance of AP2 (Figure 2.9). The final pattern of expression is similar to the observed WT but we are not able to reproduce the dynamics of the AG expression.

Parameter $n_6 = 2$ corresponds to a case of an auto-activation by a dimer of AG (that can also be part of a bigger complex). It leads to a bistable dynamics, i.e. existence of one unstable and two stable steady states: 0and $\frac{a_6+\sqrt{a_6^2-4\delta_3^2}}{2\delta_3}$; the latter requires an additional condition $a_6 > 2\delta_3$. The unstable steady state defines here a separatrix between basins of attraction of the two stable states. In other words, there exists a threshold that needs to be reached to allow a sharp transition of AG to a high value. The simulations of the model indicate that this is achieved in the course of the system evolution due to LFY and WUS activation. This bistability explains a sharp transition from very low to high AG expression in the central dome.

2.6.4 Role of AP2 in the spatial and temporal restriction of AG

AP2 is the main known repressor of AG during the first stages of flower development [Drews et al., 1991]. The accounts for a inhibition of AG production by AP2. Since *AP2* expression is observed only at the periphery of the domain from the beginning of stage 2, what is an effect of miR172 regulation independent of the AC-system, we prescribe the gradient-like pattern of AP2 with a maximum value for $r = 1$ and a minimum in $r = 0$.

Accounting for the AP2 inhibition in the equation for AG,

$$\frac{d}{dt}AG = a_6 \frac{AG^{n_6}}{1 + AG^{n_6}} \cdot \frac{M_8^{n_8}}{M_8^{n_8} + AP2^{n_8}} - \delta_3 AG, \tag{2.15}$$

yields a dependence of the bistable dynamics explored previously on the value of AP2.

Indeed, the positive stable steady state exists only if:

$$AP2 < \frac{M_8.(a_6 - 2\delta_3)}{2\delta_3} = 0.0937 \tag{2.16}$$

If AP2 is higher or equal to this value, then 0 is the only stable steady state. This way AP2 restricts the expression of AG in the two inner whorls of the flower. The value of AP2 influences also the threshold of transition from the low to the high activation level. This one is defined by:

$$AG_{threshold} = \frac{a_6 \cdot \frac{M_8}{M_8 + AP2} - \sqrt{a_6^2 \cdot \frac{M_8^2}{(M_8 + AP2)^2} - 4\delta_3^2}}{2\delta_3} \tag{2.17}$$

To compare this to experiments, we simulate the case of *ap2* mutant (Figure 2.10), i.e. a system with $AP2 = 0$ in the whole domain. We observe the second phase of activation of AG starting earlier than in the WT case. Nevertheless, there is still a delay between appearance of the activators (LFY and WUS) and the increase of AG expression. This pattern of expression correspond to results recently obtained by *ap2* knockout experiments showing the initial phase of AG low expression similar to the WT, i.e. with a delay comparing to LFY and WUS activation, and the switch to a high level expression (auto-activation) taking place much faster in the mutant than in the WT system. The width of the pattern changes as well. It confirms that it is AP2 that restricts AG in the center of the flower and causes a delay in the expression.

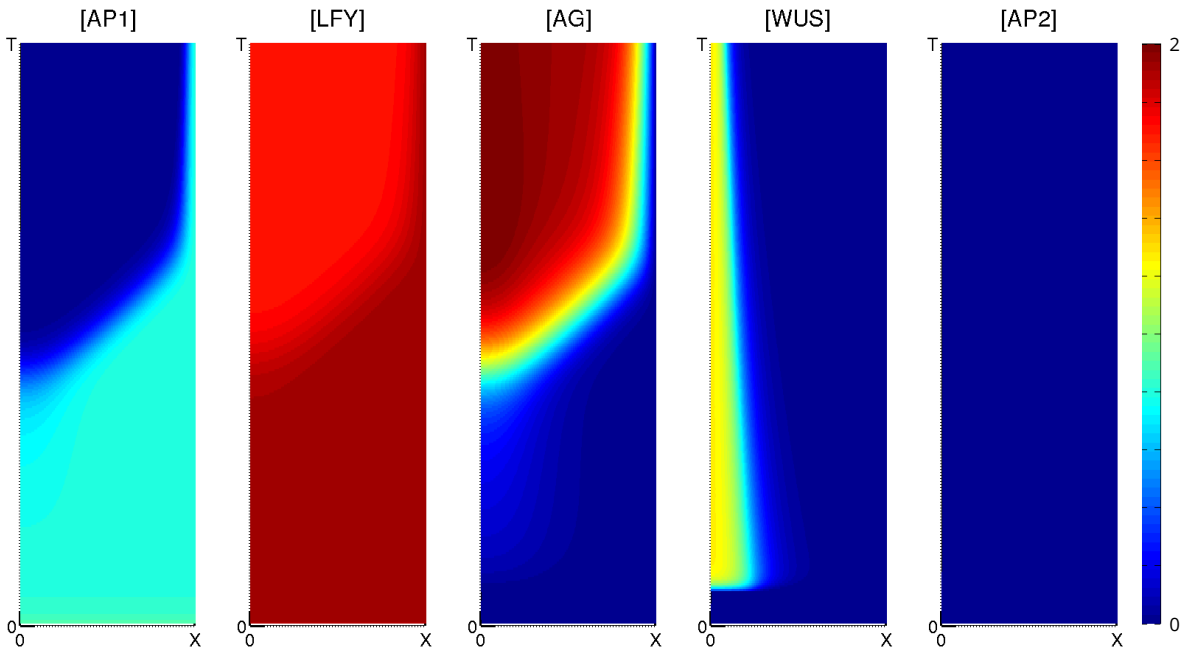


Figure 2.10: Simulation of the *ap2* mutant: The high expression of *AG* start earlier and is wider than in the WT

2.6.5 LFY and WUS are necessary for *AG* activation

Since AG is absent in the flower initially, it needs an activation by an external factor. Following the experimental evidence, we account for the activation by LFY and WUS. Setting $LFY = 0$ in the model clearly leads to no activation of AG. However, experiments on *lfy* mutant provide an evidence of a late expression of AG (around stage 4). This shows that there exist additional activation mechanisms which are not well understood and not taken into consideration in our model. An interesting question is if the unknown activation mechanism is present also in the WT flower or emerges first when the LFY activation is switched off by the knockout.

Nevertheless, the model can be compared to experiments of a partial knock-out of LFY, based on the binding sites of AG second intron [Deyholos and Sieburth, 2000]. The experiments indicate (see own ongoing experiments) a delay in the appearance of low level of AG during the first phase and also a delay in the high activation of AG for the second phase. The pattern of AG appears to be very sensitive to the LFY activation parameter. For instance, reducing a_7 of 4% from 0.5 to 0.48 leads to a significant delay in the appearance of AG, see Figure 2.11, while increasing it to 0.6 leads to a fast switch to high values of AG expression.

The model suggests that WUS has similar effects on the AG dynamics as LFY. This

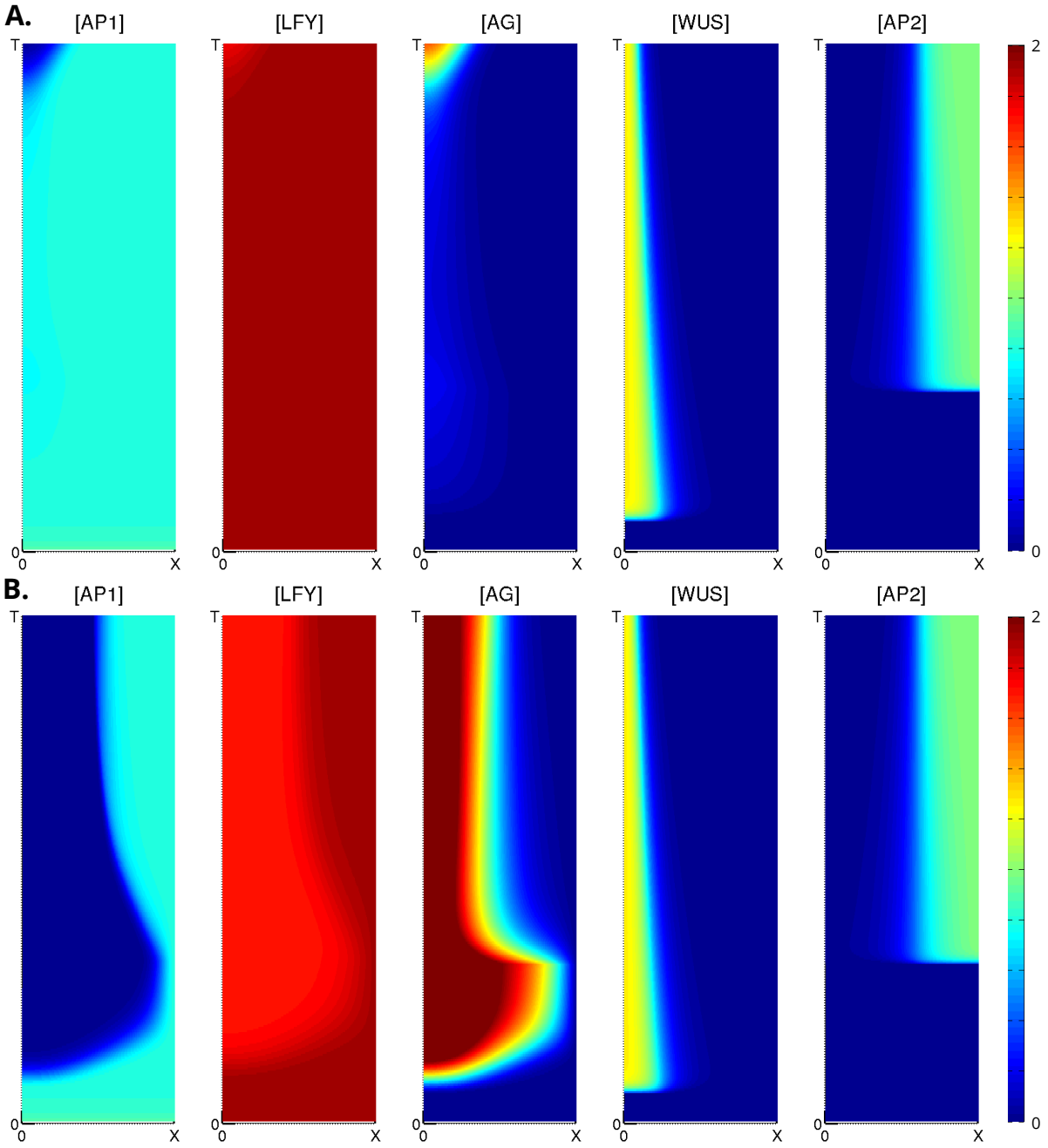


Figure 2.11: Simulation of the LFY/WUS weak interaction: (A.) Simulation with $a_7 = 0.48$ instead of 0.5 that reproduce a defect on the binding between LFY, WUS and the 2nd intron of *AG*. We see that small variation in this parameters induce strong changes in initiation of the pattern. **(B.)** Simulation with $a_7 = 0.52$ instead of 0.5. Activation of AG is earlier than in the WT. The effect of this parameter is strong on the time of AG activation.

cannot be tested experimentally due to the lack of *wus* mutant flowers, since this mutant does produce almost any flowers or flower-like organs.

2.6.6 Diffusion of AG is important to obtain a proper pattern

AG is known to be able to relocate between cells [Urbanus et al., 2010]. Our model points out the importance of this cell-to-cell communication, which we model with a diffusion operator. It has a significant influence on emergence and shape of the pattern. No diffusion of AG (Figure 2.12) results in an early activation of AG followed by a reduced expression domain. The first effect is linked to a faster accumulation of AG on a cellular level due to the lack of dispersion, which leads to a faster transition to the high expression phase. It can be compensated by changing a_7 to obtain the right time scale of the second phase.

The second result shows the importance of the protein relocation for the proper size of the AG expression pattern. A gradient of AG expression at the border of the central domain has been confirmed by the translational AG reporter [Urbanus et al., 2009]. In case of no spatial communication, the stochastic fluctuations of the initial low gene expression would would result in a large heterogeneity of the dynamics and different times of the transition to the phase 2. Relocation of the protein allows to synchronize this transition.

2.6.7 External peak of AG expression accelerates the AG activation

It has been shown previously that the LFY and WUS-induced activation of AG plays an essential role in our model in the transition from a low to a high expression level. Interestingly, spatial fluctuations in the AG expression patterns are observed experimentally at the initial stage of the development (see next chapters). To explore how much a single spike of the AG activation may influence the overall system dynamics, we perform simulations with initial data involving peaks of AG. For this, we introduce in the third equation a constant term ($p = 1$) over a short time ($0.01T$) on a space of the size corresponding to one cell ($0.7R$).

A peak of AG activation during the middle of the first phase, when AP2 is already present (at $t = 0.4T$), at a distance of 3 cells from the center of the flower disc (see Figure 2.13) yields early transition to the phase 2. Nevertheless, AG moves to the adjacent cells reducing

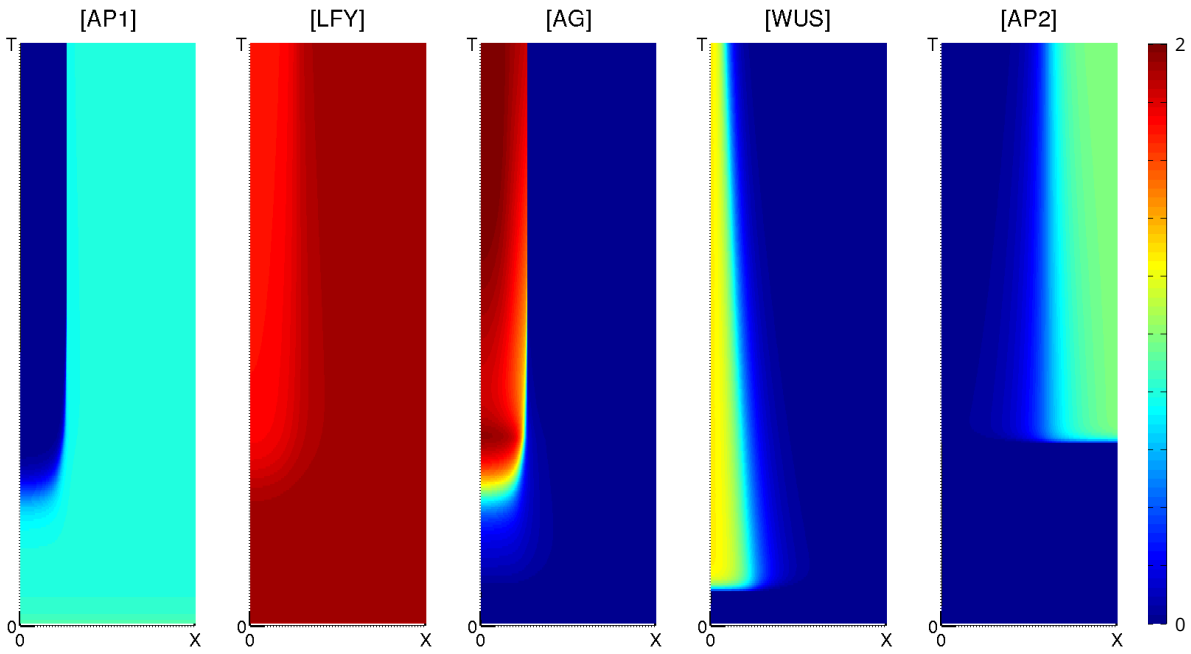


Figure 2.12: Simulation when AG is not able to diffuse: Simulation with $D_3 = 0$ reproducing the incapacity of AG to move from one cell to the other. We can observe that phase 2 start earlier, that the final domain of expression is less wide and that there is a sharp border of the AG domain of expression.

initially AG concentration in the activated cell but still leading to a faster phase 2 transition in the center of the flower due to a greater protein accumulation.

Secondly, if a high peak is introduced ($p = 5$) in the model with no LFY- or WUS-induced activation of AG ($a_7 = 0$) (Figure 2.14), we observe no deviation from a normal pattern of AG expression. This replaced activation by an AG peak needs to be sufficiently large. Otherwise, e.g. in case of $p = 1$ no patterns are produced.

This experiment shows that a peak will first diffuse in the adjacent cells and if the peak is high enough, will initiate the phase 2 of AG expression.

2.6.8 Analysis of the key parameters in robustness of the model

To obtain insights in the model robustness we perform a systematic simulation study with varying parameters.

The pattern of AP1 is robust to the changes of the AP1 dynamics parameters. On the contrary, LFY can have a stronger effect on the patterning, especially parameters a_5 and δ_2 that influence directly the level of expression of LFY and in consequence influence

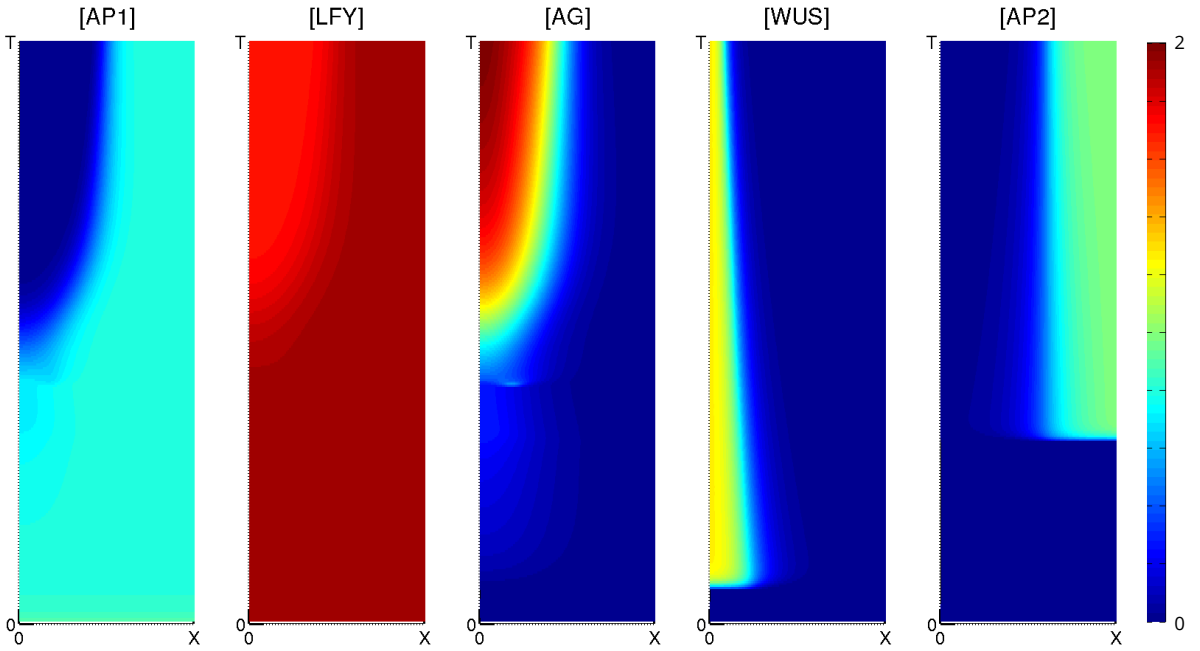


Figure 2.13: Simulation with a peak of AG after AP2 activation: Simulation of a short peak of *AG* expression starting at $t = 0.04T$ in a cell closed to the center. We observe that first the protein diffuses around this cell and then the phase 2 of AG activation starts earlier in the center

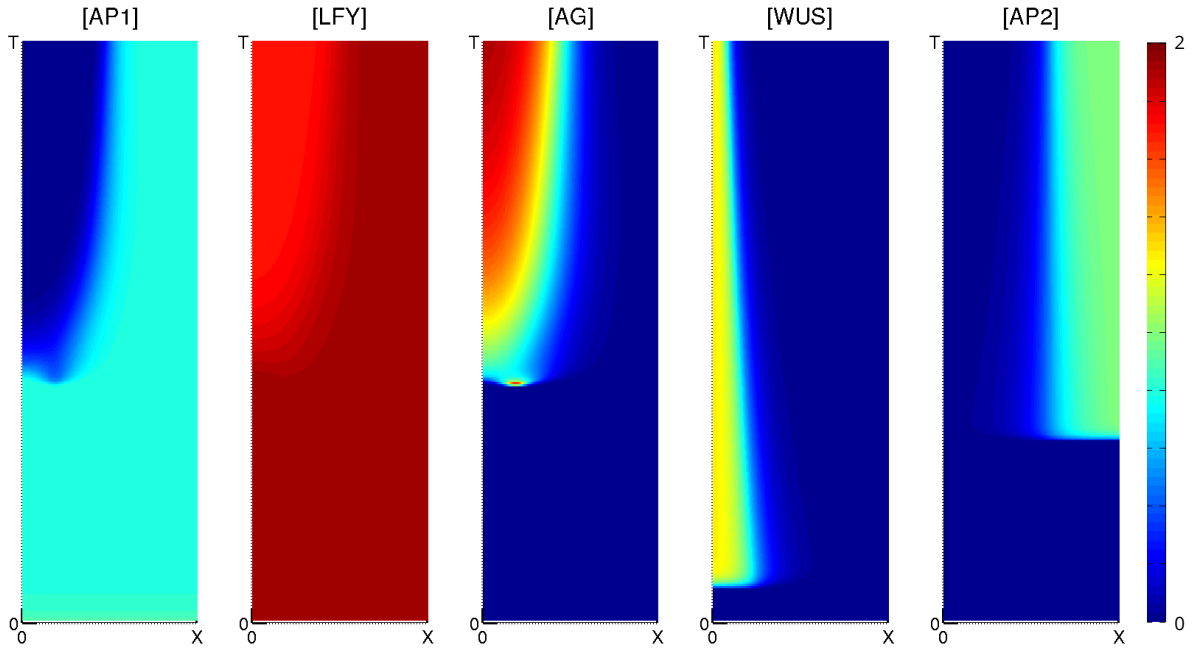


Figure 2.14: Simulation with a peak of AG when LFY and WUS are not able to activate AG: Simulation of a short peak ($p = 5$) of AG expression starting at $t = 0.04T$ in a cell closed to the center in the case where $a_7 = 0$. We observe that in the absence of the activation of AG by LFY and WUS, we can still restore the proper expression of AG with an artificial peak.

the activation of AG by LFY, may strongly change the transition to the phase 2 of AG expression. We note that values of n_2 and n_5 that correspond to a multimer form of LFY binding DNA do not influence the patterning. The only exponent of the activation process which significantly influence the model outcome is the parameter n_6.

For the third equation, we have already observed a singularity of the model dynamics in the diffusion coefficient equal to 0. We can also observe that the parameters of activation and of repression, especially a_6 and a_7, have a strong effect on the system dynamics, mainly on the beginning of the phase 2. Nevertheless, this differences can be compensated by choosing a different value of the degradation parameter δ_3.

Concerning WUS, variation of the parameters can influence the level of WUS expression and the size of its support. As for LFY, the main effect is then the activation of AG and the beginning of the phase 2. For AP2, we observe only a small effect on the size of the final AG expression domain.

To conclude, the model provides a robust mechanism of pattern formation. However, the time of transition to the phase 2 is very sensitive to changes in parameter values. Anyway, a large parameter space leads to a similar dynamics of pattern formation, which is consistent with experimental observations.

2.7 Discussion, next steps

Our model explain well the existence of the 2 whorls, with the possibility of activation of AG by WUS and the role of AP2 in relation with growth to control AG expression. We also are able to explain the presence of AP1 in the center, but we need to find an additional reaction or factor to explain the spacial dynamic.

In order to improve the results, we should use clustering to test many different set of parameters to find which one fit well with experiments. We also should implement the model on virtual tissues.

2.8 Author contribution

I created the model and did the analysis. I performed simulations using code developed by Anna Marciniack-Czochra. Anna Marciniak-Czochra advise me in the development of

the model and in the writing of this chapter. I beneficed of Pradeep Das corrections and remarks. Filip Klawe help me with the proof of the coordinates transformation.

2.9 Supplementary data

Proof of the spacial transformation

I want to detail here the proof of the spatial transformation from Eulerian to Lagrangian coordinates of the resume equation

$$\partial_t Y(x,t) = D.\Delta_x Y(x,t) + f(Y(x,t))$$

I proposed a first reduction due to the rotational symmetry of the flower to obtain equations define on a one dimensional domain $[0, P(t)]$ that is homogeneously growing. Polar transformation of the coordinates yields to the following equation:

$$\partial_t Y(\rho,t) = D(\partial_{\rho^2} Y(\rho,t) + \frac{1}{\rho} \cdot \partial_\rho Y(\rho,t)) + f(Y(\rho,t))$$

The growth of the domain is modelled by the simplest linear growth model

$$\frac{d}{dt}\rho(t) = c \cdot \rho(t)$$

with a constant growth rate $c > 0$ and the initial condition $\rho(0) = r_0$. We obtain then:

$$\rho(t) = r_0 \cdot e^{c \cdot t}$$

We apply the change of variable from Eulerian to Lagrangian coordinates. Let us denote this change with the function $\Phi : [0, r_0] \times [0, T] \to [0, \rho(t)] \times [0, T]$ such as

$$\Phi(r,t) = (\rho(r,t), t)$$

and y, the new variable of our system defined on $[0, r_0] \times [0, T]$ such as:

$$y(r,t) = Y(\rho(r,t), t) = Y \circ \Phi(r,t)$$

Then our equation become:

$$\partial_t Y \circ \Phi = D(\partial_{\rho^2} Y \circ \Phi + \frac{1}{\rho} \cdot \partial_\rho Y \circ \Phi) + f(Y \circ \Phi)$$

Applying derivative function to y, we obtain

$$\begin{aligned}
\partial_r y = \partial_r (Y \circ \Phi) &= \partial_r \rho \cdot \partial_\rho Y \circ \Phi \\
&= e^{c \cdot t} \cdot \partial_\rho Y \circ \Phi \\
\partial_{r^2} y = \partial_{r^2} (Y \circ \Phi) &= e^{2c \cdot t} \cdot \partial_{\rho^2} Y \circ \Phi \\
\partial_t y = \partial_t (Y \circ \Phi) &= \partial_t \rho \cdot \partial_\rho Y \circ \Phi + \partial_t Y \circ \Phi \\
&= cr \cdot e^{c \cdot t} \cdot \partial_\rho Y \circ \Phi + \partial_t Y \circ \Phi \\
&= cr \cdot e^{c \cdot t} \cdot \partial_\rho Y \circ \Phi + \partial_t Y \circ \Phi
\end{aligned}$$

So, with combination of these equations, we obtain

$$\begin{aligned}
\partial_\rho Y \circ \Phi &= e^{-c \cdot t} \cdot \partial_r y \\
\partial_{\rho^2} Y \circ \Phi &= e^{-2c \cdot t} \cdot \partial_{r^2} y \\
\partial_t Y \circ \Phi &= \partial_t y - cr \cdot e^{c \cdot t} \cdot \partial_\rho Y \circ \Phi \\
&= \partial_t y - cr \cdot e^{c \cdot t} \cdot e^{-c \cdot t} \cdot \partial_r y \\
&= \partial_t y - cr \cdot \partial_r y
\end{aligned}$$

We conclude that

$$\begin{aligned}
\partial_t y - cr \cdot \partial_r y &= D(e^{-2c \cdot t} \cdot \partial_{r^2} y + \frac{1}{r_0 \cdot e^{c \cdot t}} \cdot e^{-c \cdot t} \cdot \partial_r y) + f(y) \\
\partial_t y &= D \cdot e^{-2c \cdot t} (\partial_{r^2} y + \frac{1}{r} \cdot \partial_r y) + cr \cdot \partial_r y + f(y)
\end{aligned}$$

Chapter 3

Analysis of the stochastic activation of *AGAMOUS* expression during flower development

Contents

Abstract **71**

3.1 Introduction **72**

3.2 Results **73**

3.2.1 Stochastic *AG* expression is observed at stage 2 73

3.2.2 Peripheral and central cells in the *AG* expression domain present a lower concentration of AG proteins 74

3.2.3 The activation of *AG* expression is stochastic 78

3.2.4 *AG* expression spreads from one cell to its adjacent cells 79

3.2.5 *AG* activation profiles correspond to an auto-activation of AG itself as an homodimer 79

3.3 Discussion and perspectives **81**

3.4 Methods **85**

3.5 Author contributions **87**

3.6 Acknowledgement **87**

3.7 Supplementary data **88**

Abstract

AGAMOUS (*AG*), a key regulator of reproductive organ development, has been extensively described as appearing very rapidly at stage 3 of flower development, when sepals first begin

to emerge. It is expressed in the central dome of the flower during a few subsequent stages and is also responsible for terminating floral stem cell proliferation and making the flower a determinate structure. However, it is not clear how such a precise spatio-temporal expression profile is controlled. Here, we show that in fact, a few cells at stage 2 express *AG* at low levels and in apparently stochastic patches. The profile of *AG* expression is similar within each cell, with a sudden switch from no detectable expression to low expression, although different cells are activated at different times. *AG* activation spreads in adjacent cells, in part due to its capacity to move from one cell to another. Expression then rapidly undergoes a switch from low- to high-expression states. Expression is not homogeneous in each cell of the central dome, but is in fact slightly weaker in the centre than at the periphery. The border of the central dome is well defined with one row of cells showing lower AG levels. The profile of *AG* expression can be compared to mathematical models that confirm the scenario of switches and the requirement for AG to move and to bind its own promoter in dimeric or multimeric complexes.

3.1 Introduction

As described in chapter 2, I first created a model that encapsulates the key floral ABC interactions. Once the model had been implemented to describe *AG* activation satisfactorily, it was important to be able to compare the simulations of the model to observations in real tissues. Much is known about where *AG* is expressed: both its mRNA and the protein have been localized by *in situ* hybridization or by confocal imaging of an AG-GFP (green fluorescent protein) fusion chimera [Drews et al., 1991, Gómez-Mena et al., 2005, Urbanus et al., 2009]. *AG* expression was shown to be initiated very suddenly in the central dome, in whorls 3 and 4, at early stage 3, and to be maintained until stage 6-7. Recent results from whole mount *in situ* hybridization suggest that *AG* mRNA may be present slightly earlier than at stage 3 [Rozier et al., 2014].

To generate a more precise spatio-temporal analysis of *AG* expression, I have employed a recently developed method called MARS (see section 1.5), wherein a sample is imaged from different angles, which are then fused into a higher resolution image [Fernandez et al., 2010]. MARS also allows the quantification of one or more fluorescent reporters in individual cells [Fernandez et al., 2010]. Furthermore, statistical tools have been developed that utilise these

detailed data to gain a better understanding of gene expression during flower development (Legrand et al. Unpublished).

3.2 Results

3.2.1 Stochastic *AG* expression is observed at stage 2

As described above, the literature suggests that *AG* expression commences suddenly during the transition to stage 3 of flower development. However, recent data using a novel method developed in our lab, of *in situ* hybridization performed on entire inflorescences, give a more detailed tissular view of *AG* mRNA localization [Rozier et al., 2014], and suggest that AG is in fact present at stage 2, in a heterogeneous pattern in a few cells of the central dome (Figure 3.1 A.).In the past, a transgenic line carrying an *AG* translational fusion to the *GFP* reporter has been published [Urbanus et al., 2009], but does not display any visible expression prior to stage 3, possibly because of the very low level of fluorescence of this reporter. To overcome this, we created a reporter with a potentially higher level of fluorescence, by fusing two tandem copies of *Venus* to the C-terminal end of *AG* (*pAG::AG-2xVenus*, henceforth referred to as *AG-2xVenus* for simplicity). This construct is functional, as it rescues the phenotype of two strong *ag* mutant alleles and has been tested in two different ecotypes, (Landsberg *erecta* or L-*er* and *Columbia* (*Col-0*)). To be able to segment the tissue and to quantify *AG* expression in individual cells, we counterstained the membranes with FM4-64 and imaged these plants under a confocal microscope. With the *AG-2xVenus* construct and with a recent and sensitive confocal (a Leica SP8 with resonance scanning), I observed a low level of *AG* expression at stage 2 of flower development, before the sepals start to appear (Figure 3.1 B.).

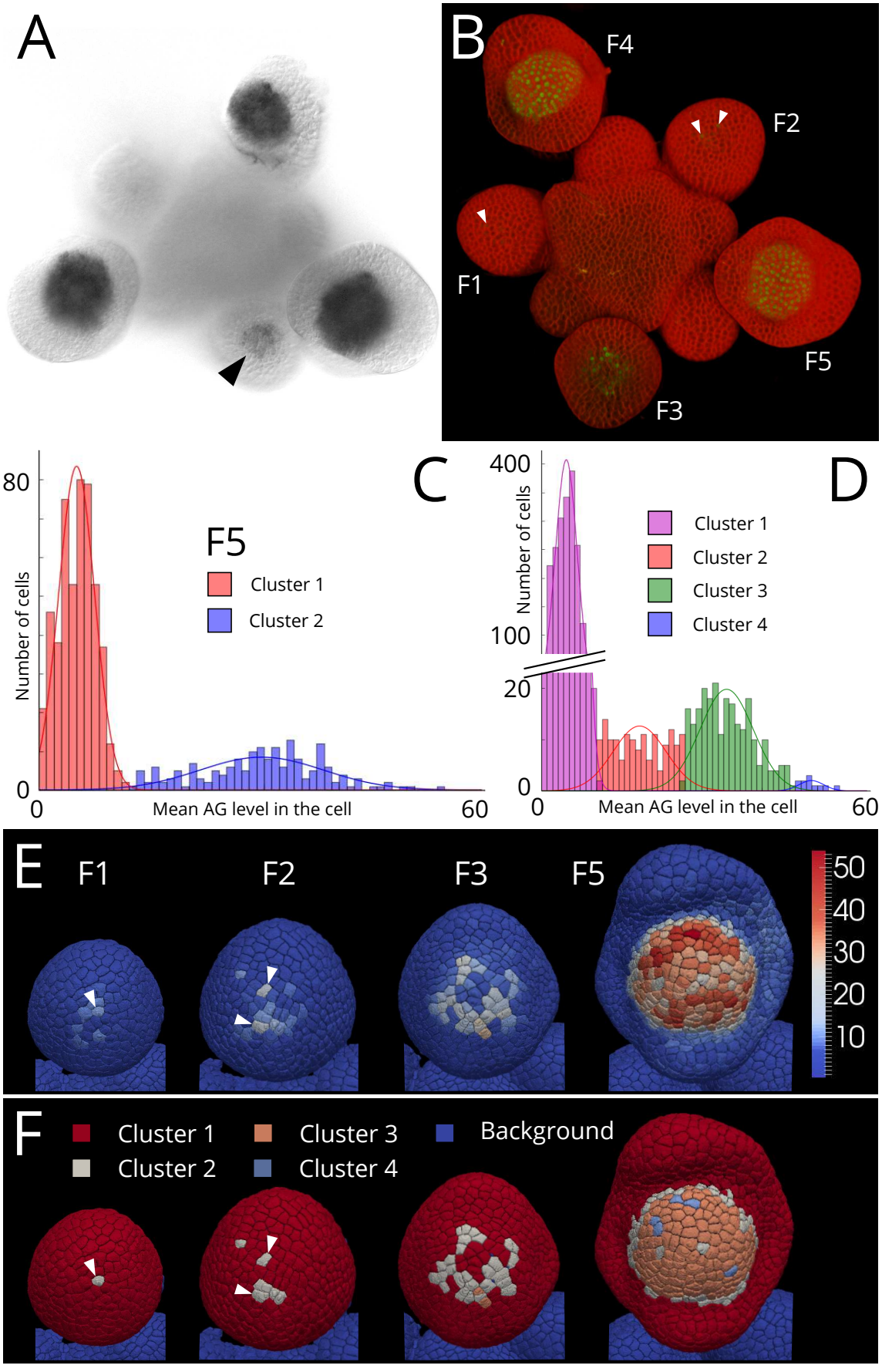
To take our analyses of the dynamics of *AG* expression further, I used the MARS pipeline (see section 1.5) to segment individual flowers (such that we obtained digitised flowers with individual cells identified and labelled) and to quantify the level of reporter fluorescence in each cell. Although confocal microscopy in general, and MARS-treated samples in particular, tend to provide clean cellular data, there is nevertheless an obvious difference between the outermost cell layer (the L1) and the inner layers, including the sub-epidermal layer (the L2). Differences include both the quality of the cell contours, so that cell volumes may be slightly incorrect, as well as the degree of fluorescence, due to the fact that the tissue tends to absorb

some of the emitted light. I thus focussed the bulk of my efforts on the L1, although I did compare L1 and L2 data in several individual flowers, where the segmentation was found to be of acceptable quality, and found the results to be comparable.

Once the reporter signal had been quantified in individual cells by simply summing the intensity of every voxel within each cell, I then normalised this against the cell volume. Because confocal images almost always have some degree of background noise, and because we were seeking to identify cells where *AG* expression initiates, and thus with potentially low levels of fluorescence, I wished to devise an unbiased method to separate 'background' cells from 'real' cells. To this end, I used the mixture model method to analyse the entire population of cells and identify those with a significant level of fluorescence (Figure 3.1 C.). In the stage 2 flowers examined, this method yields a few cells that express *AG* at a low level (Figure 3.1 E. & F.). These cells are not at the very centre of the dome, but mainly at the periphery of the AG domain, leading to *AG* expression first in whorl 3 and then in whorl 4 (Flower F3). At mid- to late-stage 3 (F4 and F5), AG is present in every cell of the central dome but not homogeneously, as described in the literature. Some cells display an AG concentration that is over two-fold higher than the others. Indeed, the distribution of cells in the flower F5 suggest that AG concentrations can vary greatly between cells (Figure 3.1 C. and Supp. table S3.2).

3.2.2 Peripheral and central cells in the *AG* expression domain present a lower concentration of AG proteins

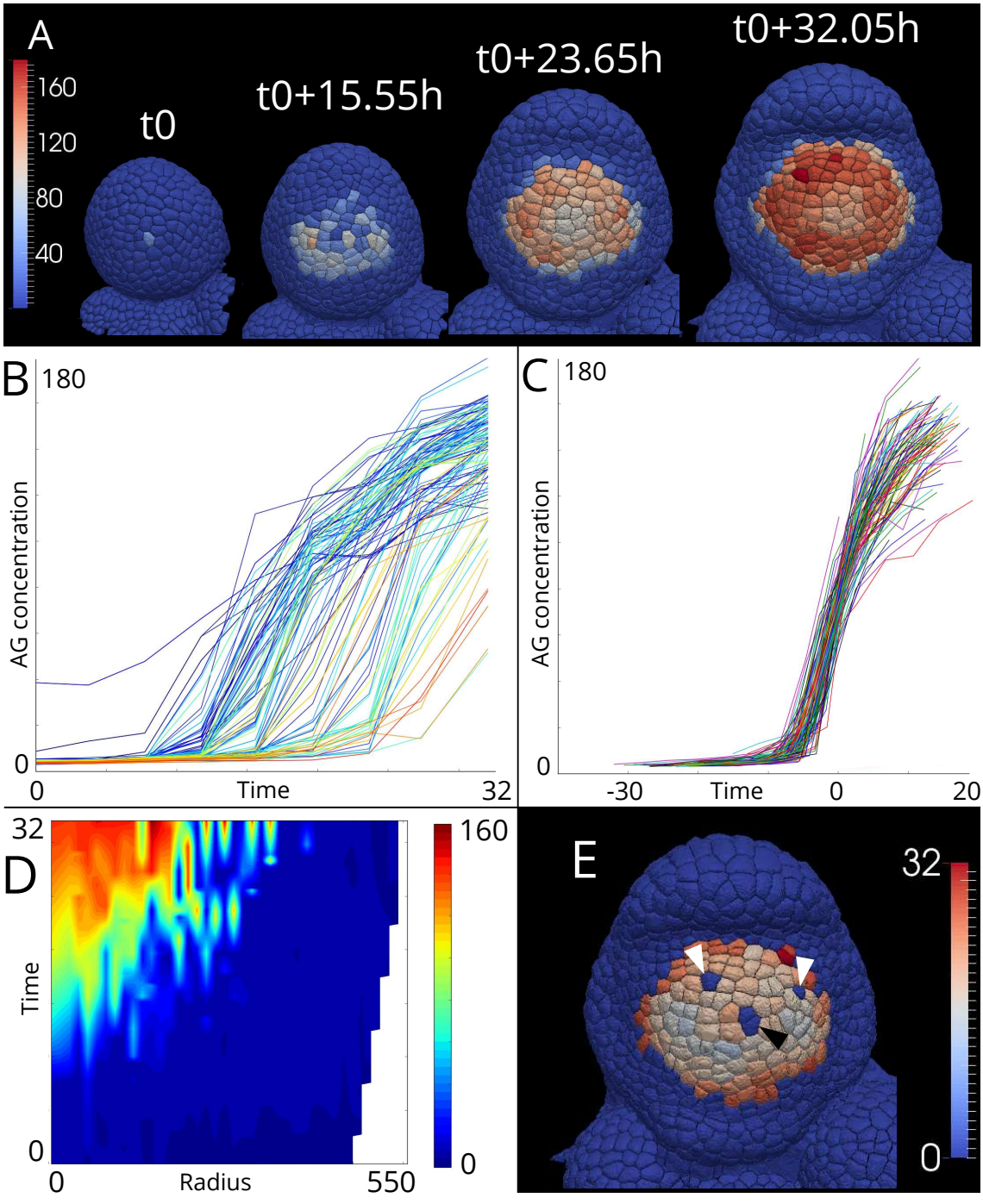
Based on the observation that AG levels vary within the tissue, I next carried out the mixture model analyses with four clusters, including data for all the cells of the L1 across all the different flowers examined (Figure 3.1 D. & F.). This yielded cell populations representing three different concentrations of AG, in addition to a fourth background cluster. Though the cells belonging to the lowest-expressing cluster (cluster 2) are mainly present in the youngest flowers, some of them are also found in a single ring of cells at the periphery of the AG domain, as well as of central dome. This layer of cells precisely defines the border between the second and the third whorls. It has been shown that AG is able to move from one cell to another. Thus when an *AG-GFP* reporter is expressed in the L1 of the flower, fluorescence was observed in the L2, though not in the underlying layers [Urbanus et al., 2010]. Thus one possible explanation for the peripheral layer of low-fluorescence cells is that



(Caption next page.)

Figure 3.1: AG expression in early *Arabidopsis* flowers: (A.) Whole mount mRNA *in-situ* localizations of *AG* mRNA in a shoot apical meristem (SAM). Black arrow indicates a stage 2 flower where AG staining is already visible. **(B.)** SAM from a plant of the L-*er* accession stained with the FM4-64 lipophilic plasma membrane dye and containing an *AG* reporter (*AG-2xVenus*). F1 and F2 are stage 2 flowers with a low stochastic expression of *AG* (white arrows). F3 is a young stage 3 flower with few cells expressing *AG*. F4 and F5 are stage 3 flowers with the expression of *AG* in all the central dome. **(C.)** A mixture model of the L1 cells of the flower F5. Cells are separated into two clusters, the first with cells that don't express *AG* and the second one with cells expressing *AG*. The two clusters can be described by a Gaussian (blue and red lines). **(D.)** A mixture model of all the cells of the L1 of the five flowers (F1 to F5). Distribution of *AG* expression in cells can be separated into 4 clusters. The first one with cells that doesn't express *AG* and the other one with the different level of *AG* expression. **(E. & F.)** Quantification of AG proteins at the cellular level of the L1 of F1, F2, F3 and F5 flowers. E. is the mean fluorescent level in each cell and F. is the representation of the 4 clusters identified in D. Few cells are detected in the two first flowers. For the F5 we can detect a ring of cells at the periphery of the central dome with a lower level of AG than in the central dome. Panel A is adapted from [Rozier et al., 2014].

they may not in fact express *AG* at all, but rather only receive AG that moves in from adjacent cells that themselves strongly express *AG*. The middle cluster (cluster 3) represents the majority of cells that contain high levels of AG and occupy the 'classical' AG domain, in the central dome of the stage 3 flower. The last cluster (cluster 4) accounts for only a few cells that contain very high levels of AG and that are found scattered within the third whorl of the oldest flowers (Figure 3.1 F. & Supp Figure S3.1 J.). In addition, *AG* expression is also lower in a few cells of the centre of the central dome (Figure 3.1). It has previously been shown that these cells have special properties, and behave differently from other cells during flower development. For instance AP1, which is initially expressed throughout the flower at stage 1 before gradually disappearing from the *AG* expressing domain, seemingly remains present in a few cells in the centre [Urbanus et al., 2009]. Additionally, in mutants of the *ULTRAPETALA1* gene, which acts as an epigenetic regulator of *AG*, a strikingly similar cluster of cells shows a visibly lower expression, which in turn leads to a delay in the repression of *WUS*, and thereby to floral patterning defects [Carles and Fletcher, 2009].



(Caption next page.)

Figure 3.2: *AG* expression dynamics in the ps143 time course: (A.) A MARS-treated AG-2xVenus floral time course (ps143), with segmented flowers of the first, fifth, seventh and ninth time points overlaid with average signal concentration per cell (total fluorescence divided by cell volume) represented in a colour scale. *AG* expression begins in a single cell at the first time point then spreads towards the periphery of the central dome. *AG* is fully expressed at stage 3 but with a lower level in the few cells in the centre. **(B.)** The evolution of signal concentration per lineage, showing that the onset of expression is broadly distributed over time. **(C.)** Signal concentration curves for each lineage, but centred on their activation inflection points, showing similar expression profiles for all lineages. **(D.)** A heat map of signal concentration over time based on the distance of each cell from the centre of the flower. **(E.)** A spatialized view of the absolute time of occurrence of the activation inflection point in each lineage, represented in a colour scale on the last time point. The colour is set to blue when an inflection point can not be defined in a given cell, either because signal was detected at the first time point (black arrow) or because a very late activation meant that those curves could be fitted with our methods (white arrows). Activation spreads to adjacent cells from early-activated lineages.

3.2.3 The activation of *AG* expression is stochastic

In order to gain a better understanding of *AG* activation dynamics, it was necessary to conduct a time course experiment with short intervals between time points. To this end, I generated the *2xmCherry-KA1* construct, which should mark cell outlines in red, and transformed this into the *AG-2xVenus* reporter line. Because this line does not require staining with the FM4-64 dye, I was able to track AG expression by imaging the same flower from multiple angles every four hours for 32 hours (Figure 3.2, S3.3 & S3.4). Two such time course experiments, ps143 and ps144 (Figure S3.5), were carried out, and each time point then treated with MARS as described previously. In addition, I manually determined the cellular lineages between time points for both time courses and analysed the signal as previously described, confirming the spatio-temporal characteristics previously deduced from the static images (Figure 3.1). I then analysed mean signal concentration per cell in each lineage from the first time point to the last (Figure 3.2 B.). Whereas most of the lineages expressing AG show a qualitatively similar profile, with expression mainly beginning at 15h, we do observe a degree of temporal stochasticity in the activation process.

To more closely examine this temporal stochasticity, I next narrowed our analysis to only the *AG*-expressing lineages (Figure 3.2 B.). We then fitted each of these curves with an exponential function, and identified the points at which *AG* expression became strongly

activated in each cell lineage, which we call the 'activation inflection points'. This point is not the onset of *AG* expression but is a more precise methods to compare activation of each lineages. Centering the original curves on their inflection points, reveals that all cells have a very similar profile of expression (Figure 3.2 C.). Once expression becomes visible within a given lineage, AG levels remain very low for several hours. Within about 4-6 hours, high levels of expression are detected in the mother or in the daughter(s), and this is maintained until the end of our analysis. This shows that the stochasticity observed in the early expression of *AG* is mainly in the initiation of its activation, and that once the activation process is engaged, all cells display a similar dynamics for *AG*.

3.2.4 *AG* expression spreads from one cell to its adjacent cells

Although *AG* is activated at different times in different cell lineages, the spatial distribution within the cell population is not strictly random. As shown above, expression is activated at stage 2, in a few cells at the periphery of the putative central dome. Furthermore, we also show that *AG* expression then 'spreads' to adjacent cells that themselves activate *AG* expression a short time later. In turn, expression then spreads to the neighbours of these cells. Thus, activation seems to spread from a few peripheral cells to the entire central dome of the flower (Figure 3.2 E.). This spreading can be linked to the capacity of AG to move through the L1 layer [Urbanus et al., 2010] as well as its capacity to autoregulate its expression by binding regulatory sequences in its own promoter [Yanofsky et al., 1990, Gómez-Mena et al., 2005].

3.2.5 *AG* activation profiles correspond to an auto-activation of AG itself as an homodimer

To gain a better understanding of *AG* regulation in early flower development, we previously developed a simplified mathematical model that summarizes the principal regulatory interactions that are known to be required for *AG* expression in the flower (Figure 3.3 A.). Given the detailed *in vivo* data described above, one interesting perspective is to compare *AG* dynamics in the model to the analysis of the time course by representing concentration of AG in each cells of the L1 by an heat map (Figure 3.2 D. and 3.3 B.). To produce this heat map, we choose the L1 cell that is at the center of the flower in the last time point of ps143. Then

we measure the distance of each cells to this one or to its mother cell for other time points. For every cells of the L1 of each time points, we plot depending of its distance to the central lineage and the corresponding time, the concentration values of AG in an heat map. We used a smooth function to obtain the result in (Figure 3.2 D.). We can compare this result with the simulation. It shows that the pattern of expression and its dynamics are very similar, nevertheless, in ps143, there is a strong stochasticity that produces an heterogeneous result and that differs from the simulation. It shows limits of using a deterministic model. Also, there is a small bias due to the fact that the pattern of expression of AG is not a perfect disc but more an ellipse that is larger at the side than at the abaxial and adaxial part of the central dome. This gives the alternation of blue and red stripes at the periphery of AG domain of expression.

In a second time I decided to compare precisely the profile of expression of each lineages to the dynamics of activation obtained with simulations. To be able to do the comparison, I took the result of the simulation of AG, select points on the X-axis homogeneously distributed and plot on the same graph the concentration of AG in function of the time in each of these points (Figure 3.3 C.). Most of the lines correspond either to x where AG is not expressed (lower limit of horizontal lines) or where AG is strongly expressed (upper limit). In the second case, the activation phase presents a fast evolution of the expression with a similar inflection point observed in the time course analysis. The curves between these two groups of lines correspond to the small border region with a low activation of AG. The main difference with observations of Figure 3.2 B. is the stochasticity of AG activation time present in our experiments that we cannot reproduce with a deterministic model.

This specific profile of activation of AG with the presence of an inflection point observed both in experiments and simulation is the characterisation of the existence of a threshold due to specific properties of the auto-activation of AG (Figure 3.3 D.). AG is known to bind its own promoter [Gómez-Mena et al., 2005]. AG belongs to the MADS family that generally form tetramers of proteins to bind specific DNA regions. For instance, it has been shown that SEPATALA3 (SEP3), another member of this family close to AG, forms homodimers and homotetramers that can bind the DNA. The composition of the auto-activation complex of AG is not known but we prove in the previous chapter that the number of molecules in this complex change the profile of expression of AG in our simulations. The profile with an inflexion point we observe here correspond to the presence of at least two molecules of AG

and gives new clues to support this hypothesis.

3.3 Discussion and perspectives

The activation of *AG* expression has been previously described as an on/off switch with a rapid activation in the two inner whorls at beginning of stage 3. Since *AG* is crucial for reproduction in angiosperms, the regulation of *AG* has been extensively researched over the past 25 years, and a lot is known about the molecular and genetic factors that contribute to its extremely precise spatial and temporal control. In the previous chapter, I described a model that I developed to summarise the interactions between these various transcription factors in order to understand the essential features required for this regulation.

AG is activated by LFY and WUS and repressed by AP2 [Drews et al., 1991, Parcy et al., 1998, **?**, Lenhard et al., 2001]. Furthermore, the maintenance of its expression requires its own activity, presumably within the quartet of TFs in which MADS family TFs function [Gómez-Mena et al., 2005]. My model suggests that this activity requires a specific composition of the quartet complex. Thus the presence of only one molecule of AG in this complex yields only one, positive, stable steady state in the system (Figure 3.3 D.). However, a small amount of AG in the cell will lead to progressively higher levels of expression. When two molecules of AG are present in this complex, it creates a threshold that falls between 0 and the positive steady state described above. The dimeric (or oligomeric) nature of the *AG* autoregulatory complex is consistent with recent observations on the MADS-box proteins [Immink et al., 2009, Puranik et al., 2014] and to results shown in the next chapter. In the model, a minimum amount of AG is necessary to trigger auto-activation and to provide stably high levels of AG in the cell. This second hypothesis yields the curves observed in Figure 3.3 C. with a long process of low activation and then a rapid increase in *AG* expression with an inflection activation point. Time course experiments with an *AG* translational reporter also yield similar results, with the presence of an inflection activation point that is not observed in the first scenario (AG monomers).

It has been previously shown that activation of *AG* is under the control of transcription factors, but in fact also involves several epigenetic factors [Saleh et al., 2007, Carles and Fletcher, 2009]. Our simulations and data suggest that the onset of expression is a long process of low activation by LFY and WUS, transcription factors that are present within

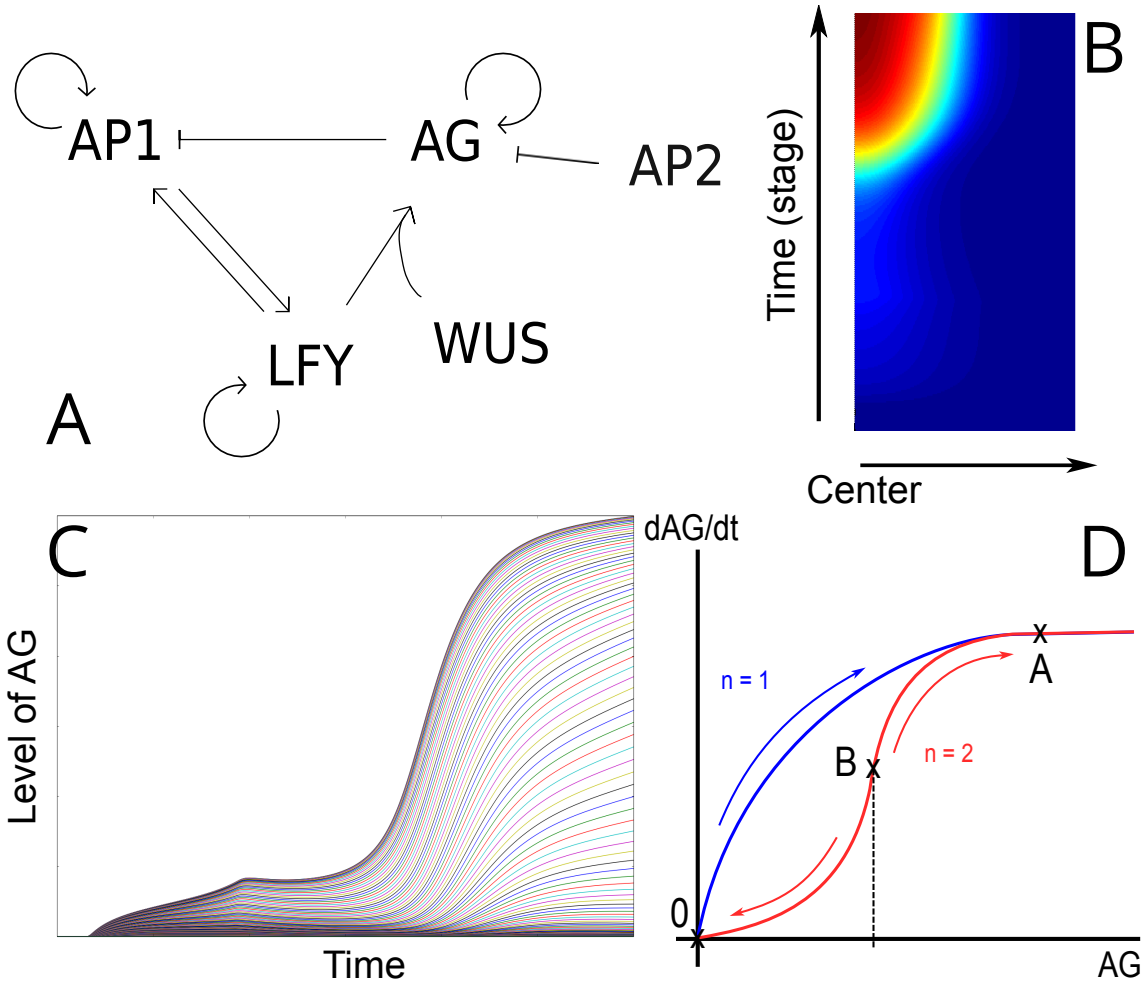


Figure 3.3: Simulation of *AG* dynamics of expression: **(A.)** Schematic representation of the principal genetic interactions used to model *AG* regulation, as described in chapter 2. **(B.)** Simulation of AG concentration during the early flower development, with the relative concentration of AG represented as a heat map (where red indicates high expression and blue indicates no expression). The X-axis represents a radius on the flower and the Y-axis represents developmental time. AG appears rapidly in the equivalent of the two inner whorls after a delay. **(C.)** AG concentration in homogeneously distributed X-axes points of the simulation in function of the time. Each curves correspond to a fixed point of the domain of study. A sudden increase in expression is clearly visible in the points that correspond to the central dome. **(D.)** Schematic illustrating how *AG* expression dynamics differ when AG acts as a monomer or a homodimer (or higher-order complex) to autoregulate its own expression (resp. n=1 or n=2).

the putative *AG* expression domain around 1-2 days prior to any detectable AG. This low-expression phase possibly coincides with change in chromatin structure at the *AG* locus. The existence of a previously-defined threshold might explain why *AG* expression takes time to be initiated, but also why, when the threshold is reached in a cell, *AG* can be rapidly activated. Confocal imaging of different *AG* reporters in two different ecotypes revealed that the initiation of *AG* expression happens in just a few cells at mid-stage 2. Thereafter, in only a few hours, full expression of *AG* can be observed in the whole central dome. The initial stochasticity is thus only with respect to the time of onset, and, the overall expression profiles are similar in all cells. These results lead to two further questions: what parameters drive the very rapid switch from low- to high-expression phases in the central dome, and what restrains *AG* expression to this precise spatial domain and prevents spillover into the first and second whorls?

To answer the first question, it has been shown that AG is able to move from cell to cell, but predominantly only by one cell layer [Urbanus et al., 2010]. I have previously discussed the requirement for AG levels to be higher than a defined threshold for the auto-activation process to be kick-started. In this manner, a cell in which *AG* expression has been activated can transfer AG to its adjacent cells thereby help activate expression in those cells, which will in turn activate expression in their own neighbours. Thus activation can spread rapidly across the flower, resulting in the homogeneous expression of *AG*. Support for this hypothesis comes from the observation of this travelling wave of activation in the time course experiments (Figure 3.2 E.), where the spreading of activation from the initially activated cells is visible. In this context, the stochastic onset of *AG* expression may present an advantage for robust patterning to occur in that, given the capacity of AG to move from cell to cell, activating expression in only a few cells is sufficient to ensure that the proper final pattern emerges.

In theory, this mechanism would lead to the expression of *AG* everywhere in the flower, and not only in the two inner whorls. In the previous chapter, I studied the effect of AP2 in defining an activation threshold. The presence of AP2 in a cell would drastically raise the value of the threshold, to the point of suppressing the positive stable steady state. Thus, given that AP2 may be present as a gradient in the flower, with high levels at the periphery and low levels at the centre [Wollmann et al., 2010], it helps define a sharp border between cells with a low threshold that are able to start auto-activation of *AG* and cells with a high

threshold that will never be able to start auto-activation. Because AG can move from cell to cell, it is present in the cells immediately surrounding those that are really expressing *AG*. These cells are likely unable to express *AG* endogenously because of the presence of AP2. This observation is another argument to support the importance of AG as a homodimer in the quartet complex for auto-activation, for without this hypothesis, the threshold would not exist and the boundary would be less sharp than what is observe in reality. To prove this, it would be necessary in the future to compare the pattern of expression of *AG* to that of *AP2* and to other genes that can repress *AG* expression, such as as *SEU* and *LUG* [Liu and Meyerowitz, 1995, Sridhar et al., 2004].

One final observation is that the cells of the central dome of the flower do not all respond identically to the transcription factors that regulate *AG*, since *AG* onset begins with only a few third whorl cells, whereas activators such as *LFY* and *WUS* are also strongly expressed in the fourth whorl, and whereas repressors such as *AP2* are not. In addition, once the final pattern has been established, a few cells in the centre display lower concentrations of AG relative to the other cells of the third and fourth whorls. This may be linked to the expression of *AP1*, which is still present in those same cells at stage 3 [Urbanus et al., 2009] and also to the lower level of *AG* expression in these cells already observed in *ult1* [Carles and Fletcher, 2009]. This small group of cells in the centre would appear to have different properties to the other cells of the flower in many different contexts [Milani et al., 2014]. While this is true at stage 3 of flower development, *AG* remains expressed during several further stages. However, because the sepals cover the flower after stage 3, it is currently technically difficult to determine whether this observation remains valid at later stages of development.

Our mathematical model shows that a low level of AG at the beginning, followed by a rapid activation of *AG*. While it is tempting to equate this low expression with the stochastic, low-expressing stage 2 cells, they are, in fact, different. In the simulations, the low-expression phase is due to the activation of *AG* by LFY and WUS prior to the initiation of the auto-activation. However, it is unknown whether this particular mechanisms really exist in the flower. Sometimes, before the rapid switch from low- to high-expression phases we observe that certain cells continue for much longer in the low-expression phase. Nevertheless, the precision of the time course, principally the fact that images are acquired every four hours, doesn't allow to have a striking observation that can be correlated to the model. More

precise experiments are necessary in order to have a clear answer.

To conclude, we can propose a new scenario for *AG* expression. Due to expression of transcription factors like *AP2*, each cell has a particular inherent capacity regarding *AG* expression. Thus some are unable to express *AG* and some in the centre can have a strong expression but of different level depending on whether they lie in the third or the fourth whorls. These cells behave like a on/off switch that is activated by the balance of epigenetic regulators and activation by LFY or WUS. The capacity of AG to move between cells is important to build up sufficient AG to flip the switch and the full central dome is rapidly activated. Once cells are activated, they maintain their expression. This scenario needs other experiments to be proved.

3.4 Methods

Reporter constructs

AG-2xVenus used in Figure 3.1 is a fragment of genomic *AG* from *Col-0* with its promoter (8.3kb) amplified with the pPD381 and pPD413 primers and transferred with XmaI digestion in *BJ36* containing *2xVenus* fluorescent reporter. *BJ36* with *2xVenus* is obtained amplified Venus from *pCS2-Venus* with pPD441 and pPD442 primers adding 5xAla at the beginning of *Venus* and transferred 2 times in BJ36 with BamHI and XmaI digestion. the *pAG::AG-2xVenus* obtain fragment is transferred in *pART* (a kanamycin resistant vector) with XmaI digestion and then transformed in L-*er Arabidospsis thaliana* using *Agrobacterium tumefaciens*.

pAG::AG-2xVenus used in Figure 3.2 and S3.5 is the same vector than previously but transformed in *Col-0*. The plants containing this construct have been crossed with a membrane marker *pRPS5a::2xmCherry-KA1*. *pRPS5a* is an ubiquitous strong promoter cloned in *pDONR P4-P1R*. *2xmCherry* is a red fluorescent protein isolated from *Discosoma sp.* cloned previously in *pDONR 221* [Simon et al., 2014]. *KA1* is a folded domain of the human microtubule affinity-regulated kinase (*MARK1*) that associates non-specifically with every anionic lipid. This domain was previously cloned in *DONR P2R-P3* [Simon et al., 2016]. The complete vector was created using gateway technology using these three vectors transferred in dpBART (a basta resistant vector for gateway triple recombinations). The vector was transformed into plants using *Agrobacterium tumefaciens*. This marker stained only the

membrane of the cell wall and allow a good segmentation.

Plant growth conditions

All plants were grown on soil at 20°C under long day conditions (16h light/8h darkness) except the one for the Figure 3.1 & the Supp Figure S3.1 that was grown during 4-5 weeks in short days (8h light/16h darkness) and then under continuous light at 14°C. All transgenic plants were selected first *in vitro* on Murashige and Skoog (MS) medium until having homozygous plants.

Confocal imaging

Sample preparation was done like described in [Fernandez et al., 2010] to obtain meristems with only stage 1, 2 and 3 flowers. Membranes were stained with FM4-64 15 minutes before imaging when plants don't have a transgenic membrane marker. Plants were observed with the SP8 confocal of Leica at three different angles to do cellular analysis. For time courses, plants were kept in sterile boxes in medium described in [Fernandez et al., 2010] in an incubator with long day growth conditions. The 2xVenus and the FM4-64 were excited with a laser of 488nm wavelength and the 2xmCherry with one of 561nm.

Image analysis

Projection of meristem and flowers are performed using '3D Viewer' plug-in of ImageJ. For cellular analysis, the 3 images with different angles were reconstructed using the membrane marker with 'blockmatching' script of MARS software. The contour of the flower was detected with LSM scripts to obtain a better segmentation using the watershed method (Unpublished method developed in the lab). Quantification of the signal in each cell and statistical analysis were performed using common tools of the Python language. Mixture model used 'mixture' library of Python. Figures were produced using paraview software.

Lineage analysis of ps143 and ps144 are performed manually. Analyses are performed with temporal graph developed previously in the lab during the thesis of J. Legrand (Unpublished). Curves of evolution of the fluorescent concentration in each lineage for cells that express AG are fitted with the curve-fit function of scipy, a python library. The choose function to fit these curves is an exponential:

$$\frac{K}{1. + a. \exp(-r.x)}$$

with K, a and r the three parameters to fit. Activation time is defined as the inflexion point of the fitted curves (Figure supp S3.2). Then original curves are centred on the activation point to obtain the Figure 3.2 C.

Heat maps are produced calculating the distance of each cell to the defined central lineage. This lineage is the one of the cell number 102 for ps143 and 27 for ps144. For each distance points, average of fluorescence of each cell at this distance is smooth to obtain the heat map in Figure 3.2 D.

Modelling

AC model has been previously described in the previous chapter. The profile of *AG* expression (Figure 3.3 C.) is an extraction of the WT *AG* simulation.

3.5 Author contributions

The *AG-2xVenus* reporter has been previously cloned by P. Das. I cloned the *2xmCherry-KA1* reporter. I performed plant selection, imaging and image analysis. P. Das, V. Mirabet and A. Kiss help me to developed new scripts for the image reconstruction and segmentation. V. Mirabet performed the statistical analyses on ps143 and ps144. I write the chapter with correction of P. Das. A. Chauvet helps me to take care of the plants.

3.6 Acknowledgement

I acknowledge Y. Jaillais and T. Vernoux for advices and the clones to produce the 2xmCherry-KA1 marker and J. Legrand for the advices in the development of statistical tools.

3.7 Supplementary data

Name	Sequence
pPD381	GTC.CCCGGG.AGTGATCCCTTCTCCAACACA
pPD413	AGT.CCCGGG.TAACTGGAGAGCGGTTTGGT
pPD441	AGT.GGATCC.GCAGCTGCCGCAGCTGCG.ATGGTGAGCAAGGGCGAG
pPD442	GTC.TCTAGA.CTAGATAGATCTCTTGTACAGCTC

Table S3.1: Used primers.

	cluster 1			cluster		
Flower	Nb of cells	Mean level	StDev	Nb of cells	Mean level	StDev
F1	7	8.85	0.78	319	3.61	1.40
F2	9	11.81	3.15	370	3.90	1.56
F3	29	16.05	4.71	478	4.31	1.87
F4	184	29.57	9.26	397	5.36	1.71
F5	175	29.69	8.24	501	5.05	2.40
All Flowers	395	28.34	9.41	2014	4.53	1.99

Table S3.2: *AG* expression in early flowering: Expression level in the 5 flowers of Figure 3.1. Cells of the L1 layer has been separated in two clusters (Figure S3.1). The mean level of fluorescence in these cluster has been fit with a Gaussian. For each cluster of every flowers we obtain the average value of mean level of fluorescence in the cells, the standard deviation of the cluster and the number of cells in the cluster. Last line of the table correspond of the analysis of L1 cells of all flowers.

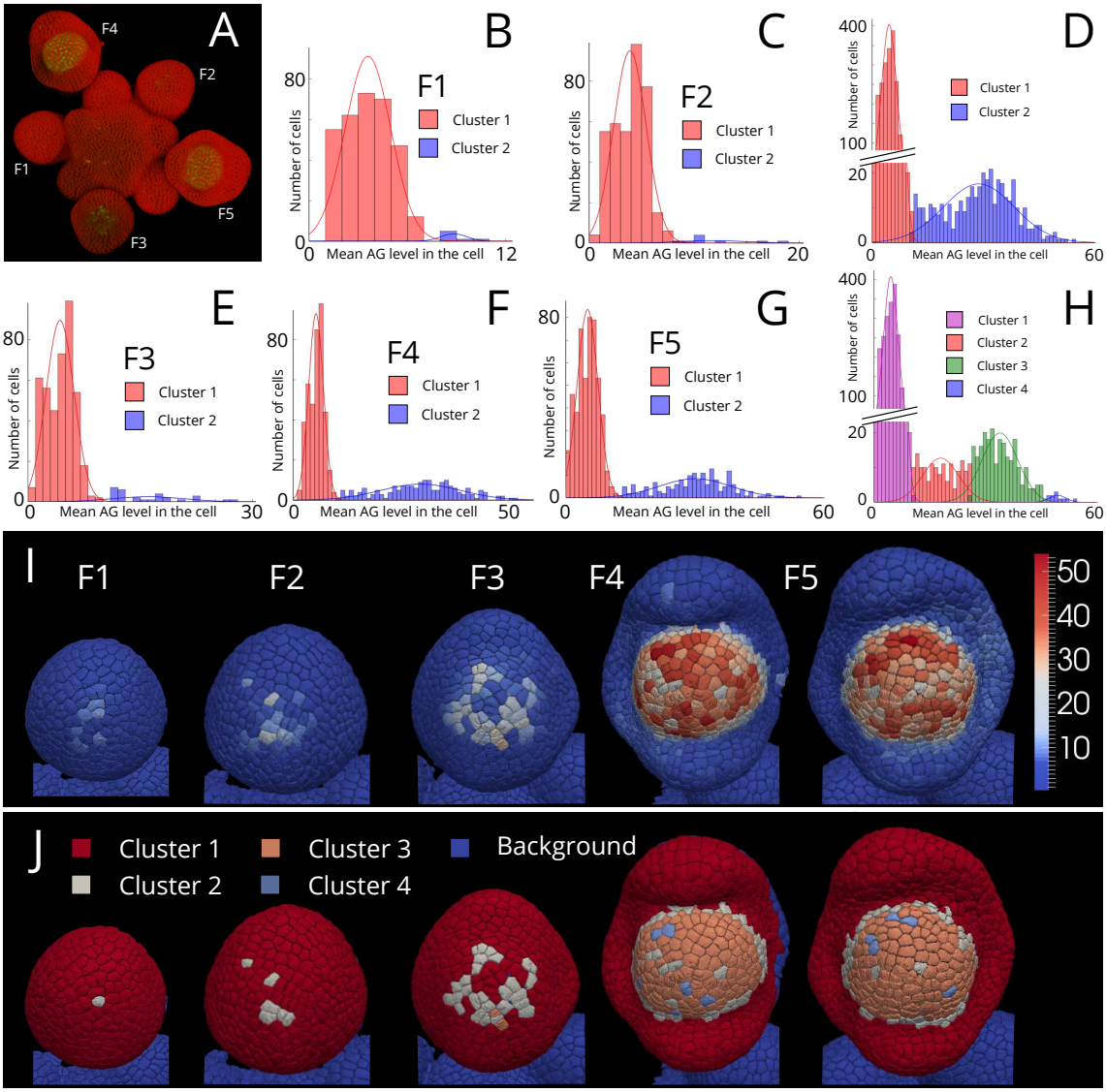


Figure S3.1: *AG* expression in early flowering: (A.) Shoot apical meristem with flowers stained with a red membrane marker (FM4-64) and containing an *AG* fluorescent reporter (*AG-2xVenus*). F1 to F5 correspond to the studied flowers. **(B. C. E. F. & G.)** Mixture model of the mean level of fluorescence of the *AG* reporter in the L1 cells of respectively the flower F1, F2, F3, F4 and F5. Two clusters are identified, the first one with only noise and the second one with significant concentration of the protein. **(D. & H.)** Mixture model of the L1 cells of the five flowers. **(D.)** Two clusters are identified (one with only noise and the second one with fluorescent cells). **(H.)** 4 clusters are identified, the fluorescent cells are separated in three clusters depending on their level. **(I. & J.)** Quantification of AG protein at a cellular level inside the L1 cells of the 5 flowers. **(I.)** is the mean fluorescent level in every cell and **(J.)** is the representation of the 4 clusters identified in H.

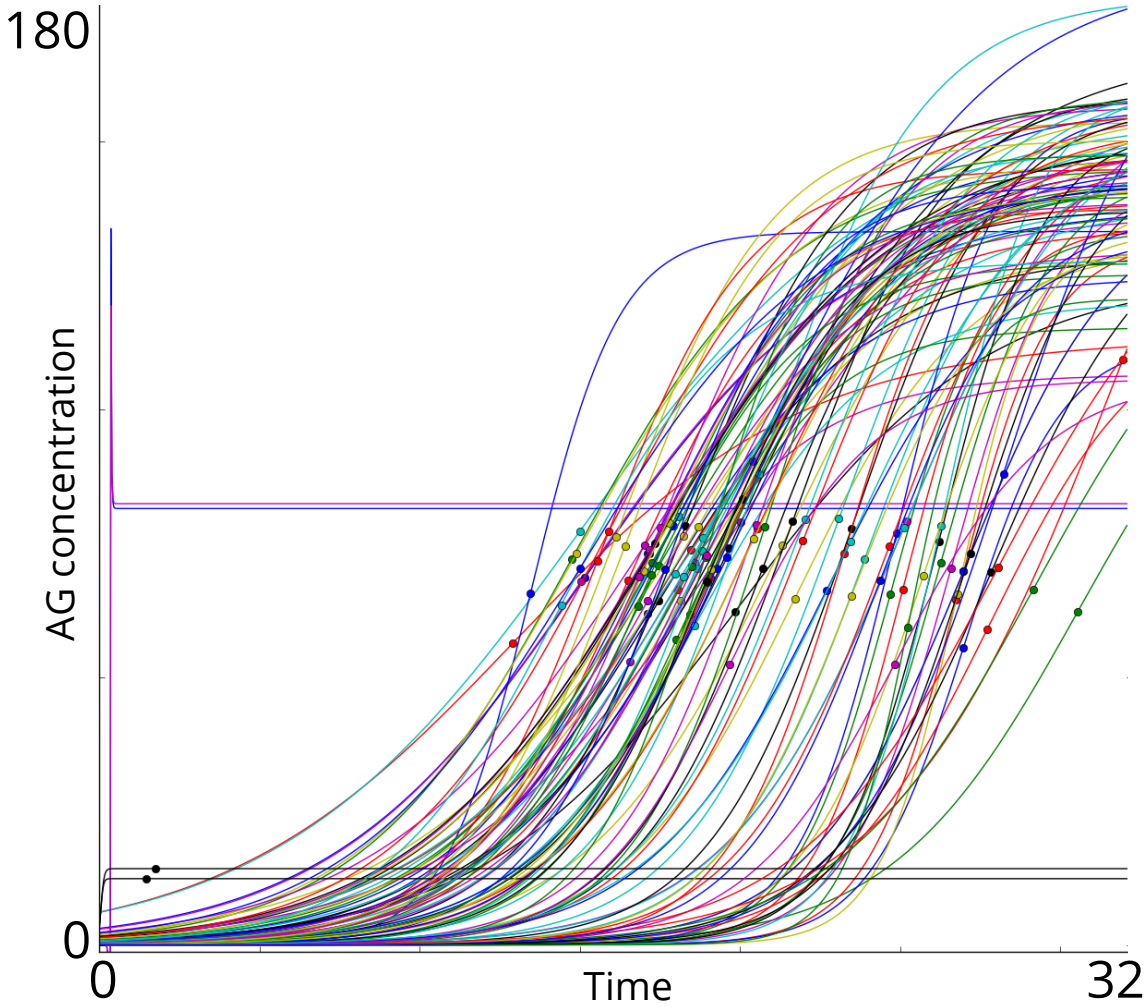


Figure S3.2: Fit of *AG* profile of expression for ps143: The few horizontals lines correspond to profile where the fitting methods cannot find a good solution.

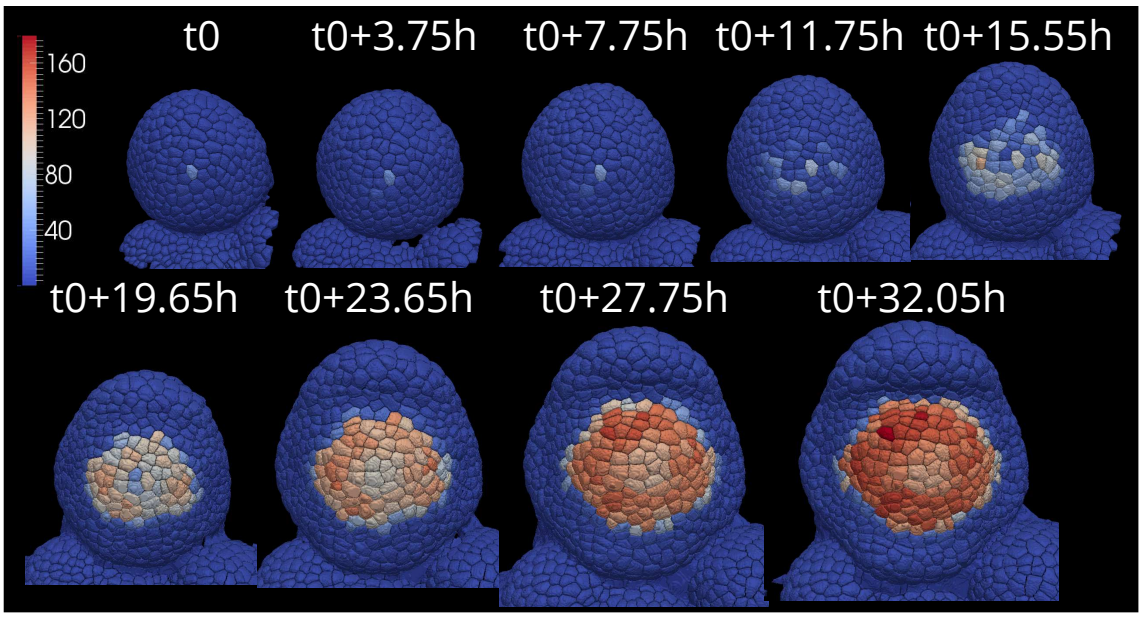


Figure S3.3: *AG* expression in ps143 time course: Quantification of fluorescent reporter of the ps143 signal.

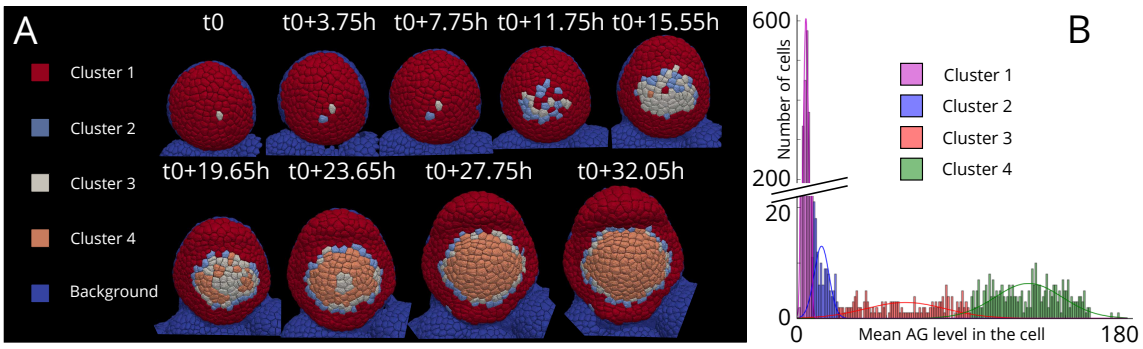
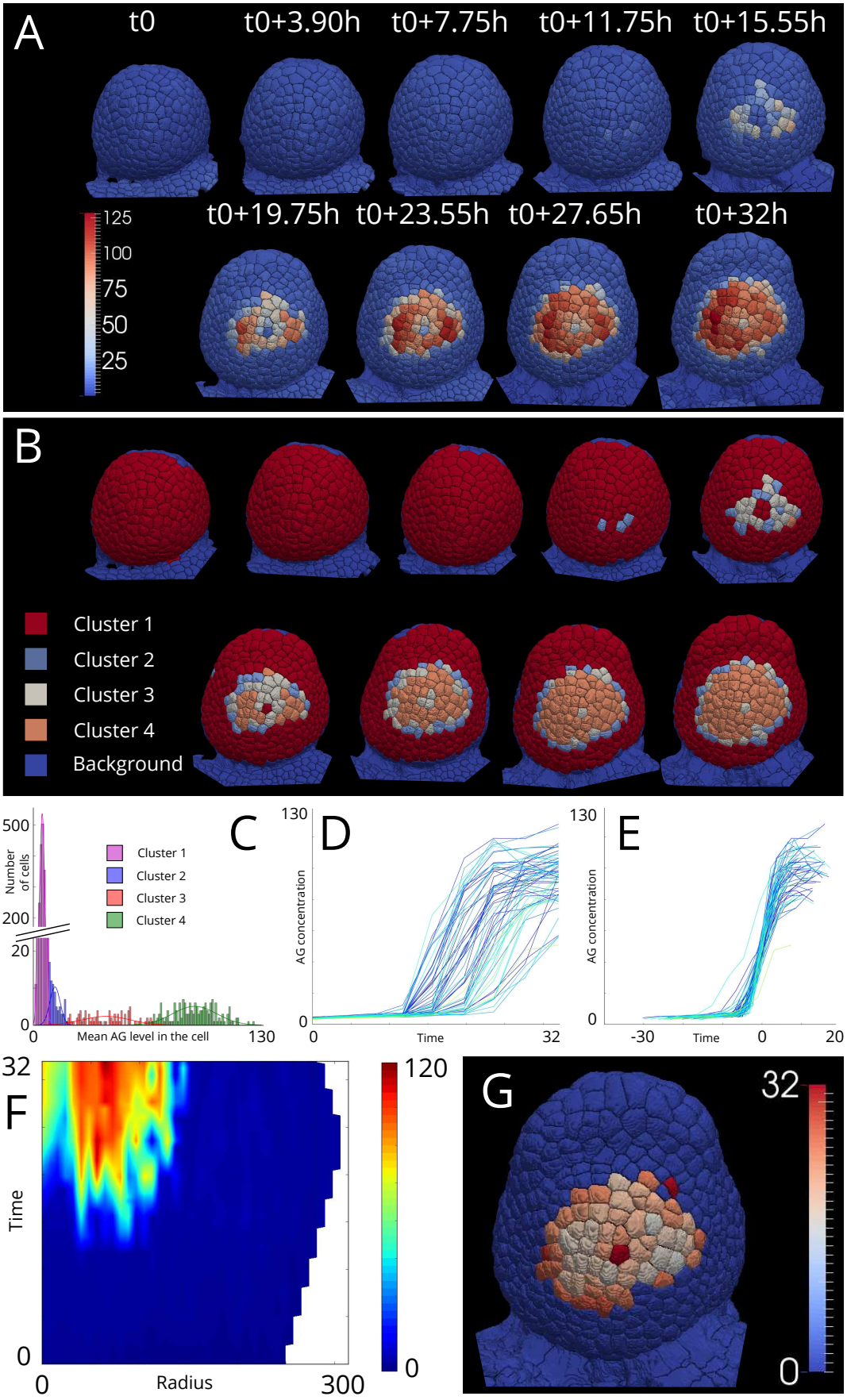


Figure S3.4: Cluster analysis of *AG* expression in ps143 time course: the four clusters are obtained with mixture model methods of the L1 cells of all the flowers of ps143.



(Caption next page.)

Figure S3.5: *AG* expression analysis of the ps144 time course: (A.) mean quantification of fluorescence of *AG-2xVenus* reporter. **(B. & C.)** Mixture models with four clusters on L1 cells of all flowers. **(D.)** Evolution of mean level of fluorescence during time in each lineages. **(E.)** Evolution of fluorescence centred on the activation time of each lineages. **(F.)** Heat map of fluorescence level. **(G.)** Representation of activation times of each lineage on the last image of the time course. Histogram of activation time of each lineages.

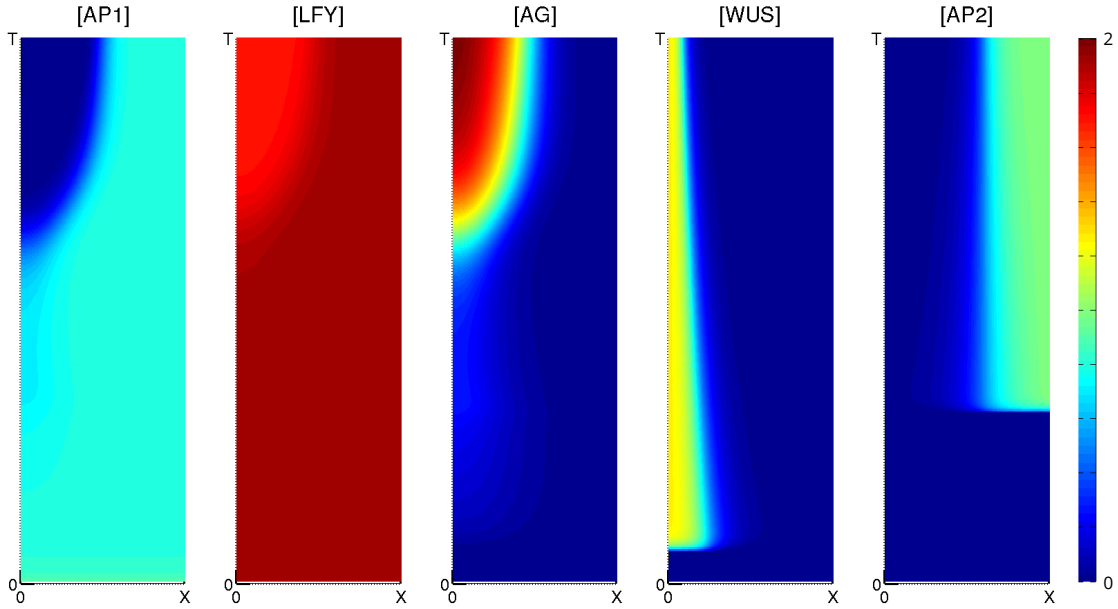


Figure S3.6: Simulation of the AC model: Simulation of the concentration of the five components of the AC model.

Chapter 4

The origins of the robust AGAMOUS pattern of expression

Contents

Abstract **95**

4.1 Introduction **96**

4.2 Results **98**

4.2.1 Over-expression of *AG* have an impact on its localization 98

4.2.2 Activation by LFY and WUS is necessary for proper *AG* expression and a WT phenotype 99

4.2.3 Auto-activation of *AG* is repressed until stage 3 of flower development 103

4.2.4 AG likely forms stable homo-dimers 104

4.2.5 The role of cell-to-cell movement in establishing the *AG* pattern . 106

4.2.6 AP2 is one of the factors that restrain the activation of *AG* in the early stages of flower development 108

4.3 Discussion and perspectives **108**

4.4 Methods **114**

4.5 Acknowledgment **116**

4.6 Author contributions **116**

4.7 Supplementary data **116**

Abstract

AGAMOUS (*AG*) is a major regulator of flower organ identity in *Arabidopsis*. While its regulation has been extensively studied over the years, how its precise spatial and temporal expression dynamics are established within the growing flower is still unclear. Based on the

hypotheses proposed in the two previous chapters, I experimentally studied the mechanisms underlying *AG* regulation during early stages of flower development. In order to understand the effect of cellular AG concentration upon its expression dynamics, I showed that altering *AG* copy number changes the onset of *AG* expression and modifies the subcellular localisation of AG. I have previously proposed the importance of nature of the AG autoactivation complex in the proper dynamics of *AG* expression. To test this, I studied the capacity of AG to form a homodimer. However, I none of my experimental conditions are able to advance the autoactivation process relative to the WT. The onset of *AG* expression is determined in part by LEAFY (LFY), and in constructs with mutatant LFY binding sites in the regulatory regions of the *AG* locus, I observe a delay in *AG* onset. Mutating the majority of binding sites produces phenotypic defects. In earlier chapters, I have noted that the movement of AG between cells appears to have an important role in the rapid increase in *AG* expression levels in the cells of the central dome. However, when this movement is blocked, only minor alterations in *AG* expression are detected, and these are not accompanied by phenotypic changes. Lastly, it is known that APETALA2 (AP2) is a key repressor of *AG* in the two inner whorls of the flower. I show here that it also plays a role in restraining the strong activation of *AG* at mid- to late-stage 2 of flower development.

4.1 Introduction

Previously in this manuscript, I have described my analysis of the published data regarding the mechanisms that regulate *AG* activity during early flower development (chapter 2). I further used them to propose a mathematical model that investigates the dynamics of *AG* expression. Using simulations and mathematical analyses, I then proposed several hypotheses that address different aspects of the key regulatory mechanisms. I followed this up with a detailed study of *AG* expression dynamics *in vivo* (chapter 3), I was able to provide experimental support for, and greater precision about, these hypotheses. This, in turn, has helped me design further experiments to better examine *AG* regulation.

My *in vivo* analyses revealed that *AG* expression commences stochastically, in a few scattered cells, during stage 2. I showed that this stochasticity pertains only to the beginning of the process of activation and that each cell displays the same dynamics of expression once they are activated. However, neither the origins of AG stochasticity, nor the key determinants

of gene activation are understood. In the general introduction to this manuscript, I describe how stochasticity in biology can have various origins. The tools used in the previous chapter allow us to study these origins by analysing subtle modifications of *AG* regulation.

AG is activated by the direct binding of LFY and WUS to multiple binding sites located in the second intron of the locus [Busch et al., 1999, Parcy et al., 1998, Lenhard et al., 2001, Lohmann et al., 2001]. Certain sites are bound by LFY alone, whereas some other sites have been shown to bind both LFY and WUS in very close proximity; the consensus binding sites are less than 10 bp apart [Lohmann et al., 2001, Hong and Hamaguchi, 2003, Moyroud et al., 2011]. The model suggest that these interactions are necessary for the initiation, but not the maintenance, of *AG* expression. This result has been previously described biologically [Lohmann et al., 2001, Lenhard et al., 2001], although, the precise role of each interaction was addressed qualitatively at the tissue level, and not with cellular quantifications during the initiation of *AG* expression. The study of these interactions, their redundancy, and their precise role in the dynamics of *AG* expression, alongside a close *in vivo* analysis and correlations to their induced phenotypes, could help us understand the robustness of flower development and the role of the multiplicity of binding sites.

Once activation is stochastically initiated by LFY and WUS, *AG* is able to maintain its own expression [Gómez-Mena et al., 2005]. It has been shown that AG is able to bind its own second intron, on consensus sequences called CArG boxes in a complex that is probably a tetramer. Based on the studies of the model shown earlier and on the *in vivo* analyses of the *AG* dynamics of expression in the previous chapters, I proposed that the composition of this complex can be important in the dynamics of *AG* expression and that we have probably two or more proteins of AG in this complex. This hypothesis is coherent with recent results on a close member of the same family, SEPATALA3 (SEP3) that is able to form stable homodimers and homotetramers and bind these CArG boxes [Puranik et al., 2014]. Furthermore, there are at least three CArG boxes within the second intron of *AG* that are known to be involved in *AG* autoregulation with a redundant activity [Busch et al., 1999, Hong and Hamaguchi, 2003, Gómez-Mena et al., 2005]. As for the activation of *AG* by LFY and WUS, precise analyses of these interactions have not been performed yet.

The Activation of *AG* is a stochastic process. Despite this however, *AG* is fully expressed in the central dome at beginning of stage 3. In the model, I hypothesised that AG movement is important to obtain this robust pattern. It has previously shown that AG is able to move

cell to cell in the flower [Urbanus et al., 2010]. When *AG* is expressed only in the L1, protein can be detected in the L2 but not in the more inner cells meaning that the range of AG movement is likely restricted to local. In the previous chapter, analysis of the precise dynamics of *AG* expression gives us some clues about the movement of AG and its importance in the activation of each cell of the two inner whorls. More generally, the role of such cell to cell movement has never been studied in the literature.

The restriction of AG to the two inner whorls is mainly due to AP2 [Drews et al., 1991]. Indeed, AG is present in the outer whorls in a strong *ap2* mutant. In the model, I confirm this result but I also propose that AP2 is able to restrict *AG* expression between mid-stage 2 and stage 3 flower in the central dome. This is due to the influence of AP2 on the level of a threshold that a cell needs to reach before auto-activation of *AG*. Experiments performed on *ap2* are not precise enough to be able to confirm this hypothesis.

Based on these observations and on the hypotheses of our model, we can ask 6 questions:

- What is the origin of AG stochasticity?
- What is the specific role of LFY and WUS in *AG* activation?
- Does *AG* auto-activation is a key factor in *AG* dynamic pattern of expression?
- Does the complex involved in *AG* auto-activation contain at least two proteins of AG?
- How the movement of AG influence its pattern of expression?
- Does AP2 has a role in early *AG* activation?

In this chapter, I will give mainly preliminary results to help to answer these questions and propose new experiments and perspective to complete my results.

4.2 Results

4.2.1 Over-expression of *AG* have an impact on its localization

In chapter 4, I showed that AG appears stochastically in a few cells during stage 2. Stochasticity can have many origins [Collaudin and Mirabet, 2014], but one factor is the number of copies of a gene in the genome [Holloway et al., 2011]. Using computational modelling, it was shown that a several copies of the same gene can buffer noise during its transcription.

Since *AG* is a single-copy gene in *Arabidopsis*, one way to test whether this hypothesis holds true for *AG* is by adding copies of *AG* to the genome and observing the effect on the pattern of expression and on floral architecture. To this end, I added one copy of a roughly 8-kb fragment of the *AG* locus in WT plants that already contained the previously-described *AG-2xVenus* reporter. Several independent transgenic lines (3 over 14) present a strong *ag* mutant phenotype, which is likely due to silencing, a reasonably common occurrence. The other lines do not display any significant phenotypic abnormalities, but the pattern of expression of *AG* is different (Figure 4.1). First, adding one copy of *AG* changes its intracellular localisation in young flowers and in independent lines from a sharp, predominantly nuclear signal to a blurry pattern with higher levels of AG in the cytoplasm (Figure S4.1). Secondly, whereas WT mid-stage 2 flowers bear some scattered cells that expressing *AG*, the lines with an extra copy of *AG* only bear flowers with reporter expression at stage 3. These flowers appear to transition very quickly from low- to high-expression phases. Thirdly, I also observe that AG is clearly lower in the centre of the 2 inner whorls than at its periphery. This difference is more visible than in the WT.

4.2.2 Activation by LFY and WUS is necessary for proper *AG* expression and a WT phenotype

We have consequent knowledges about how proteins I used in my model directly interact with the second intron of *AG*. The binding sites corresponding to these interactions are mainly known, and analysis of part of this intron gives us few insights about their roles (see subsection 1.2.3). To have precisions about these interactions and specifically to understand the origin of the stochastic pattern of *AG*, I decided to mutate the most important known binding sites of the second intron of my reporter *AG-2xVenus* for LFY-WUS, AP2 and the main CArG-boxes, and to transform these constructs in the strong *ag-salk* mutant. These experiments are not finished yet but through a collaboration with François Parcy, I could use similar work performed on the same promoter.

LFY and WUS are the two main activators of *AG* [Lenhard et al., 2001, Lohmann et al., 2001]. Together they bind five sites of the second intron of the *AG* locus, and are at least partially redundant [Deyholos and Sieburth, 2000, Hong and Hamaguchi, 2003, Moyroud et al., 2011]. Four of them was proposed by Deyholos and Hong based on consensus sequences [Deyholos and Sieburth, 2000, Hong and Hamaguchi, 2003] (consensus sequences of the

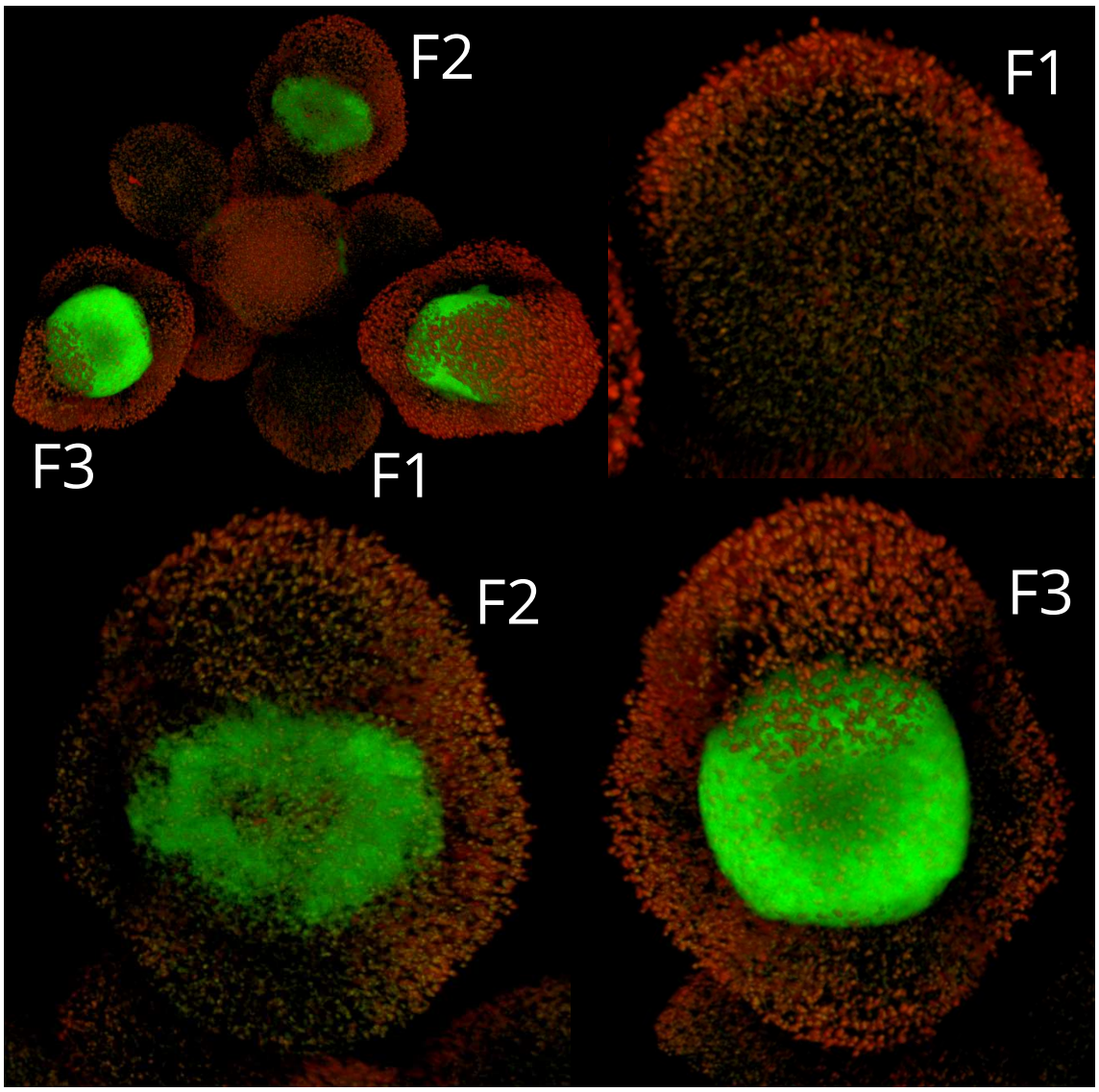
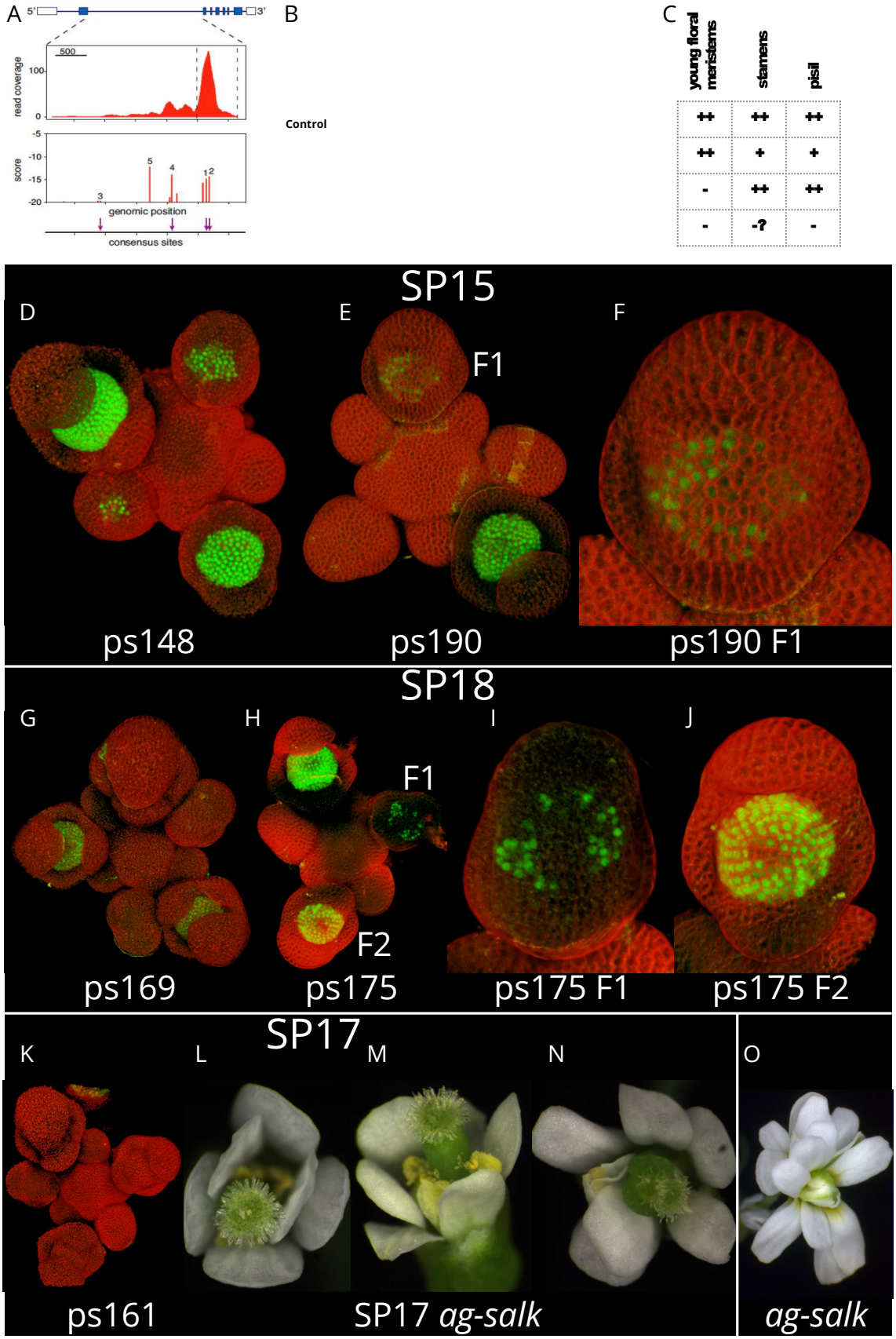


Figure 4.1: Effect of supernumerary copies of *AG* on its own expression : Confocal image of an inflorescence meristem from a line of the genotype *AG/AG AG-2xVenus/AG-2xVenus pAG-AG/+*, with Venus shown in green and autofluorescence shown in red. This WT *Col-0* plant is homozygous for the *AG-2xVenus* reporter and also bears an additional copy of the *AG* locus. No *AG* expression is observed in stage 2 flowers (F1). The onset of *AG* expression occurs very suddenly at stage 3 (F2). A large proportion of AG is visible in the cytoplasm of individual cells, and the centre of the floral dome displays less fluorescence (F3).

Figure 4.2 A.), and confirmed with ChIP-seq (second panel of the Figure 4.2 A.). The fifth one was proposed by Moyroud using SYM-T model [Moyroud et al., 2011] (third panel of Figure 4.2 A.). They can be separated in two groups, the three first ones in the 5' part of the intron that are more important for later expression in stamens and the two last ones in the 3' part that are more important for early activation and for expression in the carpels [Deyholos and Sieburth, 2000]. In *lfy* mutants, *AG* is still expressed, though later in development, and even strong *lfy* mutant flowers still bear carpel-like structures [Weigel and Meyerowitz, 1993]. However we do not know precisely how this late activation impacts on the phenotype. Given the presence of the five LFY binding sites at the *AG* locus (Figure 4.2 A.), I hypothesized that mutating them in different combinations, and within the context of the *AG-2xVenus* reporter, might allow us to observe subtle differences in *AG* expression dynamics and shed light on specific roles of these sites in the regulation of *AG*. We generated three constructs by mutating either one site (box 5; line SP16), three sites (boxes 1, 2 and 4; line SP18), or four sites (boxes 1, 2, 4 and 5; line SP17) (Figure 4.2 B.) and transformed them into WT *Col-0* plants. A control construct without mutations was also generated and transformed into plants, and shows the same pattern of expression as the one described in the previous chapter.

We selected transgenic lines by first examining fluorescence in whole inflorescences and in the different floral organs during development. For SP15, small differences are observed in stamens and carpels. In SP18, *AG* activation is delayed but the final expression pattern remains the same as WT. In SP17, although no AG is detected (Figure 4.2 C, K.), this construct is still able to partially rescue reproductive organ development in a strong *ag-salk* homozygous mutant background (Figure 4.2L, M & N.). A more careful examination of the *AG* pattern during early flower development reveals a small delay in *AG* activation in a few SP15 plants (Figure 4.2 D, E & F.). For SP18, this delay is even more obvious, and a few cells in the centre of the dome are not activated even at the end of stage 3 (Figure 4.2 G, H, I & J.). Nevertheless, these constructs restore a WT phenotype to flowers from the strong *ag-salk* mutant background. We can propose here that the system is robust enough in normal growth conditions to not display a different phenotype but these variations could reduce this robustness and have an impact in other growth conditions.



(Caption next page.)

Figure 4.2: The effect of LFY binding site mutations on *AG* expression dynamics: (A.) LFY binding sites at the *AG* locus. Schematic of the *AG* locus with exons (light blue boxes) and introns (black lines) represented. The middle panel contains ChIP-seq read coverage combined from both strands. The bottom part shows the scores of binding sites and the presence of the consensus binding sites (arrows). **(B.)** Schematic of the *AG* locus showing the LFY binding sites that were mutated in different combinations (red crosses). These fragments were then fused to *2x-Venus* to produce WT (control) or mutant (SP15, SP17 and SP18) versions of the *AG* reporter. **(C.)** Observed fluorescence in floral organs during development. **(D–K.)** 2D projections of confocal image stacks from the different mutant constructs in a WT background. **(D–F.)** Projections from two independent SP15 lines, showing a WT pattern (ps148) or slightly delayed, yet still stochastic, onset (ps190). **(G–J.)** Projections from two SP18 lines, *AG* activation is delayed and less present in few cells in the centre. **(K.)** Projection of an SP17 line with no detectable signal in young flowers. **(L–N.)** Flowers of the same SP17 line in the strong *ag-salk* mutant. In L and M, two pistils are replaced by petals, in N pistils are not present but we observe height petals, two of them present pollen at their extremity. Carpels are slightly different than WT ones. **(O.)** Flower of strong *ag-salk* mutant. Panel A was adapted from Moyroud et al. [Moyroud et al., 2011].

4.2.3 Auto-activation of *AG* is repressed until stage 3 of flower development

AG is able to bind regulatory sequences its own intron through a few identified CarG boxes. These interactions was proposed to maintain its expression once it has been activated [Gómez-Mena et al., 2005]. The models presented in the previous chapters showed that this is a necessary condition for the rapid transition from the low- to the high-expression phase of the *AG* pattern. Our *in vivo* data also showed that cells need to have a minimum concentration of AG to trigger auto-activation. However, this result is based on the observed expression profiles of individual *AG*-expressing cells and the spatial expansion of its activation in the central floral dome. To have a better understanding of this mechanism, I used the *PROTODERMAL FACTOR1* (*PDF1*) promoter to drive the expression of *AG* fused to a *CFP-N7* fluorescent reporter in the L1, where the N7 tag forces AG to localise to the nucleus of expressing cells, in plants that already carried the *AG-2xVenus* reporter. We hypothesised that forcing the accumulation of AG in those cells to see if the endogenous gene would be expressed earlier than in the WT thanks to this auto-activation reaction. I selected plants that did not display an ag mutant phenotype and that expressed the *CFP* tag (Figure

4.3 G.). In these plants, the flowers do not bear petals, although they are not replaced by any other organs either, and sepals present carpelloid characteristics at their extremities (Figure 4.3 D.). I then introduced this line into the *ag-salk* mutant background, in order to understand the importance of AG in the L1 layer for proper floral patterning. Plants homozygous for mutant allele display two different phenotypes, one with similar phenotype to *pPDF1::AG-CFP-N7* plants, except that the carpels are not fully developed and thus the flowers remain sterile (Figure 4.3 C.). The second one display an additional phenotype, with supernumerary pistils and unfused, indeterminate carpels with extra organs produced inside (Figure 4.3 B.). These phenotypes are comparable to artificial microRNA lines against *AG*, which bear sterile flowers with petals developing in the place of pistils, and carpels with an immature stigma (Figure S4.3).

I then selected one line showing both a strong phenotype as well as strong CFP expression and introduced it into the *AG-2xVenus* reporter line. In these plants, I still observe a few cells expressing *AG* at mid-stage 2 and its final pattern in the whole central dome at stage 3 (Figure 4.3 E, F, H & J. & Figure S4.2). Here, I am not able to distinguish differences in AG-2xVenus reporter signal of these plants from the WT.

4.2.4 AG likely forms stable homo-dimers

AG belongs to the MADS box protein family that is known to form dimers and tetramers. AG can directly interact with SEP3 [de Folter et al., 2005]. Studies have been carried out on SEP3 to fully understand this property and show its capacity to form homodimers and homotetramers [Puranik et al., 2014]. The specific composition of complexes that bind CArG boxes in the second intron of *AG* can have an effect on the dynamics of *AG* expression as I showed using the mathematical model. To see if AG is able to form homodimers, people of the group of Richard Immink previously performed FRET-FLIM experiments on AG in protoplasts. This experiment was used to show homo-dimerisation of SEP3 [Immink et al., 2009]. It shows the affinity of this protein to be present in homodimer. This experiment (Figure 4.4) confirms this capacity for AG.

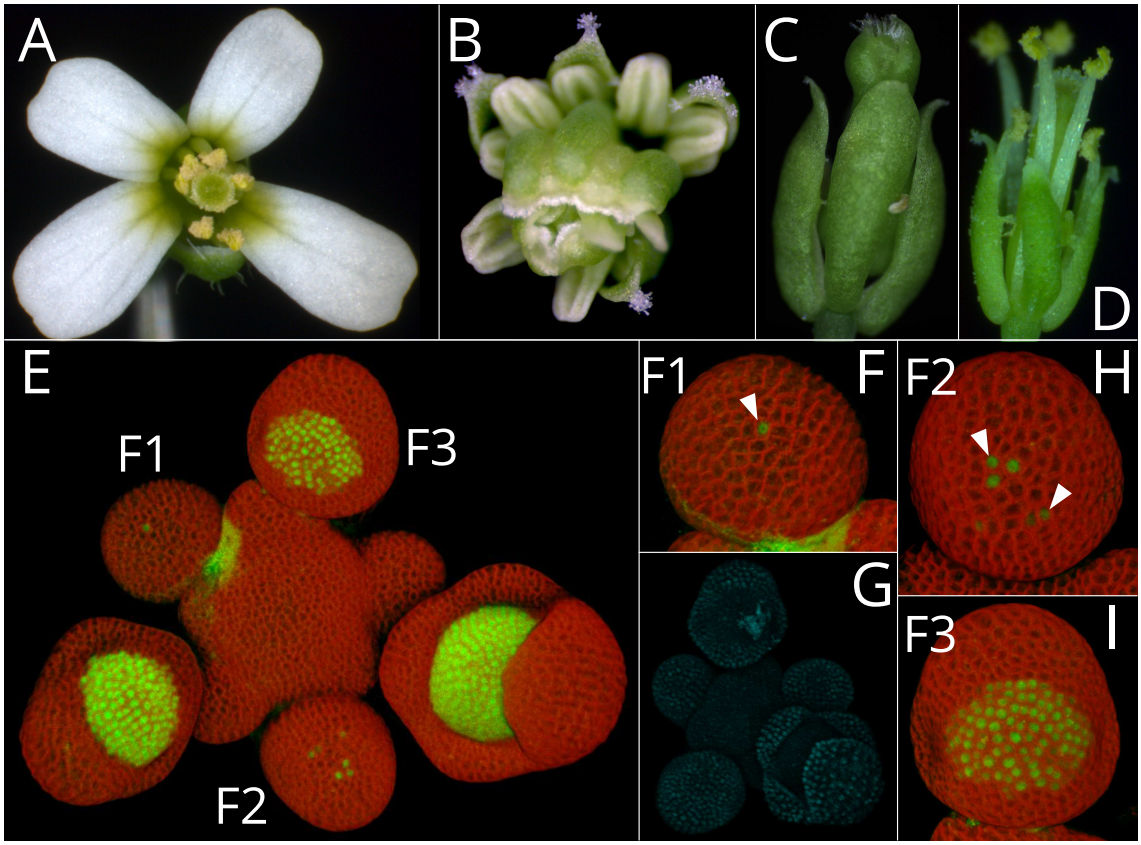


Figure 4.3: The role of auto-activation in *AG* patterning: (A.) WT *Arabidopsis* flower bearing 4 sepals, 4 petals, 6 stamens and 2 carpels. **(B.)** Flowers of *pPDF1::AG-CFP-N7* homozygous *ag-salk* mutants plants, with indeterminate carpels that are not functional petals that are transformed into stamens, and with sepals that present carpelloid structures at their tips. **(C. & D.)** Flowers of *pPDF1::AG-CFP-N7* in a WT *Col-0* background, with missing petals and sepals with carpelloid characteristics. **(E–I.)** Confocal imaging of AG-2xVenus reporter in plants containing *pPDF1::AG-CFP-N7*. **(F.)** In the youngest, mid-stage 2, flower with a visible signal, only one cell expresses *AG* (white arrow). **(H.)** At late stage 2, several cells expressing *AG* are visible. **(I.)** At stage 3, AG is present in almost all cells of the central dome. **(G.)** Confocal image of the CFP marker of the *pPDF1::AG-CFP-N7* line that confirms *AG* expression in the nuclei of the L1 cells of all the flowers and the meristem.

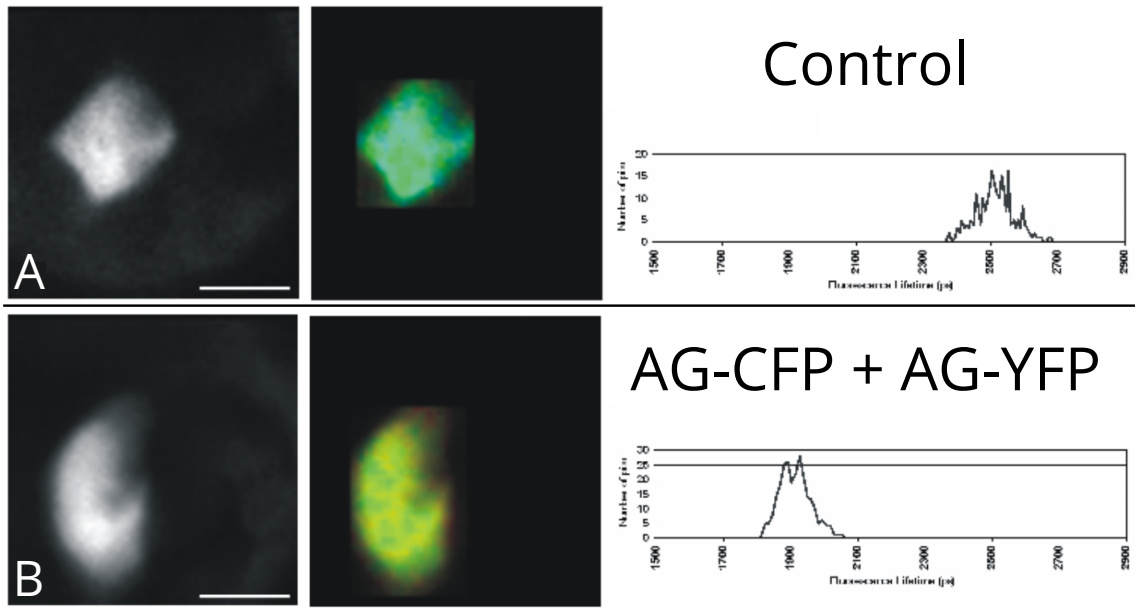


Figure 4.4: AG can form stable homo-dimers: FRET-FLIM experiments in leaf protoplasts from either control PI-CFP **(A)** or AG-CFP ; AG-YFP lines **(B)**. One protoplast is shown for each experiment. The left panel displays the intensity channel, the middle one, the fluorescence lifetime of the same nucleus in a false colour code and the right panel depicts histograms representing the distribution of fluorescence lifetime values.

4.2.5 The role of cell-to-cell movement in establishing the *AG* pattern

AG is known to be able to move from one cell to another [Urbanus et al., 2010]. When *AG* was expressed exclusively in cells of the L1, the protein was also detected in the L2, though not in the other layers. Using the mathematical model, I examined whether a minimal quantity of AG is sufficient to trigger auto-activation and lead to the establishment of the complete pattern. In the previous chapter, we saw that activation of *AG* begins stochastically in a few scattered cells and appears to spread adjacent cells. I hypothesised that it is AG protein that moves from cell to cell and facilitates the accumulation of AG necessary to trigger auto-activation. To test this hypothesis, I produced a version of *AG* fused to a *CFP-N7* tag, which, by forcing the protein into the nucleus, prevent its movement between cells. I sought to determine whether inhibiting this ability to move would generate differences in the *AG* expression pattern and/or result in a floral phenotype. I first tested this in WT *Col-0* flowers, where the *AG-CFP-N7* construct produces the same pattern of expression observed for the *AG-2xVenus* reporter and does not result in any abnormalities (data not

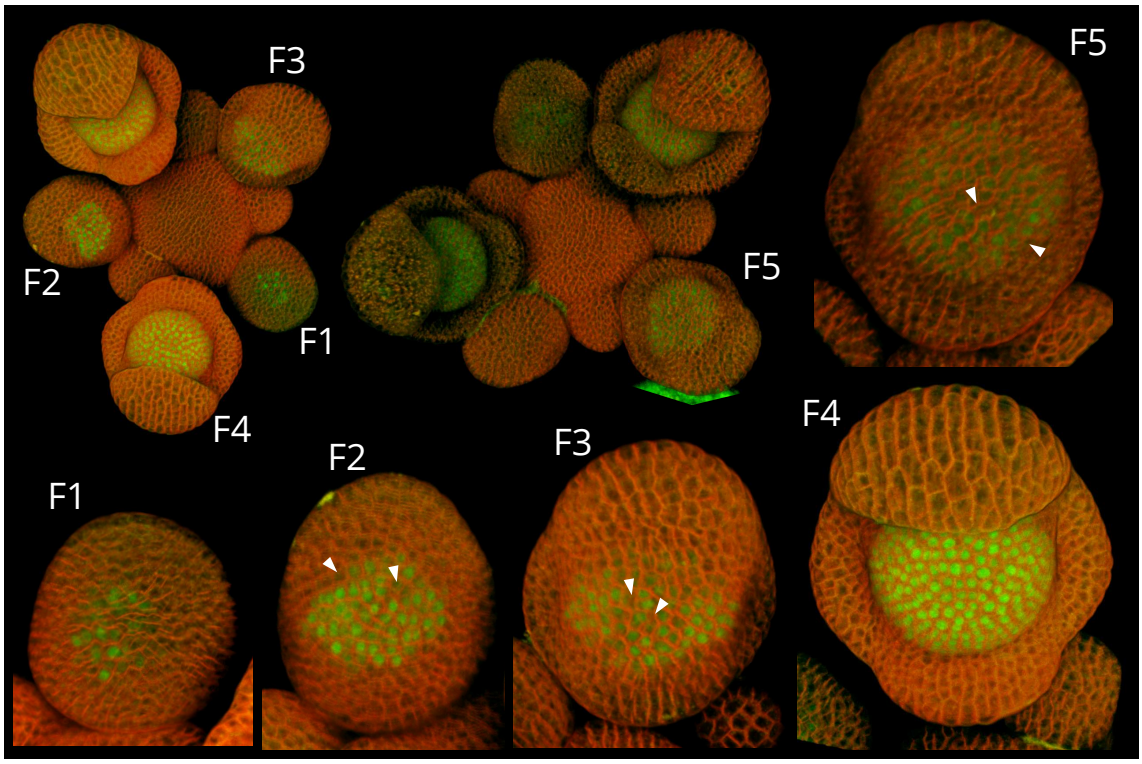


Figure 4.5: The role of movement in *AG* patterning: 2D projections of confocal images of inflorescence meristems carrying the *AG-CFP-N7* marker in the *ag-salk* mutant background. AG is localised in the nucleus and cannot move from cell to cell. As in the WT, the onset of *AG* occurs around mid-stage 2. Unlike in the WT,some cells at early stage 3 do not express *AG* (arrowheads in F5), and the establishment of complete expression in certain cells is delayed to mid-stage 3 (F4).

shown). I then moved this construct into the *ag-salk* mutant background. Preliminary data do not yield any observable phenotype in these plants. However, the pattern of expression of *AG* is different from the WT (Figure 4.5). First, the full expression of *AG* in the central dome is only present at mid-stage 3, instead of at early stage 3, which instead present a low level of expression and a few, randomly distributed cells with no expression. These small differences in expression dynamics do not appear to significantly alter the final pattern. I also observe a few plants with a level of final *AG* expression that is lower than in the WT. However, all these results need to be reproduced and further analyse to be confirmed. In an earlier chapter, I reported that one layer of cells at the periphery of the central dome does not express *AG* at high levels, but instead has a low concentration, likely due to its movement. I am not able to do the comparison with these images because of the low quality of these preliminary images.

4.2.6 AP2 is one of the factors that restrain the activation of *AG* in the early stages of flower development

It has been shown that AP2 plays a role in restricting *AG* expression to the two central whorls [Drews et al., 1991]. This transcription factor is expressed in the entire flower, but is restricted post-transcriptionally by the micro RNA miR172 [Drews et al., 1991, Chen, 2004, Wollmann et al., 2010]. *miR172* is expressed throughout the flower at stage 1 and early stage 2, as well as in the central dome at later stages [Wollmann et al., 2010]. In the mathematical model, I proposed the existence of a threshold for *AG* activation defined, in part, by the concentration of AP2. To gain a clearer understanding of the role of AP2 role in this mechanism, I examined *AG* expression in strong *ap2-7* mutant flowers [Kunst et al., 1989]. This allele carries nonsense mutation that generates a stop codon immediately after the start of transcription (data not shown) and mutant flowers bear no petals and present carpelloid structures at the tips of sepals (Figure 4.6 A.). I crossed the *AG-2xVenus* reporter line into the the *ap2-7* background to determine whether AP2 only restricts *AG* in the two central whorls or if it also required during stage 2. The onset of *AG* expression in *ap2-7* flowers is at mid-stage 2 as in the WT. However, while *AG* is expressed in only a few cells until beginning of stage 3 in the WT (Figure 4.6 D), in mutant flowers of the same stage, it is already present in almost the entire central dome (Figure 4.6 C. & F.). Later in development, *AG* is expressed in a much broader domain, most notably in the second whorl, although it is still excluded from most sepals. The third whorl displays a stronger expression of *AG* than the others, suggesting further support for previously-proposed hypotheses that the cells of the fourth whorl bear different properties. These results can be compared to the simulations from the model. In WT simulations, a delay between the onset of *AG* and the establishment of the final pattern is clearly evident. However, in simulations of the *ap2* null mutant (Figure 4.6 G.), the onset of *AG* expression starts at about the same time as the WT, but the transition to complete activation within the central dome is much more rapid and broader than in the WT.

4.3 Discussion and perspectives

In chapter 3, I show that *AG* expression commences in just a few cells in a stochastic pattern at mid-stage 2 of flower development. In order to gain a better understanding of the

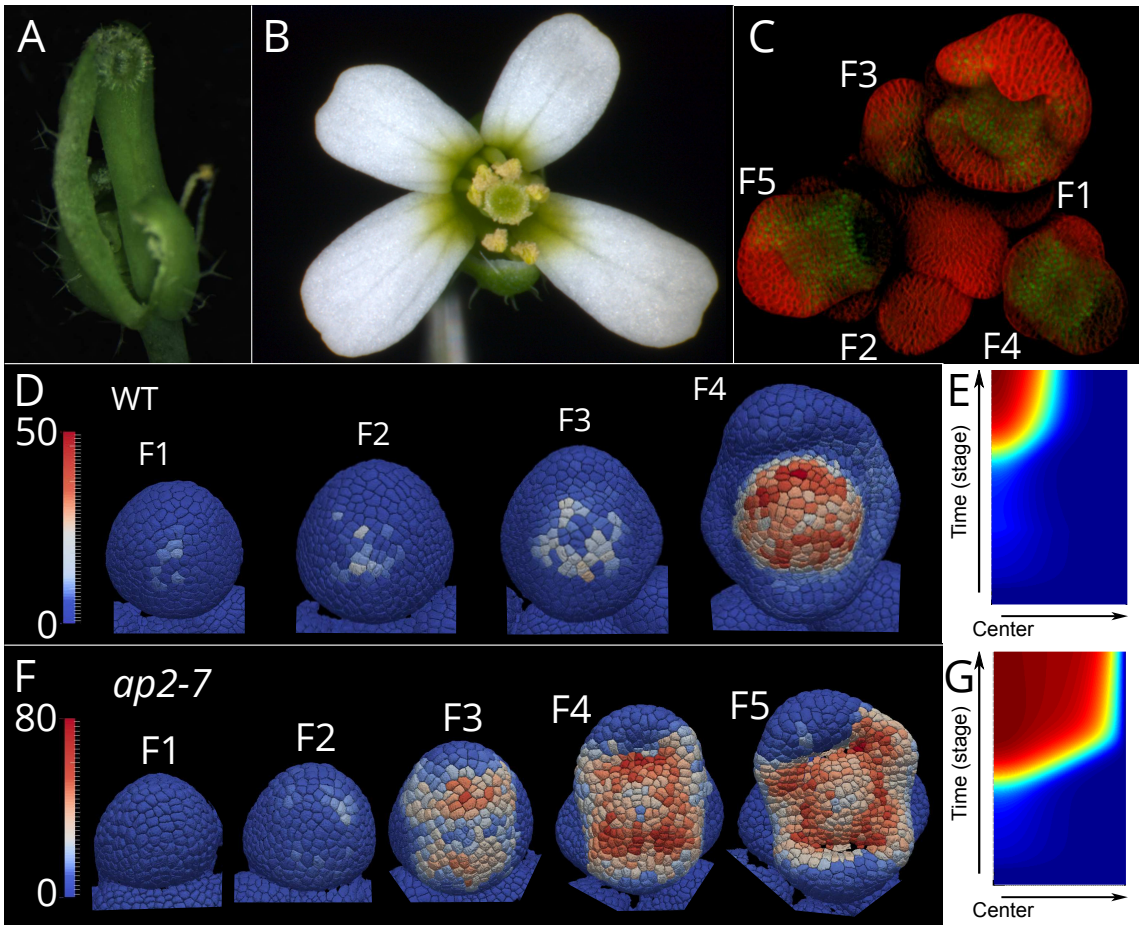


Figure 4.6: The role of AP2 in restricting *AG* expression: (A.) Flower of the *ap2-7* strong mutant. **(B.)** WT flower. **(C.)** 2D projection of a confocal image stack of an *ap2-7* inflorescence meristem carrying an *AG-2xVenus* reporter. **(D.)** MARS-treated 3D reconstructions with quantified AG fluorescence in all L1 cells of the AG-2xVenus reporter in a WT plant. **(E.)** Simulation of AG concentration in the WT. **(F.)** Quantification of the numbered flowers shown in (B). **(G.)** Simulation of *AG* expression in a null *ap2* mutant; *AG* is activated slightly earlier, and more rapidly, than in the WT and becomes more widely expressed in the flower.

importance of this stochastic process in flower development, I first studied the influence of copy number on the floral phenotype. When I add a copy of the entire *AG* genomic region or express *AG* under the *pPDF1* promoter, I obtain some lines that display an *ag* null mutant phenotype, presumably because *AG* is silenced. In lines carrying two additional copies of *AG* (thus of the genotype *AG*; *AG-2xVenus*; *pAG-AG*, and 6 copies in total)), but without the *AG* phenotype, no significant differences in the phenotype were observed. Nevertheless, the pattern of expression during early flower development presents a few differences. For instance in the WT, the vast majority of cellular AG are in the nucleus [Urbanus et al., 2009] but in the extra copy number lines, there is a very strong signal in the cytoplasm, even in cells with low *AG* expression. No satisfactory hypothesis readily explains all these observations.

The onset of *AG* expression is delayed in these plants, and the low-expression, stochastic phase is not observed prior to stage 3. It has been previously proposed that the number of copies could buffer observed noise [Holloway et al., 2011]. Thus in these plants, *AG* activation process is more rapid, without a phase of low accumulation. The sudden and rapid transition from low- to high-expression could be explained by the fact that having more copies of *AG* in the genome could produce slightly higher base levels of AG, which could in turn facilitate the start of *AG* auto-activation. Indeed, the probability of AG complexes to bind a CArG box that positively regulates *AG* transcription is higher, thus lowering the threshold defined by the model. However, this hypothesis would suggest that the low- to high-expression transition of *AG* would occur earlier in those plants than in the WT. A statistical comparison of quantified signal from these plants to the *AG-2xVenus* reporter in the *ag-salk* background, as well as the precise quantification of *AG* mRNA by qPCR in those lines may help clarify the correlation between the number of copies and the onset of *AG* activation. We also need to link these results with epigenetic regulation that can buffer *AG* activation.

AG is activated by LFY and WUS, which bind directly to sequences within the 3-kb second intron of the *AG* locus [Parcy et al., 1998, Lenhard et al., 2001, Lohmann et al., 2001, Moyroud et al., 2011]. The five known binding sites of LFY were shown to have a partially redundant role in regulating *AG* expression, such that the three sites in the 2-kb 5' region and the two in the 800-bp 3' part can act independently to produce qualitatively similar expression patterns [Busch et al., 1999]. However, further analyses showed that the

5' sites more important for later expression in stamens, and the 3' sites for early activation and for expression in carpels [Deyholos and Sieburth, 2000]. It was further shown that one of the 5' sites (site 4) (Figure 4.2 A.) is only involved in activating *AG*, while all LFY binding sites are also necessary for AG maintenance at later developmental stages [Hong and Hamaguchi, 2003]. Mutating the different LFY binding sites in the context of the *AG-2xVenus* reporter allows me to test whether the 5' enhancer region is indeed necessary for early *AG* activation. Mutating only site 5 (line SP15) induces a delay in *AG* activation but also a lower level of expression in stamens and carpels later in development. However, no significant phenotypic differences are visible in these plants. As in the experiments examining the effect of extra copies of the *AG* locus, a small delay in *AG* activation doesn't appear to affect the robust phenotype of the flower. These subtle variations in expression levels with mutant binding site constructs were not observed in previous publications, perhaps due to the fact that detection methods were less sensitive [Busch et al., 1999, Deyholos and Sieburth, 2000, Hong and Hamaguchi, 2003].

The three 3' LFY binding sites (1-3) are also necessary for early *AG* activation [Deyholos and Sieburth, 2000]. Mutating them induces a delay that is stronger in the centre of the *AG* expression domain. However, late activity in stamens and carpels is not modified and there is no variation in floral phenotypes. In the reporter construct with four of the five LFY binding sites mutated (line SP17) no expression was detected in young flowers. In addition, in mature flowers, stamens are converted into petals and carpels display developmental defects, including transformation of stamens in petals. Some of these new petals presents stamenoid structures at their tips. This shows that without those four binding sites, *AG* is expressed later and that this delay has important consequences for flower development. Taken together, my data show that LFY/WUS binding site 3) is not sufficient to activate *AG*. So that activation of *AG* is enhanced by two LFY binding sites and two LFY/WUS ones. Without the two LFY/WUS sites, the *AG* is still activated, albeit with a delay, notably in the center of the flower, near the *WUS* domain. All the sites need to be mutated to have any real effects on flower development. However, these results are preliminary and need to be repeated. More precise quantifications of the *AG* expression levels in flowers would also be helpful in deepening our analyses. Specific mutations of the two LFY/WUS binding sites (sites 1-2) could be carried out to confirm the role of WUS in *AG* activation.

The construct with four mutated sites (SP17) showed no reporter, but was able to res-

cue the mutant phenotype of *ag* flowers, though not completely, suggesting that AG is still present in the flower. This result was observed in three independent lines. Further experiments, such as whole mount *in situ* hybridisation to determine where mRNA is localised and/or AG mRNA quantification by qPCR, need to be performed to confirm these results and understand why it occurs.

Based on the analyses of the model and on the *in vivo* observations of *AG* dynamics, I hypothesised that once AG is able to accumulate in a given cell (due to the activity of LFY and WUS), it is able to rapidly increase its expression levels due to its ability to bind CArG-boxes in its own promoter. However, over-expressing *AG* in the L1 layer with a *CFP-N7* tag did not induce an earlier *AG* activation of the *AG-2xVenus* reporter. Two hypotheses may be put forward to explain this result. Firstly, the presence of factors such as epigenetic modifiers that repress *AG* expression at early stages stage 1 of flower development, such that even the overexpression of *AG* is insufficient to overcome their effects. However, even if this were the case, one would expect the more or less simultaneous activation of expression in all the cells of the central dome, rather than the observed stochastic aspect in a few scattered cells. Another possibility is that *AG* is only weakly expressed, an idea supported by the fact that AG has a non-canonical translation start site and that simple GFP reporters are difficult to image on the confocal. Furthermore, when *AG* expression is driven by the PDF1 promoter, it is also expressed in the meristem. As AG repress *WUS*, if *AG* becomes too strongly expressed, *WUS* may be turned off in the meristem, which would induce a strong seedling phenotype and prevent such transformant lines from being selected. One solution would be to use a different promoter, such as *LFY*, to restrain this overexpression to the flower. To further understand the auto-activation process, the study of the reporter with mutated CArG-boxes would also be important.

The specific composition of the AG MADS complex that binds the CarG-boxes within its own intron can also have an effect on its expression dynamics. The model suggests that this complex must have two molecules of AG in order for the rapid change from low- to high-expression phases observed *in vivo* to occur. Support for this hypothesis comes from comparing AG to another MADS-box protein, SEP3, which is able to form both homodimers and homotetramers [Puranik et al., 2014]. As in the case of SEP3, FRET-FLIM experiments in protoplasts show that AG is able to form stable homodimers. However, this experiment was performed in protoplasts and without the presence of other proteins that could have

a positive or negative effect on MADS complex formation. *in vivo* FRET-FLIM experiments would be ideal, though this is likely to be technically very challenging. To prove the importance of AG homodimers in the MADS complex that binds CArG-boxes in the *AG* second intron, one possible approach might be to generate reporters with mutations in the AG protein that would be important for dimerisation, based on detailed crystal structure and molecular analyses carried out on SEP3 [Puranik et al., 2014].

In the preceding chapters, I also suggest that the movement of AG plays a role in its transition from low- to high-expression in all cells of the central dome at late-stage 2 and early-stage 3. In constructs that restrain AG to the nucleus and block its movement from cell to cell, some cells in the central dome still do not express *AG*, even at the beginning of stage 3. Nevertheless, at the mid-stage 3, *AG* is fully expressed in all cells of the central dome, similarly to the WT. Similarly to SP18, a delay in *AG* activation in lab conditions does not affect the phenotype of the flowers. Another hypothesis mentioned in the previous chapter was that the requirement for AG movement leads to one layer of cells at the periphery of the central dome that does not express *AG* but where AG is able to accumulate due to its movement. I have not been able to test this hypothesis, but with the precise quantification of reporter expression in the plants where movement is blocked, we should be able to gain insights to this question.

AP2 is one of the key transcription factors that represses *AG* during early flower development. It has previously been shown that *AG* expressions expands to most of the cells of the flower in strong *ap2* mutants [Drews et al., 1991]. In this manuscript, I show that in the null *ap2-7* mutant as well, AG is present in a wider domain than in the WT, with expression in the second whorl as well as in the adaxial part of the presumptive sepals. It can be correlated with phenotypes of flowers that don't present petals and that have sepals with carpeloid structures. I also see that the morphology of stage 3 and 4 are different in this mutant than in the WT. Indeed, the two lateral sepal primordia stop their development after stage 3 and will become small filamentous sepals [Kunst et al., 1989]. Meanwhile, the medial (abaxial and adaxial) sepal primordia continue to grow out and sometimes we see a second primordium appear on the abaxial side that may develop into stamen.

The onset of *AG* expression begins at around the same stage, mid-stage 2, in both the WT and in *ap2-7* flowers. However, while only a few cells express *AG* until early stage 3 in the WT, in *ap2-7* flowers the activation of the central dome is much more rapid, which

correlates very well with simulations of our model. In chapter 2, I showed that a threshold delays *AG* activation and that the value of this threshold depend on the concentration of AP2. This mechanism restrains AG activity in cells with low level of AP2. In the absence of AP2, this threshold is lowered, and *AG* is able to be rapidly activated. The simulations of the model yield similar results, with *AG* activated earlier and more strongly. Based on these results, I can propose that AP2 is not only necessary to restrain *AG* in the central dome but also to delay its activation until the beginning of stage 2, although this result needs to be confirmed with the precise pattern of expression of *AP2*. In this mutant, the morphology of the flower is different than in the WT even at stage 3. It is difficult to link this phenotype to the overexpression of *AG*, but further experiments, as specific mutations of selected AP2 binding sites of the second intron of *AG*, could shed light on the importance of *AG* expression during early stages of development.

4.4 Methods

Cloning

The adding of one copy of *AG* in the genome was performed by transforming *pAG::AG-3'OCS* construct in *Col-0* plants that contain *AG-2xVenus* reporter described in the previous chapter with the help of *Agrobacterium tumefaciens*. The clone was obtained by gateway cloning, using 3.7kb of the *AG* promoters amplified with pPD421 and pPD422 primers and recombined in *DONR P4-P1R*, the genomic sequence of *AG* amplified with pPD411 and pPD412 primers and recombined in *DONR221*, and *3'OCS* amplified with pV1 and pV2 primers and recombined in *DONR P2R-P3*. The gateway reaction was done using pH7m34gw empty vector.

SP15, SP17 and SP18 are obtained by direct mutation of the *AG-2xVenus* reporter previously described. They are transformed into *Col-0* plants by François Parcy lab.

pPDF1::AG-CFP-N7 is obtained by gateway cloning in pH7m34gw with the promoter of *PDF1* amplified with pPD471 and pPD472 primers and recombined in *DONR P4-P1R*, the coding sequence of *AG* amplified from *Col-0* cDNA with pSC92 and pSC93 primers and recombined in *DONR221* (the start codon, ACG is modified in ATG) and the *CFP-N7* is obtained by amplification of *CFP*, transfer in BJ36 containing an *N7* tag with BamHI and BgIII and then by transfer of the *CFP-N7* in *E3MCS* with BamHI and SpeI. *pPDF1::amiRAG* is

obtained by gateway recombination in pH7m34gw of the *PDF1* promoter described before, the amiRNA sequence of *AG* from [ÓMaoiléidigh et al., 2013] amplified with pamiRNA-f and pamiRNA-r primers and recombined in *DONR221* and the *3'OCS* described before.

The *AG_T-2xVenus* reporter was obtained by gateway reaction in *dpBART*, a basta resistant gateway vector, with the same vectors than our previous reporter except that the start codon of *AG* is changed from an ACG to an ATG. This reporter is transformed in *ap2-7* strong mutant. This mutant presents a single mutation in the second mutant modifying the TGG in TGA stopping the transcription.

AG-CFP-N7 is obtained using gateway recombination with the promoter of *AG*, the genomic sequence of *AG* starting with ATG and the *CFP-N7* reporter previously used. The recombination was performed in *dpBART*. Plants were first selected in *Col-0* and have the same pattern of expression then the *AG-2xVenus* reporter. Then selected plants were crossed with *ag-salk* to study differences in the homozygous mutant.

Plant growth conditions

All plants were grown on soil at 20°C under long day conditions (16h light/8h darkness). All transgenic plants were selected first *in vitro* on Murashige and Skoog (MS) medium until having homozygous plants.

Confocal imaging

Sample preparation was done like described in [Fernandez et al., 2010] to obtain meristems with only stage 1, 2 and 3 flowers. Membranes were stained with FM4-64 15 minutes before imaging when plants don't have a transgenic membrane marker. Plants of the Figure 4.1 and S4.1 were not stained with FM4-64, then the red component is due to auto-fluorescence. Plants were imaged with the SP8 confocal of Leica. The 2xVenus and the FM4-64 were excited with a laser of 488nm wavelength. The CFP with a laser of 458nm. produce

Image analysis

Projections of shoot apical meristems and of flowers are performed using the 3D viewer plugin of ImageJ.

FRET-FLIM

The Fluorescence Energy Transfer by Fluorescence Lifetime Imaging Microscopy (FRET-FLIM) experiment has been performed by Isabella A. Nougalli Tonaco during her PhD in Richard Immink lab, the results of Figure 4.4 are from her thesis.

4.5 Acknowledgment

I thank R. Immink for giving me the FRET-FLIM data of one of her formed PhD, Isabella A. Nougalli Tonaco and for scientific exchanges about the project. I also thank F. Parcy for the plants with LFY binding site mutations and for advices.

4.6 Author contributions

I did the cloning with the help of A. Chauvet and previous cloning done by P. Das. The mutations of the LFY binding sites of the second intron were performed by the lab of F. Parcy. I cultivate and select plants with the help of A. Chauvet and C. Mormin. I did all the imaging and the analysis. FRET-FLIM experiments were done by I. A. Nougalli Tonaco in the lab of R. Immink and G. Angenent during her PhD. Identification of *ap2-7* mutation was done by C. Mormin under my supervision. I write this chapter with the help of P. Das.

4.7 Supplementary data

produce

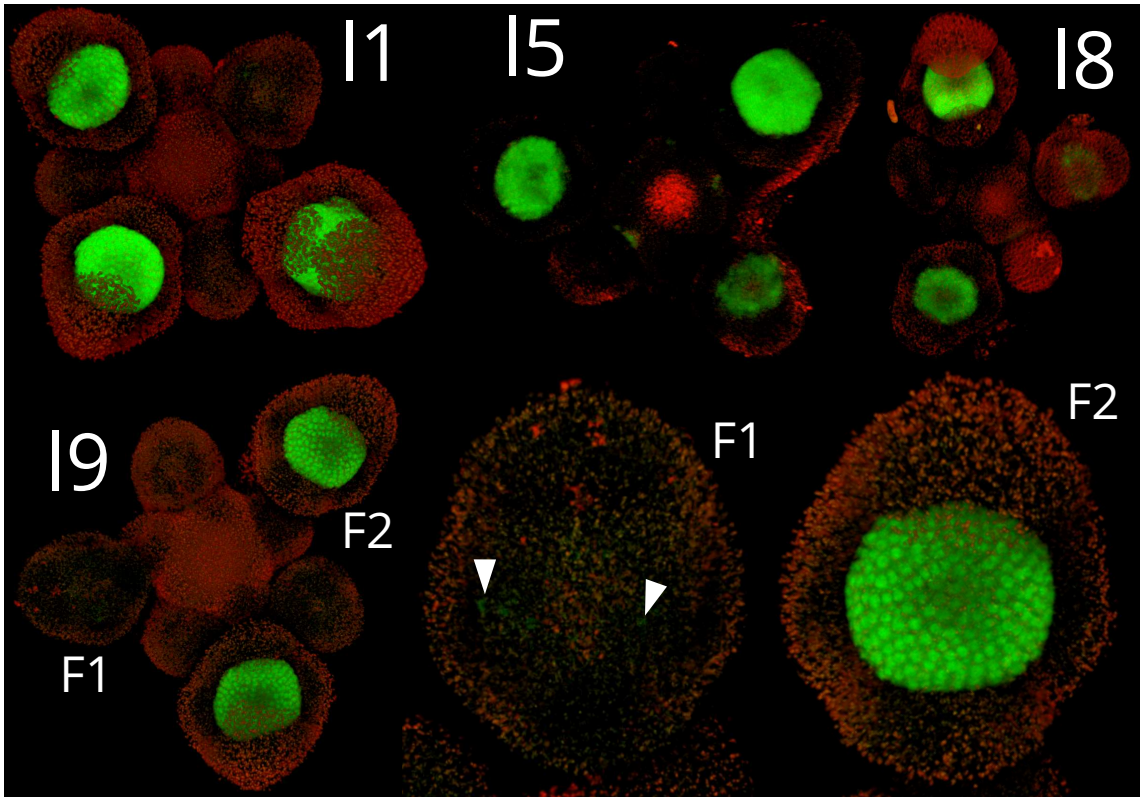


Figure S4.1: *AG* pattern of expression in plants with 6 copy of *AG*: *AG-2xVenus* reporter in four independent lines. AG is first visible at a low level in two cells of inners layer pointed with white arrows at beginning of stage 3.

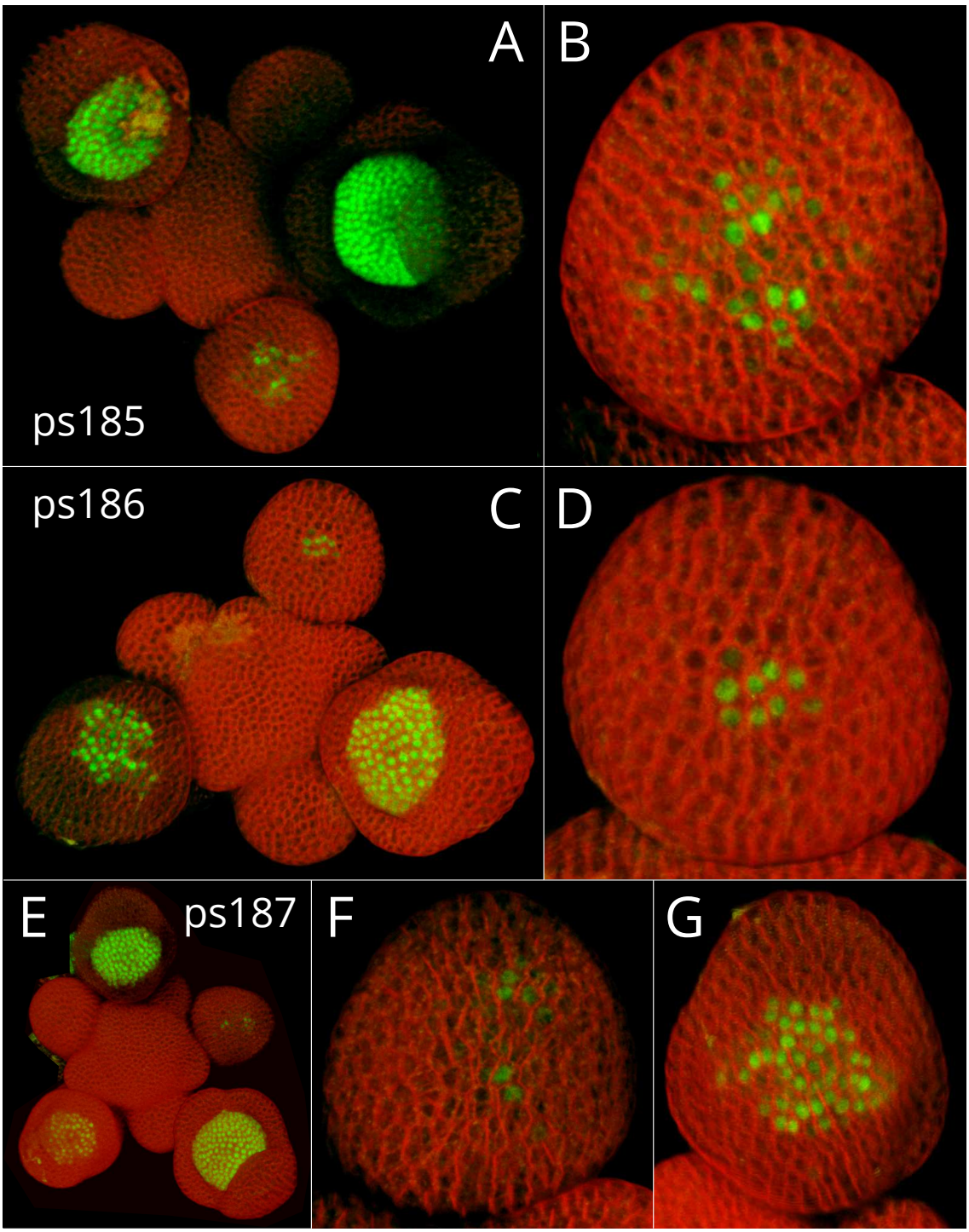


Figure S4.2: Auto-activation of *AG*: Confocal imaging of AG-2xVenus in three meristem containing *pPDF1::AG-CFP-N7*

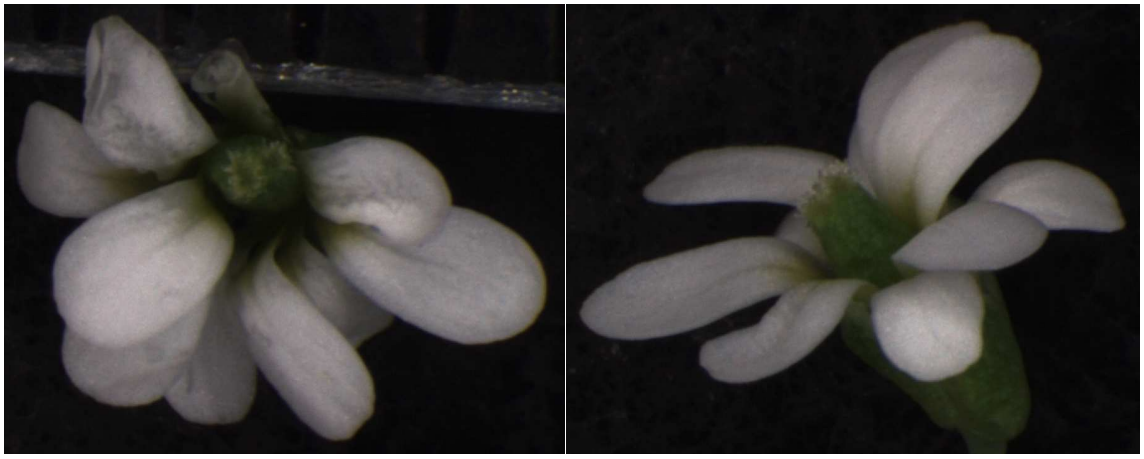


Figure S4.3: Auto-activation of *AG*: Flowers of *Col-0* plants containing *pPDF1::amiRAG*. Petals appear at the place of pistils and carpels are not well developed and not functional.

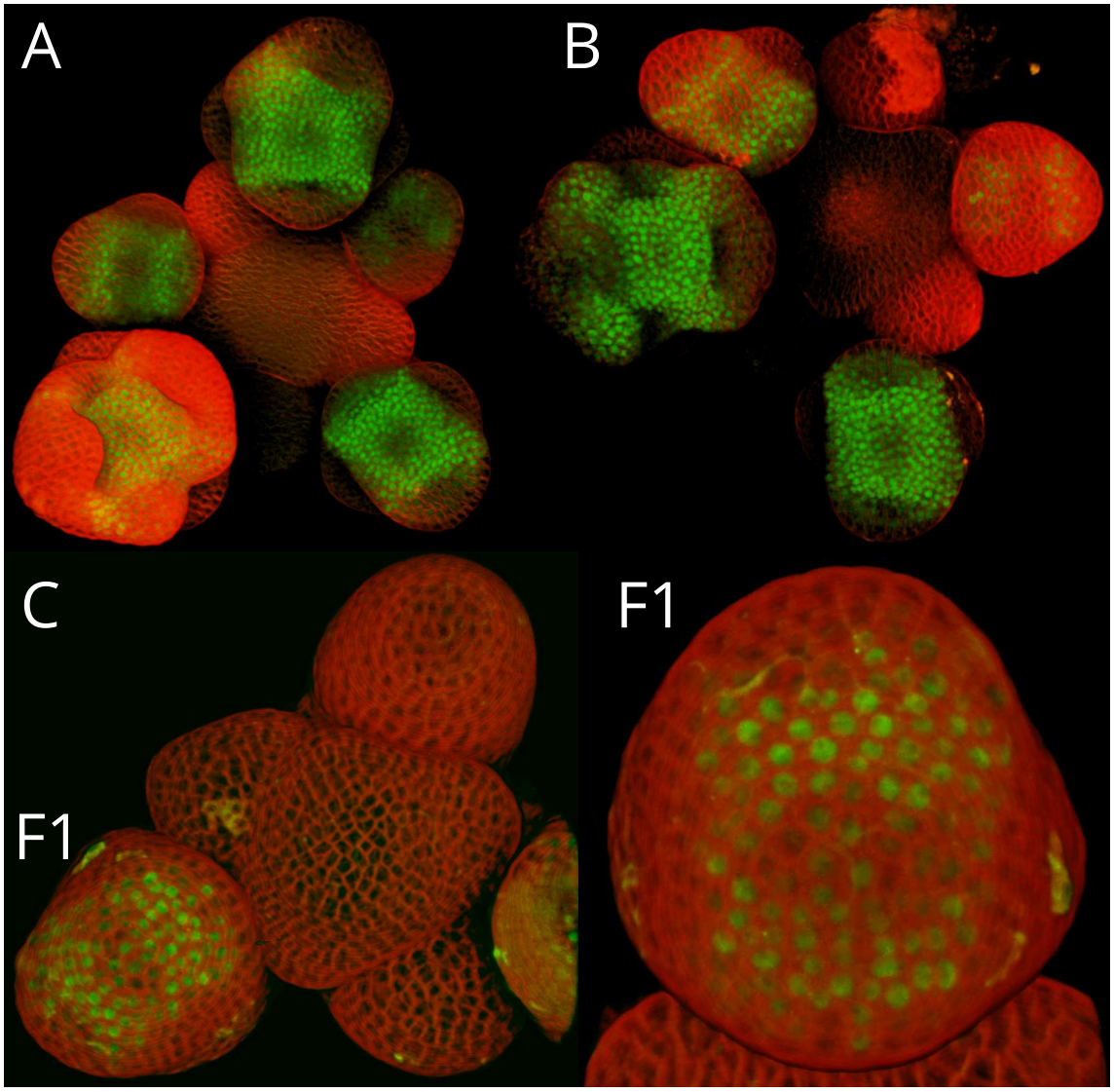


Figure S4.4: Restriction of *AG* expression by AP2: Three projections of meristem of *ap2-7* mutant containing *AG_T-2xVenus* reporter.

Name	Sequence
pamiRNA-f	GGGGACAAGTTTGTACAAAAAAGCAGGCT.GAATTCCTGCAGCCCCAAACAC
pamiRNA-r	GGGGACCACTTTGTACAAGAAAGCTGGGTC.ATCCCCCCATGGCGATGC
pPD411	GGGGACAAGTTTGTACAAAAAAGCAGGCT.CTTTGGAGCAGCAATCACG
pPD412	GGGGACCACTTTGTACAAGAAAGCTGGGTC.TAACTGGAGAGCGGTTTGGT
pPD421	GGGGACAACTTTGTATAGAAAAGTTG.AGTGATCCCTTCTCCAACACA
pPD422	GGGGACTGCTTTTTTGTACAAACTTG.CCAAAAACGTTTAGGGCAAA
pPD471	GGGGACAACTTTGTATAGAAAAGTTG.ATAGCGGAATAGCTGGCAAC
pPD472	GGGGACTGCTTTTTTGTACAAACTTG.GAGAGAAGTTTGTTGCAAATGG
pSC92	GGGGACAAGTTTGTACAAAAAAGCAGGCT.CC.CTTTGGAGCAGCAATCATG
pSC93	GGGGACCACTTTGTACAAGAAAGCTGGGT.C.TAACTGGAGAGCGGTTTGGT
pV1	GGGGACAGCTTTCTTGTACAAAGTGGCTACCATGGGTCCTGCTTTAATGAGAT
pV2	GGGGACAACTTTGTATAATAAAGTTGCTAGTAAGCTAGCTTGCATGCCGGT

Table S4.1: Used primers.

Chapter 5

General discussion and perspectives

Improvement of the mathematical model to answer new questions

The objective of this thesis was to research the combination of mechanisms required to generate the gene expression patterns within a network of transcription factors, with respect not only on the final pattern, but also to the temporal dynamics of these components. I focussed on the ABC model in *Arabidopsis thaliana* because of the large amount of data already available and the existence of tools that permit a precise analysis of these dynamics. The previous models of the ABC systems were useful to understand the existence and the stability of the four whorls but cannot help us to decipher its dynamics [Mendoza and Alvarez-Buylla, 1998]. To study this dynamic, we decided to use reaction-diffusion systems to model the main interactions. As the B-class genes don't have a strong effect on the regulation of A and C-class genes, we focus in this thesis on the AC model. This model qualitatively reproduces the dynamics of AG expression during the early stages of flower development. The model is a good tool to analyse the mechanisms behind AG activation and to propose hypotheses that may be experimentally tested. To go further with this approach, these data may then be incorporated into the model, to answer to new questions.

I focussed my analyses on the dynamics of AG activation until the end of stage 3 of flower development, when a stable AG expression pattern has been previously described. However, this model can also propose hypotheses for AG expression at later stages of development, once the appropriate adjustments are made.

The equations describing our system are solved in a deterministic manner, meaning that the results of our simulations depend entirely on the parameters of the model and the initial

conditions, and does not take into account randomness. This provides a model that can be mathematically analysed and used to propose hypotheses, such as the existence of an auto-activation threshold for *AG*, whose value is defined by AP2. It is also a simple method to find parameters that correspond to our experiments. Nevertheless, as I show in chapter 3, the process of *AG* activation appears to begin stochastically. My deterministic models are insufficient to truly understand which mechanisms are behind this apparent stochasticity. For this reason, we have recently begun a collaboration with the Kimmel lab [Bertolusso et al., 2014] to add randomness to our equations using Gillespie algorithms. This should provide a greater understanding about the possible origins of the observed stochasticity, as well as about the role of this stochasticity in producing the robust final pattern of expression of *AG*.

The employed model uses a continuous segment as a representation of the L1 cell layer. This simplification facilitates the analyses and simulations of the model, and is justifiable with respect to the hypotheses we want to explore. Nevertheless, it is important to consider that the L1 is composed of many tens to hundreds of cells (between 50 and 300 through the first three stages of development). Creating a discrete model with individual cells enormously complexifies the simulations. In addition, at stage 3, the sepals begin to grow and the geometry of the flower is more complicated – incorporating the effects of such an altered 3D landscape into the model would also be a challenge. With the use of time-course images with short intervals between each time point we are able to precisely describe the growth of the flower during the first few stages of its development. One solution to developing realistic cell-based simulations would be to use these data and adapt our models to the platform developed by the Jönsson lab [Yadav et al., 2011, Gruel et al., 2016]. While I did attempt to use this platform to simulate the model on a growing virtual tissue, I was unfortunately not able to propose any new hypotheses with this approach. However, we can now compare the results of our model to real tissular patterns of expression of AG and perform parameter optimisation steps to analyse the importance of particular parameters in the robustness of the pattern.

My model was developed to study the appearance of AG in the central dome of the stage 3 flower, as well as the disappearance of AP1 in the same region. Using the same model, but with the addition of other proteins and interactions, we could ask the same questions about AP3 and PI. These TF appear concomitantly in the second and the third whorls of

the flower. As of today, the mechanisms behind this activation, in particular the observed spatial pattern, is not completely understood [O'Maoileidigh et al., 2014] and a similar model to ours could help shed light on these mechanisms.

Flowers of different plant species have a similar architecture. The boolean models that were first developed for *Arabidopsis* have also been applied to other species and are an interesting tool to understanding the evolution of flower development [Chaos et al., 2006]. Nevertheless, those tools are not spatial. Creation of whorls in development is common and can be conserved between species like in *drosophila* [Giorgianni and Mann, 2011]. Our model was an example of that is apply to a specific system, nevertheless, a generalisation of this model could be used to help to understand general mechanisms behind this typical pattern formation.

In the chapter 2, I proposed few hypotheses based on mathematical analysis of the model and on simulations linked to questions we were asking about AG regulation. These hypotheses can be resumed by few sentences:

- Delay in AG strong expression is due to the existence of a threshold.
- This threshold and the fast expression of AG can be explained by the composition of the complex of auto-activation. Two proteins of AG in this complex are necessary.
- AP2 not only restrains AG in the central dome but also reduces its expression before stage 3 of flower development.
- AG movement is necessary to have a full expression of AG in all the central dome.

Then, I analyse precisely the pattern of AG expression to compare to our simulations and in a second time, I performed new experiments based on these hypotheses to prove them biologically.

Analysis of *AG* pattern of expression and future directions

In this thesis, I have proposed a new method to quantify and analyse the expression of a fluorescent protein in the flower through a time course with small time steps based on work already done in our lab [Fernandez et al., 2010]. We obtained the profile of expression of AG in each cell in the growing flower and performed analyses such as the creation of heat maps or defining the activation dynamics within each cell. This precise analysis allows us to have

a better understanding of *AG* activation process. We can resume the results obtained with this analysis:

- *AG* expression is stochastic during early flower development.
- The stochasticity of *AG* expression is defined by variability in onset of this expression in each cell.
- Profile of *AG* expression with a rapid activation correspond to the existence of a threshold produced by the presence of at least two proteins of AG in the complex involved in the auto-activation mechanism.
- One ring of cells at the periphery of the central dome shows a lower level of AG probably due to AG movement.
- *AG* activation spreads in adjacent cells of the central dome.
- Few cells at the center of the domain of expression show a lower level of *AG* expression.

This method of analysis uses the temporal graph developed by J. Legrand during his Ph.D. (Unpublished). In his Ph.D., he also developed tools to correlate the morphometry of cells with gene expression patterns using clustering tools. A greater exploitation of these tools could be used to further analyse our short-interval time-course experiments. The correlation between the ABC patterns and cellular morphometries in mutants like *ap2-7*, which show morphological differences in early development, could help to understand the role of these genes in the production of organs. This could potentially reveal a new role for these TF, which are principally known to be involved in floral organ identity.

The analysis of our time course experiments reveals that the activation of *AG* is rapid, within around 4-6 hours, though our time course experiments were carried out with a step time of 4 hours, so it was not possible to have an extremely precise description of the expression profile within each cell. We can potentially reduce this step to one or two hours and perform a more precise study of the few hours surrounding *AG* activation. However, this is hard work and the automation of processes such as lineage tracing is necessary to quicken the process. Since variations of cell morphology should be very small with short time intervals, it should be easier to develop this automation.

It is evident that the classically used stages of development, which depend on morphological differences as the flower grows [Smyth et al., 1990], are insufficient to address the

data we are now able to generate to analyse TF pattern of expression. For example, stage 2 corresponds to more than 2 days of growth, and detailing this stage could be useful in providing a clear description of our observations. To address this issue, I attempted to compare different floral meristems by measuring the size of the flowers and correlating these sizes with the number of cells expressing *AG* or the mean level of expression of these cells. However, I was unable to draw any conclusive inferences in this regard, perhaps because plants were not always grown in the same conditions, and thus displayed differences in *AG* expression dynamics. Nevertheless, being able to compare meristems with different conditions would be a powerful tool and would maybe permit us to discover results we are not able to detect today. We also observe that the size of the flower can influence the results of our simulations. Thus the size of the flower could be a parameter that plays a role in when *AG* is activated. In the future, it would be interesting to modify the growth of flowers by altering growth conditions or by using growth mutants, and analysing the role of growth on gene expression. These experiments can lead to new discoveries linking gene expression to the size of the tissue but also to see the impact of environmental factors on patterning of the ABC genes.

To complete the experiments performed in the chapter 4

In the two previous subsections, I resumed the hypotheses proposed by our model and the precise analysis of *AG* dynamics of expression. In chapter 4, we test some of these and obtain preliminary results that confirm our hypotheses. However, additional experiments are necessary to validate our conclusions and complete the project.

The first question concerns the origin of the stochastic patterns of *AG* expression at stage 2 and its impact on robustness. To study the impact of this stochasticity that we previously described, I decided to analysis flowers with the addition of *AG* copies in the genome to buffer the stochasticity as it was proposed in Holloway [Holloway et al., 2011]. With this experience, AG signal was present mainly in the cytoplasm and the onset of the activation was slightly delayed and activation was faster with less stochasticity. Nevertheless, this delay doesn't have visible effects on the phenotype. However, clearer quantifications will be needed to confirm that the number of copies of *AG* influences the duration of the activation phase. An analysis of the dynamics of *AG* expression compared to the number of *AG* copies could be interesting in understanding this better. However, the delay in *AG* expression observed in this experiment cannot be easily explained. One way to address this question could be

through the use of stochastic models. By mutating the binding sites of the second intron of *AG* we could also detect which interactions are at the origin of this stochasticity.

By the study of few combination of LFY and WUS binding site of the second intron of *AG*, we produce a delay in *AG* expression and a final pattern that can differ to the WT. But only mutations that lead to strong delays in *AG* expression give a different phenotype. We could further mutate other combinations of LFY binding sites to gain a full understanding of these interactions. It would be interesting to see the precise effect of WUS in these interactions by inducing the inhibition of *WUS* after the flower is initiated.

AG is able to maintain its own activity by directly binding its own promoter [Gómez-Mena et al., 2005]. Our model suggests the existence of a threshold of minimum AG concentration that allows the auto-activation process to occur. However, overexpression of *AG* (for instance under the promoter of *PDF1*), doesn't yield any differences in the pattern of *AG* expression. Using a strong promoter to drive *AG* expression can produce plants with an *ag* phenotype, likely due to silencing. Strong expression of *AG* in the meristem can also repress WUS activity and lead to plants without meristem. These remarks make us think that the lines used in this experiment was weak. So another way to have a strong overexpression of *AG* before stage 2 would to use an inducible system that strongly overexpresses *AG* only in the flowers. We are also producing constructs of our reporter with mutations of CarG boxes. These constructs are currently in the process of selection in the *ag-salk* mutant background and could give us some insights about the importance of the autoregulatory interactions in the maintenance of AG expression.

With the model, I proposed that the composition of the MADS complex that bind *AG* CArG boxes is composed of at least two molecules of AG. We show in unpublished data of Immink lab that AG is able to form homodimers in protoplasts cells. This result need to be confirmed in *in vivo* flowers. Furthermore, the precise role of this homodimerization should be studied by mutating the dimerization sites proposed by comparisons with SEP3 [Puranik et al., 2014].

AG is known to be able to move between cells of the flower. To study its role, I looked at the pattern of expression of the *AG-CFP-N7* reporter in a strong homozygous *ag* mutant. This construction force the protein to stay in the nucleus and then are not able to move from one cell to the other. In these flowers, few cells were still not expressing *AG* at beginning of stage 3 when they are all activated in the WT. Nevertheless, we cannot distinguish any

differences in the phenotype. To confirm this preliminary result, precise quantification of this line is necessary.

AP2 was previously known to restrain *AG* activation in the two inner whorls. I showed by studying *AG* expression pattern in the strong *ap2-7* mutant that AP2 is also able to produce a threshold that delay *AG* expression during end of stage 2. These results would be more convincing if the precise pattern of expression were known. While a qualitative description exists [Wollmann et al., 2010], expression in young flowers is not precisely described. The expression of *AP2* at stage 2 repressing *AG* could be one explanation of the result we obtain with our experiment with overexpression of *AG* under *pPDF1*. We also see that in strong *ap2* mutants, we have different morphologies of flowers at stage 3. These mutants could be a model to study the influence of cellular and tissular morphometries on the proper emergence of the ABC patterns.

Bibliography

[Albert and Othmer, 2003] Albert, R. and Othmer, H. G. (2003). The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster. *Journal of Theoretical Biology*, 223(1):1–18.

[Aldana, 2003] Aldana, M. (2003). Boolean dynamics of networks with scale-free topology. *Physica D: Nonlinear Phenomena*, 185(1):45–66.

[Alvarez-Buylla et al., 2010] Alvarez-Buylla, E. R., Azpeitia, E., Barrio, R. Á., Benítez, M., and Padilla-Longoria, P. (2010). From ABC genes to regulatory networks, epigenetic landscapes and flower morphogenesis: making biological sense of theoretical approaches. *Seminars in Cell & Developmental Biology*, 21(1):108–17.

[Alvarez-Buylla et al., 2008] Alvarez-Buylla, E. R., Chaos, Á., Aldana, M., Benítez, M., Cortes-Poza, Y., Espinosa-Soto, C., Hartasánchez, D. a., Lotto, R. B., Malkin, D., Escalera Santos, G. J., and Padilla-Longoria, P. (2008). Floral morphogenesis: stochastic explorations of a gene network epigenetic landscape. *PLoS One*, 3(11):e3626.

[Asthagiri and Lauffenburger, 2000] Asthagiri, A. and Lauffenburger, D. (2000). Bioengineering Models of Cell Signaling. *Annu. Rev. Biomed. Eng*, pages 31–53.

[Balleza et al., 2008] Balleza, E., Alvarez-Buylla, E. R., Chaos, Á., Kauffman, S., Shmulevich, I., and Aldana, M. (2008). Critical dynamics in genetic regulatory networks: examples from four kingdoms. *PLoS One*, 3(6):e2456.

[Barrio et al., 2010] Barrio, R. Á., Hernández-Machado, A., Varea, C., Romero-Arias, J. R., and Alvarez-Buylla, E. R. (2010). Flower development as an interplay between dynamical physical fields and genetic networks. *PLoS One*, 5(10):e13523.

[Barton, 2010] Barton, M. K. (2010). Twenty years on: The inner workings of the shoot apical meristem, a developmental dynamo. *Developmental Biology*, 341(1):95–113.

[Benítez et al., 2008] Benítez, M., Espinosa-Soto, C., Padilla-Longoria, P., and Alvarez-Buylla, E. R. (2008). Interlinked nonlinear subnetworks underlie the formation of robust cellular patterns in Arabidopsis epidermis: a dynamic spatial model. *BMC Systems Biology*, 2:98.

[Bertolusso et al., 2014] Bertolusso, R., Tian, B., Zhao, Y., Vergara, L., Sabree, A., Iwanaszko, M., Lipniacki, T., Brasier, A. R., and Kimmel, M. (2014). Dynamic cross talk model of the epithelial innate immune response to double-stranded RNA stimulation: Coordinated dynamics emerging from cell-level noise. *PLoS One*, 9(4).

[Blázquez et al., 2006] Blázquez, M. a., Ferrándiz, C., Madueño, F., and Parcy, F. (2006). How floral meristems are built. *Plant molecular biology*, 60(6):855–870.

[Bowman et al., 1991] Bowman, J. L., Drews, G. N., and Meyerowitz, E. M. (1991). Expression of the Arabidopsis Floral Homeotic Gene Agamous Is Restricted to Specific Cell Types Late in Flower Development. *The Plant cell*, 3:749–758.

[Bowman et al., 1989] Bowman, J. L., Smyth, D. R., and Meyerowitz, E. M. (1989). Genes directing flower development in Arabidopsis. *The Plant cell*, 1(1):37–52.

[Busch et al., 1999] Busch, M. A., Bomblies, K., and Weigel, D. (1999). Activation of a Floral Homeotic Gene in Arabidopsis. *Science*, 285(5427):585–587.

[Carles and Fletcher, 2009] Carles, C. C. and Fletcher, J. C. (2009). The SAND domain protein ULTRAPETALA1 acts as a trithorax group factor to regulate cell fate in plants service The SAND domain protein ULTRAPETALA1 acts as a trithorax group factor to regulate cell fate in plants. (510):2723–2728.

[Carles and Fletcher, 2010] Carles, C. C. and Fletcher, J. C. (2010). Missing links between histones and RNA Pol II arising from SAND? *Epigenetics*, 5(5):381–385.

[Carles et al., 2004] Carles, C. C., Lertpiriyapong, K., Reville, K., and Fletcher, J. C. (2004). The ULTRAPETALA1 gene functions early in Arabidopsis development to restrict shoot apical meristem activity and acts through WUSCHEL to regulate floral meristem determinacy. *Genetics*, 167(4):1893–1903.

[Castillejo et al., 2005] Castillejo, C., Romera-Branchat, M., and Pelaz, S. (2005). A new role of the Arabidopsis SEPALLATA3 gene revealed by its constitutive expression. *The Plant Journal*, 43(4):586–596.

[Causier et al., 2009] Causier, B., Bradley, D., Cook, H., and Davies, B. (2009). Conserved intragenic elements were critical for the evolution of the floral C-function. *The Plant Journal*, 58(1):41–52.

[Chaos et al., 2006] Chaos, Á., Aldana, M., Espinosa-Soto, C., León, B. G. P., Arroyo, A. G., and Alvarez-Buylla, E. R. (2006). From Genes to Flower Patterns and Evolution: Dynamic Models of Gene Regulatory Networks. *Journal of Plant Growth Regulation*, 25(4):278–289.

[Chen, 2004] Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science*, 303(5666):2022–5.

[Chisholm et al., 2010] Chisholm, R. H., Hughes, B. D., and Landman, K. a. (2010). Building a morphogen gradient without diffusion in a growing tissue. *PLoS One*, 5(9).

[Christenhusz and Byng, 2016] Christenhusz, M. J. and Byng, J. W. (2016). The number of known plants species in the world and its annual increase. *Phytotaxa*, 261(3):201.

[Coen and Meyerowitz, 1991] Coen, E. and Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature*, 353(6339):31–37.

[Collaudin and Mirabet, 2014] Collaudin, S. and Mirabet, V. (2014). Models to reconcile plant science and stochasticity. *Frontiers in Plant Science*, 5(November):1–4.

[Cox et al., 1984] Cox, K. H., DeLeon, D. V., Angerer, L. M., and Angerer, R. C. (1984). Detection of mRNAs in sea urchin embryos by in situ hybridization using asymmetric RNA probes. *Developmental Biology*, 101(2):485–502.

[Crampin et al., 1999] Crampin, E. J., Gaffney, E. a., and Maini, P. K. (1999). Reaction and diffusion on growing domains: scenarios for robust pattern formation. *Bulletin of Mathematical Biology*, 61(6):1093–120.

[Cranfill et al., 2016] Cranfill, P. J., Sell, B. R., Baird, M. A., Allen, J. R., Lavagnino, Z., de Gruiter, H. M., Kremers, G.-J., Davidson, M. W., Ustione, A., and Piston, D. W. (2016). Quantitative assessment of fluorescent proteins. *Nature Methods*, (May):1–7.

[Daum et al., 2014] Daum, G., Medzihradszky, A., Suzaki, T., and Lohmann, J. U. (2014). A mechanistic framework for noncell autonomous stem cell induction in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 2014(19).

[Davila-Velderrain et al., 2015] Davila-Velderrain, J., Villarreal, C., and Alvarez-Buylla, E. R. (2015). Reshaping the epigenetic landscape during early flower development: induction of attractor transitions by relative differences in gene decay rates. *BMC Systems Biology*, 9(20):1–14.

[de Folter et al., 2005] de Folter, S., Immink, R. G. H., Kieffer, M., Parenicová, L., Henz, S. R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M. M., Davies, B., and Angenent, G. C. (2005). Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. *The Plant cell*, 17(5):1424–1433.

[de Jong, 2002] de Jong, H. (2002). Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *Journal of Computational Biology*, 9(1):67–103.

[Deyholos and Sieburth, 2000] Deyholos, M. K. and Sieburth, L. E. (2000). Separable whorl-specific expression and negative regulation by enhancer elements within the AGAMOUS second intron. *The Plant Cell*, 12(10):1799–1810.

[Dinh et al., 2012] Dinh, T. T., Girke, T., Liu, X., Yant, L., Schmid, M., and Chen, X. (2012). The floral homeotic protein APETALA2 recognizes and acts through an AT-rich sequence element. *Development*, 139(11):1978–86.

[Ditta et al., 2004] Ditta, G. S. 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. *Current Biology*, 14:1935–1940.

[Drews et al., 1991] Drews, G. N., Bowman, J. L., and Meyerowitz, E. M. (1991). Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the APETALA2 product. *Cell*, 65(6):991–1002.

[Editors, 2000] Editors (2000). Can biological phenomena be understood by humans? *Nature*, 403(6768):345.

[Engelhorn et al., 2014] 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. *Annals of botany*, 1:1497–1505.

[Espinosa-Soto et al., 2004] Espinosa-Soto, C., Padilla-Longoria, P., and Alvarez-Buylla, E. R. (2004). A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. *The Plant Cell*, 16(11):2923–2939.

[Faure et al., 2016] Faure, E., Savy, T., Rizzi, B., Melani, C., Stašová, O., Fabrèges, D., Špir, R., Hammons, M., Čúnderlík, R., Recher, G., Lombardot, B., Duloquin, L., Colin, I., Kollár, J., Desnoulez, S., Affaticati, P., Maury, B., Boyreau, A., Nief, J.-Y., Calvat, P., Vernier, P., Frain, M., Lutfalla, G., Kergosien, Y., Suret, P., Remešíková, M., Doursat, R., Sarti, A., Mikula, K., Peyriéras, N., and Bourgine, P. (2016). A workflow to process 3D+time microscopy images of developing organisms and reconstruct their cell lineage. *Nature Communications*, 7(in press):8674.

[Félix and Barkoulas, 2015] Félix, M.-A. and Barkoulas, M. (2015). Pervasive robustness in biological systems. *Nature Reviews. Genetics*, 16(8):483–496.

[Fernandez et al., 2010] Fernandez, R., Das, P., Mirabet, V., Moscardi, E., Traas, J., Verdeil, J.-L., Malandain, G., and Godin, C. (2010). Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. *Nature Methods*, 7(7):547–553.

[Ferrándiz et al., 1999] Ferrándiz, C., Pelaz, S., and Yanofsky, M. F. (1999). Control of Carpel and Fruit Development in Arabidopsis. *Annu. Rev. Biochem.*, pages 321–354.

[Fletcher, 2001] Fletcher, J. C. (2001). ULT function in Arabidopsis meristems. 1333:1323–1333.

[Friis et al., 2010] Friis, E. M., Pedersen, K. R., and Crane, P. R. (2010). Diversity in obscurity: fossil flowers and the early history of angiosperms. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 365(1539):369–82.

[Gierer and Meinhardt, 1972] Gierer, A. and Meinhardt, H. (1972). A theory of biological pattern foramtion.

[Giorgianni and Mann, 2011] Giorgianni, M. W. and Mann, R. S. (2011). Establishment of medial fates along the proximodistal axis of the Drosophila leg through direct activation of dachshund by Distalless. *Developmental Cell*, 20(4):455–68.

[Giurumescu et al., 2006] Giurumescu, C. A., Sternberg, P. W., and Asthagiri, A. R. (2006). Intercellular coupling amplifies fate segregation during Caenorhabditis elegans vulval development. *Proceedings of the National Academy of Sciences of the United States of America*, 103(5):1331–6.

[Gómez-Mena et al., 2005] Gómez-Mena, C., de Folter, S., Costa, M. M. R., Angenent, G. C., and Sablowski, R. (2005). Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. *Development*, 132(3):429–438.

[Gordon and Beloussov, 2006] Gordon, R. and Beloussov, L. (2006). From observations to paradigms; the importance of theories and models: An interview with Hans Meinhardt. *The International Journal of Developmental Biology*, 50(2-3):103–111.

[Gruel et al., 2016] Gruel, J., Landrein, B., Tarr, P., Schuster, C., Refahi, Y., Sampathkumar, A., Hamant, O., Meyerowitz, E. M., and Jo nsson, H. (2016). An epidermis-driven mechanism positions and scales stem cell niches in plants. *Science Advances*, 2(1):e1500989–e1500989.

[Gustafson-Brown et al., 1994] Gustafson-Brown, C., Savidge, B., and Yanofsky, M. F. (1994). Regulation of the arabidopsis floral homeotic gene APETALA1. *Cell*, 76(1):131–143.

[Hanna et al., 2009] Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C. J., Creyghton, M. P., van Oudenaarden, A., and Jaenisch, R. (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature*, 462(7273):595–601.

[Härting et al., 2016] Härting, S., Marciniak-czochra, A., and Takagi, I. (2016). Stable patterns with jump discontinuity in systems with turing instability and hysteresis. *Disc. Cont. Dyn. Syst. A. Preprint.*

[Hasty et al., 2001] Hasty, J., McMillen, D., Isaacs, F., and Collins, J. J. (2001). Computational studies of gene regulatory networks: in numero molecular biology. *Nature Reviews. Genetics*, 2(4):268–279.

[Holloway et al., 2011] Holloway, D. M., Lopes, F. J. P., da Fontoura Costa, L., Travençolo, B. a. N., Golyandina, N., Usevich, K., and Spirov, A. V. (2011). Gene expression noise in spatial patterning: hunchback promoter structure affects noise amplitude and distribution in Drosophila segmentation. *PLoS Computational Biology*, 7(2):e1001069.

[Hong and Hamaguchi, 2003] Hong, R. L. and Hamaguchi, L. (2003). Regulatory elements of the floral homeotic gene AGAMOUS identified by phylogenetic footprinting and shadowing. *The Plant Cell*, 15(June):1296–1309.

[Immink et al., 2009] Immink, R. G. H., Tonaco, I. a. N., de Folter, S., Shchennikova, A. V., van Dijk, A. D. J., Busscher-Lange, J., Borst, J. W., and Angenent, G. C. (2009). SEPALLATA3: the 'glue' for MADS box transcription factor complex formation. *Genome biology*, 10(2):R24.

[Ingolia, 2004] Ingolia, N. T. (2004). Topology and robustness in the Drosophila segment polarity network. *PLoS Biology*, 2(6):805–815.

[Ito et al., 2007] Ito, T., Ng, K.-H., Lim, T.-S., Yu, H., and Meyerowitz, E. M. (2007). The homeotic protein AGAMOUS controls late stamen development by regulating a jasmonate biosynthetic gene in Arabidopsis. *The Plant Cell*, 19(11):3516–29.

[Kauffman, 1969] Kauffman, S. (1969). Homeostasis and Differentiation in Random Genetic Control Networks. *Nature*, 224(5215):177–178.

[Kaufmann et al., 2010] Kaufmann, K., Wellmer, F., Muiño, J. M., Ferrier, T., Wuest, S. E., Kumar, V., Serrano-Mislata, A., Madueño, F., Krajewski, P., Meyerowitz, E. M., Angenent, G. C., and Riechmann, J. L. (2010). Orchestration of floral initiation by APETALA1. *Science*, 328(5974):85–9.

[Köthe and Marciniak-Czochra, 2013] Köthe, A. and Marciniak-Czochra, A. (2013). Multistability and hysteresis-based mechanism of pattern formation in biology. *Pattern Formation in Morphogenesis*, 15:153–173.

[Krizek and Fletcher, 2005] Krizek, B. a. and Fletcher, J. C. (2005). Molecular mechanisms of flower development: an armchair guide. *Nature*, 6(9):688–98.

[Krogan et al., 2012] Krogan, N. T., Hogan, K., and Long, J. a. (2012). APETALA2 negatively regulates multiple floral organ identity genes in Arabidopsis by recruiting the corepressor TOPLESS and the histone deacetylase HDA19. *Development*, 139(22):4180–90.

[Kunst et al., 1989] Kunst, L., Klenz, J. E., Martinez-Zapater, J., and Haughn, G. W. (1989). AP2 Gene Determines the ldentity of Perianth Organs in FIowers of Arabidopsis thaliana. *The Plant cell*, 1(December).

[Landrein et al., 2015] Landrein, B., Refahi, Y., Besnard, F., Hervieux, N., Mirabet, V., Boudaoud, A., Vernoux, T., and Hamant, O. (2015). Meristem size contributes to the robustness of phyllotaxis in Arabidopsis. *Journal of Experimental Botany*, 66(5):1317–1324.

[Lenhard et al., 2001] 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(6):805–814.

[Li et al., 2007] Li, B., Carey, M., and Workman, J. L. (2007). The Role of Chromatin during Transcription. *Cell*, 128(4):707–719.

[Liljegren et al., 1999] 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. *The Plant Cell*, 11(6):1007–1018.

[Liu and Meyerowitz, 1995] Liu, Z. and Meyerowitz, E. M. (1995). LEUNIG regulates AGAMOUS expression in Arabidopsis flowers. *Development*, 121(4):975–91.

[Lohmann et al., 2001] 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(6):793–803.

[Macneil and Walhout, 2011] Macneil, L. T. and Walhout, A. J. M. (2011). Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. pages 645–657.

[Mandel and Yanofsky, 1995] Mandel, M. A. and Yanofsky, M. F. (1995). A gene triggering flower formation in Arabidopsis. *Nature*, 377:522–524.

[Marciniak-Czochra, 2006] Marciniak-Czochra, A. (2006). Receptor-based models with hysteresis for pattern formation in hydra. *Mathematical Biosciences*, 199(1):97–119.

[Marciniak-Czochra et al., 2016] Marciniak-Czochra, A., Karch, G., and Suzuki, K. (2016). Instability of turing patterns in reaction-diffusion-ODE systems. *Journal of Mathematical Biology*, pages 1–36.

[Marciniak-Czochra et al., 2013] Marciniak-Czochra, A., Nakayama, M., and Takagi, I. (2013). Pattern formation in a diffusion-ODE model with hysteresis. *Diff. Integral Eq.*, (28):655–694.

[Meinhardt, 1982] Meinhardt, H. (1982). Models of biological pattern formation. *Academic Press, London*, pages 1–10.

[Mendoza and Alvarez-Buylla, 1998] Mendoza, L. and Alvarez-Buylla, E. R. (1998). Dynamics of the genetic regulatory network for Arabidopsis thaliana flower morphogenesis. *Journal of Theoretical Biology*, 193(2):307–319.

[Mendoza and Alvarez-Buylla, 2000] Mendoza, L. and Alvarez-Buylla, E. R. (2000). Genetic Regulation of Root Hair Development in Arabidopsis Thaliana: A Network Model. *Journal of Theoretical Biology*.

[Mendoza et al., 1999] Mendoza, L., Thieffry, D., and Alvarez-Buylla, E. R. (1999). Genetic control of flower morphogenesis in Arabidopsis thaliana: a logical analysis. *Bioinformatics*, 15(7-8):593–606.

[Mestek Boukhibar and Barkoulas, 2015] Mestek Boukhibar, L. and Barkoulas, M. (2015). The developmental genetics of biological robustness. *Annals of Botany*, page mcv128.

[Milani et al., 2014] Milani, P., Mirabet, V., Cellier, C., Rozier, F., Hamant, O., Das, P., and Boudaoud, A. (2014). Matching Patterns of Gene Expression to Mechanical Stiffness at Cell Resolution through Quantitative Tandem Epifluorescence and Nanoindentation. *Plant Physiology*, 165(4):1399–1408.

[Ming, 2009] Ming, Y. (2009). A model for floral organ development in Arabidopsis. pages 1–13.

[Monfared et al., 2013] Monfared, M. M., Carles, C. C., Rossignol, P., Pires, H. R., and Fletcher, J. C. (2013). The ULT1 and ULT2 trxG genes play overlapping roles in arabidopsis development and gene regulation. *Molecular Plant*, 6(5):1564–1579.

[Moore, 2015] Moore, D. S. (2015). *The Developing Genome: An Introduction to Behavioral Epigenetics*. Oxford uni edition.

[Moyroud et al., 2011] 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. *The Plant Cell*, 23(4):1293–306.

[Murray, 2002] Murray, J. D. (2002). *Mathematical Biology I. An introduction*, volume 17.

[Murray, 2003] Murray, J. D. (2003). *Mathematical Biology II*.

[Ng and Yanofsky, 2001] Ng, M. and Yanofsky, M. F. (2001). Function and evolution of the plant MADS-box gene family. *Nature reviews. Genetics*, 2(3):186–195.

[Nurrish and Treisman, 1995] Nurrish, S. J. and Treisman, R. (1995). DNA binding specificity determinants in MADS-box transcription factors. *Molecular and Cellular Biology*, 15(8):4076–4085.

[O'Maoileidigh et al., 2014] O'Maoileidigh, D., Graciet, E., and Wellmer, F. (2014). Gene network controlling Arabidopsis thaliana flower development. *Tansley review*.

[ÓMaoiléidigh et al., 2013] ÓMaoiléidigh, D. S., Wuest, S. E., Rae, L., Raganelli, A., Ryan, P. T., Kwasniewska, K., Das, P., Lohan, A. J., Loftus, B., Graciet, E., and Wellmer, F. (2013). Control of reproductive floral organ identity specification in Arabidopsis by the C function regulator AGAMOUS. *The Plant cell*, 25(7):2482–503.

[Parcy et al., 1998] Parcy, F., Nilsson, O., Busch, M. A., Lee, I., and Weigel, D. (1998). A genetic framework for floral patterning. *Nature*, 395(6702):561–6.

[Pastore et al., 2011] Pastore, J. J., Limpuangthip, A., Yamaguchi, N., Wu, M.-F., Sang, Y., Han, S.-K., Malaspina, L., Chavdaroff, N., Yamaguchi, A., and Wagner, D. (2011). LATE MERISTEM IDENTITY2 acts together with LEAFY to activate APETALA1. *Development*, 138(15):3189–98.

[Pollock and Treisman, 1991] Pollock, R. and Treisman, R. (1991). Human SRF-related proteins: DNA-binding properties and potential regulatory targets. *Genes & Development*, 5(12 A):2327–2341.

[Prusinkiewicz and Runions, 2012] Prusinkiewicz, P. and Runions, A. (2012). Computational models of plant development and form. *New Phytologist*, 193(3):549–569.

[Puranik et al., 2014] Puranik, S., Acajjaoui, S., Conn, S., Costa, L., Conn, V., Vial, A., Marcellin, R., Melzer, R., Brown, E., Hart, D., Theissen, G., Silva, C. S., Parcy, F., Dumas, R., Nanao, M., and Zubieta, C. (2014). Structural Basis for the Oligomerization of the MADS Domain Transcription Factor SEPALLATA3 in Arabidopsis. *The Plant Cell*, 26(9):3603–3615.

[Raspopovic et al., 2014] Raspopovic, J., Marcon, L., Russo, L., and Sharpe, J. (2014). Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science*, 566:10–15.

[Reddy et al., 2004] Reddy, G. V., Heisler, M. G., Ehrhardt, D. W., and Meyerowitz, E. M. (2004). Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of Arabidopsis thaliana. *Development*, 131(17):4225–4237.

[Riechmann et al., 1996a] Riechmann, J. L., Krizek, B. A., and Meyerowitz, E. M. (1996a). Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. *Proceedings of the National Academy of Sciences of the United States of America*, 93(May):4793–4798.

[Riechmann et al., 1996b] Riechmann, J. L., Wang, M., and Meyerowitz, E. M. (1996b). DNA-binding properties of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. *Nucleic Acids Research*, 24(16):3134–3141.

[Roeder et al., 2012] Roeder, a. H. K., Cunha, a., Burl, M. C., and Meyerowitz, E. M. (2012). A computational image analysis glossary for biologists. *Development*, 139(17):3071–3080.

[Rozier et al., 2014] Rozier, F., Mirabet, V., Vernoux, T., and Das, P. (2014). Analysis of 3D gene expression patterns in plants using whole-mount RNA in situ hybridization. *Nature Protocols*, 9(10):2464–2475.

[Saleh et al., 2007] Saleh, A., Al-Abdallat, A., Ndamukong, I., Alvarez-Venegas, R., and Avramova, Z. (2007). The Arabidopsis homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish 'bivalent chromatin marks' at the silent AGAMOUS locus. *Nucleic Acids Research*, 35(18):6290–6296.

[Sánchez-Corrales et al., 2010] Sánchez-Corrales, Y.-E. E., Alvarez-Buylla, E. R., and Mendoza, L. (2010). The Arabidopsis thaliana flower organ specification gene regulatory network determines a robust differentiation process. *J Theor Biol*, 264(3):971–983.

[Savageau, 1979] Savageau, M. a. (1979). Growth of complex systems can be related to the properties of their underlying determinants. *Proceedings of the National Academy of Sciences of the United States of America*, 76(11):5413–7.

[Sayou et al., 2014] Sayou, C., Monniaux, M., Nanao, M. H., Moyroud, E., Brockington, S. F., Thévenon, E., Chahtane, H., Warthmann, N., Melkonian, M., Zhang, Y., Wong, G. K.-S., Weigel, D., Parcy, F., and Dumas, R. (2014). A promiscuous intermediate underlies the evolution of LEAFY DNA binding specificity. *Science*, 343(6171):645–8.

[Sayou et al., 2016] Sayou, C., Nanao, M. H., Jamin, M., Posé, D., Thévenon, E., Grégoire, L., Tichtinsky, G., Denay, G., Ott, F., Peirat Llobet, M., Schmid, M., Dumas, R., and Parcy, F. (2016). A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nature Communications*.

[Sessions et al., 1999] Sessions, A., Yanofsky, M. F., and Weigel, D. (1999). Cell-Cell Signaling and Movement by the Floral Transcription Factors LEAFY and APETALA1. *Science*, 1001:12–15.

[Shimomura et al., 1962] Shimomura, O., Johnson, F., and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol*, 59:223–39.

[Shore and Sharrocks, 1995] Shore, P. and Sharrocks, a. D. (1995). The MADS-box family of transcription factors. *European Journal of Biochemistry*, 229(1):1–13.

[Sieburth and Meyerowitz, 1997] Sieburth, L. E. and Meyerowitz, E. M. (1997). Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. *The Plant cell*, 9(3):355–365.

[Silva et al., 2016] Silva, C. S., Puranik, S., Round, A., Brennich, M., Jourdain, A., Parcy, F., Hugouvieux, V., and Zubieta, C. (2016). Evolution of the Plant Reproduction Master Regulators LFY and the MADS Transcription Factors: The Role of Protein Structure in the Evolutionary Development of the Flower. *Frontiers in Plant Science*, 6(January):1–18.

[Simon et al., 2014] Simon, M. L. A., Platre, M. P., Assil, S., Van Wijk, R., Chen, W. Y., Chory, J., Dreux, M., Munnik, T., and Jaillais, Y. (2014). A multi-colour/multi-affinity marker set to visualize phosphoinositide dynamics in Arabidopsis. *The Plant Journal*, 77(2):322–337.

[Simon et al., 2016] Simon, M. L. A., Platre, M. P., Marquès-Bueno, M. M., Armengot, L., Stanislas, T., Bayle, V., Caillaud, M.-C., and Jaillais, Y. (2016). A PtdIns(4)P-driven electrostatic field controls cell membrane identity and signalling in plants. *Nature Plants*, 2(June):16089.

[Smaczniak et al., 2012] Smaczniak, C., Immink, R. G. H., Angenent, G. C., and Kaufmann, K. (2012). Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development*, 139(17):3081–3098.

[Smith et al., 2006] Smith, R. S., Guyomarc'h, S., Mandel, T., Reinhardt, D., Kuhlemeier, C., and Prusinkiewicz, P. (2006). A plausible model of phyllotaxis. *Proceedings of the National Academy of Sciences of the United States of America*, 103(5):1301–1306.

[Smyth et al., 1990] Smyth, D. R., Bowman, J. L., and Meyerowitz, E. M. (1990). Early flower development in Arabidopsis. *The Plant Cell*, 2(8):755–767.

[Soltis et al., 2007] Soltis, D. E., Chanderbali, A. S., Kim, S., Buzgo, M., and Soltis, P. S. (2007). The ABC model and its applicability to basal angiosperms. *Annals of Botany*, 100(2):155–163.

[Sridhar et al., 2004] Sridhar, V. V., Surendrarao, A., Gonzalez, D., Conlan, R. S., and Liu, Z. (2004). Transcriptional repression of target genes by LEUNIG and SEUSS, two interacting regulatory proteins for Arabidopsis flower development. *Proceedings of the National Academy of Sciences of the United States of America*, 101(31):11494–11499.

[Tomlin and Axelrod, 2007] Tomlin, C. J. and Axelrod, J. D. (2007). Biology by numbers: mathematical modelling in developmental biology. 8(May):331–340.

[Torii, 2012] Torii, K. U. (2012). Two-dimensional spatial patterning in developmental systems. *Trends in Cell Biology*, 22(8):438–446.

[Traas and Vernoux, 2002] Traas, J. and Vernoux, T. (2002). The shoot apical meristem: the dynamics of a stable structure. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 357(1422):737–47.

[Turing, 1952] Turing, A. M. (1952). The Chemical Basis of Morphogenesis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 237(641):37–72.

[Urbanus et al., 2009] Urbanus, S. L., de Folter, S., Shchennikova, A. V., Kaufmann, K., Immink, R. G. H., and Angenent, G. C. (2009). In planta localisation patterns of MADS domain proteins during floral development in Arabidopsis thaliana. *BMC Plant Biology*, 9:5.

[Urbanus et al., 2010] Urbanus, S. L., Martinelli, A. P., Dinh, Q. D. P., Aizza, L. C. B., Dornelas, M. C., Angenent, G. C., and Immink, R. G. H. (2010). Intercellular transport of epidermis-expressed MADS domain transcription factors and their effect on plant morphology and floral transition. *The Plant Journal*, 63(1):60–72.

[Uzkudun et al., 2015] Uzkudun, M., Marcon, L., and Sharpe, J. (2015). Data-driven modelling of a gene regulatory network for cell fate decisions in the growing limb bud. *Molecular Systems Biology*, 11(7):815.

[von Dassow et al., 2000] von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M. (2000). The segment polarity network is a robust developmental module. *Nature*, 406(6792):188–192.

[Von Dassow and Odell, 2002] Von Dassow, G. and Odell, G. M. (2002). Design and constraints of the Drosophila segment polarity module: Robust spatial patterning emerges from intertwined cell state switches. *Journal of Experimental Zoology*, 294(3):179–215.

[Waddington, 1942] Waddington, C. (1942). Canalization of develoment and the inheritance of acquired characters. *Nature*, 3811.

[Weigel et al., 1992] Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., and Meyerowitz, E. M. (1992). LEAFY controls floral meristem identity in Arabidopsis. *Cell*, 69(5):843–859.

[Weigel and Meyerowitz, 1993] Weigel, D. and Meyerowitz, E. M. (1993). Activation of Floral Homeotic Genes in Arabidopsis. *Science*, 361.

[Weigel and Nilsson, 1995] Weigel, D. and Nilsson, O. (1995). A developmental switch sufficient for flower initiation in diverse plants.

[Wellmer et al., 2014] Wellmer, F., Graciet, E., and Riechmann, J. L. (2014). Specification of floral organs in Arabidopsis. *Journal of Experimental Botany*, 65(1):1–9.

[Wernet et al., 2006] Wernet, M. F., Mazzoni, E. O., Celik, A., Duncan, D. M., Duncan, I., and Desplan, C. (2006). Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature*, 440(7081):174–80.

[Wollmann et al., 2010] Wollmann, H., Mica, E., Todesco, M., Long, J. a., and Weigel, D. (2010). On reconciling the interactions between APETALA2, miR172 and AGAMOUS with the ABC model of flower development. *Development*, 137(21):3633–42.

[Wu et al., 2003] Wu, X., Dinneny, J. R., Crawford, K. M., Rhee, Y., Citovsky, V., Zambryski, P. C., and Weigel, D. (2003). Modes of intercellular transcription factor movement in the Arabidopsis apex. *Development*, 130(16):3735–3745.

[Yadav et al., 2011] Yadav, R. K., Perales, M., Gruel, J., Girke, T., Jönsson, H., and Reddy, G. V. (2011). WUSCHEL protein movement mediates stem cell homeostasis in the Arabidopsis shoot apex. *Genes & Development*, 25:2025–2030.

[Yanofsky et al., 1990] Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A., and Meyerowitz, E. M. (1990). The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. *Nature*.

[Yant et al., 2010] 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. *The Plant cell*, 22(7):2156–70.

[Zahn et al., 2006] Zahn, L. M., Feng, B., and Ma, H. (2006). Beyond the ABC-Model: Regulation of Floral Homeotic Genes. *Advances in Botanical Research*, 44(06):163–207.

List of Figures

1.1 **Flower architecture and the ABCE model: (a.)** An *Arabidopsis thaliana* flower with the following four organ types labelled: sepals, petals, stamens and carpels. **(b.)** Floral diagram showing the numbers and relative positions of floral organs. **(c.-m.)** Summary of the ABCE model showing WT and mutant flowers **(c.-g.)**, floral diagrams **(h.-l.)** and the relative expression domains of patterning genes (m.). **(c. & h.)** The ABCE model in the WT, with a WT flower (c) its architecture and the description of the ABCE model. Superposition of A and E gives sepal identity; A, B and E gives petal identity B, C and E gives stamen identity; and C and E gives carpel identity. **(d. & i.)** A class A mutant, where sepals are transformed into carpels and petals into stamens. **(e. & j.)** A class B mutant, with petals transformed into sepals and stamens into carpels. **(f. & k.)** A class C mutant, with stamens are transformed into petals and a new flower emerges from the center. **(g. & l.)** A class E mutant, where all organs of the flower are replaced by leaf-like structures. **(m.)** Expression patterns of the ABC genes in young flowers. The numbers 1 to 4 indicate the stage of flower development. The A class genes are in red, the B in yellow and the C in blue. The fourth panel provides the superposed image in the four whorls of the flower. **(n.)** The quartet model that leads to the identity of organs in the flower. Two molecules of AP1 and two of SEP act to give sepals; one molecule each of AP1, PI, AP3 and SEP give petals; one each of AG, PI, AP3 and SEP give stamens and two AG and two SEP give carpels. Panel a was adapted fromwww.weigelworld.org/, panel b was produced by P. Das, panel c - l are adapted from [Krizek and Fletcher, 2005], panel m from [Alvarez-Buylla et al., 2010] and panel n from [Wellmer et al., 2014]. 5

1.2 **Role of MADS-box proteins in the *Arabidopsis thaliana* development** Adapted from [Smaczniak et al., 2012]. 8

1.3 **Type II MADS-box protein** The type II MADS-box proteins have an N-terminal MADS DNA binding domain that resembles the animal myocyte enhancer factor 2 (MEF2), followed by a weakly conserved intervening region (I) and a Keratin-like coiled-coil domain (K) that are both involved in the dimerisation and the tetramerisation of these proteins. These are followed by a C-terminal region that is important for the transactivation process and to stabilize interactions with other proteins. [Ng and Yanofsky, 2001] 9

1.4 **Regulation by and of *AGAMOUS***: **(A & B.)** Activation of transcription factors by AG during flower development. Expression levels of selected genes by an inducible *35S::AG-GR* construct, where a rat glucocorticoid receptor (GR) moiety is fused in frame to the AG protein, rendering it inactive in the absence of dexamethasone (DEX). **(A.)** Expression detected by oligonucleotide arrays. M1 to M7 and D1 to D7 indicate the number of days after mock or DEX treatment respectively (1, 3 and 7 days). Genes listed in the grey box are targets, including AG itself, that show sustained activation. **(B.)** Validation of the 12 selected targets by RT-PCR. *APT1* is a constitutive control. **(C.)** Principal regulators of *AG* during early and late flower development. Panels A and B are adapted from [Gómez-Mena et al., 2005], and panel C is adapted from [Zahn et al., 2006]. 10

1.5 **Regulatory sequences at the *AG* locus:** Schematic of the *Arabidopsis thaliana AG* 5.68-kb coding strand (top row). Empty rectangles are the 5' and 3' UTR regions, black rectangles are exons, solid lines are introns. Zoomed view of the 3-kb second intron where many key regulatory sequences are to be found (lower panel). The coloured bars represent regions that are thought to contain redundant regulatory elements[Busch et al., 1999]. Evolutionarily conserved binding sites for LFY, WUS and MADS family protein are represented as solid or empty shapes and are labelled [Deyholos and Sieburth, 2000, Hong and Hamaguchi, 2003, Moyroud et al., 2011] and AP2 binding sites based on Yant and Dinh [Yant et al., 2010, Dinh et al., 2012] 11

1.6 **Epigenetic regulation of *AG*: (1)** PcG factors, such as CLF, are evicted from DNA to allow a PHD-domain containing protein to bind unmethylated H3K4 chromatin. PHD proteins interact with ULT1 to recruit **(2)** co-activators and transcription complexes. ULT1 also interacts with the trxG complex containing ATX1 that deposits active histone marks, such as H3K4me3. **(3)** ORC1 directly bind these marks to increase transcriptional activity. This figure was adapted from [Carles and Fletcher, 2010]. 14

1.7 **Structural properties of SEP3: (1.)** FRET-FLIM experiment in protoplasts shows the capacity of SEP3 to form dimers. **(a.)** Control with pECFP and PI-YFP. **(b.)** SEP3-CFP with SEP3-YFP showing that SEP3 is able to form stable homodimers [Immink et al., 2009]. **(2.)** Sequence of residues 75-178 of SEP3 showing its homology with other MADS box proteins. The I domain is indicated in yellow and the K domain in blue. The various structural motifs are indicated above: The two main alpha helices are represented based on the crystal structure of the SEP3 protein. Residues involved in dimerisation and tetramerisation are highlighted in light blue and light green respectively. **(3.)** Tetramer of $SEP3^{75-178}$. The two helices are at the base of the two dimers in blue and green and the tretramer formation. **(4.)** Atomic Force Microscopy of SEP3 tetramers bound to DNA. Full length SEP3 protein forms complexes with 1-kb of a target promoter DNA at 10 to 15nM protein and 5nM DNA before dilution to 1nM DNA for imaging. Arrows indicate DNA looping due to SEP3 interactions. Scale bars indicate 200nm for the right and 100nm for the left. [Puranik et al., 2014] 18

1.8 **Turing models** : **(a.)** a Turing model is composed of at least two components, an inhibitor I that diffuse faster than an activator A. **(b.)** Depending of the parameters, obtained patterns can be different. In these simulations we can observe stripes like spots described by the same system of equations. **(c.)** The model can be described by two reaction-diffusion equations that summarize interactions between the components and its capacity to move. [Torii, 2012] 20

1.9 **Model of wnt concentration in the *hydra*: (A.)** Kinetic functions of a model (Equation 1.2) with hysteresis. **(B.)** Representation of the experiment performed on hydra that show that organ definition depend on the initial concentration of a morphogen. **(C.)** Simulation of the morphogen concentration that reproduce the experiments. [Torii, 2012] 24

1.10 **Model of the drosophila segment polarity network: (a.)** Expression of the segment polarity genes *wingless* (*wg*) in green and *Engrailed* (*EN*) in red. **(b.)** Segment polarity that result on models of von Dassow et al.. The model reproduce segments by modelling the interactions between genes of the network c. The proteins are in upper-case and the mRNA in lower-case. The anterior (A) posterior (P) axis is represented by the arrow. **(c.)** The network used in the model of segment polarity. [Tomlin and Axelrod, 2007] 25

1.11 **Model of the limb development: (A.)** Study of interactions necessary to be able to reproduce the pattern formation that lead to digits development. **(B.)** Once the model was selected, protein have been find to correspond of the components of the model. **(C.)** Simulation of the *sox9* protein expression in the limb. [Raspopovic et al., 2014] 26

1.12 **Floral organ specification gene regulatory network (FOS-GRN** (reproduced from [Davila-Velderrain et al., 2015]) based mainly on mRNA expression patterns. It controls the early differentiation of inflorescences and floral organs in *Arabidopsis Thaliana*. Positive and negative regulatory interactions are represented by arrows and perpendicular lines respectively. This figure is adapted from [Espinosa-Soto et al., 2004]. 28

1.13 **Determination of AG expression patterns by different approaches**: **(A.)** *in situ* hybridisation of AG showing the localisation of AG mRNA in the central dome at stage 3 flowers . **(B.)** Projection of a confocal image of the translational fusion of AG with a Green Fluorescent Protein (GFP) reporter, counterstained with FM4-64 to mark plasma membranes showing a protein pattern similar to the mRNA. Panel A is adapted from [Gómez-Mena et al., 2005] and panel B from [Urbanus et al., 2009] 31

1.14 **3D reconstruction of a stage 2 flower**. Projection of a stage 2 flower with a membrane marker (on the left). After imaging from multiple angles and the fusion of those image stacks, the resultant high-resolution image is segmented to provide an image where each cell is assigned a number (on the right). This data can be used to determine cellular metrics such as volume, anisotropy, curvature etc. 33

1.15 **Time course of a flower during early stages of development:** Time course of a single flower stained with FM4-64 and image at multiple time points from multiple angles. The images for each time point are fused and segmented and the lineage is then determined across the entire time course. The colors represent these lineages. The time interval between images is indicated below the projections. 34

1.16 **Definition of the temporal graph:** The topological graph summarizes the geometrical properties of the cells of a single image. It can include all available data for each cell, such as protein signal or other properties. Using the lineage, a spatio-temporal graph is created to link properties of mother cells to daughter cells. 34

2.1 **The ABC model: (A.)** *Arabidopsis thaliana* flower that show the 4 kinds of organs: sepals, petals, carpels and stamens. **(B.)** Representation of the 4 whorls of flower organs. **(C.)** representation of the dynamics of expression of the A, B and C-class genes during the 3 first stages of flower development. The superposition of these 3 class of genes gives the 4 whorls of different gene expression that induce the organ identity 47

2.2 **A simple model to understand *AG* dynamic of expression:** The scheme summarize the known regulations of gene transcription (positive or negative) by direct or undirect binding of the protein to the DNA. The red proteins are known to be able to move between cells. 48

2.3 **Reduction of the spatial domain** 53

2.4 **Visualization of the results:** Simulations are represents by an heat map that correspond to protein concentrations in function of the time and the space variables. 56

2.5 **Expression pattern of the AC model** based on literature and recent experiments 57

2.6 **Simulation in the WT condition:** The graph represents the level of expression of the 5 components of the AC model 57

2.7 **Description of the phases of *AG* activation:** The scheme represent the 3 phases of *AG* activation. *AG* expression level is represented during the time of the 3 first stages of flower development and is also represented with the same color ode than in our simulations. Correspondence with experiments are present below the scheme. 58

2.8 **Nullclines and steady states of the AP1/LFY system:** the three graphs represent the nullclines of the simplified system 2.13 with different values of AG. The stability of the steady states has been calculated numerically with pplane8 module of matlab 59

2.9 **Simulation of the auto-activation with a single protein:** we use $n_6 = 1$ and we rescale $a_6 = 0.78$ to obtain a similar final pattern of expression. We see that *AG* activation appears earlier than expected and in the full domain before *AP2* expression. 60

2.10 **Simulation of the *ap2* mutant:** The high expression of *AG* start earlier and is wider than in the WT 62

2.11 **Simulation of the LFY/WUS weak interaction: (A.)** Simulation with $a_7 = 0.48$ instead of 0.5 that reproduce a defect on the binding between LFY, WUS and the 2nd intron of *AG*. We see that small variation in this parameters induce strong changes in initiation of the pattern. **(B.)** Simulation with $a_7 = 0.52$ instead of 0.5. Activation of AG is earlier than in the WT. The effect of this parameter is strong on the time of AG activation. 63

2.12 **Simulation when AG is not able to diffuse:** Simulation with $D_3 = 0$ reproducing the incapacity of AG to move from one cell to the other. We can observe that phase 2 start earlier, that the final domain of expression is less wide and that there is a sharp border of the AG domain of expression. . . . 65

2.13 **Simulation with a peak of AG after AP2 activation:** Simulation of a short peak of *AG* expression starting at $t = 0.04T$ in a cell closed to the center. We observe that first the protein diffuses around this cell and then the phase 2 of AG activation starts earlier in the center 66

2.14 **Simulation with a peak of AG when LFY and WUS are not able to activate AG:** Simulation of a short peak ($p = 5$) of AG expression starting at $t = 0.04T$ in a cell closed to the center in the case where $a_7 = 0$. We observe that in the absence of the activation of AG by LFY and WUS, we can still restore the proper expression of AG with an artificial peak. 66

3.1 **AG expression in early *Arabidopsis* flowers: (A.)** Whole mount mRNA *in-situ* localizations of *AG* mRNA in a shoot apical meristem (SAM). Black arrow indicates a stage 2 flower where AG staining is already visible. **(B.)** SAM from a plant of the L-*er* accession stained with the FM4-64 lipophilic plasma membrane dye and containing an *AG* reporter (*AG-2xVenus*). F1 and F2 are stage 2 flowers with a low stochastic expression of *AG* (white arrows). F3 is a young stage 3 flower with few cells expressing *AG*. F4 and F5 are stage 3 flowers with the expression of *AG* in all the central dome. **(C.)** A mixture model of the L1 cells of the flower F5. Cells are separated into two clusters, the first with cells that don't express *AG* and the second one with cells expressing *AG*. The two clusters can be described by a Gaussian (blue and red lines). **(D.)** A mixture model of all the cells of the L1 of the five flowers (F1 to F5). Distribution of *AG* expression in cells can be separated into 4 clusters. The first one with cells that doesn't express *AG* and the other one with the different level of *AG* expression. **(E. & F.)** Quantification of AG proteins at the cellular level of the L1 of F1, F2, F3 and F5 flowers. E. is the mean fluorescent level in each cell and F. is the representation of the 4 clusters identified in D. Few cells are detected in the two first flowers. For the F5 we can detect a ring of cells at the periphery of the central dome with a lower level of AG than in the central dome. Panel A is adapted from [Rozier et al., 2014]. 76

3.2 ***AG* expression dynamics in the ps143 time course: (A.)** A MARS-treated AG-2xVenus floral time course (ps143), with segmented flowers of the first, fifth, seventh and ninth time points overlaid with average signal concentration per cell (total fluorescence divided by cell volume) represented in a colour scale. *AG* expression begins in a single cell at the first time point then spreads towards the periphery of the central dome. *AG* is fully expressed at stage 3 but with a lower level in the few cells in the centre. **(B.)** The evolution of signal concentration per lineage, showing that the onset of expression is broadly distributed over time. **(C.)** Signal concentration curves for each lineage, but centred on their activation inflection points, showing similar expression profiles for all lineages. **(D.)** A heat map of signal concentration over time based on the distance of each cell from the centre of the flower. **(E.)** A spatialized view of the absolute time of occurrence of the activation inflection point in each lineage, represented in a colour scale on the last time point. The colour is set to blue when an inflection point can not be defined in a given cell, either because signal was detected at the first time point (black arrow) or because a very late activation meant that those curves could be fitted with our methods (white arrows). Activation spreads to adjacent cells from early-activated lineages. 78

3.3 **Simulation of *AG* dynamics of expression: (A.)** Schematic representation of the principal genetic interactions used to model *AG* regulation, as described in chapter 2. **(B.)** Simulation of AG concentration during the early flower development, with the relative concentration of AG represented as a heat map (where red indicates high expression and blue indicates no expression). The X-axis represents a radius on the flower and the Y-axis represents developmental time. AG appears rapidly in the equivalent of the two inner whorls after a delay. **(C.)** AG concentration in homogeneously distributed X-axes points of the simulation in function of the time. Each curves correspond to a fixed point of the domain of study. A sudden increase in expression is clearly visible in the points that correspond to the central dome. **(D.)** Schematic illustrating how *AG* expression dynamics differ when AG acts as a monomer or a homodimer (or higher-order complex) to autoregulate its own expression (resp. n=1 or n=2). 82

S3.1 ***AG* expression in early flowering: (A.)** Shoot apical meristem with flowers stained with a red membrane marker (FM4-64) and containing an *AG* fluorescent reporter (*AG-2xVenus*). F1 to F5 correspond to the studied flowers. **(B. C. E. F. & G.)** Mixture model of the mean level of fluorescence of the *AG* reporter in the L1 cells of respectively the flower F1, F2, F3, F4 and F5. Two clusters are identified, the first one with only noise and the second one with significant concentration of the protein. **(D. & H.)** Mixture model of the L1 cells of the five flowers. **(D.)** Two clusters are identified (one with only noise and the second one with fluorescent cells). **(H.)** 4 clusters are identified, the fluorescent cells are separated in three clusters depending on their level. **(I. & J.)** Quantification of AG protein at a cellular level inside the L1 cells of the 5 flowers. **(I.)** is the mean fluorescent level in every cell and **(J.)** is the representation of the 4 clusters identified in H. 89

S3.2 **Fit of *AG* profile of expression for ps143:** The few horizontals lines correspond to profile where the fitting methods cannot find a good solution. 90

S3.3 ***AG* expression in ps143 time course:** Quantification of fluorescent reporter of the ps143 signal. 90

S3.4 **Cluster analysis of *AG* expression in ps143 time course:** the four clusters are obtained with mixture model methods of the L1 cells of all the flowers of ps143. 91

S3.5 ***AG* expression analysis of the ps144 time course: (A.)** mean quantification of fluorescence of *AG-2xVenus* reporter. **(B. & C.)** Mixture models with four clusters on L1 cells of all flowers. **(D.)** Evolution of mean level of fluorescence during time in each lineages. **(E.)** Evolution of fluorescence centred on the activation time of each lineages. **(F.)** Heat map of fluorescence level. **(G.)** Representation of activation times of each lineage on the last image of the time course. Histogram of activation time of each lineages. 93

S3.6 **Simulation of the AC model**: Simulation of the concentration of the five components of the AC model. 93

4.1 **Effect of supernumerary copies of *AG* on its own expression :** Confocal image of an inflorescence meristem from a line of the genotype *AG/AG AG-2xVenus/AG-2xVenus pAG-AG/+*, with Venus shown in green and autofluorescence shown in red. This WT *Col-0* plant is homozygous for the *AG-2xVenus* reporter and also bears an additional copy of the *AG* locus. No *AG* expression is observed in stage 2 flowers (F1). The onset of *AG* expression occurs very suddenly at stage 3 (F2). A large proportion of AG is visible in the cytoplasm of individual cells, and the centre of the floral dome displays less fluorescence (F3). 100

4.2 **The effect of LFY binding site mutations on *AG* expression dynamics: (A.)** LFY binding sites at the *AG* locus. Schematic of the *AG* locus with exons (light blue boxes) and introns (black lines) represented. The middle panel contains ChIP-seq read coverage combined from both strands. The bottom part shows the scores of binding sites and the presence of the consensus binding sites (arrows). **(B.)** Schematic of the *AG* locus showing the LFY binding sites that were mutated in different combinations (red crosses). These fragments were then fused to *2x-Venus* to produce WT (control) or mutant (SP15, SP17 and SP18) versions of the *AG* reporter. **(C.)** Observed fluorescence in floral organs during development. **(D–K.)** 2D projections of confocal image stacks from the different mutant constructs in a WT background. **(D–F.)** Projections from two independent SP15 lines, showing a WT pattern (ps148) or slightly delayed, yet still stochastic, onset (ps190). **(G–J.)** Projections from two SP18 lines, *AG* activation is delayed and less present in few cells in the centre. **(K.)** Projection of an SP17 line with no detectable signal in young flowers. **(L–N.)** Flowers of the same SP17 line in the strong *ag-salk* mutant. In L and M, two pistils are replaced by petals, in N pistils are not present but we observe height petals, two of them present pollen at their extremity. Carpels are slightly different than WT ones. **(O.)** Flower of strong *ag-salk* mutant. Panel A was adapted from Moyroud et al. [Moyroud et al., 2011]. 103

4.3 **The role of auto-activation in *AG* patterning: (A.)** WT *Arabidopsis* flower bearing 4 sepals, 4 petals, 6 stamens and 2 carpels. **(B.)** Flowers of *pPDF1::AG-CFP-N7* homozygous *ag-salk* mutants plants, with indeterminate carpels that are not functional petals that are transformed into stamens, and with sepals that present carpelloid structures at their tips. **(C. & D.)** Flowers of *pPDF1::AG-CFP-N7* in a WT *Col-0* background, with missing petals and sepals with carpelloid characteristics. **(E–I.)** Confocal imaging of AG-2xVenus reporter in plants containing *pPDF1::AG-CFP-N7*. **(F.)** In the youngest, mid-stage 2, flower with a visible signal, only one cell expresses *AG* (white arrow). **(H.)** At late stage 2, several cells expressing *AG* are visible. **(I.)** At stage 3, AG is present in almost all cells of the central dome. **(G.)** Confocal image of the CFP marker of the *pPDF1::AG-CFP-N7* line that confirms *AG* expression in the nuclei of the L1 cells of all the flowers and the meristem. 105

4.4 **AG can form stable homo-dimers:** FRET-FLIM experiments in leaf protoplasts from either control PI-CFP **(A)** or AG-CFP ; AG-YFP lines **(B)**. One protoplast is shown for each experiment. The left panel displays the intensity channel, the middle one, the fluorescence lifetime of the same nucleus in a false colour code and the right panel depicts histograms representing the distribution of fluorescence lifetime values. 106

4.5 **The role of movement in *AG* patterning:** 2D projections of confocal images of inflorescence meristems carrying the *AG-CFP-N7* marker in the *ag-salk* mutant background. AG is localised in the nucleus and cannot move from cell to cell. As in the WT, the onset of *AG* occurs around mid-stage 2. Unlike in the WT,some cells at early stage 3 do not express *AG* (arrowheads in F5), and the establishment of complete expression in certain cells is delayed to mid-stage 3 (F4). 107

4.6 **The role of AP2 in restricting *AG* expression: (A.)** Flower of the *ap2-7* strong mutant. **(B.)** WT flower. **(C.)** 2D projection of a confocal image stack of an *ap2-7* inflorescence meristem carrying an *AG-2xVenus* reporter. **(D.)** MARS-treated 3D reconstructions with quantified AG fluorescence in all L1 cells of the AG-2xVenus reporter in a WT plant. **(E.)** Simulation of AG concentration in the WT. **(F.)** Quantification of the numbered flowers shown in (B). **(G.)** Simulation of *AG* expression in a null *ap2* mutant; *AG* is activated slightly earlier, and more rapidly, than in the WT and becomes more widely expressed in the flower. 109

S4.1 ***AG* pattern of expression in plants with 6 copy of *AG*:** *AG-2xVenus* reporter in four independent lines. AG is first visible at a low level in two cells of inners layer pointed with white arrows at beginning of stage 3. 117

S4.2 **Auto-activation of *AG*:** Confocal imaging of AG-2xVenus in three meristem containing *pPDF1::AG-CFP-N7* 118

S4.3 **Auto-activation of *AG*:** Flowers of *Col-0* plants containing *pPDF1::amiRAG*. Petals appear at the place of pistils and carpels are not well developed and not functional. 119

S4.4 **Restriction of *AG* expression by AP2:** Three projections of meristem of *ap2-7* mutant containing *AG_T-2xVenus* reporter. 119

List of Tables

S3.1 **Used primers.** 88
S3.2 ***AG* expression in early flowering:** Expression level in the 5 flowers of Figure 3.1. Cells of the L1 layer has been separated in two clusters (Figure S3.1). The mean level of fluorescence in these cluster has been fit with a Gaussian. For each cluster of every flowers we obtain the average value of mean level of fluorescence in the cells, the standard deviation of the cluster and the number of cells in the cluster. Last line of the table correspond of the analysis of L1 cells of all flowers. 88

S4.1 **Used primers.** 120